**An artificial neural network (ANN) expert system enhanced with the
electromagnetism-based firefly algorithm (EFA) for predicting the energy
consumption in buildings**

Dac-Khuong Bui¹, Tuan Ngoc Nguyen¹, Tuan Duc Ngo^{1,*}, H. Nguyen-Xuan^{2,3*}

¹*Department of Infrastructure Engineering, the University of Melbourne, Parkville, VIC 3010,
Australia*

²*Center for Interdisciplinary Research in Technology, Ho Chi Minh City University of
Technology (Hutech), Ho Chi Minh City, Vietnam*

³*Department of Architectural Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-
gu, Seoul 05006, Republic of Korea*

Abstract

In this study, a new hybrid model, namely the Electromagnetism-based Firefly Algorithm - Artificial Neural Network (EFA-ANN), is proposed to forecast the energy consumption in buildings. The model is applied to evaluate the heating load (HL) and cooling load (CL) using two given datasets. Each dataset was obtained by monitoring the effect of the façade system and dimensions of the building, respectively, on energy consumption. The performance of EFA-ANN is validated by comparing the obtained results with other methods. It is shown that EFA-ANN provides a faster and more accurate prediction of HL and CL. A sensitivity analysis is conducted to identify the impact of each input on the energy performance of the building. From the results of this study, it is evident that EFA-ANN can assist civil engineers and construction managers in the early designs of energy-efficient buildings.

Keywords: Electromagnetism-based firefly algorithm; artificial neural network; machine learning; energy consumption.

* Corresponding author. Email address: dtngo@unimelb.edu.au (Tuan Duc Ngo) and ngx.hung@hutech.edu.vn (H. Nguyen-Xuan)

1. Introduction

There are three major economic sectors in the world, namely energy, including transportation, industry, and building [1]. A substantial share of global energy is consumed by buildings, which is expected to increase to 32.4% by 2040 [2]. In participating nations of the Organization for Economic Cooperation and Development (OECD), including Australia, New Zealand, United Kingdom and the United States, energy consumption in buildings has grown by 1.5% per year from 2012 to 2040. In non-OECD nations, being mostly developing countries, the growth rate is 2.1% per year in the same period [3]. Therefore, the development of energy efficient building systems is essential and consequently, many efforts have been devoted to this area [4, 5].

In the majority of cases, heating and cooling energy demands mostly account for building energy consumption [6]. Therefore, the early prediction and reduction of heating and cooling loads plays a vital role in designing an energy-efficient building. This provides designers with access to various building designs or Heating Ventilation and Air Conditioning (HAVC) system control options to find the optimal solution for reducing energy consumption in buildings. For instance, Zemella *et al.* optimized the design of façades of energy-efficient buildings by making early predictions of energy consumption due to heating, cooling, and lighting [7]. Also, Magnier and Haghighat applied a method to predict the energy consumption due to heating, cooling, and fan systems, and optimized the building design based on these predictions [8]. In another study, Ferreira *et al.* proposed a model-based predictive control methodology to control the HAVC system in a building and reported savings of around 50% in energy consumption [9]. Ghahramani *et al.* provided a systematic approach to optimize the setpoint and deadband parameter of the HVAC system by pre-calculating the energy

1 consumption of HVAC [10]. Therefore, the early prediction of heating and cooling loads
2
3
4 is critical to reducing the total energy consumption of a building.
5

6 However, many aspects, including temperature, sunlight equipment, occupant
7
8 behavior, wall materials, glazing area, surface, height and volume of the building, have
9
10 interactions between the overall energy requirements of the building [11-13]. For instance,
11
12 Ihara *et al.* concluded that all façade properties including solar reflectance, U-value, solar
13
14 heat gain coefficient (SHGC) have different effects on the energy efficiency of buildings.
15
16 They suggested that the reduction of SHGC is the most effective method for reducing
17
18 energy consumption [14]. In contrast, Liu *et al.* showed that reducing SHGC does not
19
20 improve the energy efficiency of a building without an appropriate U-value [15]. Many
21
22 other factors affect the total energy consumption in a building, including weather
23
24 conditions, building dimensions, and the behavior of occupants. Therefore, it is quite
25
26 challenging to calculate energy consumption given that all the above parameters and their
27
28 interactions should be considered.
29
30
31
32
33
34

35 There are three main categories in building energy assessment, including
36
37 engineering calculations, numerical simulations, and machine learning. The first
38
39 approach focuses on using physical laws to calculate the energy consumption of an entire
40
41 building. This approach is only suitable for preliminary analysis as it is mathematically
42
43 intensive. In this research, the numerical simulation approach was used to simulate the
44
45 energy performance of a building and overcome the limitations of engineering
46
47 calculations. Several building energy simulation programs, including Energy Plus, DOE-
48
49 2, Window, Autodesk Ecotect, TRNSYS, and eQUEST, were used to simulate and
50
51 predict energy consumption of a building. However, this method employs physics-based
52
53 simulations, which are often time-consuming and resource-intensive, and the complexity
54
55
56
57
58
59
60
61
62
63
64
65

1 and demand increase with the size and complexity of the project [16]. Also, the
2
3 optimization process must be manual based on user experience [17]. In many cases,
4
5 energy models cannot reflect the actual performance of a building in reality as they lack
6
7 the required details and need to make simplifications [18]. Thus, the conventional
8
9 procedures are often unsatisfactory for designing a façade because they are mainly based
10
11 on a particular design condition and the experience of experts. Accordingly, a significant
12
13 development of an advanced approach is required.
14
15
16

17
18 To this effect, machine learning (ML), was proposed to design energy-efficient
19
20 building services. This method applies a learning process to infer a relationship between
21
22 the building data and energy consumption of buildings. In recent years, several
23
24 advantages of ML have been demonstrated over conventional approaches [19, 20].
25
26 Consequently, many ML techniques have been applied to solve energy problems [21-23].
27
28 Robinson *et al.* successfully applied various ML techniques to predict the energy
29
30 consumption of a commercial building based on various building features [21]. Ahmad
31
32 *et al.* proposed four ML approaches to forecast short, medium and long-term energy
33
34 consumption in a building [22]. In another study, Chou and Tran used a hybrid ML model
35
36 to estimate the energy consumption of residential householders with high accuracy [23].
37
38 ML can also be used to support several adaptive systems in a building [24, 25]. For
39
40 example, Ghahramani *et al.* developed a novel adaptive hybrid metaheuristic algorithm
41
42 based on ML and other smart components to optimize the energy of an HVAC system
43
44 [24]. Moon *et al.* proposed an ML model for controlling the temperature of an adaptive
45
46 double skin envelope [25].
47
48
49
50
51
52
53

54
55 Recently, among many ML techniques, an artificial neural network (ANN) has
56
57 been widely used in optimizing the design of an energy-efficient building. Jin *et al.* used
58
59
60
61
62
63
64
65

1 an ANN-based thermal control logic model to optimize the initial conditions and heating
2
3 system operations in a building [26]. In another study, Wang *et al.* forecasted the dynamic
4
5 building cooling load by combining ANN and an ensemble model [27]. Chung *et al.*
6
7 proposed an ANN model to design a comfortable indoor thermal environment in an
8
9 energy-efficient manner [28].
10

11
12 Notwithstanding many advantages, the ANN model depends on several initial
13
14 parameters, including weights and biases [29]. Therefore, many studies were carried out
15
16 to improve the performance of ANN by combining it with optimization algorithms. For
17
18 example, Yam and Chow optimized the initial weights of ANN by using linear algebraic
19
20 methods [30]. Liu *et al.* enhanced the generalizations and accuracies of ANN by using an
21
22 ensemble method [31]. Also, Chang *et al.* [32] and Lee *et al.* [33] proposed a genetic
23
24 algorithm and a harmony search algorithm, respectively, to optimize the initial weights
25
26 of the ANN model.
27
28
29
30
31

32
33 Among many optimization algorithms, our approach is to propose an
34
35 Electromagnetism-based Firefly Algorithm (EFA), which is found to be an efficient
36
37 optimization tool that can solve complicated problems. For instance, Shammari *et al.*
38
39 combined the firefly algorithm (FA) and support vector machine model to predict the
40
41 heating load of a heating system [34]. Coelho and Mariani proposed FA and a Gaussian
42
43 distribution function to optimize the loading of a chiller for energy conservation [35].
44
45 Chu and Chang used the Electromagnetism Algorithm (EA) to solve the resource
46
47 allocation problem in stochastic networks [36]. However, both FA and EA still show
48
49 several disadvantages, including premature convergence and divergence [37-39], which
50
51 will be discussed in the following section. Therefore, in this study, we propose a new
52
53
54
55
56
57
58
59
60
61
62
63
64
65

approach based on a hybrid model of FA and EA, called EFA, which aims to enhance the capability of ANN for predicting the energy consumption of a building.

Section 2 and 3 describe the theory behind of EFA and ANN, respectively. Details of the proposed EFA-ANN are provided in Section 4. Section 5 introduces the two datasets used to train EFA-ANN. The proposed method is validated in Section 6. Concluding remarks and recommendations for future research are provided in Section 7.

2. Electromagnetism-based Firefly Algorithm – Artificial Neural Network Model

2.1 An electromagnetism-based firefly algorithm

The Electromagnetism-based Firefly Algorithm (EFA) is a new hybrid optimization algorithm that incorporates the advantages of the firefly algorithm (FA) and Electromagnetism Algorithm (EA). Firstly, FA, which was proposed by Yang [37], is a swarm intelligence method and is based on the flashing patterns and behavior of tropical fireflies. In FA, the brightness of a firefly is determined by an objective function. A firefly will tend to move closer to the brighter firefly and is not affected by a darker one. On the other hand, the EA was introduced by Birbil and Fang [39]. It imitates the attraction-repulsion mechanism to solve global optimization problems. In EA, all points converge to the highly attractive valleys and move further away from steeper hills. This idea is the same as the attraction-repulsion mechanism described by the theory of electromagnetism. However, better points are easily distracted if there are many worse points in the population. Also, EA is not effective in a local search because it randomly moves all points without any advanced technique.

In EFA, all fireflies are assumed to be magnetized, and the charge of a firefly depends on its objective value. Therefore, fireflies will move close to attractive fireflies and move away from the repulsive ones. Thus, all fireflies can contribute to the search

process and improve the performance of EFA. Based on their charges, all fireflies are ranked in each iteration, and a firefly only has attraction and repulsion by an attractive or repulsive neighboring firefly, respectively. In this way, fireflies are not distracted by many forces as in EA. Consequently, EFA can quickly find promising areas in the exploration phase and then perform local searches in these areas to find the optimal solution in the exploitation phase. Local searches in EFA are performed using Lévy flight inspired by FA.

2.1.1 Initialization

Firstly, EFA is initialized in the same manner as FA, when the coordinates of the initial population of fireflies are uniformly distributed between the corresponding lower and upper bounds as follows

$$x^o = lb_k + rand(0,1)(ub_k - lb_k) \quad (1)$$

where lb_k , ub_k are the lower and upper boundaries of the k^{th} coordinate, respectively; $rand(0,1)$ is a random number following a normal distribution within $[0,1]$. This study uses a logistic map to improve the performance of EFA. The logistic map generates the initial coordinates of fireflies using Eq. (2). It can provide an initial diverse population and also reduce the probability of premature occurrence [40]:

$$x_o^{\bar{t}+1} = ax_o^{\bar{t}}(1 - x_o^{\bar{t}}) \quad (2)$$

where $x_o^{\bar{t}}$ is a chaotic number at $\bar{t}$ iteration of the initialization process; a is fixed at 4 [41].

2.1.2 Local search

EFA inherits the exploitation capacity of FA and EA for searching conducting a local search in the vicinity of each coordinate. In other words, the local search of EFA

follows Eq. (3) and the improvement of each firefly is sought by coordinate. In this way, the probability of finding a better point is increased.

$$x_i^{t+1} = x_i^t + \alpha^t \varepsilon L(s) \quad (3)$$

where ε is a vector of random numbers and is defined by Eq. (4) as follows:

$$\varepsilon = rand - 1/2 \quad (4)$$

where $rand$ is a random number generated by a uniform distribution in $[0, 1]$.

In addition, α^t is a trade-off constant at t iteration [37] and is calculated as follows:

$$\alpha^t = \alpha_0 \theta^t \quad (5)$$

where α_0 and α^t are the initial trade-off coefficient and the trade-off coefficient at the t^{th} iteration, respectively; and θ is the adaptive parameter ($0 < \theta < 1$) [37].

Finally, $L(s)$ is the Lévy distribution, which can be defined as follows:

$$L(s) \sim s = \frac{u}{|v|^{1/\tau}} \quad (6)$$

where s is a power-law distribution, τ is an index, and v and u are calculated to follow a normal distribution as follows:

$$v \sim N(0, \sigma_v^2) \quad (7)$$

$$u \sim N(0, \sigma_u^2) \quad (8)$$

where:

$$\sigma_v = 1 \quad (9)$$

$$\sigma_u = \left\{ \frac{\Gamma(1+\tau) \sin(\pi\tau/2)}{\Gamma[(1+\tau)/2] \tau 2^{(\tau-1)/2}} \right\}^{1/\tau} \quad (10)$$

where $\Gamma(z)$ is the Gamma function, which is determined by:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt \quad (11)$$

2.1.3 Movement

EFA benefits from the advantages of both EA and FA. Firstly, EFA has inspired the concept of attraction and repulsion forces from EA. However, EFA ranks all fireflies based on their objective function values and a firefly i is only affected by the attraction (F_g) and repulsion force (F_r) from the next better ($i-1$) and worse firefly ($i+1$). In contrast, all fireflies in the population are considered in EA. In this way, EFA can mitigate the interference from many fireflies and quickly find the optimal solution. The movement of a magnetic firefly is written as:

$$x_i^{t+1} = x_i^t + \left(x_{i-1}^t - x_i^t \right) \frac{F_g}{\|F_g\|} (1 - \alpha) + \left(x_{i+1}^t - x_i^t \right) \frac{F_r}{\|F_r\|} \alpha + \alpha \varepsilon L(s) \quad (12)$$

where F_g and F_r are attraction and repulsion forces, which are defined by Eq. (13) and Eq. (14), respectively:

$$F_g = |x_{i-1} - x_i| \frac{q_{i-1} q_i}{\|x_{i-1} - x_i\|^2} \quad (13)$$

$$F_r = |x_i - x_{i+1}| \frac{q_i q_{i+1}}{\|x_i - x_{i+1}\|^2} \quad (14)$$

where q_i is the charge of the point i , which is defined as follows:

$$q_i = \exp \left(- \frac{f(x^i) - f(x^{best})}{\sum_{k=1}^n (f(x^i) - f(x^{best}))} \right) \quad (15)$$

where $f(x)$ is the objective function. A firefly with better objective values has a higher charge. The objective function used in EFA-ANN model will be discussed in more detail in Section 2.3.

EFA uses the weights $1-\alpha$ and α to modify the effects of the attraction (F_g) and repulsion (F_r) forces during the optimization process. As alpha decreases from 1 to 0 during the optimization process in the early stages, the fireflies are significantly influenced by F_r to explore the promising area. In the final stages, the effect of F_g and F_r is increased and decreased, respectively, to help the fireflies exploit the optimal solution in the best promising area. Therefore, these improvements help EFA find a better solution as opposed to the separate applications of EA and FA. The efficiency of EFA will be validated in the following sections.

In each iteration, EFA chooses the best firefly based on the value of the objective function. The best firefly then performs a local search around its place by Levy flight while the rest of the population searches for other areas. At the end of the iteration, the best firefly is chosen, and the process continues until the termination criteria are reached. The termination criteria helps to reduce the computing time without affecting the solution quality. The objective function and termination criteria in this study will be discussed in Section 2.3. Figure 1 summarizes the pseudo-code of EFA, while Figure 2 illustrates a flowchart of the search process of EFA.

```

1  \\Start EFA
2
3  Eq. (1)    \\ Generate initial population of fireflies by logistic map
4
5  Define objective function  $f(x)$  for all fireflies
6
7  Rank the fireflies based on objective function and find the current best
8
9  while ( $t < MaxGeneration$ )
10
11     for  $i = 1:n$  \\  $n$  is number of fireflies
12
13         if  $i \neq 1$     \\ the current best firefly is not attracted by others
14
15             Eq. (12)    \\Move firefly  $i$  in  $d$ -dimension (Exploration phase)
16
17         end if
18
19         Eq. (3) \\Local search (Exploitation phase)
20
21         Evaluate new solutions and update the coordinate and the objective function
22
23     end for  $i$ 
24
25      $t = t + 1$ 
26
27 end while
28
29 Post process results and visualization
30
31 \\End EFA

```

Figure 1. Pseudocode for EFA.

Figure 2. The flowchart of EFA.

2.2 Artificial Neural Network

Technical details of ANN can be found in a previous study by [42, 43]. This study outlines the main concepts of the ANN model. Figure 3 illustrates the layout of an ANN model with a set of artificial neurons. Each neuron in a layer sends a signal to another

neuron in the next layer by a connection, which is assigned a weight, and the weight represents the strength of the signal [44, 45].

Figure 3. The layout of an ANN model.

In the hidden layer, the signals from the input layer are calculated by a linear function (see Eq. 16) and a transfer function (see Eq. 17) to generate the output signal of a hidden node [46] as depicted in Figure 4.

$$net_i = \sum_{p=1}^P w_{i,p} I_p + b_i \quad (16)$$

where net_i is the value of the i^{th} net; w_{ip} and b_i are the weight of the p^{th} input to the i^{th} hidden node and the bias parameter of the i^{th} hidden node, respectively; I_p is the value of the p^{th} input node.

The transfer function is defined as:

$$y_i = f(net_i) = \frac{1}{1 + \exp(-net_i)} \quad (17)$$

where the transfer function in this study is a sigmoid function and y_i is the output signal of the i^{th} hidden node.

Figure 4. The structure of a neuron in an ANN model.

During the learning process, ANN uses the Mean Square Error (MSE), to evaluate the performance of the model as follows:

$$MSE = \frac{1}{NN_{out}} \sum_{n=1}^N \sum_{o=1}^{N_{out}} (e_{n,o})^2 \quad (18)$$

where N and N_{out} are the number of instances and the number of outputs, respectively;

$e_{n,o} = \bar{y}_{n,o} - y_{n,o}$ is the training error at the o^{th} output with n^{th} instance; y is the actual output and $\bar{y}$ is the predicted output by ANN.

The Levenberg–Marquardt algorithm [47, 48] is used to update the weights and biases to minimize the MSE. The calculation of the Levenberg–Marquardt algorithm is presented as:

$$w^{k+1} = w^k - \left(J^{k^T} J^k + \mu \bar{I} \right)^{-1} J^{k^T} e^k \quad (19)$$

where w^k is the weight and bias matrix at the k^{th} iteration; $\bar{I}$ is the identity matrix; μ is the combination coefficient ($\mu > 0$); and J is the Jacobian matrix [49]:

$$J = \begin{matrix} & \begin{matrix} \text{Node 1} & & \dots & \text{Node } i & & \dots \end{matrix} \\ \left[\begin{array}{cccc} \frac{\partial e_{1,1}}{\partial w_{1,1}} & \frac{\partial e_{1,1}}{\partial w_{1,2}} & \dots & \frac{\partial e_{1,1}}{\partial w_{i,1}} & \frac{\partial e_{1,1}}{\partial w_{i,2}} & \dots \\ \frac{\partial e_{1,2}}{\partial w_{1,1}} & \frac{\partial e_{1,2}}{\partial w_{1,2}} & \dots & \frac{\partial e_{1,2}}{\partial w_{i,1}} & \frac{\partial e_{1,2}}{\partial w_{i,2}} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial e_{1,N_{out}}}{\partial w_{1,1}} & \frac{\partial e_{1,N_{out}}}{\partial w_{1,2}} & \dots & \frac{\partial e_{1,N_{out}}}{\partial w_{i,1}} & \frac{\partial e_{1,N_{out}}}{\partial w_{i,2}} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial e_{n,1}}{\partial w_{1,1}} & \frac{\partial e_{n,1}}{\partial w_{1,2}} & \dots & \frac{\partial e_{n,1}}{\partial w_{i,1}} & \frac{\partial e_{n,1}}{\partial w_{i,2}} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial e_{n,o}}{\partial w_{1,1}} & \frac{\partial e_{n,o}}{\partial w_{1,2}} & \dots & \frac{\partial e_{n,o}}{\partial w_{i,1}} & \frac{\partial e_{n,o}}{\partial w_{i,2}} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial e_{N,1}}{\partial w_{1,1}} & \frac{\partial e_{N,1}}{\partial w_{1,2}} & \dots & \frac{\partial e_{N,1}}{\partial w_{i,1}} & \frac{\partial e_{N,1}}{\partial w_{i,2}} & \dots \\ \frac{\partial e_{N,2}}{\partial w_{1,1}} & \frac{\partial e_{N,2}}{\partial w_{1,2}} & \dots & \frac{\partial e_{N,2}}{\partial w_{i,1}} & \frac{\partial e_{N,2}}{\partial w_{i,2}} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial e_{N,N_{out}}}{\partial w_{1,1}} & \frac{\partial e_{N,N_{out}}}{\partial w_{1,2}} & \dots & \frac{\partial e_{N,N_{out}}}{\partial w_{i,1}} & \frac{\partial e_{N,N_{out}}}{\partial w_{i,2}} & \dots \end{array} \right] & \begin{matrix} o = 1 \\ o = 2 \\ \dots \\ o = N_{out} \\ \dots \\ o = 1 \\ \dots \\ o = o \\ \dots \\ o = 1 \\ o = 2 \\ \dots \\ o = N_{out} \end{matrix} \end{matrix} \quad \begin{matrix} \left. \begin{matrix} \dots \\ \dots \\ \dots \end{matrix} \right\} n = 1 \\ \left. \begin{matrix} \dots \\ \dots \end{matrix} \right\} n = n \\ \left. \begin{matrix} \dots \\ \dots \end{matrix} \right\} n = N \end{matrix} \quad (20)$$

where the error vector e at each neuron is written as [46]:

$$e = \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ \dots \\ e_{1,N_{out}} \\ \dots \\ e_{N,1} \\ e_{N,2} \\ \dots \\ e_{N,N_{out}} \end{bmatrix} \quad (21)$$

2.3 EFA-ANN implementation

A flowchart of EFA-ANN is shown in Figure 5. Firstly, historical data is divided into learning and test data following k-fold cross-validation. The learning data is then divided into training and validation data with a ratio of 90% and 10%, respectively. The training data are trained by the Levenberg–Marquardt algorithm, while the validation data are used to validate the trained ANN. After finding an optimal ANN model, the test data is used to evaluate its performance. The EFA-ANN will be validated in Section 5.

This study also applied initial weights and biases in the range of $[-0.5, 0.5]$ at the beginning of the training process. Chang *et al.*[32] showed that this is the optimal range for finding initial weights and biases of ANN. Figure 5 shows that the weights and biases are first updated inside the ANN. EFA automatically memorizes and optimizes these parameters to minimize prediction errors. As a result, the computing time can be significantly reduced. The Root Mean Square Error (RMSE) is used to calculate the objective function of EFA-ANN as follows:

$$f = RMSE_{Validation-data} \quad (22)$$

1 It is essential to clarify that RMSE is the objective function of EFA-ANN while the
2
3
4 Mean Square Error (MSE) mentioned in Section 2.2 is the error function inside the ANN
5
6 model. Also, the maximum generation is used as the termination criterion in this study.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 5. A flowchart of the EFA-ANN model.

3. Model performance evaluation methods

Five performance measures were used to evaluate the performance of the proposed model. These include the linear correlation coefficient (R), determination coefficient (R^2), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). Details of these performance measures can be found in the literature [42, 50, 51] and are calculated as follows:

$$R = \frac{NN_{out} \sum_{n=1}^N \sum_{o=1}^{N_{out}} y_{n,o} \cdot \bar{y}_{n,o} - \left(\sum_{n=1}^N \sum_{o=1}^{N_{out}} y_{n,o} \right) \left(\sum_{n=1}^N \sum_{o=1}^{N_{out}} \bar{y}_{n,o} \right)}{\sqrt{NN_{out} \left(\sum_{n=1}^N \sum_{o=1}^{N_{out}} y_{n,o}^2 \right) - \left(\sum_{n=1}^N \sum_{o=1}^{N_{out}} y_{n,o} \right)^2} \sqrt{NN_{out} \left(\sum_{n=1}^N \sum_{o=1}^{N_{out}} \bar{y}_{n,o}^2 \right) - \left(\sum_{n=1}^N \sum_{o=1}^{N_{out}} \bar{y}_{n,o} \right)^2}} \quad (23)$$

$$R^2 = RR \quad (24)$$

$$RMSE = \sqrt{\frac{1}{NN_{out}} \sum_{n=1}^N \sum_{o=1}^{N_{out}} (\bar{y}_{n,o} - y_{n,o})^2} \quad (25)$$

$$MAE = \frac{1}{NN_{out}} \sum_{n=1}^N \sum_{o=1}^{N_{out}} |y_{n,o} - \bar{y}_{n,o}| \quad (26)$$

$$MAPE = \frac{1}{NN_{out}} \sum_{n=1}^N \sum_{o=1}^{N_{out}} \left| \frac{y_{n,o} - \bar{y}_{n,o}}{y_{n,o}} \right| \quad (27)$$

4. Data collection

4.1 Dataset 1

Dataset 1 was generated by modeling a typical single-family house with a lightweight wood-frame structure in Istanbul, Turkey [52]. In this case study, the effect of the façade system on the total heating and cooling energy were investigated by changing its properties. Fig. 6 shows the details of the façade system including glass wool

insulation (layer iii), two layers of oriented strand board (OSB) applied as sheathing (layer ii and v), two layers of gypsum board (layer i and iv) and a layer of cement-bonded particleboard used as the exterior finishing of the wall.

Figure 6. The details of the façade system in dataset 1.

The façade system was modeled using different properties for the insulation layer, namely thickness and thermal conductivity (K-value), which are listed in Table 1. Five types of façades with different layer thicknesses were also investigated. The thickness of layer i, ii, iv, v and vi of type 1 are 1.5 cm, 2.5 cm, 2 cm, 2.5 cm and 1.5 cm, respectively. The details of all five types are listed in Table 1. A total of 180 simulations were carried out to evaluate how the thermal conductivity and thickness of the façade affect the total heating and cooling energy in the building.

Parameter	Unit	Values	Variable
Insulation K-value (iii)	W/m-K	0.03, 0.04, 0.05, 0.08	
Insulation thickness (iii)	cm	4, 6, 8, 10, 12, 14, 16, 18, 20	
		1 [i, ii, iv, v, vi: 1.5, 2.5, 2, 2.5, 1.5 (cm)]	
		2 [i, ii, iv, v, vi: 2, 2, 2.5, 2, 1.5 (cm)]	Input
Façade type	N/A	3 [i, ii, iv, v, vi: 1, 3, 1.5, 3, 1.5 (cm)]	
		4 [i, ii, iv, v, vi: 2.5, 1.5, 3, 1.5, 1.5 (cm)]	
		5 [i, ii, iv, v, vi: 1, 3.5, 1.9, 1, 1.5 (cm)]	
Total heating and cooling energy	kWh	Min: 6094.24; Max: 11095.06	Output

Table 1. Statistical parameters for dataset 1.

After determining the simulated cases, the building models were drawn in Sketchup, and the thermal properties of the wall components, including window location, direction and orientation, were then inputted into EnergyPlus to simulate the energy consumption of the building. As the primary purpose of the simulation is to validate the effect of the façade system on the total heating and cooling energy in the building, the variables of these simulations are restricted to the type of façade material. Other factors of the building were maintained constant, and the properties used in the simulation are listed in Table 2.

	Unit	Value
Building type	-	Residential building
Building location	-	Istanbul, Turkey
Floor area		
First story	m ²	81.7
Second story	m ²	48.4
Run period	year	1
Electric equipment	W/m ²	10
Lighting	W/m ²	12

Thermostat

Heating setpoint (constant)	°C	21
-----------------------------	----	----

Cooling setpoint (constant)	°C	26
-----------------------------	----	----

Table 2. Design information in EnergyPlus simulation.

4.2 Dataset 2

Dataset 2 consisted of 768 entries that were generated from twelve building types using Ecotect simulation software [53]. These twelve building types were represented by 18 simple cubes (3.5m x 3.5m x 3.5m) and the shape of each type is shown in Fig.7. Therefore, the buildings have the same volume but different surface areas and dimensions. These buildings were each simulated as a residential building in Athens, Greece. This dataset is used to investigate the effect of dimension on the cooling load (CL) in a building. Hence, the same material properties of the façade system were used for all twelve buildings, including U-value of the wall (1.78 W/m²K), window (2.26 W/m²K) floors (0.86 W/m²K) and roofs (0.50 W/m²K). The lighting level and latent heat were set to 300 lux and 2 W/m², respectively. The cooling load of the residential building was simulated by using eight features including the relative compactness (RC), surface area, wall area, roof area, overall height, orientation, glazing area and glazing distribution.

Each input parameter represents a property of the building. For instance, the *RC* indicator [54] represents the type of building, and the *RC*-value of each building is shown in Fig. 7. The *RC* is calculated by Eq. (27) as follows:

$$RC = 6V^{2/3}A^{-1} \quad (28)$$

where *V* and *A* are the volume and surface area of the building, respectively.

Figure 7. Shapes of the buildings in dataset 2.

The experiments simulated the building using two configurations, namely with and without glazing. In a glazing system, three glazing-to-floor area ratios were used, including 10%, 25%, and 40%. Five glazing distributions are considered including: (1) 25% glazing for each faces; (2) 55% for the north side and 15% for the other faces; (3) 55% for the east face and 15% for the remaining faces; (4) 55% for the south face and 15% for the other faces; and (5) 55% for the west face and 15% for the other faces [53]. Finally, all building shapes were rotated to four orientations, namely north, south, east and west. Overall, 768 building configurations were simulated, and the details of the input and output parameters are provided in Table 4. By applying this process, all input values are discretized and calculated from the twelve building types. The number of possible values of each input are also listed in Table 4.

Parameter	Unit	No. of possible values	Min.	Max.	Variable
Relative compactness	N/A	12	0.62	0.98	Input
Surface area	m ²	12	514.50	808.50	
Wall area	m ²	7	245.00	416.50	

Roof area	m ²	4	110.25	220.50	
Overall height	m	2	3.50	7.00	
Orientation	N/A	4	-	-	
Glazing area	%	4	0.00	40.00	
Glazing distribution	N/A	6	0.00	5.00	
Cooling load (CL)	kW	768	10.90	48.03	Output

Table 4. Descriptions of dataset 2.

5. Performance Evaluation and Discussion

5.1 Data Preprocessing and Model Application

This study used the K-fold cross-validation method to alleviate problems with over-fitting data [55]. The historical data was divided into 10 folds, which is the optimal number determined by other researchers [56]. Also, the K-fold cross-validation method can help to provide equal weightings to the results to obtain a fair comparison. The models that were used to assess the performance of our proposed model also utilized the K-fold cross-validation method [50, 53, 57].

In this method, each dataset is randomly divided into 10 separate folds. There are 10 rounds in the entire process. In each round, nine folds are used for training the EFA-ANN and the remaining fold is used for testing in each round. Therefore, all data is guaranteed to be used in both the learning and testing phases. Finally, the average results in ten rounds are obtained to assess the performance of the model. The flowchart for each round is shown in Figure 5. In addition, all parameters of EFA-ANN used in this study are listed in Table 5.

	Parameters	Setting
EFA	Number of fireflies	15
	Number of Maximum generations	15
	α_o	0.9
	Adaptive inertial weight (θ)	$(10^{-2} / 0.9)^{1/MaxGeneration}$
	Percentage of training data	90%
	Percentage of validation data	10%
	Number of folds	10
ANN	Objective function	RMSE
	Transfer function	Sigmoid
	Learning algorithm	Levenberg–Marquardt
	Number of hidden layers	1
	Number of nodes	20

Table 5. The parameter setting in EFA-ANN.

At the beginning of the training process, the data is normalized to avoid numerical difficulties, i.e. inputs in higher numeric ranges may dominate those in smaller numeric ranges [58]. Therefore, this study normalizes the data to the range of [-1, 1]. The normalized value (x') is calculated from the original value (x) as follows:

$$x' = \frac{2(x - \min(x))}{\max(x) - \min(x)} - 1 \quad (29)$$

5.2 Results and discussion

Table 6 lists the predictive performance measures of the proposed model for dataset 1. Several models are used for comparing the predictive accuracy of the proposed model for the same dataset. For example, Naji *et al.* proposed an extreme learning machine (ELM) method to forecast the energy consumption in a building [52]. Their model performs quite well ($RMSE = 74.02$ (kWh), $R = 0.999$ and $R^2 = 0.997$). Their comparison showed that ELM was superior to genetic programming (GP) and ANN' models, which were run in their study, in terms of accuracy and computing time.

Method	Performance measure					CT	Improvement rates by EFA-ANN				
	R	R^2	$RMSE$	MAE	$MAPE$		R	R^2	$RMSE$	MAE	$MAPE$
	(-)	(-)	(kWh)	(kWh)	(%)	(s)	(%)	(%)	(%)	(%)	(%)
Dataset 1											
ELM [52]	0.999	0.997	74.02	-	-	330	0.15**	0.30**	93.28*	-	-
ANN' [52]	0.971	0.943	331.57	-	-	424	2.97*	6.02*	98.50*	-	-
GP [52]	0.977	0.954	314.35	-	-	436	2.38*	4.81*	98.42*	-	-
ANN	0.999	0.998	39.69	28.29	0.05	66	0.15**	0.20**	87.47*	88.76*	18.84*
EFA-ANN	1.000	1.000	4.97	3.18	0.04	66	-	-	-	-	-

Note: CT is computing time which is calculated based on 1 running time.

*, ** indicates significance levels higher than (1%, 5%), respectively;

Table 6. Performance measures and improvement rates of the EFA-ANN model for dataset 1.

However, Table 6 indicates that the performance of EFA-ANN is superior to ELM and other methods. Moreover, EFA-ANN produces the smallest error rate (RMSE), and it is 93.28% - 98.50% better than those reported by other studies. Notably, in terms of computational cost, EFA-ANN is 5 times faster than ELM and 7 times faster than ANN' and GP. Moreover, this research also compares the performance of EFA-ANN with the ANN model to confirm that there is a prominent improvement over the combined EFA-ANN model. The ANN model, in this case, is configured with the same setting as the ANN in the EFA-ANN model. The computing time of ANN is modified to be consistent with that of EFA-ANN for a fair comparison. The comparative results show that the error rates of EFA-ANN are 18.84% - 87.47% better than those of ANN. The results indicate that EFA-ANN can provide an alternative approach for predicting the energy consumption in a building.

The predicted results from EFA-ANN and other methods for dataset 2 are listed in Table 7. The proposed approach is shown to be superior to other methods available in the literature. For example, EFA-ANN obtains the lowest RMSE (0.51 kW) compared to other models: iteratively reweighted least squares (IRLS) (3.39 kW) [53], random forests

(RF) (2.57 kW) [53], ensemble model (1.57 kW), smart artificial firefly colony algorithm-based support vector regression (SAFCA-SVR) (0.68 kWh) [57] and ANN model. Similarly, Table 7 shows that EFA-ANN has the lowest MAE and MAPE, which is equal to 0.38 (kW) and 1.71 (%), respectively. Also, the R and R^2 of EFA-ANN are equal to SAFCA-SVR, which is better than the ensemble model. With these results, EFA-ANN has improved the error rate from 16.18% to 84.84% when compared with other models. Additionally, the proposed model is around 5 times faster than the SAFCA-SVR model in predicting CL of this dataset (47 minutes compared to 240 minutes). Table 7 shows that the error rates of EFA-ANN are 31.87% - 53.27% better than ANN. Details of the comparison between ANN and EFA-ANN will be discussed hereafter.

Method	Performance measure					CT	Improvement rates by EFA-ANN				
	R	R^2	RMSE	MAE	MAPE		R	R^2	RMSE	MAE	MAPE
	(-)	(-)	(kW)	(kW)	(%)	(s)	(%)	(%)	(%)	(%)	(%)
Dataset 2											
IRLS [53]	N/A	N/A	3.39	2.21	9.41	N/A	-	-	84.84*	83.02*	81.83*
RF [53]	N/A	N/A	2.57	1.42	4.62	N/A	-	-	80.00*	73.58*	62.99*
Ensemble model [50]	0.99	0.97	1.57	0.97	3.46	N/A	1.27**	2.55**	67.18*	61.44*	50.51*
SAFCA-SVR [57]	1.00	0.99	0.68	0.47	2.04	1440	-	0.30	24.41**	20.17**	16.18***
ANN	1.00	0.99	1.10	0.56	2.51	280	-	0.59	53.27**	33.00**	31.87**
EFA-ANN	1.00	1.00	0.51	0.38	1.71	280	-	-	-	-	-

Note: CT is computing time which is calculated based on 1 running time.;

*, ** indicates significance levels higher than (1%, 5%), respectively;

Table 7. Performance measures and improvement rates of the EFA-ANN model for dataset 2.

This study compares the EFA-ANN model with the single ANN model to investigate the effects of EFA on the RMSE value and computing time. In order to make a fair comparison, the parameters in the single ANN model were set to be identical with the ANN model in EFA-ANN. Only the number of iterations of the single ANN model,

in this case, was increased to extend the computing time (to be consistent with EFA-ANN). Both single ANN and EFA-ANN were run on the same computer. Figure 8 shows that the single ANN is stuck at the local optima after around 20 and 50 seconds for Dataset 1 (a) and Dataset 2 (b), respectively. In contrast, EFA-ANN approaches a better result because EFA helps ANN find better weights and bias values. This comparison confirms that EFA is effective in optimizing the weights and biases of ANN.

Figure 8. Convergence results of ANN and EFA-ANN for (a) Dataset 1 and (b) Dataset 2

The statistical relationship between the predicted outputs obtained by EFA-ANN and the actual outputs of the two datasets is shown in Figure 9. The R -values for the two datasets are almost equal to 1, which indicates that the predicted values from EFA-ANN have a strong correlation with the actual values. Notably, the MAPE in dataset 1 is equal to 0.04%, which indicates that the predicted values are approximately equal to the actual values. In other words, the proposed model can forecast the exact energy consumption for dataset 1. Meanwhile, the MAPE for dataset 2 is higher than that for dataset 1 but remains negligible (1.71%). The results of EFA-ANN are slightly different from the actual values, but these results are predicted with better accuracy than the aforementioned

methods. Therefore, the proposed EFA-ANN is an efficient model for predicting the energy consumption (including HL and CL) of a building.

Figure 9. The correlation between the actual and the predicted outputs for dataset 1 (a) and dataset 2 (b).

5.3 Sensitivity Analysis

A sensitivity analysis is performed to quantify the effects of different inputs on the predicted energy consumption. Among the different approaches, this study used the method proposed by Garson [59] because it is suitable for discrete input data such as wall type, relative compactness, orientation, and glazing distribution. This method is also suitable for the ANN model as the sensitivity analysis deconstructs the weights of the connections between neurons to quantify the influence of the various inputs. The influence of a specified input on the output can be determined by assessing all the connecting weights between the nodes of interest. Therefore, all connecting weights between a specific input and output are identified, and the importance of all the inputs is calculated (as a percentage) by the following equation:

$$Q_p = \sum_{i=1} \left(\frac{w_{ip}}{\sum_{p=1} w_{ip}} \times \frac{w_{oi}}{\sum_{i=1} w_{oi}} \right) \quad (30)$$

Where Q_p is the impact of the input p^{th} on the output in percentage; w_{ip} is the weight of the p^{th} input to the i^{th} hidden node; w_{oi} is the weight of the i^{th} hidden node to the output.

The results from the sensitivity analysis on datasets 1 and 2 are shown in Figure 10. For dataset 1, the insulation thickness has the highest impact on the total heating and cooling load (43.8%), followed by the facade type and K-value of the insulation (34.1% and 22.1%, respectively). This result is reasonable because the insulation thickness has a prominent effect on the energy consumption in a building, which was shown in previous studies [60, 61]. Meanwhile, the glazing area (29.4%) and the relative compactness (27.5%) are the most critical parameters that affect the total cooling load in dataset 2. This result is expected because the solar heat gain will increase with a larger glazing area, which thereby affects the cooling load in the building. Also, several studies confirmed that an increasing window-to-wall ratio causes an increase in the cooling load [62, 63]. In contrast, the orientation of the façade and the roof area have the least impact with just under 5%. The results of the sensitivity analysis can help a designer to quickly identify which input should be modified to improve the energy performance of the building. For example, the insulation thickness in dataset 1 or the glazing area and relative compactness in dataset 2 should be modified before the other variables to effectively improve the energy performance of the building.

Figure 10. The effect of several design parameters on the energy consumption of the building.

6. Conclusions

A novel method was developed for predicting the energy consumption in a building. The model, namely EFA-ANN, integrates EFA into ANN to improve its performance by optimizing the set of initial parameters. The proposed approach was compared with several published models in terms of its speed and accuracy in forecasting heating and cooling energy. Two datasets with various material and isolation properties (dataset 1), and building characteristics (dataset 2) were used to validate EFA-ANN. Additionally, a 10-fold cross-validation method was applied to mitigate the over-fitting problem when comparing the performance of these models.

In dataset 1, EFA-ANN showed a 93.28% - 98.50% improvement of other methods reported in the literature. Notably, EFA-ANN is 5 times faster than the ELM method and 7 times faster than GP in terms of computing time. Similarly, EFA-ANN not only obtained the lowest RMSE, MAE and MAPE, but also the highest *R*-value compared to the other techniques for dataset 2. Also, the computing time of EFA-ANN is less 5 times than that of SAFCA-SVR. The obtained results demonstrated the strong capabilities of machine learning in predicting the energy consumption of buildings.

A sensitivity analysis was also performed to identify the input with the most critical impact on the output of each dataset. The results of the sensitivity analysis indicate that the insulation thickness and the glazing area have the most significant impact for datasets 1 and 2, respectively. This result can help designers to quickly validate their design of a building and improve its energy performance by focusing on these essential inputs.

This study showed that the proposed EFA-ANN model achieved both good results and a short computational time relative to other methods. Therefore, EFA-ANN can help

1 energy engineers to design energy-efficient buildings whilst reducing experimental
2 requirements and it could assist civil engineers and construction managers in the early
3 design phase of energy-efficient buildings. Besides, the proposed approach can be used
4 as a useful tool for quickly and accurately solving many problems in engineering,
5 including energy-efficient buildings, construction material strength, and structure
6 strength.
7
8
9
10
11
12
13
14
15

16 **Acknowledgments**

17
18
19 The first author would like to thank the University of Melbourne for offering the
20 Melbourne Research Scholarship. This work was mainly supported by the CRC-P for
21 Advanced Manufacturing of High Performance Building Envelope project, funded by the
22 CRC-P program of the Department of Industry, Innovation and Science, Australia, and
23 the Asia Pacific Research Network for Resilient and Affordable Housing (APRAH) grant,
24 funded by the Australian Academy of Science, Australia. This work was also supported
25 by the ARC Training Centre for Advanced Manufacturing of Prefabricated Housing
26 (CAMP.H) at the University of Melbourne.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

References

- [1] Al-Homoud, M.S., *Computer-aided building energy analysis techniques*. Building and Environment, 2001. **36**(4): p. 421-433.
- [2] IEA, *World Energy Outlook 2015*. 2015.
- [3] Administration, U.S.E.I., *INTERNATIONAL ENERGY OUTLOOK 2016*. 2016.
- [4] Zanchini, E. and C. Naldi, *Energy saving obtainable by applying a commercially available M-cycle evaporative cooling system to the air conditioning of an office building in North Italy*. Energy, 2019. **179**: p. 975-988.
- [5] Ahmadi-Karvigh, S., et al., *Real-time activity recognition for energy efficiency in buildings*. Applied Energy, 2018. **211**: p. 146-160.
- [6] Pérez-Lombard, L., J. Ortiz, and C. Pout, *A review on buildings energy consumption information*. Energy and Buildings, 2008. **40**(3): p. 394-398.
- [7] Zemella, G., et al., *Optimised design of energy efficient building façades via Evolutionary Neural Networks*. Energy and Buildings, 2011. **43**(12): p. 3297-3302.
- [8] Magnier, L. and F. Haghighat, *Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network*. Building and Environment, 2010. **45**(3): p. 739-746.
- [9] Ferreira, P.M., et al., *Neural networks based predictive control for thermal comfort and energy savings in public buildings*. Energy and Buildings, 2012. **55**: p. 238-251.
- [10] Ghahramani, A., et al., *Energy savings from temperature setpoints and deadband: Quantifying the influence of building and system properties on savings*. Applied Energy, 2016. **165**: p. 930-942.
- [11] Kumar, A. and B.M. Suman, *Experimental evaluation of insulation materials for walls and roofs and their impact on indoor thermal comfort under composite climate*. Building and Environment, 2013. **59**: p. 635-643.
- [12] Lechner, N., *Heating, cooling, lighting: Sustainable design methods for architects*. 2014: John Wiley & sons.
- [13] Yang, Z., A. Ghahramani, and B. Becerik-Gerber, *Building occupancy diversity and HVAC (heating, ventilation, and air conditioning) system energy efficiency*. Energy, 2016. **109**: p. 641-649.
- [14] Ihara, T., A. Gustavsen, and B.P. Jelle, *Effect of facade components on energy efficiency in office buildings*. Applied Energy, 2015. **158**: p. 422-432.
- [15] Liu, Z., et al., *Application and suitability analysis of the key technologies in nearly zero energy buildings in China*. Renewable and Sustainable Energy Reviews, 2019. **101**: p. 329-345.
- [16] Naganathan, H., W.O. Chong, and X. Chen, *Building energy modeling (BEM) using clustering algorithms and semi-supervised machine learning approaches*. Automation in Construction, 2016. **72**(Part 2): p. 187-194.
- [17] Hong, T., et al., *Automatic calibration model of a building energy simulation using optimization algorithm*. Energy Procedia, 2017. **105**(Supplement C): p. 3698-3704.
- [18] Ryan, E.M. and T.F. Sanquist, *Validation of building energy modeling tools under idealized and realistic conditions*. Energy and Buildings, 2012. **47**(Supplement C): p. 375-382.

- [19] Mousavi, S.M., et al., *A new predictive model for compressive strength of HPC using gene expression programming*. Advances in Engineering Software, 2012. **45**(1): p. 105-114.
- [20] Platon, R., V.R. Dehkordi, and J. Martel, *Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis*. Energy and Buildings, 2015. **92**: p. 10-18.
- [21] Robinson, C., et al., *Machine learning approaches for estimating commercial building energy consumption*. Applied Energy, 2017. **208**: p. 889-904.
- [22] Ahmad, T., et al., *Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment*. Energy, 2018. **158**: p. 17-32.
- [23] Chou, J.-S. and D.-S. Tran, *Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential households*. Energy, 2018. **165**: p. 709-726.
- [24] Ghahramani, A., S.A. Karvigh, and B. Becerik-Gerber, *HVAC system energy optimization using an adaptive hybrid metaheuristic*. Energy and Buildings, 2017. **152**: p. 149-161.
- [25] Moon, J.W., et al., *Determining optimum control of double skin envelope for indoor thermal environment based on artificial neural network*. Energy and Buildings, 2014. **69**: p. 175-183.
- [26] Moon, J.W., et al., *Preliminary performance tests on artificial neural network models for opening strategies of double skin envelopes in winter*. Energy and Buildings, 2014. **75**: p. 301-311.
- [27] Wang, L., E.W.M. Lee, and R.K.K. Yuen, *Novel dynamic forecasting model for building cooling loads combining an artificial neural network and an ensemble approach*. Applied Energy, 2018. **228**: p. 1740-1753.
- [28] Chung, M.H., et al., *Application of artificial neural networks for determining energy-efficient operating set-points of the VRF cooling system*. Building and Environment, 2017. **125**: p. 77-87.
- [29] Kolen, J.F. and J.B. Pollack, *Back propagation is sensitive to initial conditions*, in *Proceedings of the 1990 conference on Advances in neural information processing systems 3*. 1990, Morgan Kaufmann Publishers Inc.: Denver, Colorado, USA. p. 860-867.
- [30] Yam, Y.F. and T.W.S. Chow, *Determining initial weights of feedforward neural networks based on least squares method*. Neural Processing Letters, 1995. **2**(2): p. 13-17.
- [31] Liu, Q., et al., *Ensemble artificial neural networks applied to predict the key risk factors of hip bone fracture for elders*. Biomedical Signal Processing and Control, 2015. **21**: p. 146-156.
- [32] Chang, Y.-T., et al., *Optimization the Initial Weights of Artificial Neural Networks via Genetic Algorithm Applied to Hip Bone Fracture Prediction*. Advances in Fuzzy Systems, 2012. **2012**: p. 9.
- [33] Lee, A., W.Z. Geem, and K.-D. Suh, *Determination of Optimal Initial Weights of an Artificial Neural Network by Using the Harmony Search Algorithm: Application to Breakwater Armor Stones*. Applied Sciences, 2016. **6**(6).
- [34] Al-Shammari, E.T., et al., *Prediction of heat load in district heating systems by Support Vector Machine with Firefly searching algorithm*. Energy, 2016. **95**: p. 266-273.

- [35] Coelho, L.d.S. and V.C. Mariani, *Improved firefly algorithm approach applied to chiller loading for energy conservation*. Energy and Buildings, 2013. **59**: p. 273-278.
- [36] Chu, W.-M. and K.-Y. Chang, *Improving electromagnetism algorithm for solving resource allocation problem in stochastic networks*. Computers & Operations Research, 2017. **86**(Supplement C): p. 30-40.
- [37] Yang, X.-S., *Firefly algorithm, stochastic test functions and design optimisation*. International Journal of Bio-Inspired Computation, 2010. **2**(2): p. 78-84.
- [38] Yelghi, A. and C. Köse, *A modified firefly algorithm for global minimum optimization*. Applied Soft Computing, 2018. **62**: p. 29-44.
- [39] Birbil, Ş.İ. and S.-C. Fang, *An electromagnetism-like mechanism for global optimization*. Journal of global optimization, 2003. **25**(3): p. 263-282.
- [40] Li, Y., S. Deng, and D. Xiao, *A novel Hash algorithm construction based on chaotic neural network*. Neural Computing and Applications, 2011. **20**(1): p. 133-141.
- [41] Hong, W.-C., et al., *SVR with hybrid chaotic genetic algorithms for tourism demand forecasting*. Applied Soft Computing, 2011. **11**(2): p. 1881-1890.
- [42] Bui, D.-K., et al., *A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete*. Construction and Building Materials, 2018. **180**: p. 320-333.
- [43] Nguyen, T., et al., *Deep neural network with high - order neuron for the prediction of foamed concrete strength*. Computer - Aided Civil and Infrastructure Engineering, 2019. **34**(4): p. 316-332.
- [44] Singh, T.N., S. Sinha, and V.K. Singh, *Prediction of thermal conductivity of rock through physico-mechanical properties*. Building and Environment, 2007. **42**(1): p. 146-155.
- [45] Lee, S., et al., *Background Information of Deep Learning for Structural Engineering*. Archives of Computational Methods in Engineering, 2017: p. 1-9.
- [46] Yu, H. and B. Wilamowski, *Levenberg-Marquardt Training*, in *Intelligent Systems*. 2011, CRC Press. p. 1-16.
- [47] Levenberg, K., *A METHOD FOR THE SOLUTION OF CERTAIN NON-LINEAR PROBLEMS IN LEAST SQUARES*. Quarterly of Applied Mathematics, 1944. **2**(2): p. 164-168.
- [48] Marquardt, D., *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*. Journal of the Society for Industrial and Applied Mathematics, 1963. **11**(2): p. 431-441.
- [49] Wilamowski, B.M. and Y. Hao, *Neural Network Learning Without Backpropagation*. IEEE Transactions on Neural Networks, 2010. **21**(11): p. 1793-1803.
- [50] Chou, J.-S. and D.-K. Bui, *Modeling heating and cooling loads by artificial intelligence for energy-efficient building design*. Energy and Buildings, 2014. **82**: p. 437-446.
- [51] Chou, J.-S., W.K. Chong, and D.-K. Bui, *Nature-Inspired Metaheuristic Regression System: Programming and Implementation for Civil Engineering Applications*. Journal of Computing in Civil Engineering, 2016. **30**(5).
- [52] Naji, S., et al., *Estimating building energy consumption using extreme learning machine method*. Energy, 2016. **97**(Supplement C): p. 506-516.

- [53] Tsanas, A. and A. Xifara, *Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools*. Energy and Buildings, 2012. **49**: p. 560-567.
- [54] Mahdavi, A. and B. Gurtekin. *Shapes, Numbers, and Perception: Aspects and Dimensions of the Design Performance Space*. in *Proceedings of the 6th International Conference: Design and Decision Support Systems in Architecture, The Netherlands*. 2002.
- [55] Bishop, C., *Pattern Recognition and Machine Learning*. Information Science and Statistics. 2006, New York: Springer-Verlag.
- [56] Kohavi, R., *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. 1995. 1137-1143.
- [57] Chou, J.S. and A.D. Pham, *Smart Artificial Firefly Colony Algorithm - Based Support Vector Regression for Enhanced Forecasting in Civil Engineering*. Computer-Aided Civil and Infrastructure Engineering, 2015. **30**(9): p. 715-732.
- [58] Lin, C.-w.H.a.C.-c.C.a.C.-j., *A practical guide to support vector classification*. 2010: Department of Computer Science, National Taiwan University, Taipei, Taiwan.
- [59] Garson, G.D., *Interpreting neural-network connection weights*. AI Expert, 1991. **6**(4): p. 46-51.
- [60] Jie, P., et al., *Optimizing the insulation thickness of walls and roofs of existing buildings based on primary energy consumption, global cost and pollutant emissions*. Energy, 2018. **159**: p. 1132-1147.
- [61] Olivieri, F., et al., *An experimental method to quantitatively analyse the effect of thermal insulation thickness on the summer performance of a vertical green wall*. Energy and Buildings, 2017. **150**: p. 132-148.
- [62] Marino, C., A. Nucara, and M. Pietrafesa, *Does window-to-wall ratio have a significant effect on the energy consumption of buildings? A parametric analysis in Italian climate conditions*. Journal of Building Engineering, 2017. **13**: p. 169-183.
- [63] Alghoul, S.K., H.G. Rijabo, and M.E. Mashena, *Energy consumption in buildings: A correlation for the influence of window to wall ratio and window orientation in Tripoli, Libya*. Journal of Building Engineering, 2017. **11**: p. 82-86.