Article

Estimating the Heating Load of Buildings for Smart City Planning Using a Novel Artificial Intelligence Technique PSO-XGBoost

Le Thi Le [1],*, Hoang Nguyen [2],* , Jian Zhou [3] , Jie Dou [4] and Hossein Moayedi [5]

[1] Thanh Hoa University of Culture, Sports and Tourism, Thanh Hoa 440000, Vietnam
[2] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
[3] School of Resources and Safety Engineering, Central South University, Changsha 410083, China
[4] Civil and Environmental Engineering, Nagaoka University of Technology, 1603-1, Kami-Tomioka, Nagaoka, Niigata 940-2188, Japan
[5] Centre of Tropical Geoengineering (Geotropik), School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
* Correspondence: lethiledhvhttdl@gmail.com (L.T.L.); nguyenhoang23@duytan.edu.vn (H.N.)

Abstract: In this study, a novel technique to support smart city planning in estimating and controlling the heating load (HL) of buildings, was proposed, namely PSO-XGBoost. Accordingly, the extreme gradient boosting machine (XGBoost) was developed to estimate HL first; then, the particle swarm optimization (PSO) algorithm was applied to optimize the performance of the XGBoost model. The classical XGBoost model, support vector machine (SVM), random forest (RF), Gaussian process (GP), and classification and regression trees (CART) models were also investigated and developed to predict the HL of building systems, and compared with the proposed PSO-XGBoost model; 837 investigations of buildings were considered and analyzed with many influential factors, such as glazing area distribution (GAD), glazing area (GA), orientation (O), overall height (OH), roof area (RA), wall area (WA), surface area (SA), and relative compactness (RC). Mean absolute percentage error (MAPE), root-mean-squared error (RMSE), variance account for (VAF), mean absolute error (MAE), and determination coefficient (R^2), were used as the statistical criteria for evaluating the performance of the above models. The color intensity, as well as the ranking method, were also used to compare and evaluate the models. The results showed that the proposed PSO-XGBoost model was the most robust technique for estimating the HL of building systems. The remaining models (i.e., XGBoost, SVM, RF, GP, and CART) yielded more mediocre performance through RMSE, MAE, R^2, VAF, and MAPE metrics. Another finding of this study also indicated that OH, RA, WA, and SA were the most critical parameters for the accuracy of the proposed PSO-XGBoost model. They should be particularly interested in smart city planning as well as the optimization of smart cities.

Keywords: energy efficiency; PSO-XGBoost; meta-heuristic algorithm; heating load; smart city

1. Introduction

Smart cities are the development goals of many countries around the world [1]. Intelligent systems have been widely researched and applied in smart cities to provide a better quality of life, as well as to bring higher economic efficiency [2–6]. One of the critical issues of smart cities is the efficient use of energy by buildings [7–9]. Of those, energy for cooling or heating the buildings is significant since they take a considerable part of the total energy [10]. In winter, the demand for energy for the heating load (HL) is significant [11]. The ineffective use of HL not only results in economic losses but also a threat

to the surrounding environment [12]. Therefore, accurately predicting the HL of the buildings is a challenge for smart cities.

For estimating HL of building systems, many efforts of scientists have been implemented to research and propose artificial intelligence (AI) systems, as advanced alternative techniques to control the HL of buildings. Yang et al. [13] attempted to develop an artificial neural network (ANN) to predict the heating system in building with a promising result. An ANN model with the back-propagation algorithm was optimized based on the learning rate, moment, hidden layer number, nodes of the hidden layer, bias, and learning factors, for their aim. Finally, the heating system was accurately predicted with a determination coefficient (R^2) of 0.988 in their study. In another study, Braun et al. [14] used a regression analysis to evaluate energy consumption in the UK. Unlike the previous study, their study investigated the energy consumption of a supermarket based on electricity and gas used. The humidity ratio was analyzed and used based on the relative humidity and the dry-bulb temperature for their aim. A long-term prediction of energy consumption for the supermarket was implemented from 2030 to 2059. Their findings showed that whereas gas consumption reduces 13%, electricity use will increase by 2.1%. Based on ANN techniques and their ensemble, Jovanović et al. [15] successfully developed a model to predict heating energy consumption, using feedforward backpropagation neural network (FFNN), adaptive neuro-fuzzy interference system (ANFIS), and radial basis function network (RBFN); 35 buildings with the total area of 300,000 m^2 were considered in their study. Different input variables were used for each type of neural networks. Subsequently, three neural network models were developed (i.e., FFNN, RBFN, and ANFIS). Eventually, the outcome predictions of the three neural network models were integrated to generate the new ensemble model with high accuracy (i.e., root-mean-squared error—RMSE = 8169.1, MAPE = 5.3204, and R^2 = 0.9845). A dynamic ANN model with the use of the Taguchi method was also developed by Sholahudin, Han [16], for predicting the HL of a building with high accuracy. Dew point temperature, diffuse horizontal radiation, dry-bulb temperature, wind speed, and direct normal radiation, were used as the input variables in their study for the development of an ANN model to predict HL. Then, the Taguchi method was applied to define the effect of the input variables on HL. As a result, the wind speed and outdoor temperature were determined as the most influential parameters on HL in their study, and an ANN model was developed based on the selected parameters. Finally, a promising result with an RMSE of 37.3 and R^2 of 0.999. Gunay et al. [17] also developed an inverse black-box technique to model HL and the cooling load of five office buildings. Electrical load data, hourly cooling, heating, as well as the solar irradiance, wind speed, temperature, and humidity, were collected for their aim. They concluded that their black-box technique could model the HL and cooling load of the buildings very well, and it was introduced as an alternative operational decision-making process. In another study, Ahmad, Chen [18] used six data-mining methods to predict the cooling load and HL of a building in a short-term and medium-term, including Gaussian process (GP), bagged tree (BT), ANN, boosted tree (BST), multiple linear regression (MLR), and tree bagger (TB). Ultimately, the GP model was found as the best model in their study with an RMSE of 0.782, 0.771, 2.102; R^2 of 0.997, 0.996, and 0.987, for 7-day, 14-day, and 1-month periods, respectively. Kim et al. [19] used a reduced-order method for modeling heating and cooling energy demand with an excellent result. An error of about 0.43% for the annual loads was found in their study. By the optimization approach, Bui et al. [20] developed a hybrid intelligence model, namely M5Rules-GA, using a genetic algorithm to optimize the M5Rules model for estimating HL of buildings using 768 experimental datasets. A variety of Ai techniques, such as GP, multi-layer perceptron (MLP), radial basis function regression (RBFR), and M5-Rules models were also developed to estimate HL. As a result, they found that their proposed M5Rules-GA model provided the best performance with an RMSE of 0.0548 and R^2 of 0.998. Using another optimization algorithm, namely the evolutionary grey wolf algorithm (GWA), Jitkongchuen, Pacharawongsakda [21] predicted heating and cooling loads of building with high reliability. Six other AI techniques were also considered and developed to estimate the HL for their comparison purposes, including geometric semantic genetic programming (GSGP), ANN, evolutionary multivariate adaptive regression splines (EMARS), support

vector regression (SVR), MLP, and random forests (RF). The same dataset (i.e., 768 experimental datasets) was used to develop the HL predictive models in their study. Their results indicated that the GWA technique can predict HL of buildings more accurately than the other models (i.e., GSGP, ANN, EMARS, SVR, MLP, and RF). Similar works for predicting HL of buildings can be found in the following literatures [22–29].

Although soft computing models for estimating the HL of the building system were developed; however, they have not been evaluated comprehensively on subjective and objective factors, such as glazing area distribution (GAD), glazing area (GA), orientation (O), overall height (OH), roof area (RA), wall area (WA), surface area (SA), and relative compactness (RC). Furthermore, new smart systems with high efficiency are always the target of engineers and scientists to optimize building systems, as well as smart cities planning. Therefore, this study developed and proposed a new technique to estimate HL, based on an evolutionary algorithm (particle swarm optimization—PSO) and an extreme gradient boosting (XGBoost) model, namely PSO-XGBoost model. Careful consideration of GAD, GA, O, OH, RA, WA, SA, and RC were implemented in the present study for estimating HL of building systems. Five other AI techniques include XGBoost, support vector machine (SVM), random forest (RF), Gaussian process (GP), and classification and regression trees (CART), were also investigated and developed to predict the HL of building systems and compared with the proposed PSO-XGBoost model.

The structure of this study was organized as follows: Section 1 presents the reason for conducting this study and related works; Section 2 presents the details of data collection and properties of the database used; Section 3 presents the background of the methods used; Section 4 proposes the framework of the new technique to estimate the HL of buildings systems (i.e., PSO-XGBoost); Section 5 introduces several performance indices for evaluating the accuracy of the developed models; The results and discussion are presented in Section 6; finally, conclusions and remarks are given in Section 7.

2. Experimental Database

To implement this study, a database includes 768 buildings was analyzed by Tsanas, Xifara [30], which was used with nine parameters, i.e., glazing area distribution (GAD), glazing area (GA), orientation (O), overall height (OH), roof area (RA), wall area (WA), surface area (SA), relative compactness (RC), and heating load (HL). Sixty-nine other buildings with similar parameters were also analyzed and investigated in Vietnam during the winter (2018), using Ecotect computer software. Ultimately, in total, we used 837 simulated buildings in this study for estimating the HL of buildings; 12 forms of building shape were surveyed with the different RC, as shown in Figure 1. All buildings have different dimensions and surface areas but have the same materials for each building.

For this aim, GAD, GA, O, OH, RA, WA, SA, and RC were considered as the input variables to estimate the HL of buildings according to the recommendations of the previous studies [30–33]. As introduced above, most buildings have shapes of orthogonal polyhedral, as shown in Figure 1. Different types of building were interpreted through the RC, and is calculated as follows:

$$RC = \frac{6V^{\frac{2}{3}}}{A} \tag{1}$$

where V denotes the volume of the building, m^3; A denotes the surface area (i.e., SA) of the buildings, m^2.

The SA parameters were calculated based on the floor, roof, wall areas, and the overall building height (i.e., OH). To implement this study, four primary orientations were surveyed, including east, west, south, and north. A data encryption procedure under the numerical for orientations has been performed with the values of 1, 2, 3, 4, for east, west, south, north, respectively (Table 1). Six glazing area (i.e., GA) percentages were recorded, including 0%, 10%, 15%, 25%, 40%, and 50%. Besides, the glazing area distribution (i.e., GAD) also investigated five forms of distribution, including south, uniform, west, north, and east. They were also encrypted as 1, 2, 3, 4, 5, for east, uniform, south, north, and west, respectively. The remaining parameters (i.e., RA and WA) were calculated based on their

dimensions in the AutoCAD environment. To determine the HL of the buildings, the Ecotect computer software was used to simulate the energy efficiency of the 837 buildings. Box and whisker plots of the dataset are shown in Figure 2.
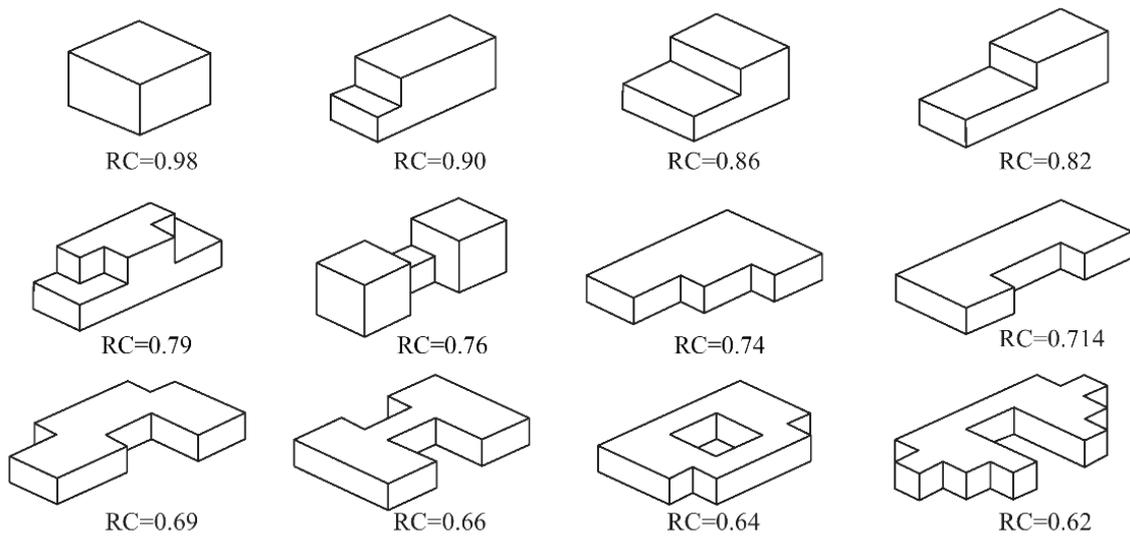

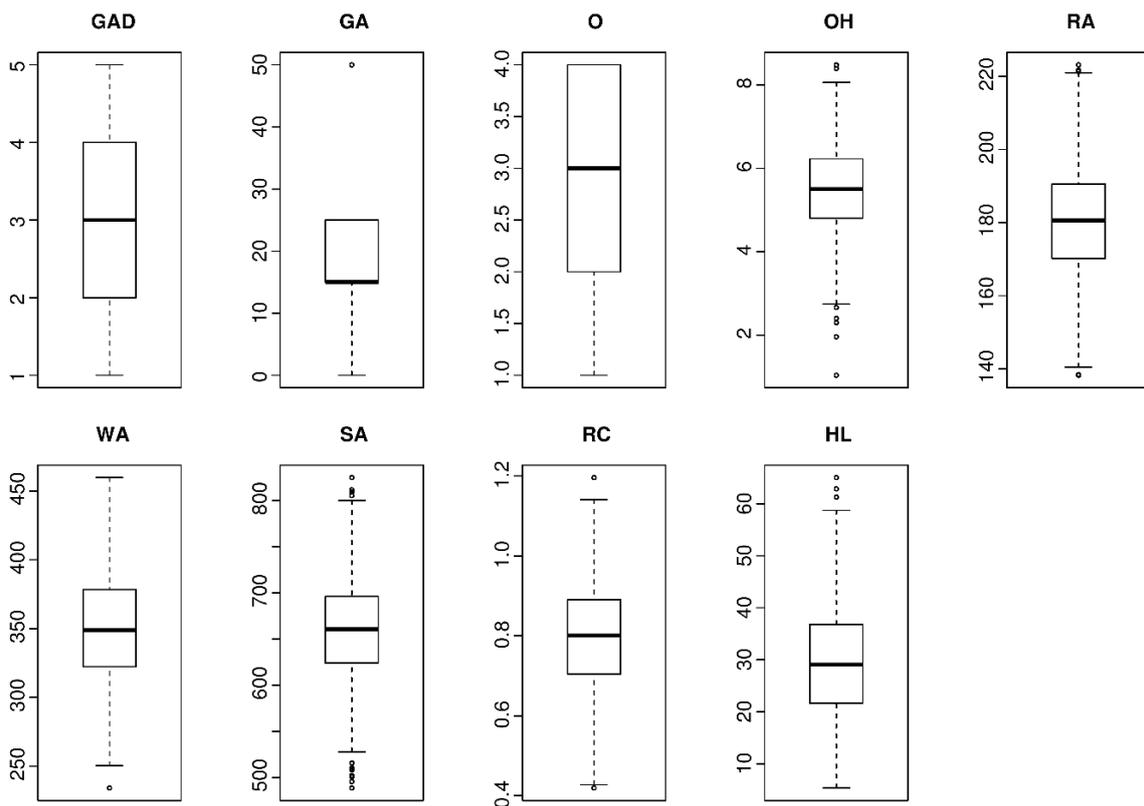**Figure 1.** Individual building shapes with different relative compactness [31].

Figure 2. Box and whisker plots of the dataset used for estimating the heating load (HL) in this study.

3. Background of the Methods Used

As stated above, this study performs an HL estimation of buildings by the use of six techniques, i.e., PSO-XGBoost, XGBoost, SVM, RF, GP, and CART. However, due to the details of SVM, RF, GP, and CART were introduced in many previous kinds of literature [34–42]; so, some description of them

were added in the present study. Since the main objective of this study was to develop the new hybrid technique, i.e., PSO-XGBoost; therefore, the details of PSO and XGBoost were presented in this section.

3.1. Particle Swarm Optimization (PSO) Algorithm

PSO is a swarm algorithm inspired by the behavior of the particles/social animals, such as fish, or birds. It is a stochastic optimization method that was introduced and developed by Eberhart, Kennedy [43]. It was also classified as one of the metaheuristic techniques. The main idea of the PSO algorithm is to make better social information sharing among individuals in a crowd. Each individual act as a particle in the swarm. Then, they implement a searching procedure in a searching space. During the search process, they share information and experience to update better locations [44]. Thus, it was also considered as an evolutionary computation technique in the statistical community [44–49]. The PSO algorithm implements five steps for optimal searching:

- Step 1: Initialize the aboriginal population and velocity of particles. Subsequently, compute the particles fitness and detect the most logical place as local best and global best.

- Step 2: Every particle flies roundly in the space of search with the initial velocity as established in the first step. The speed depends on the local best and global best. For each loop, the best solution is corresponding to the local best, and the best spot of the particle obtained so far is corresponding to the global best. In other words, corresponding to the local best and global best in each loop, the velocity is updated in this step. It is described as:

$$v_j^{i+1} = w v_j^{(i)} + (c_1 \times r_1 \times (local\ best_j - x_j^{(i)})) + (c_2 \times r_2 \times (\text{global } best_j - x_j^{(i)})), v_{\min} \leq v_j^{(i)} \leq v_{\max}$$

where $x_j^{(i)}$ indicates the position of the particles; $v_j^{(i)}$ denotes the speed of the jth particle at the ith iteration; w represents the coefficient of inertial weight; i is the number of repetitions; and r_1 and r_2 symbolize numbers in the interval [0,1].

- Step 3: After the new velocity is calculated and updated, the particles fly in the search space with the new speed. Corresponding to each position, the fitness of them is determined and updated through a fitness function (i.e., RMSE).

- Step 4: Update local best and global best for the better position with lower RMSE. The local best can be updated as:

$$x_j^{i+1} = x_j^{(i)} + v_j^{(i+1)}; \ j = 1, 2, \ldots, n \tag{2}$$

- Step 5: Check the satisfaction of the searching. If the fitness of the particle is the best (i.e., lowest RMSE), stopping the searching. Otherwise, return to step 2.

The pseudo code of the PSO algorithm for optimization of searching is shown in Figure 3.

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

Figure 3. The particle swarm optimization (PSO) pseudo-code for the optimization process [50].

3.2. Extreme Gradient Boosting Machine (XGBoost)

Based on the ideas of a gradient boosting machine [51,52], Chen, He [53] improved and introduced the XGBoost algorithm as a robust decision tree. Unlike the gradient boosting machine, XGBoost can

run in parallel based on the constructed boosted trees. It can handle complex data at high speed and accuracy. The XGBoost algorithm can be described as follows [54]:

Given a dataset with n examples and m features $D = \{(x_i, y_i)\}$ $(|D| = n, x_i \in R^m, y_i \in R)$, K additive functions will be used to predict the output values of a tree ensemble model as follows:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \ f_k \in F \tag{3}$$

where F is the regression trees space. It is calculated as:

$$F = \left\{ f(x) = \omega_{q(x)} \right\} (q : R^m \to T, \ \omega \in R^T) \tag{4}$$

where q denotes for the structure of each tree; T denotes for the number of leaves in the tree, and f_k is a function that corresponds to an independent tree structure q and leaf weights w.

To reduce errors of ensemble trees, the objective function is calculated in the XGBoost model:

$$L^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{5}$$

where l is a differentiable convex objective function to determine the error between predicted and measured values; y_i and $\hat{y}_i$ is regulated and predicted values, respectively; t denotes the repetitions in order to minimize the errors; and Ω is the complexity penalize with the regression tree functions:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|w\|^2 \tag{6}$$

3.3. Support Vector Machine (SVM)

SVM is a machine learning algorithm based on the principle of minimizing structural risk to generalize a limited number of samples better and was proposed by Cortes, Vapnik [55]. It can solve both classification and regression problems [56,57]. Since the objectives of this study was numeric, so, support vector machine for regression problems were applied. In the general regression learning problem, the learning machine-based training data from which it attempts to learn the input-output relationship (dependency, mapping, or function) $f(x)$. A training data set $D = \{[x(i), y(i)] \in \mathfrak{R}^n * \mathfrak{R}, i = 1, \dots, l\}$. However, in the case of SVM's regression, the error of approximation was measured rather than the margin used in classification. Therefore, a linear regression hyperplane by minimizing Vapnik's insensitivity loss function is shown in formula (7).

$$\mathfrak{R}_{w,\psi,\psi^*} = \left[\frac{1}{2}\|W\|^2 + C\left(\sum_{i=1}^{l} \psi_i + \sum_{i=1}^{l} \psi_i^*\right)\right] \tag{7}$$

From the above function it can be seen the slack variables ψ_i and ψ_i^* also related to Lagrange multipliers v_i and v_i^*, and note also the trade-off between an approximation error and the weight vector norm $\|W\|$ were affected by constant C. While, in the case of nonlinear regression, solving the maximum value of dual Lagrange function is equivalent to solving the learning problem, the dual Lagrangian in input space is:

$$L_d(v) = -0.5v^T H v + \sum_{i=1}^{l} v_i \tag{8}$$

Subject to : $\quad v = \left[v_1, v_2, \dots v_l, v_1^*, v_2^*, \dots, v_1^*\right]^T$

$$H = [G \ -G; -G \ G]$$

$$G = \begin{bmatrix} G_{11} & \cdots & G_{1l} \\ \vdots & G_{ii} & \vdots \\ G_{l1} & \cdots & G_{ll} \end{bmatrix}$$

$$\text{where } G_{ij} = \Phi^T(x_i)\Phi(x_j)$$

After calculation of Lagrange multipliers v_i and v_i^*, an optimal weight vector of the regression hyperplane could be expressed with the formula (9).

$$w_0 = \sum_i^l (v_i - v_i^*)x_i \tag{9}$$

In order to create the best nonlinear regression function, the most crucial method is using the kernel function, there are some formulas kernel functions had been proposed, as listed in Table 1.

Table 1. Kernel functions.

Kernel Functions	Type
$K(x, y) = x \cdot y$	Linear kernel
$K(x, y) = [(x \cdot y) + 1]^d;\ d = (1, 2, \ldots)$	Polynomial kernel
$K(x, y) = \exp[\frac{-\|x-y\|^2}{\sigma^2}]$	Radial primary kernel function
$K(x, y) = \tanh[a(x \cdot y) - \delta]$	Two-layer neural kernel

We replaced x_i by the corresponding feature vector in a feature space and made some changes with a transform $G_{ij} = x_i^T x_j$ to $G_{ij} = \Phi^T(x_i)\Phi(x_j)$; then it follows that $G_{ij} = K(x_i, x_j)$ based the above kernel function $(K(x_i, x_j) = \Phi^T(x_i)\Phi(x_j))$.

Besides, an optimal weighting vector of the kernel's expansion, as shown in formulas (10), can be calculated with Lagrange multiplier.

$$V_0 = v - v^* \tag{10}$$

In this way, the regression function combined weighting vector and kernels function as follows,

$$f(x, w) = w_0^T \Phi(x) + b = \sum_{i=1}^l (v_i - v_i^*)\Phi^T(x_i)\Phi(x) + b$$
$$= \sum_{i=1}^l (v_i - v_i^*)K(x_i, x) + b \tag{11}$$

3.4. Random Forest (RF)

Breiman [58] provided RF that is an ensemble machine learning method, which is a decision trees algorithm and has a proper performance for classification and regression purposes. The RF algorithm consists of many decision trees through bootstrap aggregation (bagging) [59,60]. For arriving at a final decision, this algorithm includes the combination of a result set of various decision trees, and each tree is commonly trained with choosing variables as well as data samples, randomly, from an initial training database [61].

RF should be applied as follows to estimate the HL of buildings:

(a): For ensuring the forest richness, the number of trees indicated was determined.

(b): The remaining amount that had been utilized for validation were named out-of-bag (OOB) data, bootstrap with replacement by the principal HL training database.

(c): A non-pruning regression was expanded to enhance each node, for each bootstrap sample.

(d): OOB should be utilized to estimate the HL and mean of whole trees at each bootstrap iteration.

(e): Performance indices including RMSE, R^2, MAE, VAF, and MAPE could be used to analyze the predicted HL amounts on OOB.

3.5. Gaussian Process (GP)

AI has different non-parametric models like GP. This non-parametric model consists of random parameters along with finite element by a Gaussian distribution [62]. According to the covariance function, and also mean function, $h(x)$, GP can be determined as follow:

$$f(x) \sim GP(h(x), c(x, x'))$$ (12)

The GP starts for encrypting the uncertainty before performing the training in the case of regression. There is a relation between the function and the data that can be indicated by it. The Bayes' rule can be employed for updating the confidences by the function. Moreover, the posterior distribution can be calculated by the Bayes' rule [63].

In previous works, the non-parametric model of GP was not used for the prediction of HL. Therefore, in this paper, we studied the usages of GP associated with the radial basis function that can be described in Table 2.

3.6. Classification and Regression Tree (CART)

As one of the most popular statistical methods, the CART (classification and regression tree) has been widely employed to handle classification and regression problems [64]. Inspired by the growth process of trees, the construction of CART trees generally consists of roots, leaves, branches, and nodes. By means of binary recursive classification techniques, CART algorithms divide each sample set into two sub-sample sets, and thus, there are two branches in each none-leaf node [65,66].

The CART was first proposed and developed by Breiman et al. (1984). Different from traditional statistical methods, the CART is mainly established by many binary decision trees and is easy to interpret and understand. Especially for the complex data and significant variables, the CART tends to show better prediction accuracy compared with previous prediction methods. The most prominent advantages of the CART are the keen ability to distinguish the importance of variables and eliminating outliers.

4. Proposing the PSO-XGBoost Framework for Estimating HL

In this section, the PSO-XGBoost model was proposed to predict the HL of the building systems. For this stage, an initial XGBoost model was developed first; then, its hyper-parameters were optimized by the PSO algorithm. In the initial XGBoost model, seven hyper-parameters were considered and optimized, including subsample ratio of columns (δ), boosting iterations (k), minimum loss reduction (γ), max tree depth (d), shrinkage (η), subsample percentage (ς), and a minimum sum of instance weight (μ). For determination of the optimal values of these parameters, the particles fly in the search space and exchange the experiences. For each position, they calculate the fitness of particles via a fitness function, i.e., RMSE in equation 13. For each value of hyper-parameters, a corresponding RMSE value was computed, and the best fit model is corresponding to the lowest RMSE. The scheme of the development of the PSO-XGBoost model for estimating the HL of buildings is shown in Figure 4.
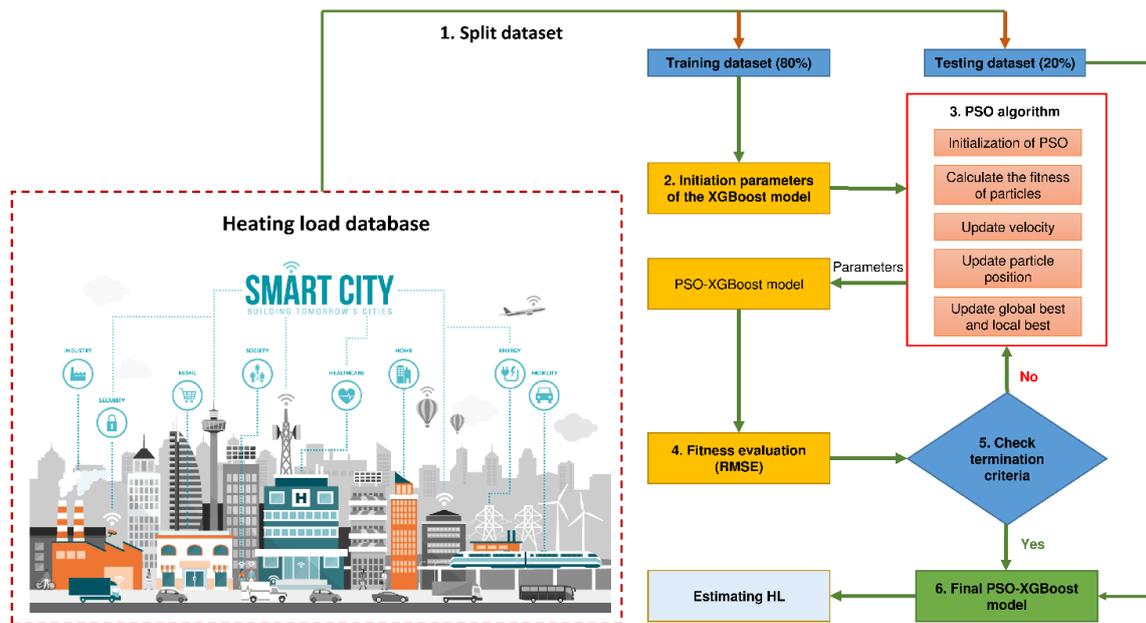
Figure 4. Scheme of the development of the PSO and extreme gradient boosting (PSO-XGBoost) model for estimating the HL of building systems.

5. Performance Evaluation Indices

To evaluate the quality of the PSO-XGBoost, XGBoost, SVM, RF, GP, and CART models, five indices of performance were used, including mean absolute percentage error (MAPE), root-mean-squared error (RMSE), variance account for (VAF), mean absolute error (MAE), and determination coefficient (R^2). The calculation of RMSE, R^2, MAE, VAF, and MAPE was described in Equations (13)–(17):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{13}$$

$$\text{R}^2 = 1 - \frac{\sum_i \left(y_i - \hat{y}_i\right)^2}{\sum_i \left(y_i - \overline{y}\right)^2} \tag{14}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{15}$$

$$\text{VAF} = \left(1 - \frac{\text{var}\left(y_i - \hat{y}_i\right)}{\text{var}\left(y_i\right)}\right) \times 100 \tag{16}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{17}$$

where n stands for the number of instances; $\overline{y}$, y_i, and $\hat{y}_i$ consider as average, calculated, and modeled amounts of the response variable.
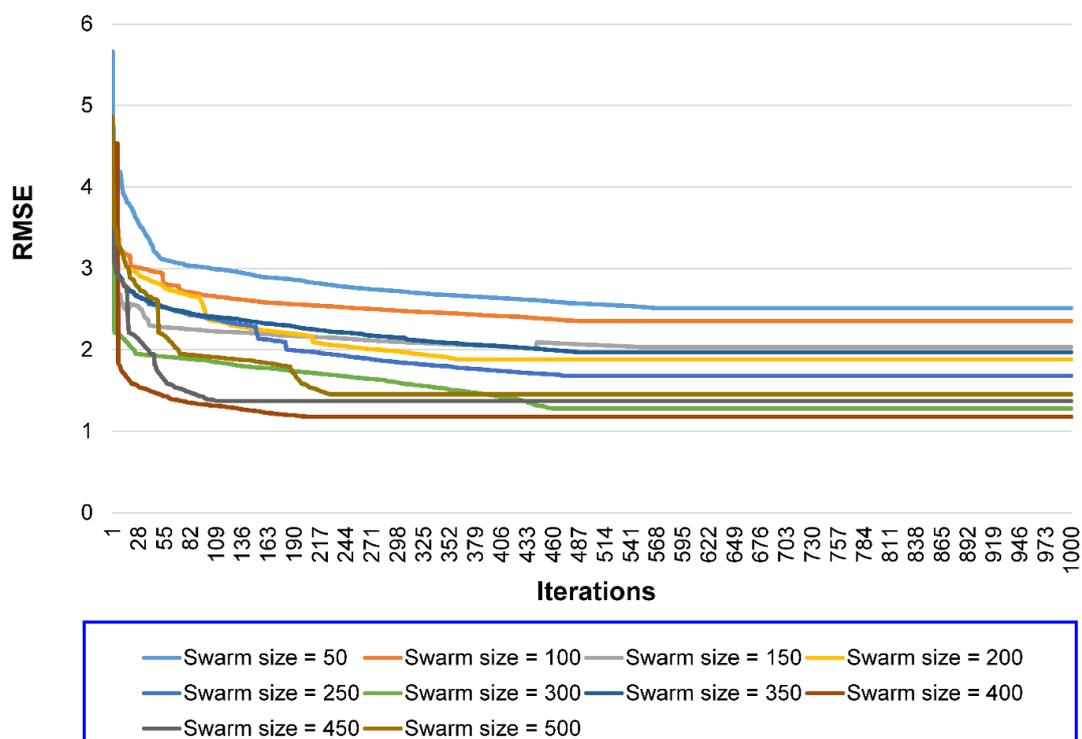
6. Results and Discussions

In this stage, 80% of the whole dataset was randomly selected to develop the HL of buildings forecasting models; the remaining 20% of the dataset was used to test and re-evaluate the accuracy, as well as the performance of the developed models. The re-sampling method of 10-fold cross-validation was applied to reduce the error of the models. Note that, the same training/testing datasets, as well as

the re-sampling techniques, were used for all models. It is of interest to consider the feasibility of the proposed PSO-XGBoost model in estimating the HL of buildings systems. Indeed, a combined process between the XGBoost model and the PSO algorithm was made based on the framework introduced in Figure 4. The parameters of the PSO algorithm has been optimally set before performing the optimization of the XGBoost model, as shown in Table 2. After the parameters of the PSO algorithm have been set, the process of searching and optimizing for the hyper-parameters of the XGBoost was performed. Its performance is shown in Figure 5.

Table 2. The optimal values of the PSO algorithm in this study.

Parameters	Acronym	Value
The number population	p	50,100,150,200,250,300,350,400,450,500
Maximum particle's velocity	V_{max}	2.00
Individual cognitive	ϕ_1	1.8
Group cognitive	ϕ_2	1.8
Inertia weight	w	0.95
Maximum number of iteration	m_i	1000

Figure 5. Performance of the proposed PSO-XGBoost technique in estimating the HL of building systems.
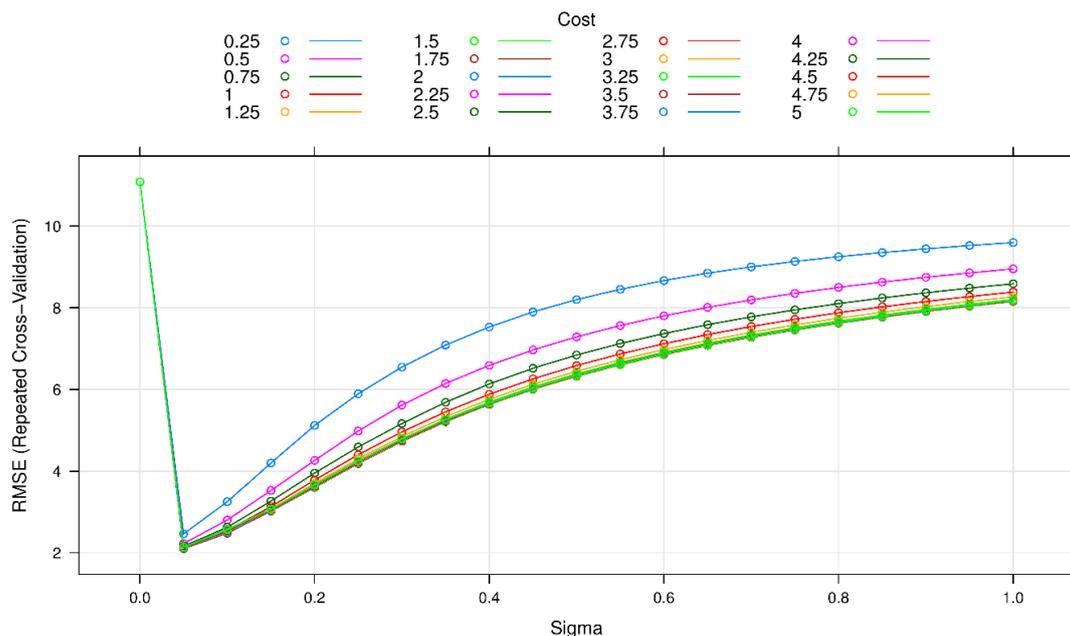
As shown in Figure 5, an optimal PSO-XGBoost model was found with the swarm size of 400 and stopped at the iteration of 209 with the lowest RMSE (i.e., RMSE = 1.776). For comparison and overall performance evaluation of the developed PSO-XGBoost model, the remaining models have also been developed as previously introduced, including XGBoost, SVM, RF, GP, and CART models.

For the development of the classical XGBoost model, the same parameters were used to control the accuracy of the classical XGBoost model, like those used for the PSO-XGBoost model. However, a grid search technique was applied to determine the hyper-parameters of the classical XGBoost model. The parameters of the grid were set as follows: η = [0.1, 0.4]; δ = 0.6 and 0.8; k = 50, 100, 150; γ = 0; d = [1,3]; ς = 0.5, 0.75, 1; and μ = 1. Subsequently, 10-fold cross-validation was applied during the development of the classical XGBoost model. Finally, the optimal XGBoost model was found with k = 50; d = 3; η = 0.3, γ = 0, δ = 0.6, ς = 1, and μ = 1 (Figure 6).

Figure 6. The development process of the classical XGBoost model for estimating HL.

For the development of the SVM model, the radial basis function (RBF) was used as a kernel function to estimate the HL in the present study. Two hyper-parameters of the SVM model with the RBF were selected and considered to construct the SVM model, including Sigma (σ) and cost (C). A grid search was established to find the optimal values for σ and C with σ lies in the range of [0,1]; C lies in the range of [0.25,5]. The stepwise for σ and C in the grid search is 0.05 and 0.25, respectively. The "scale" method was applied to reduce the skewness of the data in this study. The ten-fold cross-validation technique was also applied to increase the accuracy of the model. Ultimately, the optimal SVM model for estimating HL in this study was defined with the following hyper-parameters: σ = 0.05 and C = 1.75, as shown in Figure 7. Note that, the same training dataset (i.e., 672 experimental datasets) was used to develop the SVM model, like those used for the development of the PSO-XGBoost and XGBoost models.

Figure 7. The development process of the SVM model for estimating HL

For the development of the RF model, the number of the tree in the forest (*n*) and randomly selected predictor (ϖ), were used to adjust the accuracy/performance of the RF model. According to the recommendation of Nguyen, Bui [67], *n* should be set equal to 2000 to ensure the enrichment of the forest. Then, ϖ was tested to check the accuracy of the RF model. Since the predictors used in this study was eight, ϖ was set in the range of 1 to 8. Ultimately, the performance of the RF model based on the 2000 trees and randomly predictors was computed in Figure 8. As a result, the best RF model for predicting HL in this study was the RF3 (i.e., ϖ = 3). Note that, the same techniques, as well as the same training dataset, were used for the development of the RF model, like those used to develop the previous models (i.e., PSO-XGBoost, XGBoost, SVM).
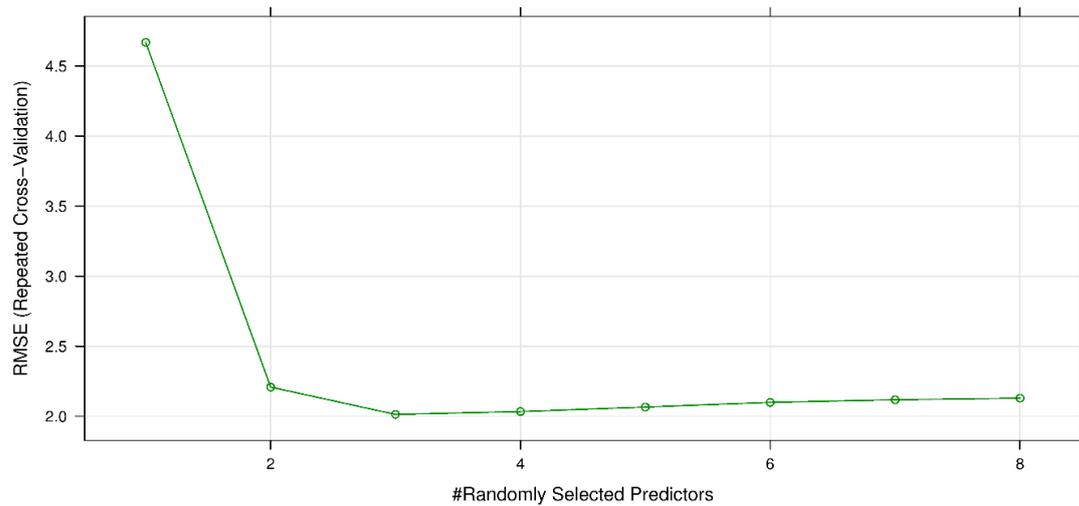

Figure 8. The development process of the random forest (RF) model for estimating HL.

Like to the SVM model, kernel functions can also be applied to develop the GP model. RBF was also used to develop the GP model for estimating HL in this work with σ was the only parameter used to control the accuracy of the GP model. A grid search, as well as the same techniques, like those used for the previous models, was also applied to develop the GP model. As a result, an optimal GP model was developed in this study with an σ of 0.009, for estimating the HL of buildings (Figure 9).
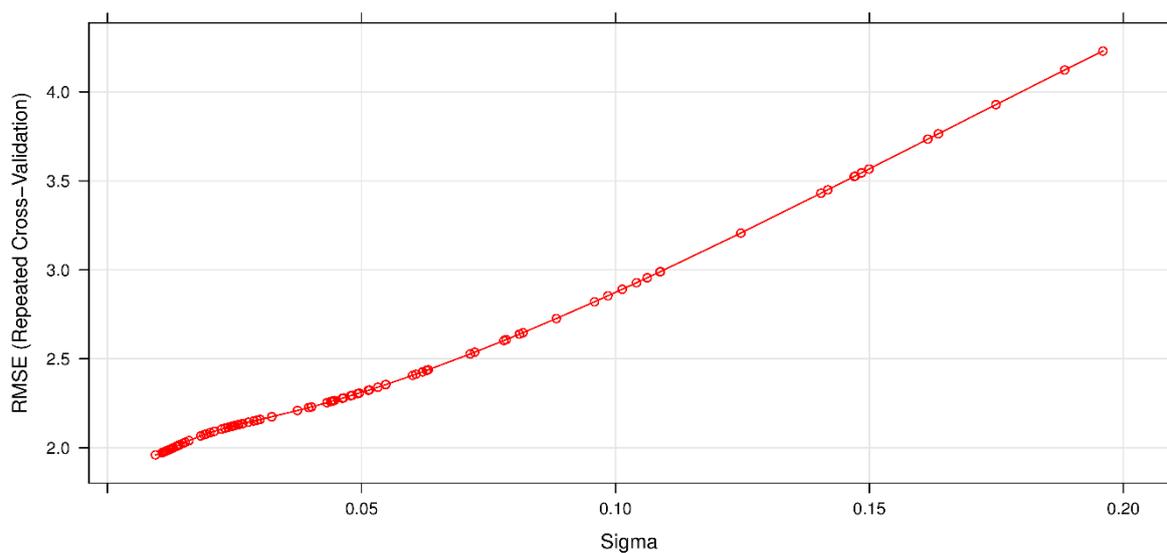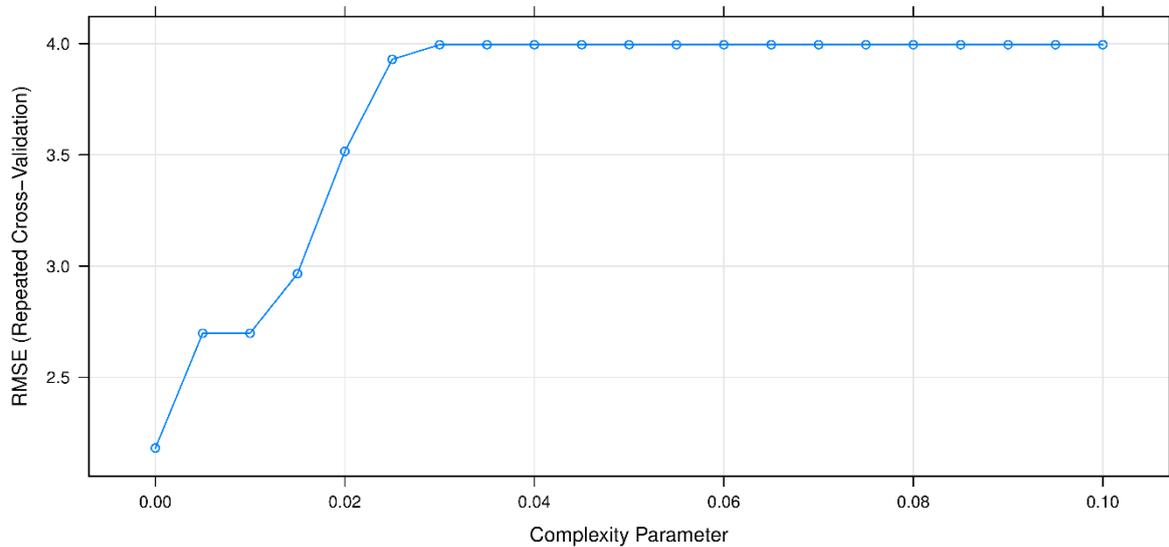

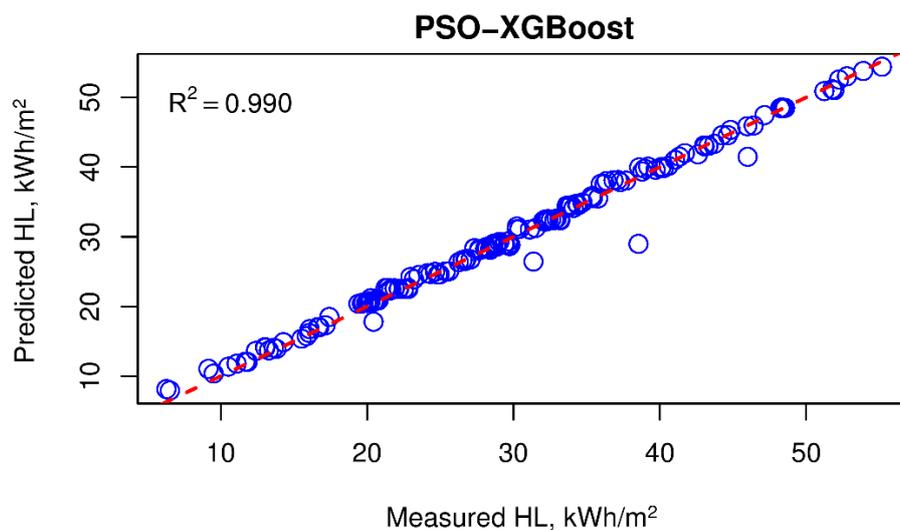**Figure 9.** The development process of the Gaussian process (GP) model for estimating HL.

Finally, an optimal CART model was also developed based on the same techniques and the training dataset, like those used for the previous models. Note that, only the complexity parameter (ψ) was used to develop the CART model with the grid search lies in the range of [0,0.1] (Figure 10).

Figure 10. The development process of the classification and regression trees (CART) model for estimating HL.

After the HL predictive models were developed, 165 observations of the testing dataset were used to evaluate the performance of the model through the statistical criteria, i.e., RMSE, R^2, VAF, MAE, and MAPE, as shown in Table 3. The intensity of color and ranking method were also applied to evaluate the models.

The results in Table 3 confirmed the perfect predictability of the proposed AI techniques for the HL of buildings systems in the present study. Of those AI techniques, the proposed PSO-XGBoost model was defined as the most superior model in predicting the HL of buildings with an RMSE of 1.124, R^2 of 0.990, MAE of 0.615, VAF of 98.934, MAPE of 0.024, and total ranking of 29. Figure 11 illustrates the exact predictability of the proposed PSO-XGBoost model on the testing dataset.

Figure 11. The results of HL prediction from the proposed PSO-XGBoost model.

Table 3. The results of the developed model on the testing dataset through several statistical indices.

Technique	RMSE	R^2	MAE	VAF	MAPE	Rank for RMSE	Rank for R^2	Rank for MAE	Rank for VAF	Rank for MAPE	Total Ranking
PSO-XGBoost	1.124	0.990	0.615	98.934	0.024	6	6	5	6	6	29
XGBoost	1.651	0.977	0.720	97.664	0.028	3	3	4	3	4	17
SVM	1.776	0.973	0.910	97.315	0.037	2	1	1	2	1	7
RF	1.589	0.978	0.557	97.835	0.026	5	4	6	5	5	25
GP	1.632	0.978	0.798	97.726	0.033	4	4	2	4	2	16
CART	1.779	0.973	0.773	97.286	0.031	1	1	3	1	3	9

Compared with the classical XGBoost model (without being optimized by the PSO algorithm), the proposed PSO-XGBoost model provided superior performance in estimating the HL of buildings. The results of the XGBoost model (i.e., RMSE = 1.651, R^2 = 0.977, MAE = 0.720, VAF = 97.664, MAPE = 0.028, total ranking of 17) recognized the significant optimization capabilities of the PSO algorithm in this study. Observing closely in Table 3, it can be seen that the XGBoost model had provided lower performance than the RF model (RMSE = 1.589, R^2 = 0.978, MAE = 0.557, VAF = 97.835, MAPE = 0.026, total ranking of 25). However, the XGBoost model had become even more potent than the RF model when optimized by the PSO algorithm (i.e., PSO-XGBoost). The remaining models provided a lower performance than the proposed PSO-XGBoost model. Figures 12–16 illustrate the accuracy level of the XGBoost, SVM, RF, GP, and CART models, respectively, on the testing dataset.
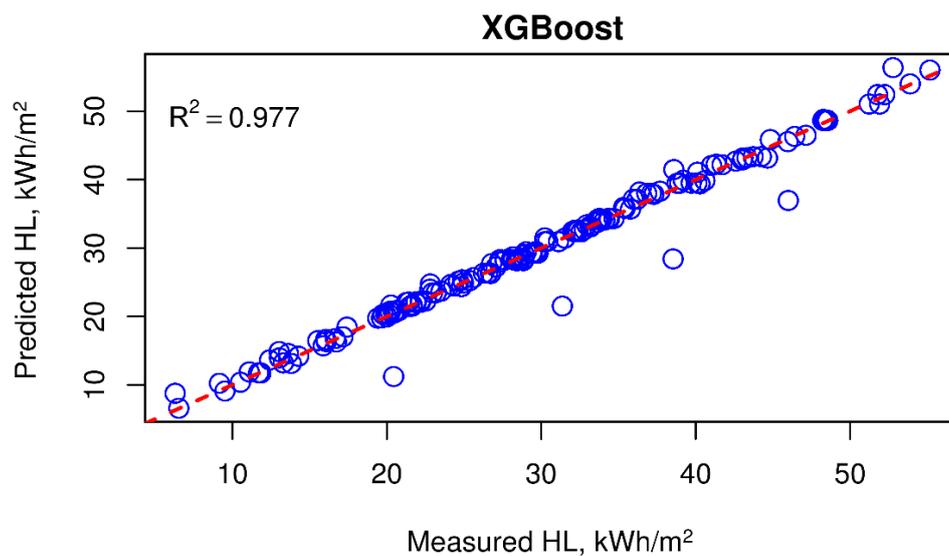

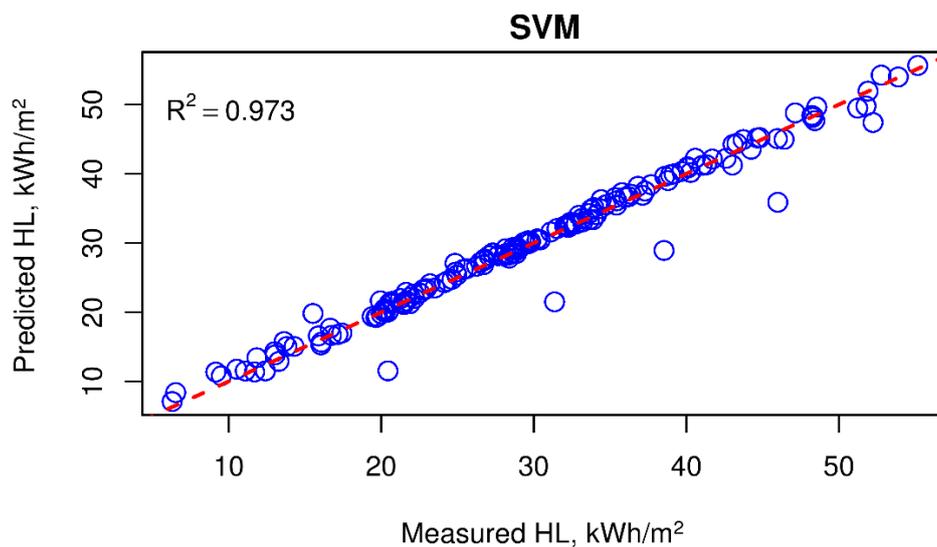**Figure 12.** The results of HL prediction from the XGBoost model.

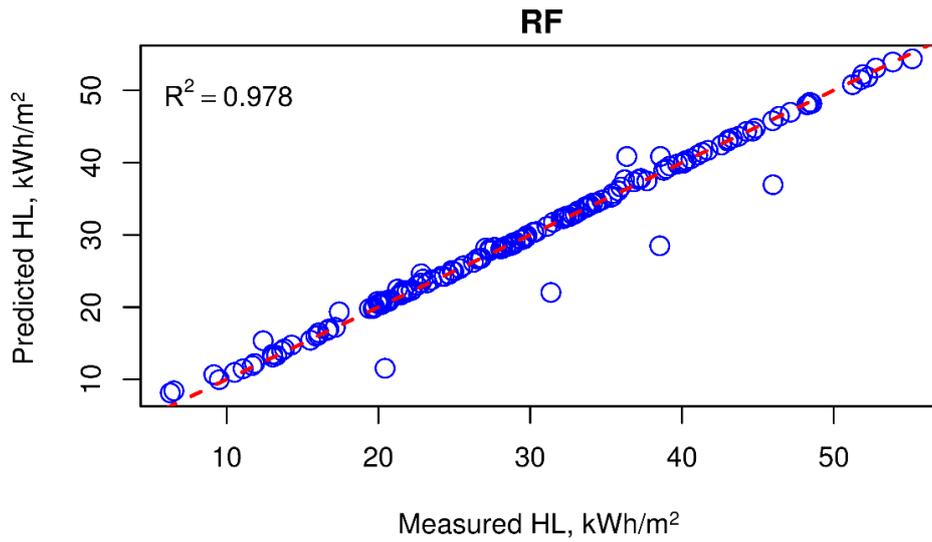**Figure 13.** The results of HL prediction from the support vector machine (SVM) model.

RF

Figure 14. The results of HL prediction from the RF model.
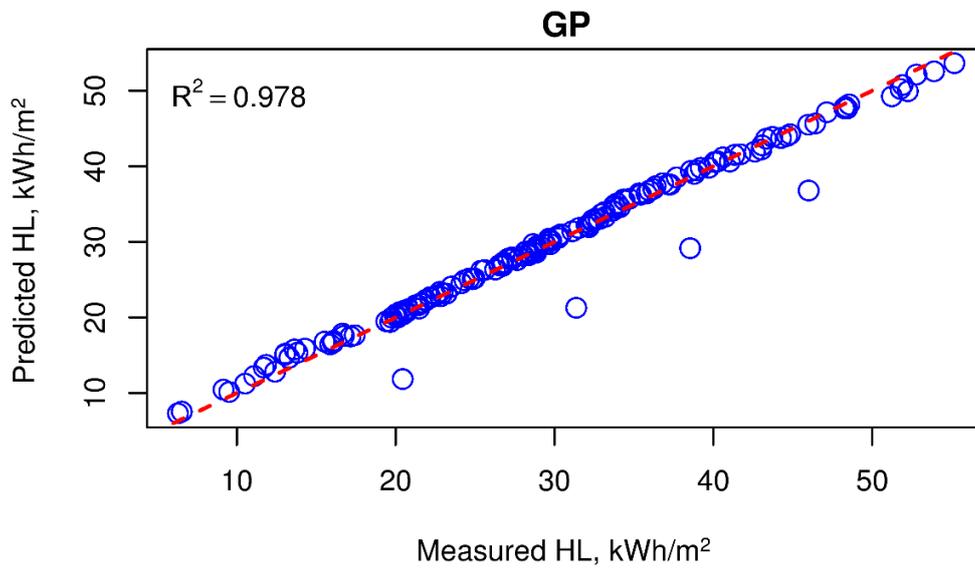
GP

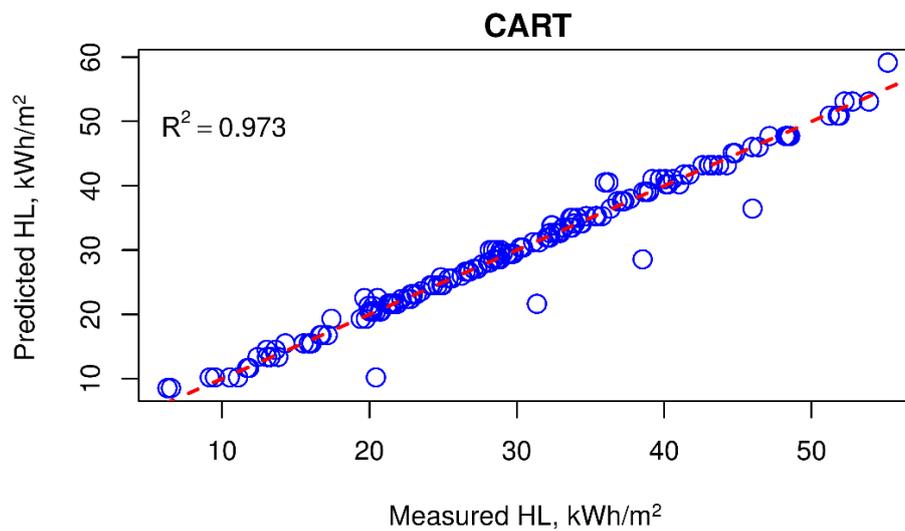**Figure 15.** The results of HL prediction from the GP model.
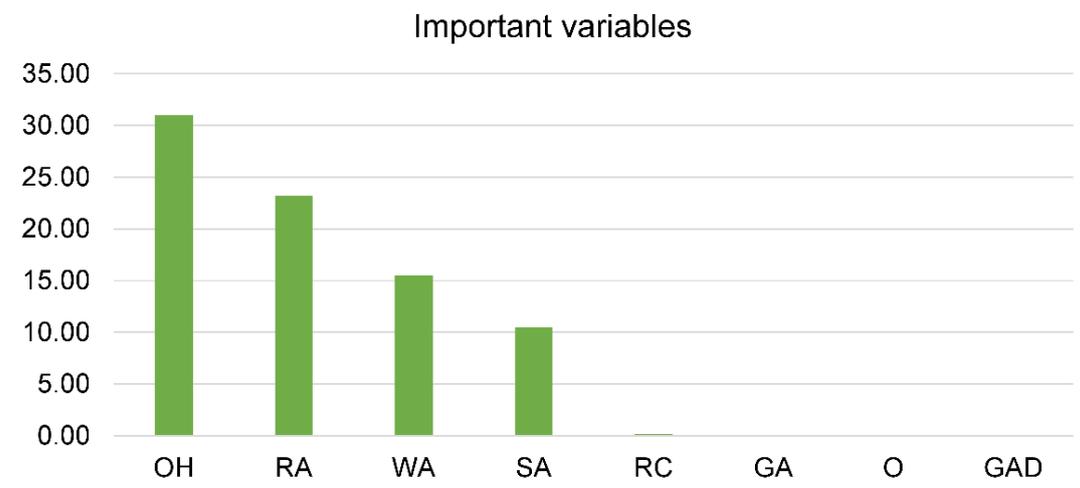
CART

Figure 16. The results of HL prediction from the CART model.

Considering the advantages and disadvantages of the other methods (i.e., XGBoost, SVM, RF, GP, and CART), it can be seen that the development of these models was more straightforward than the proposed PSO-XGBoost model, especially the GP and CART models with only one parameter used to build the models. The RF and SVM models are a bit more complicated when using two parameters to build predictive models. Most notably, the XGBoost model used seven parameters to build the model (i.e., k, d, η, γ, δ, ς, and μ). The higher the number of variables used, the more processing time and model construction. This is one of the disadvantages of complex models. Besides, although the construction of the single models (i.e., XGBoost, SVM, RF, GP, and CART) was more straightforward than the proposed PSO-XGBoost model; however, their performance was proven to be lower than the proposed PSO-XGBoost model.

Herein, the proposed PSO-XGBoost model had performed very well in estimating the HL of buildings systems with very high accuracy based on input variables. However, the number of input variables used in the present study was high; therefore, a thorough consideration of the relationship, as well as the importance of the input variables used in this study is necessary. The sensitivity indices based on the Csiszar f-divergence method was used to determine the significance of the inputs [68–70]. This technique implements a sensitivity analysis based on the density. The influence of the input variables was established in terms of the difference between the density function of the entire output and the density function of output. Note that the input variables were fixed. The difference between density functions was measured with Csiszar f-divergences. Finally, an evaluation was performed through the estimation of kernel density. The results of the estimation were shown in Figure 17. The results revealed that OH, RA, WA, and SA were the most critical variables in estimating the HL of buildings systems.

Figure 17. The importance of the input variables used in this study.

7. Conclusions and Remarks

Optimizing and designing HL systems of the buildings is one of the crucial tasks in smart cities. Effective use of HL enables buildings to be more energy efficient, reduce economic losses, as well as reduce adverse environmental impacts. This study proposed a new technique (PSO-XGBoost) to predict HL of buildings system with high reliability (RMSE = 1.124, R^2 = 0.990, MAE = 0.615, VAF = 98.934, MAPE = 0.024). Based on the obtained results of this study, some conclusions and remarks are drawn as:

- The AI techniques in this study, included PSO-XGBoost, XGBoost, SVM, RF, GP, and CART, were bright candidates for estimating the HL of building systems in practice. They could predict the HL of building systems with high reliability, especially the proposed PSO-XGBoost model.

- The proposed PSO-XGBoost model was a robust technique, which could accurately predict the HL of building systems with a promising result (RMSE = 1.124, R^2 = 0.990, MAE = 0.615, VAF = 98.934, MAPE = 0.024). It should be used as an alternative tool for experimental measurements. Furthermore, building design optimization methods can also be applied based on the proposed PSO-XGBoost model to minimize heat loss for buildings.

- Although the SVM, RF, GP, and CART models were bringing acceptable performance in this study. However, they need further research to improve the accuracy of estimating the HL of building systems, uniquely combining them with optimization algorithms.

- OH, RA, WA, and SA are the input variables that have the most influence on the accuracy of the HL of building systems forecasting model. It should be carefully collected and used as essential variables in the development of HL forecasting models.

Although the results of this study was perfect for evaluating and predicting the HL of building systems; however, they need to be further researched in the future works, such as improving the accuracy of the other models (i.e., XGBoost, SVM, RF, GP, and CART), or try to build several novel hybrid artificial intelligence systems based on these models and optimization algorithms. The optimization of building design is also one of the challenges of engineers in the future aim of energy-efficiency. It can be conducted based on the models of this study.

References

1. Lin, Y.-H. Design and Implementation of an IoT-Oriented Energy Management System Based on Non-Intrusive and Self-Organizing Neuro-Fuzzy Classification as an Electrical Energy Audit in Smart Homes. *Appl. Sci.* **2018**, *8*, 2337. [CrossRef]

2. De Paz, J.F.; Bajo, J.; Rodríguez, S.; Villarrubia, G.; Corchado, J.M. Intelligent system for lighting control in smart cities. *Inf. Sci.* **2016**, *372*, 241–255. [CrossRef]

3. Hao, L.; Lei, X.; Yan, Z.; ChunLi, Y. The application and implementation research of smart city in China. In Proceedings of the 2012 International Conference on System Science and Engineering (ICSSE), Dalian, China, 30 June–2 July 2012; pp. 288–292.

4. Alsarraf, J.; Moayedi, H.; Rashid, A.S.A.; Muazu, M.A.; Shahsavar, A. Application of PSO–ANN modelling for predicting the exergetic performance of a building integrated photovoltaic/thermal system. *Eng. Comput.* **2019**. [CrossRef]

5. Wang, D.; Pang, X.; Wang, W.; Qi, Z.; Li, J.; Luo, D. Assessment of the Potential of High-Performance Buildings to Achieve Zero Energy: A Case Study. *Appl. Sci.* **2019**, *9*, 775. [CrossRef]

6. Olszewski, R.; Pałka, P.; Turek, A.; Kietlińska, B.; Płatkowski, T.; Borkowski, M. Spatiotemporal Modeling of the Smart City Residents' Activity with Multi-Agent Systems. *Appl. Sci.* **2019**, *9*, 2059. [CrossRef]

7. Yu, Z.; Haghighat, F.; Fung, B.C.; Yoshino, H. A decision tree method for building energy demand modeling. *Energy Build.* **2010**, *42*, 1637–1646. [CrossRef]

8. Zhao, H.-X.; Magoulès, F. A review on the prediction of building energy consumption. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3586–3592. [CrossRef]

9. Aparicio-Ruiz, P.; Guadix-Martín, J.; Barbadilla-Martín, E.; Muñuzuri-Sanz, J. Applying Renewable Energy Technologies in an Integrated Optimization Method for Residential Building's Design. *Appl. Sci.* **2019**, *9*, 453. [CrossRef]

10. De Boeck, L.; Verbeke, S.; Audenaert, A.; De Mesmaeker, L. Improving the energy performance of residential buildings: A literature review. *Renew. Sustain. Energy Rev.* **2015**, *52*, 960–975. [CrossRef]

11. Eskin, N.; Türkmen, H. Analysis of annual heating and cooling energy requirements for office buildings in different climates in Turkey. *Energy Build.* **2008**, *40*, 763–773. [CrossRef]

12. Nojavan, S.; Majidi, M.; Zare, K. Optimal scheduling of heating and power hubs under economic and environment issues in the presence of peak load management. *Energy Convers. Manag.* **2018**, *156*, 34–44. [CrossRef]

13. Yang, I.-H.; Yeo, M.-S.; Kim, K.-W. Application of artificial neural network to predict the optimal start time for heating system in building. *Energy Convers. Manag.* **2003**, *44*, 2791–2809. [CrossRef]

14. Braun, M.; Altan, H.; Beck, S. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Appl. Energy* **2014**, *130*, 305–313. [CrossRef]

15. Jovanović, R.Ž.; Sretenović, A.A.; Živković, B.D. Ensemble of various neural networks for prediction of heating energy consumption. *Energy Build.* **2015**, *94*, 189–199. [CrossRef]

16. Sholahudin, S.; Han, H. Simplified dynamic neural network model to predict heating load of a building using Taguchi method. *Energy* **2016**, *115*, 1672–1678. [CrossRef]

17. Gunay, B.; Shen, W.; Newsham, G. Inverse blackbox modeling of the heating and cooling load in office buildings. *Energy Build.* **2017**, *142*, 200–210. [CrossRef]

18. Ahmad, T.; Chen, H. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy Build.* **2018**, *166*, 460–476. [CrossRef]

19. Kim, E.-J.; He, X.; Roux, J.-J.; Johannes, K.; Kuznik, F. Fast and accurate district heating and cooling energy demand and load calculations using reduced-order modelling. *Appl. Energy* **2019**, *238*, 963–971. [CrossRef]

20. Bui, X.-N.; Moayedi, H.; Rashid, A.S.A. Developing a predictive method based on optimized M5Rules–GA predicting heating load of an energy-efficient building system. *Eng. Comput.* **2019**, 1–10. [CrossRef]

21. Jitkongchuen, D.; Pacharawongsakda, E. Prediction Heating and Cooling Loads of Building Using Evolutionary Grey Wolf Algorithms. In Proceedings of the 2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON), Nan, Thailand, 30 January–2 February 2019; pp. 93–97.

22. Al-Shammari, E.T.; Keivani, A.; Shamshirband, S.; Mostafaeipour, A.; Yee, L.; Petković, D.; Ch, S. Prediction of heat load in district heating systems by Support Vector Machine with Firefly searching algorithm. *Energy* **2016**, *95*, 266–273. [CrossRef]

23. Sajjadi, S.; Shamshirband, S.; Alizamir, M.; Yee, L.; Mansor, Z.; Manaf, A.A.; Altameem, T.A.; Mostafaeipour, A. Extreme learning machine for prediction of heat load in district heating systems. *Energy Build.* **2016**, *122*, 222–227. [CrossRef]

24. Pino-Mejías, R.; Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A. Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO_2 emissions. *Energy* **2017**, *118*, 24–36. [CrossRef]

25. Xie, L. The heat load prediction model based on BP neural network-markov model. *Procedia Comput. Sci.* **2017**, *107*, 296–300. [CrossRef]

26. Protić, M.; Shamshirband, S.; Anisi, M.H.; Petković, D.; Mitić, D.; Raos, M.; Arif, M.; Alam, K.A. Appraisal of soft computing methods for short term consumers' heat load prediction in district heating systems. *Energy* **2015**, *82*, 697–704. [CrossRef]

27. Mottahedi, M.; Mohammadpour, A.; Amiri, S.S.; Riley, D.; Asadi, S. Multi-linear regression models to predict the annual energy consumption of an office building with different shapes. *Procedia Eng.* **2015**, *118*, 622–629. [CrossRef]

28. Kim, W.; Kim, Y.-K. Optimal Operation Methods of the Seasonal Solar Borehole Thermal Energy Storage System for Heating of a Greenhouse. *J. Korea Acad.-Ind. Coop. Soc.* **2019**, *20*, 28–34.

29. Wang, Z.; Srinivasan, R.S. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renew. Sustain. Energy Rev.* **2017**, *75*, 796–808. [CrossRef]

30. Tsanas, A.; Xifara, A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy Build.* **2012**, *49*, 560–567. [CrossRef]

31. TCVN. Solid minerals fuels—Determination of ash. **1995**, *173*.

32. Pessenlehner, W.; Mahdavi, A. *Building Morphology, Transparence, and Energy Performance*; Eighth international IBPSA conference: Eindhoven, Netherlands, 2003.

33. Schiavon, S.; Lee, K.H.; Bauman, F.; Webster, T. Influence of raised floor on zone design cooling load in commercial buildings. *Energy Build.* **2010**, *42*, 1182–1191. [CrossRef]

34. Nguyen, H. Support vector regression approach with different kernel functions for predicting blast-induced ground vibration: A case study in an open-pit coal mine of Vietnam. *SN Appl. Sci.* **2019**, *1*, 283. [CrossRef]

35. Bui, X.N.; Nguyen, H.; Le, H.A.; Bui, H.B.; Do, N.H. Prediction of Blast-induced Air Over-pressure in Open-Pit Mine: Assessment of Different Artificial Intelligence Techniques. *Nat. Resour. Res.* **2019**. [CrossRef]

36. Hasanipanah, M.; Faradonbeh, R.S.; Amnieh, H.B.; Armaghani, D.J.; Monjezi, M. Forecasting blast-induced ground vibration developing a CART model. *Eng. Comput.* **2017**, *33*, 307–316. [CrossRef]

37. Rasmussen, C.E. Gaussian processes in machine learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 63–71.

38. Grange, S.K.; Carslaw, D.C.; Lewis, A.C.; Boleti, E.; Hueglin, C. Random forest meteorological normalisation models for Swiss PM 10 trend analysis. *Atmos. Chem. Phys.* **2018**, *18*, 6223–6239. [CrossRef]

39. Moayedi, H.; Hayati, S. Modelling and optimization of ultimate bearing capacity of strip footing near a slope by soft computing methods. *Appl. Soft Comput.* **2018**, *66*, 208–219. [CrossRef]

40. Nguyen, H.; Drebenstedt, C.; Bui, X.-N.; Bui, D.T. Prediction of Blast-Induced Ground Vibration in an Open-Pit Mine by a Novel Hybrid Model Based on Clustering and Artificial Neural Network. *Nat. Resour. Res.* **2019**. [CrossRef]

41. Nguyen, H.; Bui, X.-N.; Tran, Q.-H.; Mai, N.-L. A new soft computing model for estimating and controlling blast-produced ground vibration based on hierarchical K-means clustering and cubist algorithms. *Appl. Soft Comput.* **2019**, *77*, 376–386. [CrossRef]

42. Nguyen, H.; Bui, X.-N.; Moayedi, H. A comparison of advanced computational models and experimental techniques in predicting blast-induced ground vibration in open-pit coal mine. *Acta Geophys.* **2019**. [CrossRef]

43. Eberhart, R.; Kennedy, J. A new optimizer using particle swarm theory. In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995; pp. 39–43.

44. Nguyen, H.; Moayedi, H.; Foong, L.K.; Al Najjar, H.A.H.; Jusoh, W.A.W.; Rashid, A.S.A.; Jamali, J. Optimizing ANN models with PSO for predicting short building seismic response. *Eng. Comput.* **2019**. [CrossRef]

45. Armaghani, D.J.; Hajihassani, M.; Mohamad, E.T.; Marto, A.; Noorani, S.A. Blasting-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization. *Arab. J. Geosci.* **2014**, *7*, 5383–5396. [CrossRef]

46. Gordan, B.; Armaghani, D.J.; Hajihassani, M.; Monjezi, M. Prediction of seismic slope stability through combination of particle swarm optimization and neural network. *Eng. Comput.* **2016**, *32*, 85–97. [CrossRef]

47. Yang, X.; Zhang, Y.; Yang, Y.; Lv, W. Deterministic and Probabilistic Wind Power Forecasting Based on Bi-Level Convolutional Neural Network and Particle Swarm Optimization. *Appl. Sci.* **2019**, *9*, 1794. [CrossRef]

48. Moayedi, H.; Mehrabi, M.; Mosallanezhad, M.; Rashid, A.S.A.; Pradhan, B. Modification of landslide susceptibility mapping using optimized PSO-ANN technique. *Eng. Comput.* **2018**, *35*, 967–984. [CrossRef]

49. Moayedi, H.; Moatamediyan, A.; Nguyen, H.; Bui, X.-N.; Bui, D.T.; Rashid, A.S.A. Prediction of ultimate bearing capacity through various novel evolutionary and neural network models. *Eng. Comput.* **2019**. [CrossRef]

50. Kulkarni, R.V.; Venayagamoorthy, G.K. An estimation of distribution improved particle swarm optimization algorithm. In Proceedings of the 2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information, Melbourne, QLD, Australia, 3–6 December 2007; pp. 539–544.

51. Friedman, J.H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [CrossRef]

52. Zhou, J.; Li, E.; Wang, M.; Chen, X.; Shi, X.; Jiang, L. Feasibility of Stochastic Gradient Boosting Approach for Evaluating Seismic Liquefaction Potential Based on SPT and CPT Case Histories. *J. Perform. Constr. Facil.* **2019**, *33*, 04019024. [CrossRef]

53. Chen, T.; He, T. *Xgboost: Extreme Gradient Boosting*; R Package Version 0.4-2; Available online: https://cran.r-project.org/web/packages/xgboost/vignettes/xgboost.pdf (accessed on 11 March 2019).

54. Nguyen, H.; Bui, X.-N.; Bui, H.-B.; Cuong, D.T. Developing an XGBoost model to predict blast-induced peak particle velocity in an open-pit mine: A case study. *Acta Geophys.* **2019**, *67*, 477–490. [CrossRef]

55. Cortes, C.; Vapnik, V. Support vector machine. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

56. Dou, J.; Paudel, U.; Oguchi, T.; Uchiyama, S.; Hayakavva, Y.S. Shallow and Deep-Seated Landslide Differentiation Using Support Vector Machines: A Case Study of the Chuetsu Area, Japan. *Terr. Atmos. Ocean. Sci.* **2015**, *26*, 227–239. [CrossRef]

57. Zhou, J.; Li, X.; Shi, X. Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. *Saf. Sci.* **2012**, *50*, 629–644. [CrossRef]

58. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

59. Effron, B.; Tibshirani, R.J. An introduction to the bootstrap. *Monogr. Stat. Appl. Probab.* **1993**, *57*, 436.

60. Breiman, L. *Random Forests*; Technical report 567; University of California-Berkeley, Statistic Department, 1999.

61. Zhou, J.; Shi, X.; Du, K.; Qiu, X.; Li, X.; Mitri, H.S. Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel. *Int. J. Geomech.* **2016**, *17*, 04016129. [CrossRef]

62. Rasmussen, C.E. Gaussian processes in machine learning. In *Advanced Lectures on Machine Learning*; Springer: Berlin, Germany, 2004; pp. 63–71.

63. Särkkä, S.; Álvarez, M.A.; Lawrence, N.D. Gaussian Process Latent Force Models for Learning and Stochastic Control of Physical Systems. *arXiv* **2017**, arXiv:1709.05409.

64. Khandelwal, M.; Armaghani, D.J.; Faradonbeh, R.S.; Yellishetty, M.; Majid, M.Z.A.; Monjezi, M. Classification and regression tree technique in estimating peak particle velocity caused by blasting. *Eng. Comput.* **2017**, *33*, 45–53. [CrossRef]

65. Myles, A.J.; Feudale, R.N.; Liu, Y.; Woody, N.A.; Brown, S.D. An introduction to decision tree modeling. *J. Chemom.* **2004**, *18*, 275–285. [CrossRef]

66. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]

67. Nguyen, H.; Bui, X.-N. Predicting Blast-Induced Air Overpressure: A Robust Artificial Intelligence System Based on Artificial Neural Networks and Random Forest. *Nat. Resour. Res.* **2018**. [CrossRef]

68. Borgonovo, E. A new uncertainty importance measure. *Reliab. Eng. Syst. Saf.* **2007**, *92*, 771–784. [CrossRef]

69. Da Veiga, S. Global sensitivity analysis with dependence measures. *J. Stat. Comput. Simul.* **2015**, *85*, 1283–1305. [CrossRef]

70. Krzykacz-Hausmann, B. Epistemic sensitivity analysis based on the concept of entropy. In Proceedings of the Sensitivity Analysis of Model Output, Madrid, Spain, 18–20 June 2001; pp. 31–35.