

## Investigating the occupant existence to reduce energy consumption by using a hybrid artificial neural network with metaheuristic algorithms

Nehal Elshaboury<sup>a\*</sup>

<sup>a</sup>Lecturer, Construction and Project Management Research Institute, Housing and Building National Research Center, Giza, Egypt

### CHRONICLE

*Article history:*

Received May 5, 2021

Received in revised format:

June 1, 2021

Accepted August 17 2021

Available online

August 17, 2021

*Keywords:*

Occupancy detection

Machine learning

Metaheuristic algorithm

Particle swarm optimization

Gravitational search algorithm

Neural network

### ABSTRACT

There is an acute need to evaluate the energy consumption of buildings in response to climate change. The “occupant” factor has been largely overlooked in building energy analysis. This research aims at investigating occupancy existence in the office environment using a hybrid artificial neural network with metaheuristic algorithms for improved energy management. It proposes and compares three classification models, namely particle swarm optimization (PSO), gravitational search algorithm (GSA), and hybrid PSO-GSA in combination with the feedforward neural network (FFNN). The inputs to these models are data related to temperature, humidity, light, and carbon dioxide emissions. Two data sets are used for testing the models while the office door is open and closed. The capabilities of the optimized models are evaluated using best, average, median, and standard deviation of the mean squared error. Most of the performance metrics indicate that the FFNN-PSO-GSA model exhibits better performance compared to the other models using the two datasets. The proposed model yields a classification accuracy ranging between 98.47-98.73% using one predictor (i.e., temperature). Besides, it yields an accuracy ranging between 85.45-94.03% using temperature and CO<sub>2</sub> predictors. It can be concluded that the FFNN combined with PSO and GSA algorithms can be a useful tool for occupancy detection modeling.

© 2022 by the authors; licensee Growing Science, Canada.

## 1. Introduction

The energy intensity is expected to improve by only 0.8% in 2020 as a result of the Covid-19 pandemic. This is considered as slow progress towards achieving the global energy and sustainability goals (International Energy Agency, 2020). Meanwhile, climate change is foreseen to be one of the greatest threats in modern times (Delzendeh et al., 2017). The public concerns regarding climate change have increased the focus on analyzing the energy consumption of buildings (Janda, 2011). In recent years, the energy consumption of buildings has increased rapidly because of the increased population growth and the global climate changes, etc (Tang et al., 2021). Buildings are responsible for consuming 40% of the global energy and emitting more than 30% of the worldwide greenhouse gas emissions (Chen et al., 2018; Ahmad et al., 2020). Distinct building types present differences in both energy usage and consumption. For instance, the major areas of energy consumption in residential and commercial buildings are heating and lighting, respectively (De Simone and Fajilla, 2018). The energy crisis and climate change challenge could be addressed by developing energy-efficient buildings and promoting sustainable construction initiatives (Chen et al., 2018; De Simone & Fajilla, 2018). Buildings consume energy based on various factors related to the construction and maintenance quality of the installed systems, thermo-physical properties of the building elements, climatic characteristics, and occupant behavior towards energy utilization (Page et al., 2007; Chen et al., 2015). The importance of evaluating the occupant energy-use has been highlighted in several research studies. For example, the occupant behavior has a larger influence on building energy performance than the building façade (Degelman, 1999). Branco et al. (2004) and Lindén et al. (2006) found that similar or identical buildings, with different occupancy numbers and behavior, are associated with different levels of energy consumption. Furthermore, Martinaitis et al. (2015)

\* Corresponding author.

E-mail address: [nehal\\_ahmed\\_2014@hotmail.com](mailto:nehal_ahmed_2014@hotmail.com) (N. Elshaboury)

referred to the occupant behavior and preferences as major contributors to the deviation between the predicted and actual building energy performance. Determining the occupant behavior has the potential for energy savings, ranging between 10-25% for residential buildings and 5-30% for office buildings (Zhang et al., 2018). With the rapid advances in technology, it has been possible to lower the energy consumption of buildings by determining the occupancy behavior and adjusting the indoor conditions (Trivedi and Badarla, 2020). The indoor conditions reflect many building properties such as temperature, humidity, and carbon dioxide levels. These properties can be modified by controlling some devices, e.g. opening and closing doors and turning on/off fans, lights, and computers (Hong et al., 2016). Therefore, the number and activity of occupants in a building significantly affect the values of these parameters (Fisk, 2000; Erickson et al., 2011; Hong and Lin, 2013; Oldewurtel et al., 2013). It has been reported that occupants who do not have control over the parameters affecting the building's internal environment are unsatisfied compared to users in control (Brager et al., 2004). Besides, it has not been confirmed that the building occupants have an energy-aware point of view regarding the equipment (De Simone & Fajilla, 2018). The occupant information could be described using six properties: presence, location, track, activity, identity, and count (Labeodan et al., 2015). The occupant behaviors are derived by a variety of factors (Tam et al., 2018): (1) subjective factors such as preferences, lifestyle, age, gender, expectations, etc. (Nisiforou et al., 2012; Kavousian et al., 2013), (2) objective factors such as climate, air temperature, velocity, wind speed, noise, building properties, etc. (Schweiker and Shukuya, 2009; Fabi et al., 2011), and (3) external factors such as economics, politics, and culture. Occupants can adapt to the conditions of the indoor environment in one of three ways (Fabi et al., 2013; Polinder et al., 2013): (1) controlling the indoor environment (e.g. thermostats, windows, or shadings), (2) unintentionally (e.g. appliances), (3) adjusting themselves to the existing environmental conditions (e.g. changing clothing or activity level) (Andersen, 2012).

This research proposes hybrid classification models to detect occupant existence based on data recorded from temperature, humidity, light, and carbon dioxide sensors as well as a digital camera. The models are feedforward neural network (FFNN) trained using particle swarm optimization (PSO), gravitational search algorithms (GSA), and hybrid PSO-GSA algorithm. The performance of the trained models is assessed using best, average, median, and standard deviation of the mean squared error (MSE). Rankings of the optimized models are aggregated using an approach based on the half-quadratic theory. This model helps building managers and occupants anticipate an accurate occupancy estimation, resulting in an automated energy management system.

## 2. Occupancy detection models

The occupancy data collection methods can be grouped into four categories: in situ, questionnaire surveys, laboratory, and virtual reality (O'Brien et al., 2017). The first method applies one sensor or sensor fusion to evaluate the presence and actions of occupants as well as the indoor environmental quality. On the other hand, the main drawbacks of this method are represented by privacy, cost, accuracy, sample size, location availability, and social acceptability concerns (Amayri et al., 2019; Ahmad et al., 2020). Questionnaire surveys and interviews aim at collecting quantitative results from a population sample regarding the lifestyles of occupants. These aspects could not be measured using sensors. The laboratory experiments create a controllable environment to study occupant behavior. However, this method might yield negative results in case the occupants feel observed. Finally, virtual reality experiments enable occupants to stimuli and interact with the virtual environment. However, this method is still limited to only visual and acoustics sectors. There has been a continuous growth of research studies interpolating the occupant's behavioral aspects into building energy analysis (Hong et al., 2015). The number of these studies has increased by 230% in the past ten years (Pereira et al., 2018). Some researchers used different in situ measurements such as cameras and sensors to achieve an accurate estimation of occupant behavior. For example, Dziedzic et al. (2019) proposed the usage of a depth registration camera to monitor the indoor occupant behavior while maintaining and satisfying privacy concerns. This technique did not require any interaction from the occupants and had proved its ability to improve the accuracy of building monitoring systems. Sembroiz et al. (2019) applied a wireless sensor network inside smart buildings. The optimum locations of different types of sensors and gateways were optimized to maintain connectivity and reduce energy consumption. Furthermore, an occupants' behavioral model was developed to ensure occupant comfort with minimum energy consumption. Rinaldi et al. (2018) examined occupant preferences and behavior in residential buildings using questionnaire surveys. Besides, multivariate regression analyses were applied to determine the correlations among building characteristics, occupant behaviors, and energy consumption. It was found that there was a strong relationship between occupant behaviors and building age. It was also recommended to conduct further studies to investigate the occupants' behavior patterns in existing and modern buildings. Zhang et al. (2020) conducted questionnaire surveys to explore the occupant behavior related to cooking, lighting, water heating, space heating and cooling, and usage of other appliances in residential buildings. The results showed that the usage behavior was efficient for most of the appliances except for air conditioners. Besides, the occupants' habitual reactions were found to be consistent across all categories. These findings could assist policymakers in improving the energy efficiency of residential buildings on a city scale. While some researchers conducted questionnaire surveys to extract the dominant factors influencing the energy behavior of occupants, others adopted building energy simulation modeling. Fabi et al. (2017) assessed the impacts of different lifestyles of occupants and automation systems on the building energy performance. The needs and behavior of occupants were understood by conducting a questionnaire survey. Besides, the effect of occupant behavior on building energy use was studied using energy simulation software. It was concluded that the application of automation systems and the conscious behavior of the occupants could lead to significant energy savings in the building. Berger and Mahdavi (2020)

evaluated the building energy and indoor environmental performance using an agent-based model. The capability of this modeling technique to capture the dynamics and complexities of occupant behavior was highlighted. On the contrary, its limitation regarding the requirement of excessive computational loads was discussed. Finally, it was recommended to integrate agent-based modeling in virtual reality frameworks. The energy simulation is used to forecast the energy consumption of buildings at the design stage. However, building simulation can only yield accurate results if the building use is predictable (Degelman, 1999). It has been reported that the actual energy consumption could reach up to three times greater than the predicted consumption. This considerable discrepancy between the energy consumption at the design and as-built stages refers to the difference in the workmanship and installations, equipment and material choice, and the occupant behavior, which is often overlooked in the energy simulation process (Fabi et al., 2013; Cali et al., 2016). Artificial intelligence models have been widely applied in many applications, because of their ability to handle complex relationships. Artificial neural network (ANN) has been widely used in modeling occupant behavior. Its performance relies on input combinations, model architecture, activation function, and training algorithm of the network (Raza and Khosravi, 2015). Candanedo and Feldheim (2016) developed a model to determine the occupant's existence in an office room. The models were developed using linear discriminant analysis (LDA), classification and regression trees (CART), and random forest (RF) algorithms. The results affirmed that the LDA and RF models yielded the best and worst prediction accuracies in the two testing datasets, respectively. For temperature as well as CO<sub>2</sub> and temperature predictors, the CART model yields accuracies of 67-87% and 79-85%, respectively. Ekwevugbe et al. (2016) investigated the number of occupants in an office building using a network of sensors. The main features were selected using symmetrical uncertainty analysis and a genetic-based search. The accuracy of occupant detection was reported to be more than 74% using ANN. This approach had proved its usefulness in predicting building occupancy for improving heating, ventilation, and air conditioning operations and promoting energy savings.

Yang et al. (2014) simulated occupancy existence in the office environment using six machine learning models. It had been found that average accuracy improved with increasing the number of sensors. For temperature predictor, the decision tree model yielded accuracies of 55-65%. Meanwhile, its accuracy ranged between 69% and 89% for temperature and CO<sub>2</sub> predictors. Salamone et al. (2020) evaluated occupant thermal comfort in real and virtual scenarios by combining the application of the internet of things and machine learning. Different machine learning techniques were fed with users' biometric and feedback data about their thermal comfort along with the environmental parameters. Extra trees classifier was found to be the best algorithm. Besides, the most significant parameters affecting the personal thermal comfort perception in real and virtual scenarios were identified with an average accuracy of greater than 0.99. After reviewing the literature, the major contributions of this research are identified as follows:

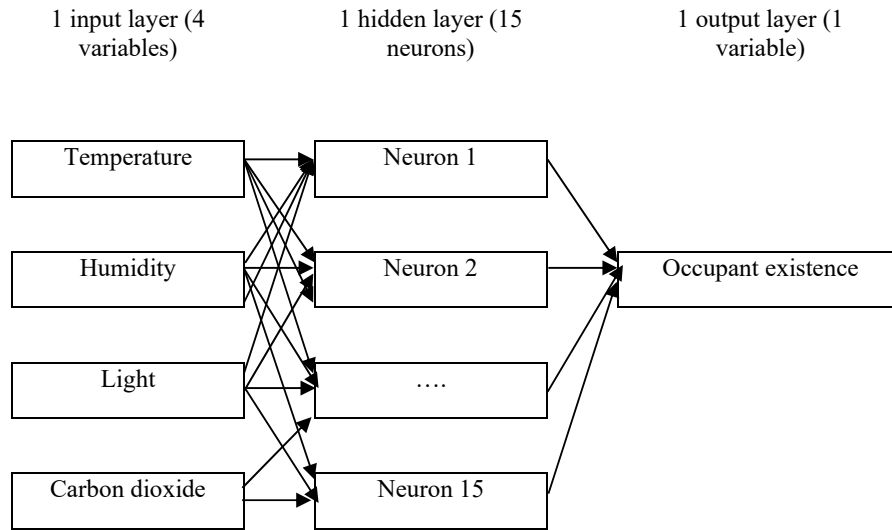
- (1) Applying hybrid classification models (i.e., machine learning models trained using metaheuristic algorithms) to investigate occupancy existence.
- (2) Comparing the performance of models using evaluation metrics.
- (3) Ranking the models by deploying an ensemble approach based on the half-quadratic theory.
- (4) Evaluating the accuracy of the proposed model to be validated against the surveyed literature using different predictor combinations.

### 3. Materials and methods

This research comprises employing a neural network model trained using a hybrid PSO-GSA algorithm for detecting occupancy existence in the office environment. The proposed model is evaluated against the neural network coupled with PSO and GSA algorithms using evaluation metrics. The rankings of the models are aggregated using a group decision making technique. The descriptions of these methods are presented in the next subsections. The author coded the hybrid neural network trained using metaheuristic algorithms and the ensemble ranking by using MATLAB R2015a software.

#### 3.1. Feedforward neural network

FFNN is one of the most common and powerful binary classification techniques (Rumelhart et al., 1985). As illustrated in Fig. 1, the topology of a neural network consists of three layers: an input layer, a hidden layer(s), and an output layer. The input layer contains the input data to the model, whereas the output layer comprises the model's outcome. The neural network has training and testing/validation phases. The learning/training phase is responsible for training the network to model the relationship between input(s) and output(s). Meanwhile, the recalling phase forecasts the output(s) based on the trained network (Elshaboury & Marzouk, 2020a).



**Fig. 1.** Neural network typology

### 3.2. Metaheuristic optimization algorithms

Two metaheuristic optimization algorithms are examined with occupancy detection, namely PSO and GSA. The descriptions of these algorithms are provided in the next subsection.

#### 3.2.1. Particle swarm optimization algorithm

PSO algorithm has many merits such as simplicity of implementation, requirement of setting few parameters, effectiveness in global search, and insensitivity to scaling of variables (Poli et al., 2007; Gong et al., 2009; Bai, 2010). It was developed to mimic the behavior of bird flocking (Kennedy & Eberhart, 1995). A flock of birds can search for food without colliding with each other. Birds share information within the flock to adjust their velocities and directions, searching for the best solution. In this algorithm, each particle represents a potential solution to the problem (Lynn and Suganthan, 2015). The objective function of the problem refers to the closeness of each particle to the best possible solution. Each particle is represented by four attributes; current position, current velocity, distance to particle best solution, and distance to the global best position to modify its position (Mirjalili et al., 2012). The particle and global best positions are automatically updated, whenever a better position is found, and stored for subsequent iterations. This process will continue until a termination criterion is met (Elshaboury et al., 2020).

#### 3.2.2. Gravitational search algorithm

GSA is a recent algorithm developed to solve non-linear problems (Rashedi et al., 2009). It is a simple algorithm that can find near-global optimum solutions (Kumar and Sahoo, 2014). It is based on Newton's law of gravity in physics, such that a mass in the space corresponds to a candidate solution. The mass of each solution is proportional to its fitness function value. There exists a gravitational force between masses such that heavier objects are characterized by bigger attraction forces compared to smaller ones (Mirjalili et al., 2012). Therefore, heavier masses (solutions close to the global minimum) attract the other masses to ensure convergence (García-Ródenas et al., 2021). The positions and velocities of particles are updated as a result of the interaction between particles and gravity. The fitness values of all particles are updated and recorded accordingly. The iteration continues until a stopping criterion is met (Mirjalili et al., 2012).

#### 3.2.3. Hybrid particle swarm optimization and gravitational search algorithm

Two inherent characteristics of population-based algorithms with social behavior, such as PSO and GSA, should be considered: the algorithm's ability to explore the whole problem area and to exploit the best solution. These two important characteristics should be included in a population-based algorithm to ensure obtaining the right solution. The basic idea of the hybrid PSO-GSA algorithm is to combine the exploitation ability of PSO with the exploration capability of GSA. It has proved its ability to solve a wide range of optimization problems (Jiang et al., 2014; Shaheen et al., 2021). In this algorithm, the agents (i.e., candidate solutions) are randomly initialized. Then, the gravitational constant, gravitational force, resultant forces, and accelerations among agents are calculated. The best solution shall be updated in each iteration. After that, the velocities and positions of agents are calculated and updated. This process terminates when meeting an end criterion. This hybrid algorithm ensures the following points: (1) considering the quality of solutions in the updating procedure, (2) attracting the agents exploring the search space to the agents near the good solutions, (3) moving slowly nearby a good

solution to exploit the global best solution, (4) storing the global best solution to be accessible at any time, (5) observing and tending towards the global best solution, and (6) balancing between the global and local searching abilities (Mirjalili et al., 2012).

### 3.3. Neural network trained using metaheuristic algorithms

As illustrated in Fig. 2, the flowchart of the proposed models comprises specifying the input and output parameters from the dataset. The dataset is then divided into training and testing groups to run the models. The neural network relies on adjusting initial weights, based on the differences between desired and actual values, using the gradient-search algorithm. However, the network performance is affected by the values of initial weights and biases. It is worth noting that the neural network can be trained in one of three ways: (1) manipulating weights and biases with fixed network architecture, (2) employing heuristic algorithms to obtain the optimum network structure, or (3) applying an evolutionary algorithm to tune the parameters of the learning algorithm (Mirjalili et al., 2012). Besides, selecting the most suitable approach depends on the classification problem and the dataset type. In this study, the first method is adopted to optimize the randomly assigned weights and biases in the FFNN model using PSO, GSA, and hybrid PSO-GSA algorithms. The output for the FFNN model is computed, and the error objective function is minimized using Eq. (1) (Lazzús, 2013). The optimization process stops when a termination criterion (i.e., minimum error function) is achieved.

$$MSE = \frac{\sum_{i=1}^{N_D} (y_i^{calc} - y_i^{exp})^2}{N_D} \quad (1)$$

where;  $MSE$  refers to the mean squared error,  $N_D$  refers to the number of data points, and  $y_i^{calc}$  and  $y_i^{exp}$  refer to the actual and desired values, respectively.

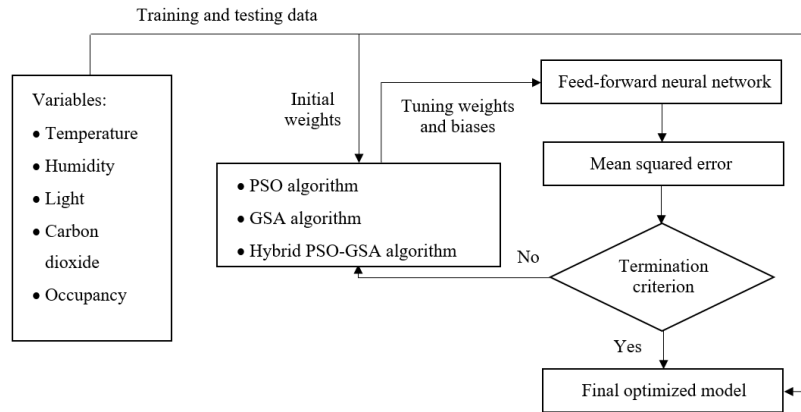


Fig. 2. Flowchart of the optimized machine learning models

### 3.4. Group decision making technique

The evaluation metrics rank the hybrid classification models in different orders. Therefore, the final rankings of the optimized neural network models are obtained using a new approach based on the half-quadratic theory (Mohammadi and Rezaei, 2020). The proposed approach uses the half-quadratic theory to calculate a weight for each of the classification models. It generates weights that satisfy the non-negativity and unit-sum properties. Since the weights are calculated without expert feedback, the recommended approach is objective. Furthermore, the proposed method, unlike averaging, is unaffected by outliers (i.e., ranks that are far from the rest of rankings) due to using stable half-quadratic functions. This implies that outliers are supposed to contribute less to the overall aggregated ranking. In addition to calculating the final ranking, a consensus index and a trust level are computed. The consensus index refers to the consensus of different methods on the final aggregated rankings while the trust level is an indicator of how reliable the final aggregated ranking is. This approach has proven its efficiency in some publications (Elshaboury & Marzouk, 2020b; Ridha et al., 2021). The proposed method develops the aggregated ranking by minimizing the Euclidean distance to each calculated ranking, as per Eqs. (2-4).

$$\alpha_m = \delta(\|R^m - R^*\|_2) \quad (2)$$

$$w_m = \alpha_m / \sum_j \alpha_j \quad (3)$$

$$R^* = \sum_m w_m \times R^m \quad (4)$$

where;  $m$  refers to the number of optimized models,  $\alpha_m$  refers to the half-quadratic auxiliary variable,  $R^m$  refers to the ranking acquired from each model,  $w_m$  refers to the weight of each model, and  $R^*$  refers to the ensemble ranking. The consensus and trust in the aggregated ranking are estimated using Equations 5 and 6, respectively.

$$C(R^*) = \frac{1}{KM} \sum_{k=1}^K \sum_{m=1}^M \frac{\mathcal{N}_\sigma(R_k^* - R_k^m)}{\mathcal{N}_\sigma(0)} \quad (5)$$

$$T(R^*) = \frac{1}{K} \sum_{k=1}^K \sum_{m=1}^M w_m \times \left( \frac{\mathcal{N}_\sigma(R_k^* - R_k^m)}{\mathcal{N}_\sigma(0)} \right) \quad (6)$$

where;  $C(R^*)$  and  $T(R^*)$  refer to the consensus and trust in the ensemble ranking, respectively,  $K$  refers to the number of alternatives, and  $\mathcal{N}_\sigma$  refers to the probability density of the Gaussian distribution, with zero mean and  $\sigma$  standard deviation.

#### 4. Model validation

The accuracy of the proposed model is evaluated using a confusion matrix (see Table 1) for the two assessed datasets. The sum of true positives (A) and true negatives (D) divided by the total number of observations determines the model accuracy, as per Equation 7 (Candanedo & Feldheim, 2016).

$$Accuracy = (A + D)/(A + B + C + D) \quad (7)$$

**Table 1.** Confusion matrix

Reference	Predicted	
	Not occupied	Occupied
Not occupied	A	B
Occupied	C	D

#### 5. Model development

This research presents the application of hybrid machine learning models for detecting occupant existence. As shown in Fig. 3, the model comprises ten major steps: 1) preparing the data matrix, 2) identifying the input and output parameters, 3) splitting the data into training and testing, 4) selecting the machine learning algorithm, 5) initializing the model parameters, 6) performing the hybrid classification models, 8) comparing the models using performance evaluation metrics, 9) aggregating the rankings for the models, and 10) suggesting the optimum classification model.

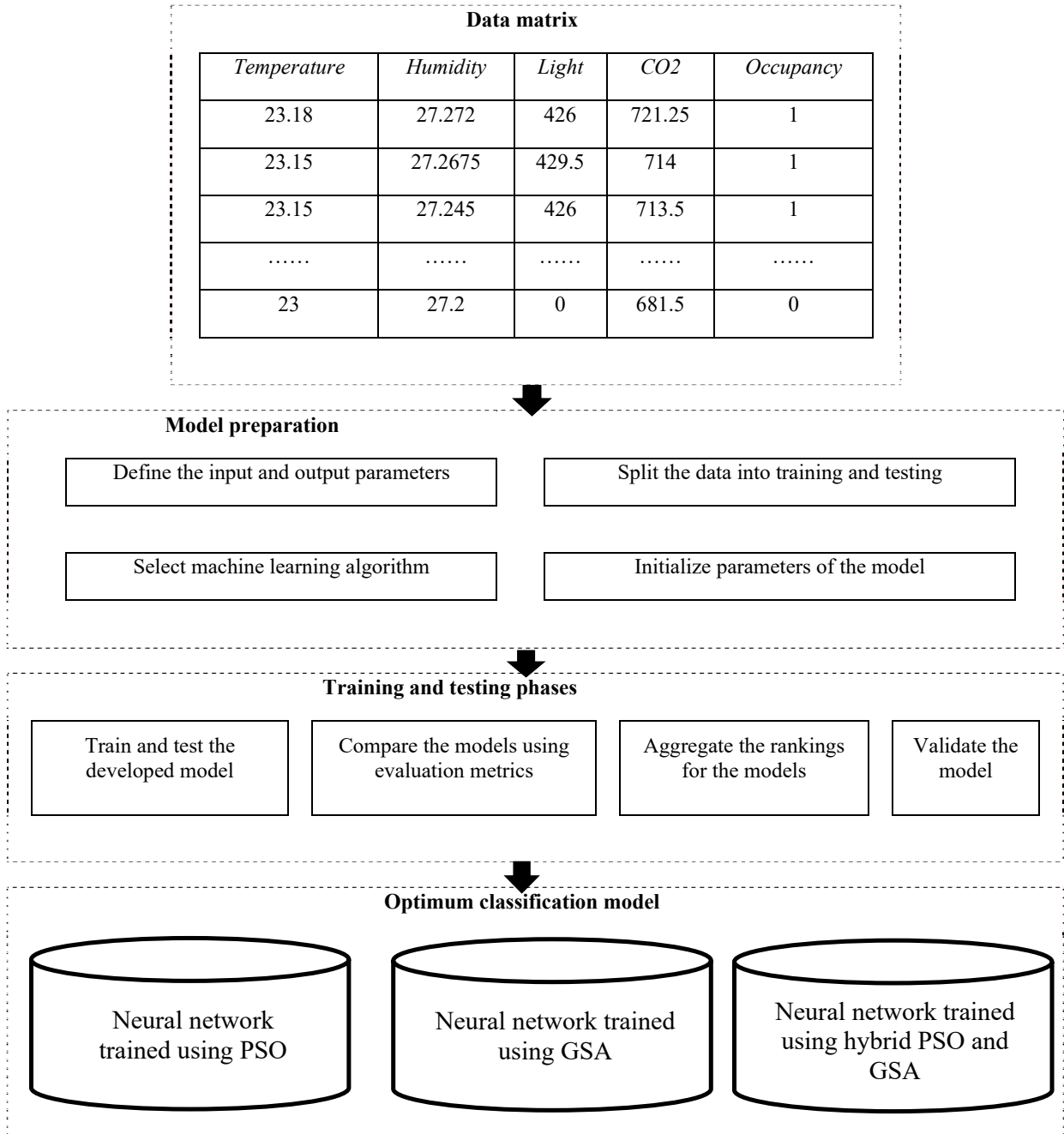
#### 6. Case study

The occupancy detection data used in this research was acquired from an office room in Belgium (Candanedo & Feldheim, 2016). The dimensions of this room are stated as follows: width of 5.85 meters, depth of 3.50 meters, and height of 3.53 meters. The data was collected in February during the winter. The input factors comprised of temperature (°C), humidity (%), light (lux), CO2 (ppm) while the output factor referred to the occupancy status (0 for non-occupied, 1 for occupied). The monitoring sensors were employed to acquire the temperature, humidity, light, and CO2 levels while the occupation status was determined using a digital camera, and then the pictures were labeled manually. As outlined in Table 2, the classification models were trained and tested using two data sets with the office door open and closed.

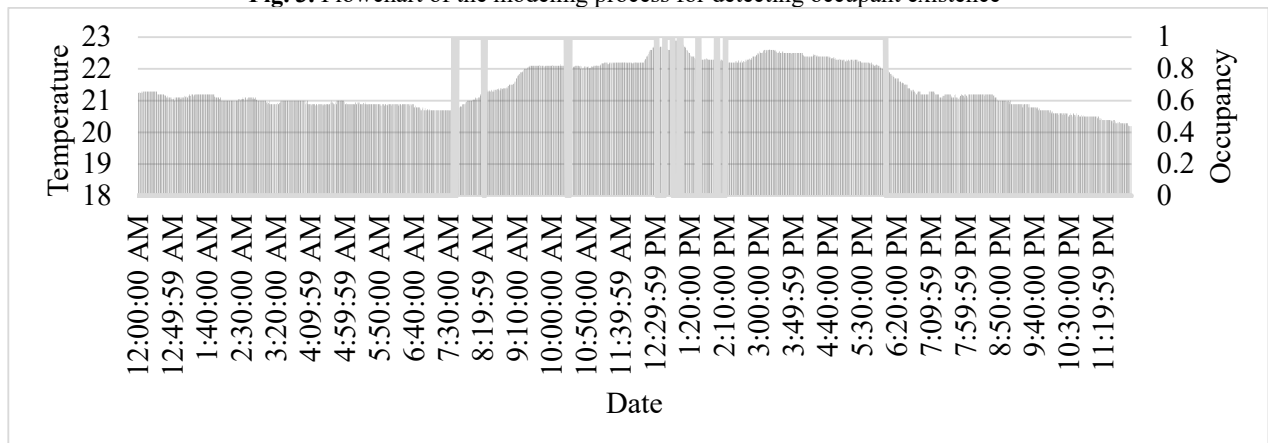
**Table 2**  
Descriptions of datasets

Dataset	Description	Number of observations	Distribution of classes	
			Non-occupied	Occupied
1	Most of the measurements were taken when the door was closed.	8143	0.79	0.21
2	Most of the measurements were taken when the door was open.	9752	0.79	0.21

As for the first dataset, Fig. 4 portrays data gathered between the evenings of two consecutive working days. As seen in the graphs, as the first person enters the building just before 8:00, all of the temperature, humidity, and CO2 sensors display a rise in their measurements. Besides, the lights in the office have been switched out. The slope of all the curves also rises as the second occupant arrives shortly after 09:00. The light and CO2 sensors detect a large drop while the room is not occupied between 13:00 and 13:30. Meanwhile, the humidity and temperature sensors record a slight reduction in their readings during this period. All of the sensors display a downward slope when space is left empty after 18:00. The most significant decline is exhibited by the light sensor. More statistical parameters about the five factors are depicted in Table 3.



**Fig. 3.** Flowchart of the modeling process for detecting occupant existence



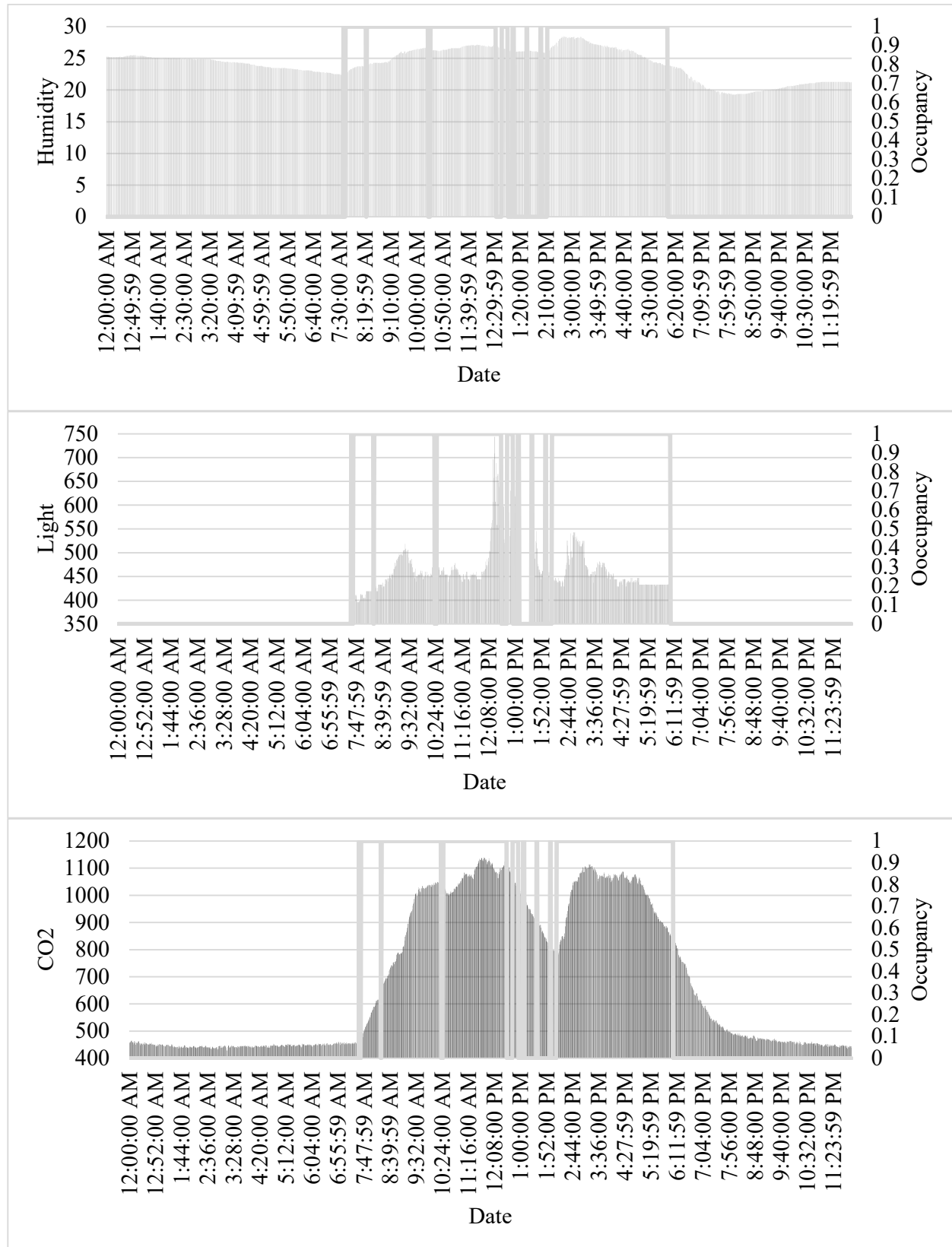


Fig. 4. Measurement profiles throughout a working day

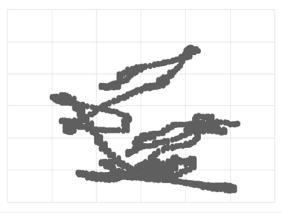
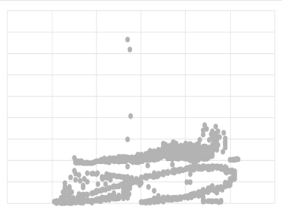
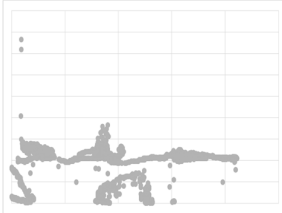
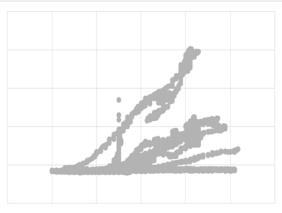
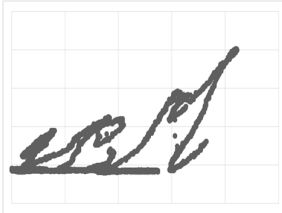
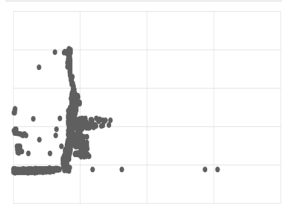


**Table 3**  
Statistical parameters of occupancy detection for the first dataset

Variable	Minimum	Maximum	Median	Mean	Standard deviation	Skewness	Kurtosis
Temperature	19.00	23.18	20.39	20.62	1.02	0.45	-0.81
Humidity	16.75	39.12	26.22	25.73	5.53	0.27	-0.93
Light	0.00	1546.33	0.00	119.52	194.76	1.24	0.12
Carbon dioxide	412.75	2028.50	453.50	606.55	314.32	2.38	5.78
Occupancy	0.00	1.00	0.00	0.21	0.41	1.41	-0.02

Table 4 displays the correlation among all the variables detecting occupancy existence. It is worth noting that variables with correlation coefficients between 0.5 and 0.7 can be deemed moderately correlated. This applies to the pairs: temperature and light, temperature and CO2, and light and CO2. Furthermore, variables with correlation coefficients between 0.3 and 0.5 have a low correlation. This applies to the pair: humidity and CO2. Correlation coefficients of less than 0.3 have little correlation and this is the case between humidity and light variables. Finally, there exists an inversely proportional relationship between the temperature and humidity variables.

**Table 4**  
Correlation between input variables for occupancy detection for the first dataset

Temperature	-0.14	0.65	0.56
	Humidity	0.04	0.44
		Light	0.66
			CO2

Statistical parameters about the input and output factors in the second dataset are depicted in Table 5. The correlation among these factors is illustrated in Fig. 5. It can be noted that light and temperature variables are strongly correlated. Furthermore, there is a little correlation between temperature and CO2, and light and CO2 pairs of variables. Finally, there exists an inversely proportional relationship between the temperature and humidity, humidity and light, and humidity and CO2 variables.

**Table 5**  
Statistical parameters of occupancy detection for the second dataset

Variable	Minimum	Maximum	Median	Mean	Standard deviation	Skewness	Kurtosis
Temperature	19.50	24.39	20.79	21.00	1.02	1.23	1.27
Humidity	21.87	39.50	30.20	29.89	3.95	0.12	-0.85
Light	0.00	1581.00	0.00	123.07	208.22	1.51	1.23
Carbon dioxide	484.67	2076.50	639.00	753.22	297.10	1.55	1.48
Occupancy	0.00	1.00	0.00	0.21	0.41	1.42	0.03

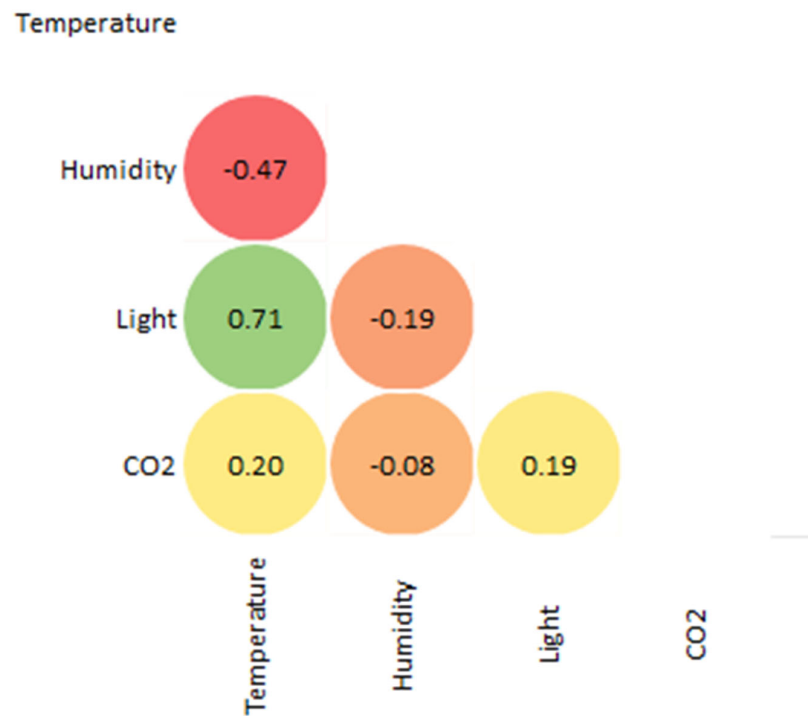


Fig. 5. Correlation between input variables for occupancy detection for the second dataset

## 7. Results and discussion

This research examines the performance of three meta-heuristic algorithms (i.e., PSO, GSA, and PSO-GSA) for training neural network model to detect occupant existence in an office room, Belgium. The parameters of these algorithms are summarized in Table 6. The models are tested using two datasets with the office door open and closed. Each dataset is classified into two groups such that the first 80% and 20% of the data points are used for training and testing purposes, respectively.

**Table 6**

Parameters of metaheuristic algorithms

PSO	GSA	PSO-GSA
Maximum iterations= 100	Maximum iterations = 100	Maximum iterations= 100
Population size= 30	Population size = 30	Population size= 30
Inertia weight= 2	Gravitational constant= 1	Inertia weight= 2
Maximum inertia weight= 0.9		Maximum inertia weight= 0.9
Minimum inertia weight= 0.5		Minimum inertia weight= 0.5
Global learning coefficient= 2		Gravitational constant= 1
Personal learning coefficient= 2		

For the first dataset, a comparison of the convergence curves of the classification models is visualized in Fig. 6. It shall be noted that the models are run ten times and the average results are computed to provide a fair comparison of the models. In each run, the number of iterations is set at 100 times. It is clear from Fig. 6 that the models seem to run horizontally after 20 iterations. PSO-GSA end up with a smaller MSE of 0.049 than PSO and GSA, which have MSEs of 0.207 and 0.279, respectively. This figure confirms that the FFNN-PSO-GSA model has the best convergence rate compared to other models.

The performance of the machine learning models is assessed using four different metrics, as summarized in Table 7. These metrics are the best MSE, average MSE, median MSE, and standard deviation MSE. These metrics evaluate whether or not the modeled values match or mismatch with the actual values. As shown in Table 7, there is a substantial improvement in the best MSE, average MSE, and median MSE values of the FFNN-PSO-GSA model compared to the other models. The results show that the FFNN-PSO-GSA model is associated with the lowest best MSE (0.049), average MSE (0.078), and median MSE (0.049) values. On the other side, the FFNN-GSA model exhibits better performance than the other two models in terms of the standard deviation of MSE (0.144). In this regard, the rankings of the three classification models are different according to the performance metrics (see Table 8). Therefore, this can be regarded as a multiple criteria decision making problem. These calculations are implemented in Microsoft Excel.

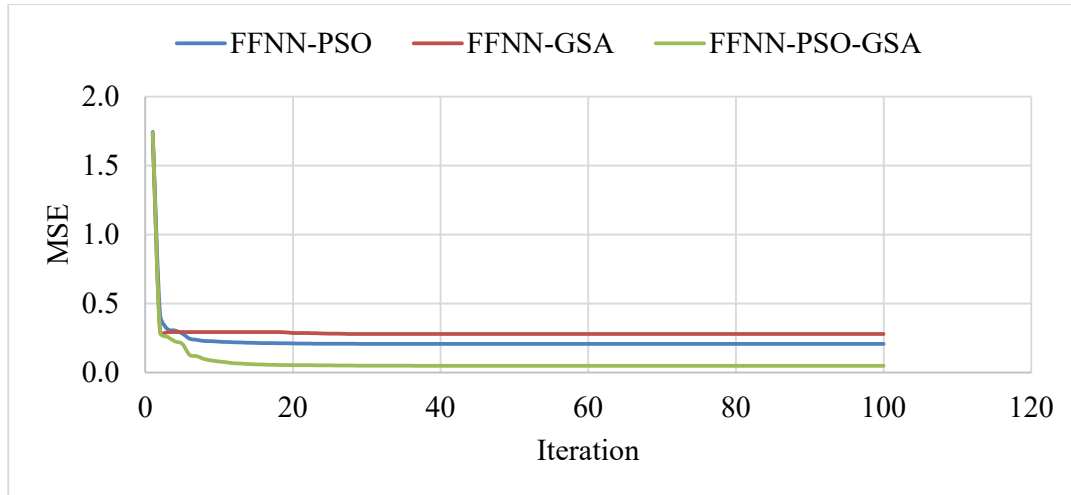


Fig. 6. Convergence curves of the classification models

Table 7

Evaluation metrics of classification models for the first dataset

Classification model	Best MSE	Average MSE	Median MSE	Standard deviation MSE
FFNN-PSO	0.207	0.230	0.207	0.155
FFNN-GSA	0.279	0.296	0.279	0.144
FFNN-PSO-GSA	0.049	0.078	0.049	0.173

In the decision making problem, the classification models are considered as the alternatives and the evaluation metrics are regarded as the attributes. Since the FFNN-PSO-GSA model obtains the minimum best, average, and median MSE, it is ranked first with respect to the referenced metrics. However, this model is associated with the highest standard deviation MSE, and it is therefore ranked as the worst model based on this metric. The rankings of the remaining models could be interpreted similarly. It is clear that the rankings of the classification models are not consistent, except for the FFNN-PSO model. Thus, the rankings of the hybrid machine learning models are aggregated using a new approach based on the half-quadratic theory. The results indicate that FFNN-PSO-GSA is the first-ranked classification model, as shown in Table 8. The final ranking obtains a consensus index of 0.83 and a trust level of 1.00, indicating consistency between the rankings.

Table 8

Final ranking of the classification models for the first dataset

Classification model	Best MSE	Average MSE	Median MSE	Standard deviation MSE	Final ranking
FFNN-PSO	2	2	2	2	2
FFNN-GSA	3	3	3	1	3
FFNN-PSO-GSA	1	1	1	3	1

For the second dataset, the FFNN-PSO-GSA model (best MSE= 7.54E-22, average MSE= 2.85E-06, median MSE= 7.70E-22, and standard deviation of MSE= 2.06E-05) exhibits superior performance compared to other classification models (see Table 9). Based on the results of all the metrics, the proposed model is ranked first, yielding a consensus index and a trust level of 1.00. It can be therefore concluded that the proposed FFNN-PSO-GSA model outperformed the other models for the two datasets. This can be attributed to its ability to obtain global minima rather than a local minimum error by being independent of the parameters of the initial weights and biases.

Table 9

Performance metrics and rankings of the classification models for the second dataset

Metrics	FFNN-PSO	FFNN-GSA	FFNN-PSO-GSA
Best MSE	1.08E-17	1.49E-07	7.54E-22
	2	3	1
Average MSE	2.98E-06	3.31E-06	2.85E-06
	2	3	1
Median MSE	1.39E-17	1.49E-07	7.7E-22
	2	3	1
Standard deviation MSE	2.14E-05	2.2E-05	2.06E-05
	2	3	1

The FFNN-PSO-GSA model is trained using different predictor combinations to be compared against the results reported in the literature. The suggested model yields a classification accuracy ranging between 98.47-98.73% using one predictor (i.e., temperature), higher than the reported value of 67-87% by Salamone et al. (2020) and substantially higher than 55-65% reported in Sembroiz et al. (2019). As for the temperature and CO<sub>2</sub> predictors, the proposed model has an accuracy

ranging between 85.45-94.03% compared to 79-85% in Salamone et al. (2020) and 69-89% in Sembroiz et al. (2019). It can be concluded that incorporating PSO and GSA algorithms into the classical FFNN model enhances the robustness of modeling occupancy detection. Therefore, the proposed optimized model can be widely applied in different applications.

## 8. Conclusion

Accurate occupancy detection is crucial in energy planning and management. It is an important factor for the facility manager to assess and then enhance the building's energy efficiency and indoor environmental quality. In this regard, this research examined occupancy existence in the office environment using a feedforward neural network (FFNN) model trained using particle swarm optimization (PSO), gravitational search algorithm (GSA), and hybrid PSO-GSA algorithm. The performance of the proposed models was examined using data obtained from an office room, Belgium. Two datasets were used for testing the models while the office door was open and closed. The data was divided into 80% and 20% for training and testing purposes, respectively. Four factors related to the temperature, humidity, light, and carbon dioxide levels were selected as input parameters for the models. The proposed classification models were evaluated using best, average, median, and standard deviation of mean squared error (MSE) metrics. For the first dataset, the results showed that the FFNN-PSO-GSA model (best MSE= 0.049, average MSE= 0.078, median MSE= 0.049, and standard deviation of MSE= 0.173) enhanced most of the performance indices compared to the other models. The rankings of the classification models were further examined using an approach based on the half-quadratic theory. The proposed model was identified as the best-ranked classification model, yielding a consensus index of 0.83 and a trust level of 1.00. The classification models were further tested and validated using a second dataset. The FFNN-PSO-GSA model (best MSE= 7.54E-22, average MSE= 2.85E-06, median MSE= 7.70E-22, and standard deviation of MSE= 2.06E-05) exhibited superior performance to that of the FFNN-PSO and FFNN-GSA models. Therefore, the suggested model was ranked first, yielding a consensus index and a trust level of 1.00. It can be concluded that the proposed FFNN-PSO-GSA model outperformed the other models for the two datasets. For temperature predictor, the FFNN-PSO-GSA model had accuracies of 98.47-98.73%, higher than the accuracies reported in the literature. Finally, it had accuracies of 85.45-94.03% which exceeded the reported range for CO<sub>2</sub> and temperature predictors. The main findings of the research could be summarized as follows:

- FNN-GSA did not perform properly because of the weak exploitation potential of the GSA algorithm.
- The performance of the FNN-PSO was better than that of the FNN-GSA owing to the precise exploitation ability of the PSO algorithm.
- FNN-PSO-GSA combined the strong exploitation and exploration abilities of the PSO and GSA algorithms, yielding outstanding performance in training neural networks.
- FNN-PSO-GSA boosted the problem of trapping in local minima and enhanced the convergence rate compared to the other models.
- FFNN-PSO-GSA improved the classification accuracy of the other models and it can be therefore used for occupancy detection modeling.

On the other hand, the limitation of the suggested model might be the data splitting technique for training and testing purposes. For future work, it is recommended to assess the performance of the developed models based on cross-validation. This helps in demonstrating the effectiveness of the classification models, particularly in addressing overfitting and underfitting.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

- Ahmad, J., Larijani, H., Emmanuel, R., Mannion, M., & Javed, A. (2020). Occupancy detection in non-residential buildings—A survey and novel privacy preserved occupancy monitoring solution. *Applied Computing and Informatics*.
- Amayri, M., Ploix, S., Kazmi, H., Ngo, Q. D., & Safadi, E. L. (2019). Estimating occupancy from measurements and knowledge using the bayesian network for energy management. *Journal of Sensors*, 2019.
- Andersen, R. (2012). The influence of occupant's behaviour on energy consumption investigated in 290 identical dwelling and 35 apartments. In: *10th International Conference on Healthy Buildings*, Brisbane, Australia.
- Bai, Q. (2010). Analysis of particle swarm optimization algorithm. *Computer and information science*, 3(1), 180.
- Berger, C., & Mahdavi, A. (2020). Review of current trends in agent-based modeling of building occupants for energy and indoor-environmental performance analysis. *Building and Environment*, 173, 106726.
- Brager, G., Paliaga, G., & De Dear, R. (2004). Operable windows, personal control and occupant comfort. *ASHRAE Transactions*, 110, 17-35.
- Branco, G., Lachal, B., Gallinelli, P., & Weber, W. (2004). Predicted versus observed heat consumption of a low energy multifamily complex in Switzerland based on long-term experimental data. *Energy and Buildings*, 36(6), 543-555.
- Cali, D., Osterhage, T., Streblov, R., & Müller, D. (2016). Energy performance gap in refurbished German dwellings: Lesson learned from a field test. *Energy and buildings*, 127, 1146-1158.

- Candanedo, L. M., & Feldheim, V. (2016). Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models. *Energy and Buildings*, 112, 28-39.
- Chen, S., Yang, W., Yoshino, H., Levine, M. D., Newhouse, K., & Hinge, A. (2015). Definition of occupant behavior in residential buildings and its application to behavior analysis in case studies. *Energy and Buildings*, 104, 1-13.
- Chen, Z., Jiang, C., & Xie, L. (2018). Building occupancy estimation and detection: A review. *Energy and Buildings*, 169, 260-270.
- De Simone, M., & Fajilla, G. (2018). Occupant behavior: A “new” factor in energy performance of buildings. Methods for its detection in houses and in offices. *Journal of World Architecture*, 2(2), 1-9.
- Degelman, L.O. (1999). A model for simulation of daylighting and occupancy sensors as an energy control strategy for office buildings. In: *Building simulation conference*, Kyoto, Japan, pp. 571-578.
- Delzendeh, E., Wu, S., Lee, A., & Zhou, Y. (2017). The impact of occupants’ behaviours on building energy analysis: A research review. *Renewable and Sustainable Energy Reviews*, 80, 1061-1071.
- Dziedzic, J. W., Da, Y., & Novakovic, V. (2019). Indoor occupant behaviour monitoring with the use of a depth registration camera. *Building and Environment*, 148, 44-54.
- Ekwevugbe, T., Brown, N., Pakka, V., & Fan, D. (2016). Advanced occupancy sensing for energy efficiency in office buildings. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 230(5), 410-423.
- Elshaboury, N., & Marzouk, M. (2020a). Comparing machine learning models for predicting water pipelines condition. In: *2nd novel intelligent and leading emerging sciences conference (NILES)*, Giza, Egypt, pp. 134-139.
- Elshaboury, N., & Marzouk, M. (2020b). Optimizing construction and demolition waste transportation for sustainable construction projects. *Engineering, Construction and Architectural Management*.
- Elshaboury, N., Attia, T., & Marzouk, M. (2020). Application of evolutionary optimization algorithms for rehabilitation of water distribution networks. *Journal of Construction Engineering and Management*, 146(7), 04020069.
- Erickson, V.L., Carreira-Perpinan, M.A., & Cerpa, A.E. (2011). Observe: Occupancy-based system for efficient reduction of HVAC energy. In: *10th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, Chicago, United States, pp. 258-269.
- Fabi, V., Barthelmes, V. M., Schweiker, M., & Corgnati, S. P. (2017). Insights into the effects of occupant behaviour lifestyles and building automation on building energy use. *Energy Procedia*, 140, 48-56.
- Fabi, V., Corgnati, S., Andersen, R., Filippi, M., & Olesen, B. (2011). Effect of occupant behaviour related influencing factors on final energy and uses in building. In: *Climamed*, Madrid, Spain, pp. 243-258.
- Fabi, V., D’Oca, S., Buso, T., & Corgnati, S. (2013). The influence of occupant's behaviour in a high performing building. In: *Climamed -7th Mediterranean Congress of Climatization*, Istanbul, Turkey.
- Fisk, W. J. (2000). Health and productivity gains from better indoor environments and their relationship with building energy efficiency. *Annual review of energy and the environment*, 25(1), 537-566.
- García-Ródenas, R., Linares, L. J., & López-Gómez, J. A. (2021). Memetic algorithms for training feedforward neural networks: an approach based on gravitational search algorithm. *Neural Computing and Applications*, 33(7), 2561-2588.
- Gong, D., Lu, L., & Li, M. (2009). Robot path planning in uncertain environments based on particle swarm optimization. In: *IEEE Congress on Evolutionary Computation*, Trondheim, Norway, pp. 2127-2134.
- Hong, T., & Lin, H.W. (2013). Occupant behavior: Impact on energy use of private offices, <https://www.osti.gov/servlets/purl/1172115#:~:text=The%20simulation%20results%20demonstrate%20that%20compared%20to%20the%20Standard%20workstyle> (accessed 15 January 2020).
- Hong, T., D’Oca, S., Turner, W. J., & Taylor-Lange, S. C. (2015). An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework. *Building and Environment*, 92, 764-777.
- Hong, T., Taylor-Lange, S. C., D’Oca, S., Yan, D., & Corgnati, S. P. (2016). Advances in research and applications of energy-related occupant behavior in buildings. *Energy and buildings*, 116, 694-702.
- International Energy Agency (2020). Energy efficiency, <https://www.iea.org/reports/energy-efficiency-2020> (accessed 17 January 2020).
- Janda, K. B. (2011). Buildings don't use energy: people do. *Architectural science review*, 54(1), 15-22.
- Jiang, S., Ji, Z., & Shen, Y. (2014). A novel hybrid particle swarm optimization and gravitational search algorithm for solving economic emission load dispatch problems with various practical constraints. *International Journal of Electrical Power & Energy Systems*, 55, 628-644.
- Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy*, 55, 184-194.
- Kennedy, J., & Eberhart, R.C. (1995). Particle swarm optimization. In: *IEEE International Conference on Neural Networks*, Perth, Australia, pp. 1942-1948.
- Kumar, Y., & Sahoo, G. (2014). A review on gravitational search algorithm and its applications to data clustering & classification. *International Journal of Intelligent Systems and Applications*, 6(6), 79-93.
- Labeodan, T., Zeiler, W., Boxem, G., & Zhao, Y. (2015). Occupancy measurement in commercial office buildings for demand-driven control applications—A survey and detection system evaluation. *Energy and Buildings*, 93, 303-314.
- Lazzús, J. A. (2013). Neural network-particle swarm modeling to predict thermal properties. *Mathematical and Computer Modelling*, 57(9-10), 2408-2418.

- Lindén, A. L., Carlsson-Kanyama, A., & Eriksson, B. (2006). Efficient and inefficient aspects of residential energy behaviour: What are the policy instruments for change?. *Energy policy*, 34(14), 1918-1927.
- Lynn, N., & Suganthan, P. N. (2015). Heterogeneous comprehensive learning particle swarm optimization with enhanced exploration and exploitation. *Swarm and Evolutionary Computation*, 24, 11-24.
- Martinaitis, V., Zavadskas, E. K., Motuzienė, V., & Vilutienė, T. (2015). Importance of occupancy information when simulating energy demand of energy efficient house: A case study. *Energy and Buildings*, 101, 64-75.
- Mirjalili, S., Hashim, S. Z. M., & Sardroudi, H. M. (2012). Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm. *Applied Mathematics and Computation*, 218(22), 11125-11137.
- Mohammadi, M., & Rezaei, J. (2020). Ensemble ranking: Aggregation of rankings produced by different multi-criteria decision-making methods. *Omega*, 96, 102254.
- Nisiforou, O. A., Poullis, S., & Charalambides, A. G. (2012). Behaviour, attitudes and opinion of large enterprise employees with regard to their energy usage habits and adoption of energy saving measures. *Energy and Buildings*, 55, 299-311.
- O'Brien, W., Gunay, B., Tahmasebi, F., & Mahdavi, A. (2017). Special issue on the fundamentals of occupant behaviour research. *Journal of Building Performance Simulation*, 10(5-6): 439-443.
- Oldewurtel, F., Sturzenegger, D., & Morari, M. (2013). Importance of occupancy information for building climate control. *Applied energy*, 101, 521-532.
- Page, J., Robinson, D., & Scartezzini, J.L. (2007). Stochastic simulation of occupant presence and behaviour in buildings. In: *Building simulation conference*, Beijing, China.
- Pereira, P. F., Ramos, N. M., Almeida, R. M., & Simões, M. L. (2018). Methodology for detection of occupant actions in residential buildings using indoor environment monitoring systems. *Building and Environment*, 146, 107-118.
- Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. *Swarm intelligence*, 1(1), 33-57.
- Polinder, H., Schweiker, M., Van der Aa, A., (2013). Occupant behavior and modeling. Total energy use in buildings, analysis and evaluation methods. Final Report Annex 53.
- Rashedi, E., Nezamabadi-Pour, H., & Saryazdi, S. (2009). GSA: a gravitational search algorithm. *Information sciences*, 179(13), 2232-2248.
- Raza, M. Q., & Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews*, 50, 1352-1372.
- Ridha, H. M., Gomes, C., Hizam, H., Ahmadipour, M., Heidari, A. A., & Chen, H. (2021). Multi-objective optimization and multi-criteria decision-making methods for optimal design of standalone photovoltaic system: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 135, 110202.
- Rinaldi, A., Schweiker, M., & Iannone, F. (2018). On uses of energy in buildings: Extracting influencing factors of occupant behaviour by means of a questionnaire survey. *Energy and Buildings*, 168, 298-308.
- Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1985). *Learning internal representations by error propagation*. In: David ER, James LM and Group CPR (eds). *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge, MA: MIT Press, pp. 318-362.
- Salamone, F., Bellazzi, A., Bellussi, L., Damato, G., Danza, L., Dell'Aquila, F., ... & Vitaletti, W. (2020). Evaluation of the visual stimuli on personal thermal comfort perception in real and virtual environments using machine learning approaches. *Sensors*, 20(6), 1627.
- Schweiker, M., & Shukuya, M. (2009). Comparison of theoretical and statistical models of air-conditioning-unit usage behaviour in a residential setting under Japanese climatic conditions. *Building and Environment*, 44(10), 2137-2149.
- Sembroiz, D., Careglio, D., Ricciardi, S., & Fiore, U. (2019). Planning and operational energy optimization solutions for smart buildings. *Information Sciences*, 476, 439-452.
- Shaheen, M. A., Hasanien, H. M., & Alkuhayli, A. (2021). A novel hybrid GWO-PSO optimization technique for optimal reactive power dispatch problem solution. *Ain Shams Engineering Journal*, 12(1), 621-630.
- Tam, V. W., Almeida, L., & Le, K. (2018). Energy-related occupant behaviour and its implications in energy use: A chronological review. *Sustainability*, 10(8), 2635.
- Tang, R., Wang, S., & Sun, S. (2021). Impacts of technology-guided occupant behavior on air-conditioning system control and building energy use. *Building Simulation*, 14(1), 209-217.
- Trivedi, D., & Badarla, V. (2020). Occupancy detection systems for indoor environments: A survey of approaches and methods. *Indoor and Built Environment*, 29(8), 1053-1069.
- Yang, Z., Li, N., Becerik-Gerber, B., & Orosz, M. (2014). A systematic approach to occupancy modeling in ambient sensor-rich buildings. *Simulation*, 90(8), 960-977.
- Zhang, Y., Bai, X., & Mills, F. P. (2020). Characterizing energy-related occupant behavior in residential buildings: Evidence from a survey in Beijing, China. *Energy and Buildings*, 214, 109823.
- Zhang, Y., Bai, X., Mills, F. P., & Pezzey, J. C. (2018). Rethinking the role of occupant behavior in building energy performance: A review. *Energy and Buildings*, 172, 279-294.



© 2022 by the authors; licensee Growing Science, Canada. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).