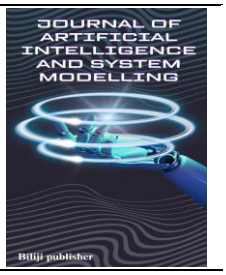
Research Article – Vol. 01, Issue 01, No. 07

Journal of Artificial Intelligence and System Modelling

Journal Web Page: https://jaism.bilijipub.com




# The Utilization of a Naïve Bayes Model for Predicting the Energy Consumption of Buildings

Behnam Sadaghat[1,*] , Ali Javadzade khiavi[2] , Babak Naeim[2], Erfan Khajavi[3] , Hadi sadaghat[4], Amir Reza Taghavi Khanghah[3]

[1] *Department of Civil Engineering, Tabriz University, Tabriz, Iran*
[2] *Department of civil engineering, Mohaghegh Ardabili university, Ardabil, Iran*
[3] *Department of civil engineering, Islamic Azad university of Ardabil branch, Ardabil, Iran*
[4] *Departement of Mechanical Engineering, University of Mohaghegh Ardabili, Ardabil, Iran*

## Highlights

- Integration of advanced optimization algorithms with heating load prediction for energy-efficient building management.
- Utilization of African Vultures Optimization Algorithm (AVOA) and Sand Cat Swarm Optimization (SCSO) with the Naïve Bayes (NB) model.
- Focus on optimizing HVAC operations, equipment sizing, energy management, and cost reduction.
- Significance of accurate heating load prediction for achieving energy efficiency, cost-effectiveness, and environmental sustainability in building operations.
- Evaluation metrics such as R2, RMSE, MSE, WAPE, and NSE used to assess the predictive efficacy, highlighting the superiority of the NBSC model in real-world applicability and accuracy.

**Article Info**

Received: 25 October 2023
Received in revised: 31 December 2023
Accepted: 31 December 2023
Available online: 31 December 2023

**Keywords**

Heating Energy Consumption;
Heating Load;
Naïve Bayes;
African Vultures Optimization Algorithm;
Sand Cat Swarm Optimization

**Abstract**

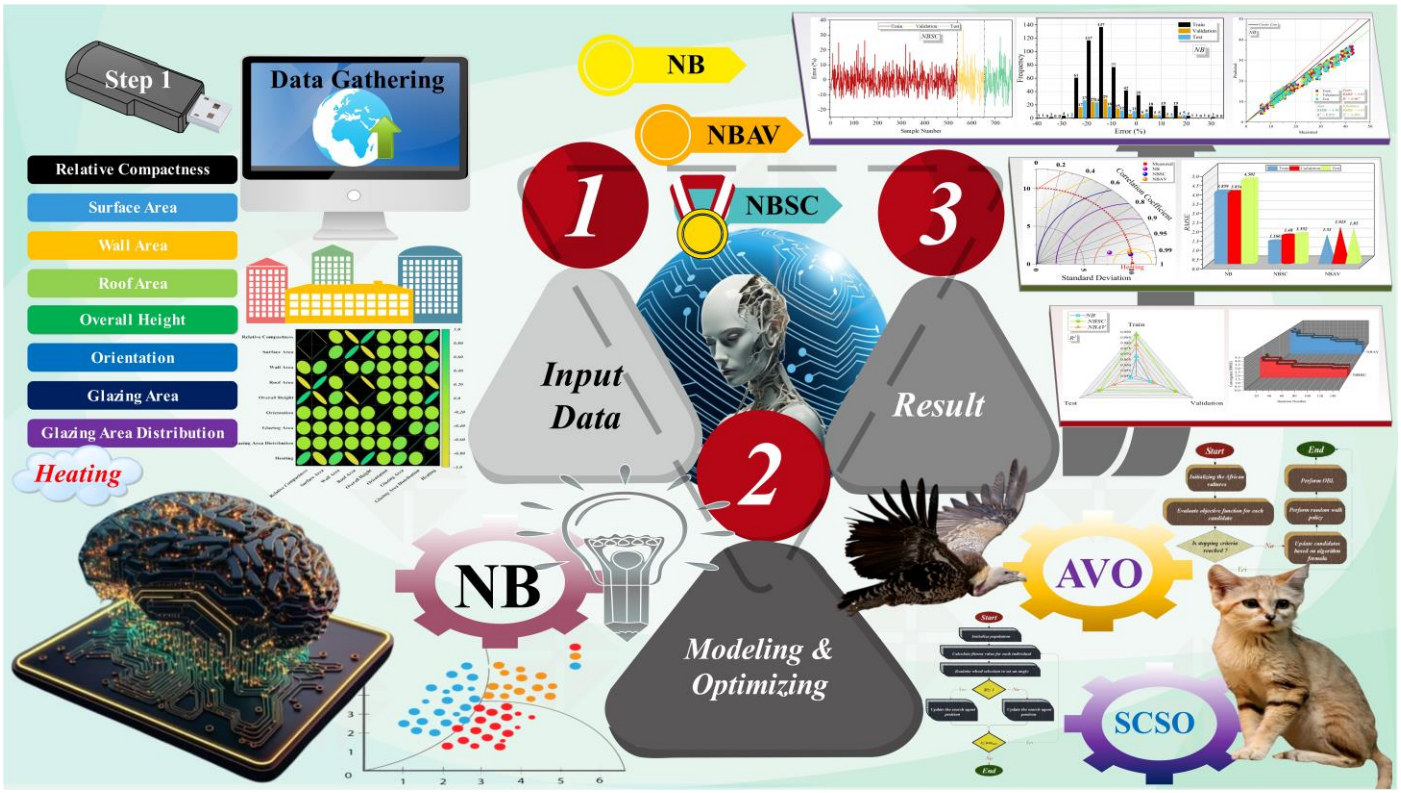
This study tackles the imperative of energy-efficient building management by marrying advanced optimization algorithms with heating load (HL) prediction within the realm of heating, ventilation, and air conditioning (HVAC) systems. Highlighting the pivotal role of HL prediction in optimizing HVAC operations, fostering energy efficiency, and realizing cost savings, this research pioneers innovative strategies. It introduces a fusion of the African Vultures Optimization Algorithm (AVOA) and the Sand Cat Swarm Optimization (SCSO) with the Naïve Bayes (NB) model, aiming to elevate heating load prediction accuracy and streamline HVAC system optimization. These algorithms are employed to improve HVAC system control, equipment sizing, energy management, and cost reduction. The significance of accurate HL prediction in achieving energy efficiency, cost-effectiveness, and environmental sustainability in building operations is showcased. To gauge the predictive efficacy of the models, an array of performance metrics, including $R^2$, RMSE, MSE, WAPE, and the NSE, were employed for assessment. These evaluations demonstrate that the NBSC model stands out as the most exceptional predictor in terms of real-world applicability and accuracy. It achieves an outstanding maximum $R^2_{train}$ value of 0.987, showcasing a high degree of explanatory power and exhibiting an impressively low $RMSE_{train}$ value of 1.166, signifying minimal prediction errors in comparison to other models. Additionally, the NBAV obtained a valuable result based on an $R^2$ value of 0.978 and an RMSE value of 1.510, indicating the model's reliable results. This study did not only produce an accurate model.

## Nomenclature

| *Variables* | *Description* | *Variables* | *Description* |
|---|---|---|---|
| *AVOA* | *African Vultures Optimization Algorithm* | *NSE* | *Nash-Sutcliffe Efficiency* |

* Corresponding Author: Behnam Sadaghat
Email: behnamsedaghat07@gmail.com

| | | | |
|---|---|---|---|
| *HL* | *Heating Load* | *R2* | *Coefficient of Determination* |
| *HVAC* | *heating, ventilation, and air conditioning* | *RMSE* | *Root Mean Squared Error* |
| *MSE* | *Mean Squared Error* | *SCSO* | *Sand Cat Swarm Optimization* |
| *NB* | *Naïve Bayes* | *WAPE* | *Weighted Absolute Percentage Error* |




*Schematic presentation of the paper*

## 1. Introduction

### *1.1. Background*

The emphasis on building energy efficiency research projects has increased noticeably during the last few years. This heightened interest can be traced back to the escalating apprehensions regarding the squandering of energy resources and the enduring adverse consequences it imposes on the natural world. Acknowledging the pivotal role that buildings assume in energy utilization and the generation of greenhouse gases, scholars have been energetically investigating strategies to augment the efficiency of buildings while curbing their ecological footprint [1]–[3]. Achieving energy conservation in buildings requires the development of diverse strategies for efficient energy management, with a crucial focus on accurately predicting energy consumption. This emphasis on precise prediction, which has gained significant attention recently, enables the creation of targeted and highly efficient energy-saving initiatives by closely monitoring building energy usage patterns.

Moreover, accurate energy consumption prediction enhances the effectiveness of energy-saving measures and deepens understanding of the underlying dynamics within buildings, facilitating the development of customized approaches to optimize energy efficiency in various building types. This precision in forecasting energy usage is a foundational element for building a sustainable and energy-efficient environment [4]. innovative approaches include demand-side management techniques [5], strategies for locating and diagnosing faults [6], and intelligent control systems [7] largely on a thorough understanding of forecasting building energy use. To optimize energy use, reduce waste, and guarantee the smooth operation of building systems, these techniques use predictive insights. They detect prospects for prospective energy-saving measures and address operational shortcomings, It promotes energy efficiency and efficient building infrastructure operation [2]. According to the paper, Even little adjustments to the building energy consumption estimates can lead to significant energy use reductions [8]. Building managers and residents can take proactive measures to maximize energy use by making well-informed decisions based on precise forecasts of energy use patterns. This may entail altering HVAC settings, improving lighting schedules, putting in place energy-efficient machinery, and changing behavior to support energy-saving goals[9].

An essential component of energy modeling is accurately projecting energy use within buildings, however, this approach frequently deviates from actual performance. Numerous studies have highlighted differences between these projections and actual energy usage, occasionally showing that real consumption can be significantly higher than initial projections, frequently by two or three times. Traditional energy models are suitable for initial evaluations but fall short in accounting for the many complexities since they rely on engineering calculations based on physical principles, and they have practical uses [10]–[12]. These limitations are successfully circumvented by the use of numerical simulation techniques to simulate building energy use by incorporating machine learning (ML) models into efforts to improve building energy efficiency. Building energy consumption can be predicted and optimized using artificial intelligence (AI) models, which generate precise projections and insightful data for effective energy management using historical data, real-time sensor inputs, and ML algorithms [13]–[15].

### *1.2.Related work*

The science of predicting energy usage has advanced significantly over time. A variety of methods and procedures have been developed by academics and industry professionals with the specific goal of accurately forecasting energy consumption patterns [16]. These efforts have resulted in the creation of numerous tools and approaches intended to increase the precision of energy usage projections [17]–[21]. Gong et al. [22] utilized Support Vector Regression ($SVR$), using the methods Multilayer Perceptron ($MLP$), Random Forest ($RF$), and Light Gradient Boosted Machine ($LGBM$), Tianjin residential buildings' heating energy usage is forecasted. The findings of their study showed that the LGBM model fared better than the other models in terms of a number of evaluation parameters. In their scholarly investigation, Moradzadeh et al. [23] applied $SVR$ and $MLP$ models for multilayer perceptrons are used to forecast cooling and heating loads [24]. Notably, the $MLP$ model displayed outstanding performance, achieving the highest R-value of 0.9993 in forecasting Heating Load, whereas the SVR model showcased its excellence by attaining the highest R-value of 0.9878 for predicting Cooling Load. Karijadi and Chou [25] introduced a novel hybrid approach for forecasting building energy consumption. Long Short-Term Memory ($LSTM$) and $RF$ techniques were used in this method. Real-world data was used to validate this method's efficacy and show how much better it performs than traditional benchmark techniques. Nebot and Mugica [26] concentrated on forecasting heating and cooling loads in residential structures. They used fuzzy inductive reasoning ($FIR$) and adaptive neural fuzzy inference system ($ANFIS$) in particular for this goal. They also contrasted thirteen machine learning techniques with these fuzzy approaches. The results showed that SVR performed better than the other techniques, along with the two fuzzy methods ($ANFIS\ and\ FIR$). $Olu-Ajayi$ et al. [27] studied the use of various ML methods to forecast the annual energy use of residential buildings, including $ANN$, decision trees ($DT$), $SVM$, $GBM$, $DNN$, $RF$, stacking, K closest neighbour ($KNN$), and linear regression ($LR$). The results show that DNN is the best accurate forecast model for energy consumption at the early design phase.

### *1.3.Objective of the study*

This research aimed to create a machine-learning model that could forecast Heating Load ($HL$) using information from trustworthy sources. The study used the Naïve Bayes (NB) technique to construct strong composite models. To forecast $HL$ values, these composite models smoothly integrated the African Vultures Optimization Algorithm (AVOA) and Sand Cat Swarm Optimization (SCSO). Using a Naive Bayes model for predicting HL can be considered when simplicity, speed, and interpretability are essential, especially in situations where the naive independence assumption aligns reasonably well with the data. A comprehensive array of five evaluation criteria was utilized to measure the efficacy of these models. These criteria evaluated how well the model estimated HL, ensuring a comprehensive assessment of its accuracy and reliability.

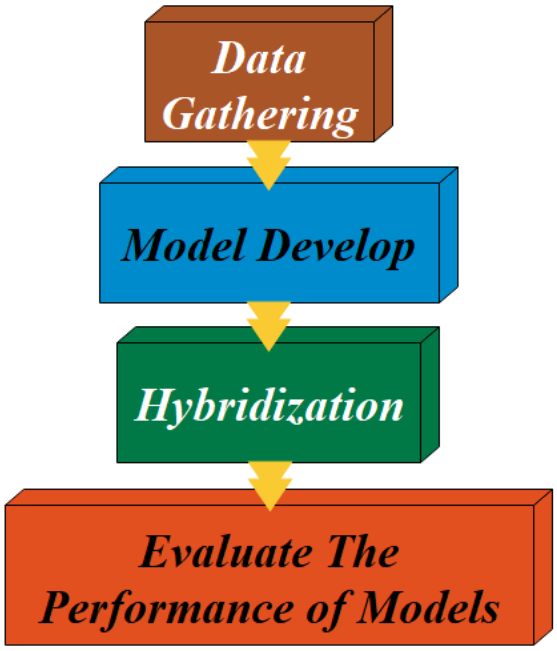



**Fig. 1.** The current study's process

## 2. Materials and Methodology

### *2.1.Data gathering*

The current study involves categorizing data into eight specific factors: Relative Compactness (RC), Surface Area (SA), Wall Area (WA), Roof Area (RA), Overall Height (OH), Orientation (Or), Glazing Area (GA), and Glazing Area Distribution (GAD). At the same time, the primary focus is on predicting Heating Load (KW) values as the desired output parameter. The dataset has been divided into three parts to simplify the analysis process. Initially, the training phase comprises a significant 70% of the entire dataset and acts as the basis for developing and training the model. Following that, the validation phase, representing 15% of the data, is used to fine-tune and verify the model's performance to ensure its applicability to different situations. Finally, the remaining 15% is allocated to the testing phase, where the model's overall efficacy and capacity to make precise predictions on new, unseen data are assessed. This partitioning enables a thorough evaluation of the model's performance across various aspects of the dataset. Table 1 provides a numerical presentation of the parameters used in building the model. This table offers a comprehensive summary of specific features, including mean values (M), maximum values (Max), standard deviation (St.), and minimum values (Min). It is important to stress that the Heating Load (HL) value is constrained within clearly defined limits. Its upper limit is firmly set at 43.1 KW, while its lower threshold is precisely determined at 6.01 KW, in strict accordance with the specifications of its output parameter.

**Table 1.** The statistical properties of the input variable of Heating

| Variables | Indicators | | | | |
|---|---|---|---|---|---|
| | Category | Min | Max | Avg | St. Dev. |
| RC | Input | 0.62 | 0.98 | 0.764 | 0.106 |
| SA ($m^2$) | Input | 514.5 | 808.5 | 671.70 | 88.086 |
| WA ($m^2$) | Input | 245 | 416.5 | 318.5 | 43.63 |
| RA (m) | Input | 110.25 | 220.5 | 176.60 | 45.165 |
| OH (m) | Input | 3.5 | 7 | 5.25 | 1.751 |
| Or | Input | 2 | 5 | 3.5 | 1.118 |
| GA (%) | Input | 0 | 0.4 | 0.235 | 0.133 |
| GAD | Input | 0 | 5 | 2.81 | 1.55 |
| Heating (kW) | Output | 6.01 | 43.1 | 22.30 | 10.09 |

The correlation between several input variables and the result in Fig. 2 using the statistical characteristics listed in Table 1, heating is represented. The associations between the heating power in kilowatts and the inputs (RC, SA, WA, RA, OH, Or, GA, and GAD) are displayed in the graph. Some inputs show larger connections with heating output, indicating significant influences on the generated heat, according to the analysis, which reveals noteworthy tendencies. Finding important factors influencing heating levels is made easier by comprehending these linkages.

Potential insights into optimizing heating systems are given by the visual representation of clear patterns or dependencies between particular input variables and the heating output.

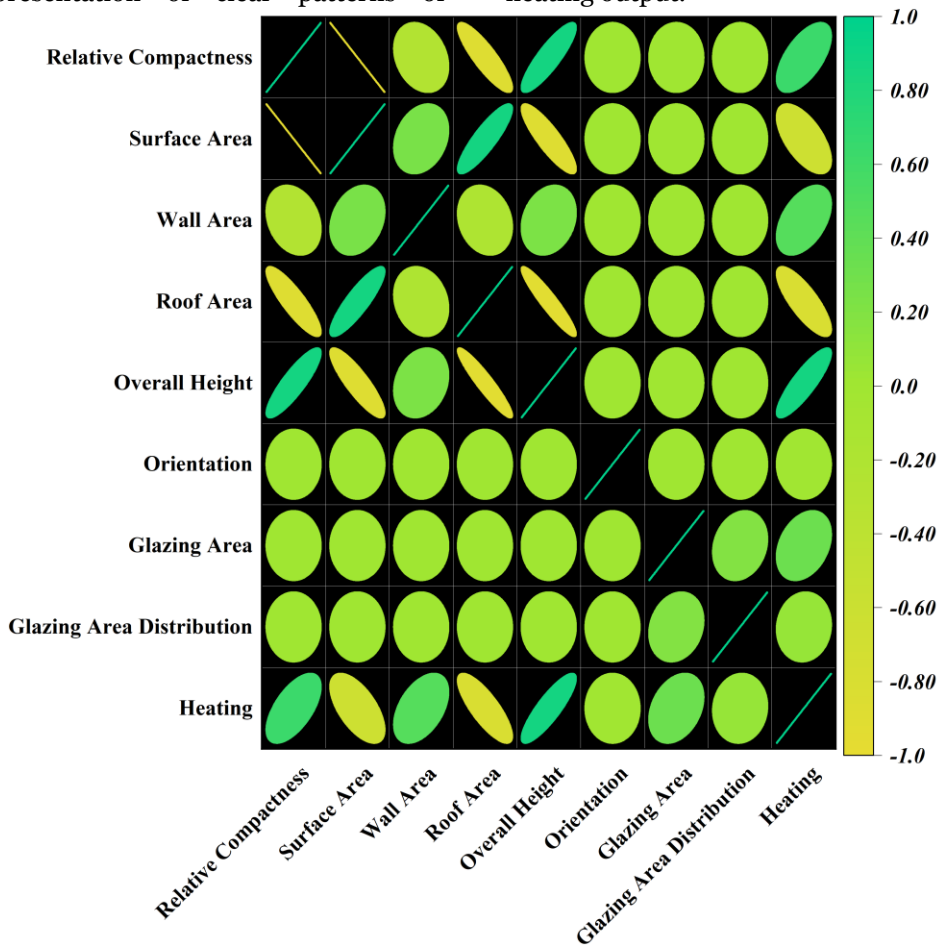



**Fig. 2.** The correlation between inputs and outputs

### *2.2. Naïve Bayes Regression (NB)*

One version of the Naive Bayes algorithm that can be applied to regression tasks is Naive Bayes regression [28]. $NB$ broadens the use of NB beyond classification problems to include continuous numerical value prediction. Addressing the task of predicting a numeric target value $Y$ is rooted in an instance $E$, which encompasses a collection of m attributes: $X_1$, $X_2$, ..., $X_m$. These characteristics can have two roles: nominal, meaning they indicate an array of unordered values, or numeric, meaning they are treated as natural numbers. Selecting $Y$ to minimize forecast errors becomes possible when the probability density function $p(Y|E)$ of the target value is known for all possible instances of $E$. But in most cases, $p(Y|E)$ is unknown, so estimation from available data is necessary. To handle this, $NB$ assumes attribute independence and applies Bayes' theorem. $X_1, X_2, \ldots, X_m$ given Y, the target value. In line with the Bayes theorem:

$$p(Y|E) = \frac{p(E,Y)}{\int p(E,Y)dY} = \frac{p(E,Y)p(Y)}{\int p(E|Y)p(Y)dY} \quad (1)$$

In this context, the probability density function (PDF) that describes instance $E$ for a given target value $Y$ is represented by the likelihood $p(E|Y)$. Moreover, before any examples are observed, the prior $p(Y)$ matches the target value's PDF. A fundamental tenet of $NB$ is the assumption of attribute independence when the model is conditioned on the target value [29], [30]. Therefore, the formulation of Eq. (1) is as follows:

$$p(Y|E) = \frac{p(X_1|Y)p(X_2|Y) \ldots p(X_m|Y)p(Y)}{\int p(X_1|Y)p(X_2|Y) \ldots p(X_m|Y)p(Y)dY} \quad (2)$$

It becomes possible to estimate the individual PDFs $p(X_i|Y)$ separately rather than estimating the PDFs $p(E|Y)$. The learning process is made much less complex by this division of $pdfs$. The estimation of $p(X_i|Y)$ is generally more reliable than estimating $p(E|Y)$ because higher problem dimensionality requires more data for accurate estimates.

#### *2.2.1. Managing Numerical Features*

Let's start by discussing how to estimate $p(X|Y)$ for numerical attributes $X$. In this case, it is assumed that these attributes have been normalized using the training dataset's range. In this case, a $PDF$ with two numerical variables is formed by $p(X|Y)$. This situation is brought about by:

$$P(X|Y) = \frac{p(X,Y)}{\int p(X,Y)dX} \quad (3)$$

A technique based on approximating the joint probability $P\ (X;\ Y)$ can be used to estimate the probability $P(X|Y)$. It is theoretically possible to represent $p\ (X;\ Y)$ using two-dimensional $PDFs$ estimated through various methods, such as mixture models. The kernel density estimator has been the option in our situation.

$$\hat{p}(X = x, Y = y) = \frac{1}{nh_x h_y}\sum_{i=1}^{n} K\left(\frac{x - x_1}{h_x}\right) K\left(\frac{y - y_1}{h_y}\right) \quad (4)$$

In this context, $x_i$ denotes the attribute value while $y_i$ corresponds to the target value of training example $i. K(.)$ signifies a chosen kernel function, and $h_x$ and $h_y$ function as the kernel widths for $X$ and $Y$. Should either $x_i$ or $y_i$ be missing, the example is not factored into the calculation. When the kernel function complies with specified smoothness attributes and the kernel widths are selected appropriately, this estimation generally approaches the true PDF. A commonly used option for $K(.)$ is the Gaussian kernel $K(t) = (2\pi)^{-\frac{1}{2}} e^{-\frac{t^2}{2}}$, and this is the kernel utilized in the experiments. Ideally, the kernel widths $h_x$ and $h_y$ should be chosen to reduce the difference between the estimated PDFs $p(X;\ Y)$ and the true PDFs $p(X;\ Y)$. Calculating the expected cross-entropy between the two PDFs is one method for assessing this difference. It is possible to get an objective estimate of this by using leave-one-out cross-validation [31].

$$CV_{CE} = -\frac{1}{n}\sum_{j=1}^{n} \log\left(\frac{1}{(n-1)h_X h_y}\sum_{i=1, i\neq 1}^{n} k\left(\frac{x_j - x_i}{h_X}\right) K\left(\frac{y - y_1}{h_y}\right)\right) \quad (5)$$

Following this, $h_x$ and $h_y$ are defined as $h_X = c_X/\sqrt{n}$, and $h_Y = c_Y/\sqrt{n}$. The values of $c_X$ and $c_Y$ are chosen to optimize the computed cross-entropy. In the practical experiments, which undertake a grid search within the interval [0.4, 0.8] for both $c_X$ and $c_Y$. For adjustments, the grid search uses a step size of 0.1. Different search parameter configurations were investigated, and there were not many noticeable variations in the results. This phenomenon has the following causes:

$$\int \hat{p}\ (X,Y)dx = \frac{1}{nhy}\sum_{i=1}^{n} K(\frac{y - y_1}{h_y}) \quad (6)$$

All the terms required for estimating are available. $\hat{p}(X|Y)$ of $p(X|Y)$ for a numeric attribute $X$.

### *2.3. African Vultures Optimization Algorithm (AVOA)*

The African vulture optimization algorithm was presented by [32]. To identify the most effective vultures in every category, the initial population's proposed solutions are evaluated for their appropriateness. The best solution is designated as the top-performing vulture for the initial and subsequent groups, serving as the optimal choice. It is essential to highlight that the fitness of all populations needs to be reassessed in every iteration. Additionally, the remaining solutions are determined in the following manner:

$$G(i) = \begin{cases} Bestvulture_i\ if\ h_i = a \\ Bestvulture_2\ if\ h_i = b \end{cases} \quad (7)$$

$a$ and $b$ are situated within the range of (0,1). A roulette wheel approach is employed to select a potential optimal solution. This technique offers a systematic way to identify the most suitable solution, and the process is delineated as follows:

$$h_i = \frac{k_i}{\sum_{i=1}^{n} k_i} \quad (8)$$

If $b$ is less than $a$, employing the $AVOA$ could increase degradation. Conversely, even when $a$ is less than $b$, the $AVOA$ might produce different outcomes. To progress from the exploration stage to the exploitation stage, Eq. (9) is employed:

$$K = (2 \times rand_1 + 1) \times y \times \left(1 - \frac{Iter_i}{Max_{Iter}}\right) \quad (9)$$

$B$ denotes the level of hunger. $Iter$ implies the occurrence of multiple iterations. $rand_1$ and $y$ represent random numbers generated within the range of 0 to 1. $Max_{Iter}$ signifies an integer value that corresponds to the maximum number of iterations. If $K$ lies in the range of values greater than 1 but less than 1, the African vulture optimization algorithm initiates the search phase. Conversely, when $K$ is less than 1, the $AVOA$ algorithm shifts to the exploitation phase, mirroring the actions of a vulture scavenging for nearby food.

In the exploration phase of the $AVOA$, the vulture employs two methods to explore different regions. If the random number generated by $\text{rand}_{h_1}$ is greater than or equal to the $h_1$ parameter, it utilizes Eq. (10)(a). Conversely, when the random number produced by

$rand_{h_1}$ is less than the $h_1$ parameter, it chooses Eq. (11). The vulture's movement in this phase can be described as follows:

$$S(i+1) = \begin{cases} G(i) - Q(i) \times \mathrm{K} & if\ h_1 \geq rand_{h_1},\ (a) \\ G(i) - K + rand_2\big((uc - lc) \times rand_3 + lc\big) & if\ h_1 < rand_{h_1},\ (b) \end{cases} \quad (10)$$

$$Q(i) = |X \times G(i) - S(i)| \quad (11)$$

$S(i)$ shows the present vector indicating the vulture's position. $S(i+1)$ denotes the vector that represents the vulture's position in the next iteration. K represents the degree of satisfaction or contentment among the vultures. $uc\ and\ lc$ refer to the upper and lower limits or boundaries of the variable, respectively. $rand$ denotes a random number within the range of 0 to 1.

$X$ signifies the unpredictable or random movement undertaken by the leader vulture

Boosting the element of randomness is accomplished by utilizing $rand_2$. This leads to increased unpredictability at the environmental level, promoting diversity and preserving unique characteristics across different domains.

When $K$ drops below 1 in the AVOA algorithm, it shifts into an exploitation phase consisting of two segments, each incorporating two distinct procedures. The choice of which procedure to utilize in each segment is determined in a deterministic manner, relying on the parameters $h_2$ and $h_3$. Two distinct rotation flight procedures are implemented in the initial segment to avoid conflicts. Furthermore, $h_2$ determines the selection rate for each strategy; if $rand_{h_2}$ is greater than or equal to $h_2$, it executes the stall and outbound strategy, while if the random number is less than the $h_2$ parameter, it selects the rotational flight process.

$$S(i+1) = \begin{cases} Q(i) \times (K + rand_4) - c(t) & if\ h_3 \geq rand_{\mathrm{h2}}\ \ (a) \\ G(i) - S(i) & if\ \ h_3 < rand_{h2}\ \ (b) \end{cases} \quad (12)$$

$$c(t) = G(i) - S(i) \quad (13)$$

$G(i)$ denotes the vector's location. $rand_4$ is a number generated randomly and falls between 0 and 1. This approach begins by determining the distance between the vulture and one of two vests using Eq. (11). Following that, a spiral equation is formulated among the vultures, and their movement is guided by the parameter B, as described in Eq. (9):

$$K = S(i) \times \left(\frac{rand_5 \times G(i)}{2\pi}\right) \times \cos\big(h(i)\big) \quad (14)$$

$rand_5$ signifies a number generated at random. When $K$ falls below 0.5, the AVOA enters its second exploitation phase. If the random value generated by $\mathrm{rand_{h_3}}$ equals or surpasses the $h_3$ parameter, the vultures employ collection strategies involving training on different food sources. Should the random number produced by $\mathrm{rand_{h_3}}$ be less than the $\mathrm{h_3}$ parameter, alternative strategies are initiated as outlined in Eqs. (13) and (14):

$$U_1 = BestVulture_1(i) - \frac{BestVulture_1(i) \times S(i)}{BestVulture_1(i) \times S(i)^2} \times K \quad (15)$$

$$U_2 = BestVulture_2(i) - \frac{BestVulture_2(i) \times S(i)}{BestVulture_2(i) \times S(i)^2} \times K \quad (16)$$

$BestVulture_1(i)$ and $BestVulture_2(i)$ represent the best vultures in the first and second groups, respectively. In the final stage of the African vulture optimization algorithm, all vultures are gathered together following the procedure described in Eq. (17)(a). During this phase, conflicts and disagreements may occur among the vultures as they circle each other, as depicted by Eq. (17)(b).

$$S(i+1) = \begin{cases} \dfrac{(U_1 + U_2)}{2} & if\ h_a \geq rand_{h_3}\ \ (a) \\ G(i) - |c(t)| \times K \times L(c) & if\ h_a < rand_{h_3}\ \ (b) \end{cases} \quad (17)$$

Levy Flight ($L$) is incorporated to enhance the effectiveness of the African vulture optimization algorithm, as explained in Eq. (18). It is combined with Eq. (17)(b) to simulate the conflicts and skirmishes that may occur among the vultures during the algorithm's final phase.

$$L(v) = 0.01 \times \frac{y \times \delta}{|w|^{1/a}} \quad (18)$$

$$\delta = \left(\frac{\tau(1+b) \times \sin\left(\frac{\pi b}{2}\right)}{\tau(1+2b) \times b \times 2 \times \frac{(b-1)}{2}}\right)^{1/b} \quad (19)$$

$v$ denotes the problem's dimensionality, indicating the number of variables or dimensions involved. $b$ is established as a constant, and it remains fixed at 1.5. $y$ represents random numbers falling within the interval of 0 to 1. The core steps of the African vulture optimization algorithm are detailed using pseudo-code in Algorithm 1. Fig. 3 shows the flowchart of the AVOA.

| **Algorithm 1: Pseudo-Code of AVOA Algorithm** |
|---|
| Inputs: The maximum number of iterations T and the population size N<br>Outputs: Where the vulture is located and how fit it is |

**Initialize** the random population $h_i$ $(i = 1,2,\ldots,N)$
**while** (stopping condition is not met) **do**
**Calculate** the fitness values of the vulture
**Set** $h$BestVulture1 as the location of Vulture (First best location Best Vulture Category 1)
**Set** $h$BestVulture2 as the location of Vulture (Second best location Best Vulture Category 2)
**for** (each vulture $(h_i)$) **do**
**Select** $G(i)$
**Update** the K
**if** $(|K| \geq 1)$ **then**
**if** $(h_1 \geq randh_1)$ **then**
**Update** the location of the vulture
**else**
**Update** the location of Vulture
**if** $(|K| < 1)$ **then**
**if** $(|K| \geq 0.5)$ **then**
**if** $(h_2 \geq randh_2)$ **then**
**Update** the location of the vulture
**else**
**Update** the location of the vulture
**else**
**if** $(h_3 \geq randh_3)$ **then**
**Update** the location of the vulture
**else**
**Update** the location of the vulture
**Return** $h_{BestVulture1}$

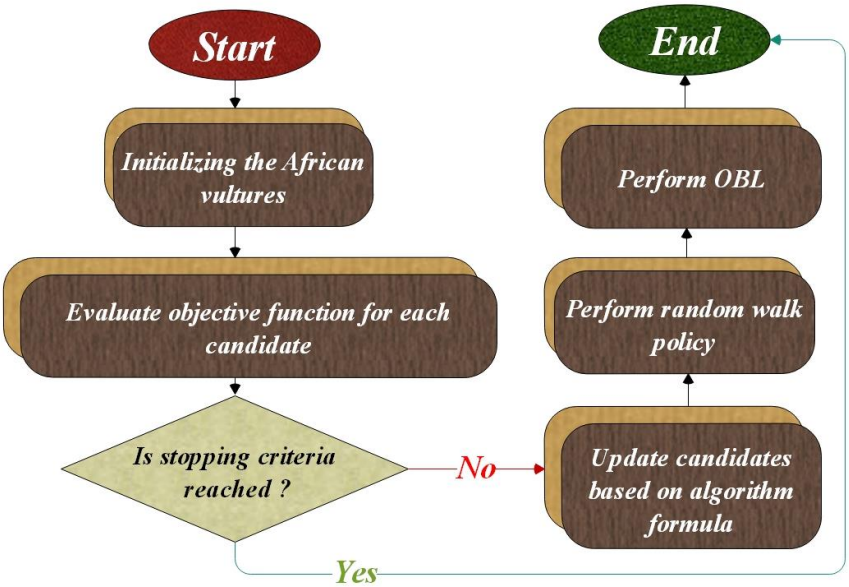



**Fig. 3.** The flowchart of AVOA

### *2.4. Sand Cat Swarm Optimization (SCSO)*

The SCSO algorithm, detailed in [33], draws inspiration from the foraging behaviors of sand cats in desert habitats. These remarkable felines can detect low-frequency sounds, enabling them to locate prey, whether it is above or below ground. The algorithm's central idea

revolves around pinpointing the optimal point within an exploration space, akin to prey in a sand cat's natural hunting context. To do this, the algorithm uses a search agent that moves steadily in the direction of the assumed position of the best value while continuously navigating the search space via periodic location updates. A prey search and assault mechanisms make up the two essential components of the SCSO algorithm, which is methodically organized. A mathematical equation that describes the population's search behavior serves as the basis for the prey search mechanism, which mimics the way by which sand cats look for prey in the wild. This equation mirrors the collective actions of sand cats as they scour their environment for potential targets, forming the algorithm's foundational approach to optimization and discovering the desired solution within the exploration space.

$$\vec{X}(t+1) = \vec{r}.\vec{X}_b(t) - rand(0,1).\vec{X}_c(t) \quad (20)$$

$\vec{X}$ represents the position vector of the search agent. $t$ signifies the current iteration's number. $\vec{X}_b(t)$ represents the position of the best candidate at iteration $t$. $\vec{X}_c(t)$ signifies the current position of the search agent at iteration $t$. $r$ represents the range of sensitivity of sand cats to low-frequency noise, and this sensitivity can be described as follows:

$$\vec{r} = \vec{r}_G \times rand(0,1) \quad (21)$$

$\vec{r}_G$ denotes the general sensitivity range, which decreases linearly from 2 to 0. This can be described as follows:

$$\vec{r}_G = s_M - \left(\frac{s_M \times iter_c}{iter_{max}}\right) \quad (22)$$

$iter_c$ represents the current iteration, and $iter_{max}$ represents the maximum number of iterations. Additionally, given that sand cats sense low frequencies of 2 $kHz$, the value of $s_M$ is set to 2. After the prey search phase, the SCSO algorithm starts the prey assault phase. The prey attack mechanism for the population of sand cats is as follows:

$$\vec{X}_{md} = \left|rand(0,1).\vec{X}_b(t) - \vec{X}_c(t)\right| \quad (23)$$

$$\vec{X}(t+1) = \vec{X}_b(t) - \vec{r}.\vec{X}_b(t).cos(\theta) \quad (24)$$

$\theta$ represents a random angle ranging from 0 to 360 degrees. As a result, the cosine function $\cos(\theta)$ yields values within the range of $-1$ to 1. $\vec{X}_{md}$ refers to the random position that is generated based on both the best position and the current position. Through this approach, every member of the population can move in distinct circular directions. Each sand cat selects a random angle, enabling them to steer clear of local optimal traps as they close in on the prey's location. The random angle mentioned in Eq. (24) plays a crucial role in influencing the hunting and search direction of the agent.

The SCSO method uses an adaptive factor known to preserve equilibrium between the exploration and exploitation phases as $\vec{r}$. This factor can be elucidated as follows:

$$R = 2 \times \vec{r}_G \times rand(0,1) - \vec{r}_G \quad (25)$$

$\vec{r}_G$ diminishes gradually from 2 to 0 in a linear fashion with the progression of iterations. The revised explanation of the sand cat's positions throughout both the exploration and exploitation phases can be stated as follows:

$$\vec{X}(t+1) = \begin{cases} \vec{r}.(\vec{X}_b(t) - rand(0,1).\vec{X}_c(t)) & |R| > 1 \\ \vec{X}_b(t) - \vec{r}.\vec{X}_b(t).cos(\theta) & |R| \leq 1 \end{cases} \quad (26)$$

When $R's$ absolute value is less than or equal to 1, the SCSO algorithm's search agent assaults the intended victim. In cases where $|R| > 1$, the search agent switches to a global search mode, exploring potential solutions across a wider range. It is worth noting that each sand cat possesses a distinct search radius during the exploration phase, thereby preventing the algorithm from getting trapped in local optimal solutions.

Algorithm 2 is described in the pseudo-code [34], and Fig. 4 displays the flowchart.

**Algorithm 2: Pseudo-Code of SCSO Algorithm**

Population initiation
Do the fitness function calculation
Initialize the $r$; $\vec{r}_G$; $R$
**while** ($t \leq iter_{max}$) **do**
**for** each agent **do**
Get a random angle ($0^\circ \leq \theta \leq 360^\circ$)
**if** ($|R| \leq 1$) **then**
Update the position of the search agent
**else**
Update the search agent position

**end if**
**end for**
$t = t + 1$
**end while**

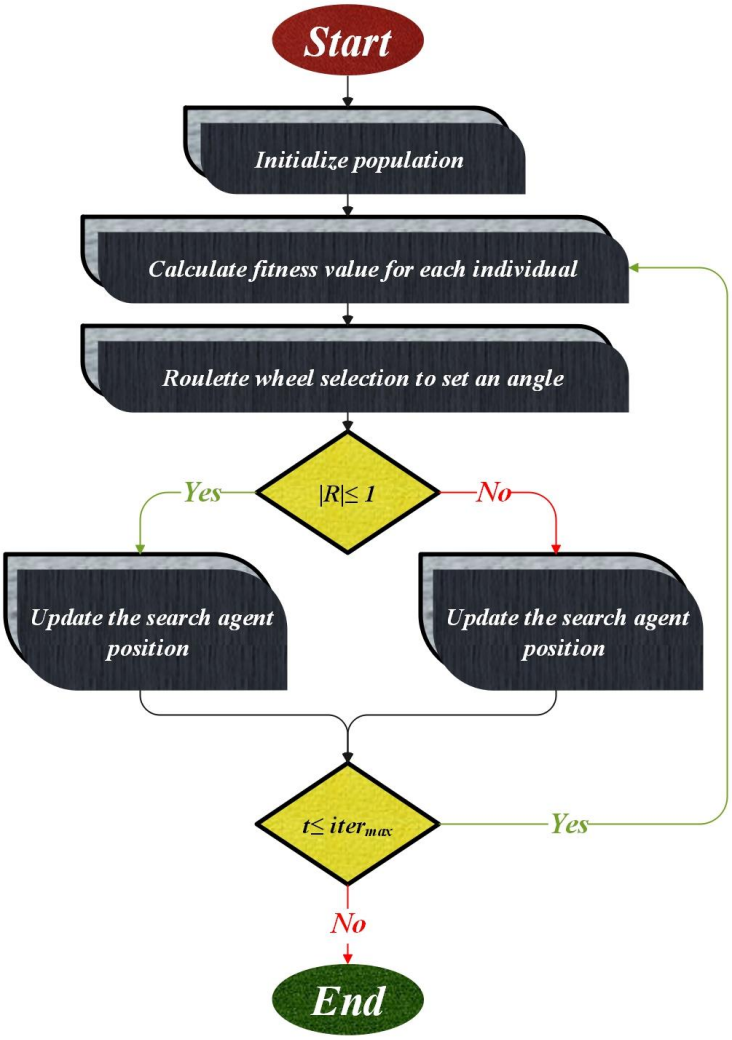



**Fig. 4.** The flowchart of SCSO

### *2.5. Performance evaluation methods*

Hybrid models undergo thorough scrutiny using diverse metrics, particularly focusing on their ability to precisely measure errors and correlations. These metrics serve as essential tools for evaluating hybrid models' effectiveness across various applications. They include Root Mean Squared Error ($RMSE$), Coefficient of Determination ($R^2$), Mean Squared Error ($MSE$), Nash-Sutcliffe Efficiency ($NSE$), and Weighted Absolute Percentage Error ($WAPE$). These measures critically assess how well hybrid models perform in different scenarios, offering valuable insights into their accuracy, reliability, and predictive capacity. By leveraging these metrics, one gains a comprehensive understanding of hybrid models' performance, crucial for optimizing their application in real-world contexts.

- **Root Mean Squared Error**

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(s_i - p_i)^2} \quad \textbf{\textit{Lower is desirable}} \tag{27}$$

- **Coefficient of Determination**

$$R^2 = \left( \frac{\sum_{i=1}^{N}(p_i - \bar{p})(s_i - \bar{s})}{\sqrt{[\sum_{i=1}^{N}(p_i - \bar{s})^2][\sum_{i=1}^{N}(s_i - \bar{s})^2]}} \right)^2 \quad \textbf{\textit{Higher is desirable}} \quad (28)$$

- **Mean Squared Error**

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(s_i - p_i)^2 \quad \textbf{\textit{Lower is desirable}} \quad (29)$$

- **Weighted Absolute Percentage Error**

$$WAPE = \frac{\sum_{i=1}^{N}|s_i - p_i|}{\sum_{i=1}^{N}|p_i|} \quad \textbf{\textit{Lower is desirable}} \quad (30)$$

- **Nash-Sutcliffe Efficiency**

$$NSE = 1 - \frac{\sum_{i=1}^{N}(s_i - p_i)^2}{\sum_{i=1}^{N}(p_i - \bar{p})^2} \quad \textbf{\textit{Lower is desirable}} \quad (31)$$

- $N$ correspond to the numbers associated with the samples.
- $s_i$ represents the estimated value.
- $\bar{s}$ represents the mean of the estimated values.
- $p_i$ refers to the experimental value,
- $\bar{p}$ shows the experimental amount's average.

## 3. Investigate the convergence

The convergence analysis, as indicated by RMSE, is visually depicted in Fig. 5 for these hybrid models over 150 iterations. The x-axis tracks the number of iterations, while the y-axis illustrates their accuracy, presenting a comprehensive view of their performance trends. Notably, among the models considered, the NBSC hybrid model achieved the lowest RMSE of 0.9, hitting its best iteration count by the 120th iteration. Conversely, the NBAV hybrid model, identified by its blue color, emerged as a leader, showcasing a remarkable RMSE of 1.4. This outstanding precision highlights its efficacy in predicting heating. What sets the NBAV hybrid model apart goes beyond its accuracy; it efficiently reaches optimal convergence by the 135th iteration. This unique characteristic underscores consistent lower RMSE and swift convergence, distinguishing it from the struggling NBSC model.

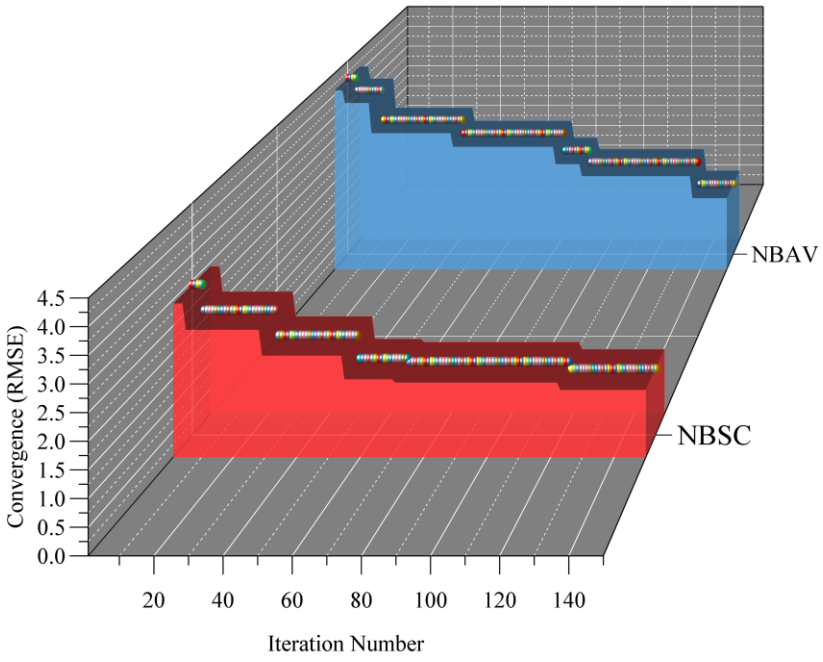



**Fig. 5.** The 3D wall convergence for the output hybrid models

## 4. Results and discussion

The results of this investigation have been intricately detailed in Table 2, providing a comprehensive numerical portrayal of the outcomes achieved through Hybrid Learning (HL). After an in-depth scrutiny of these findings, it becomes increasingly apparent that integrating optimization techniques (AVOA and SCSO) has significantly enhanced the Naive Bayes (NB) model's ability to make estimations. Adopting these optimization methodologies has yielded positive developments across a spectrum of performance metrics. These enhancements include a notable rise in the $R^2$ values, which signify the

model's explanatory power, and a simultaneous reduction in error metrics, encompassing RMSE, MSE, NSE, and WAPE. These improvements collectively highlight the substantial impact of optimizer integration on the NB model's predictive capabilities and overall performance.

- **Coefficient of determination ($R^2$):**

Based on the data in Table 2, a consistent pattern emerges where $R^2$ values for all models consistently show higher values during the training phase compared to the validation and testing phases. This suggests the models were inadequately trained, leading to subpar performance in subsequent phases. Notably, the NBSC model excelled during the training phase, achieving the highest $R^2$ value of 0.987, signifying its status as the top-performing model. Conversely, the NB model lagged with an $R^2$ value of 0.954, making it the least effective model in this study. These findings underscore the critical role of practical training in model performance and highlight the substantial performance gap between the best and worst-performing models.

- **Root Mean Squared Error (RMSE):**

Considering the results, it becomes clear that within the group of models assessed, the NBSC model demonstrated the lowest RMSE value ($RMSE = 1.166$) during the training phase, solidifying its position as the most accurate model in terms of prediction. In contrast, the NB model, with an RMSE value of 4.501, was identified as the least effective among the models, indicating its relatively inferior predictive performance.

- **Mean Squared Error (MSE):**

When examining the MSE results, it is clear that the NBSC model excelled compared to the NB and NBAV models, as it exhibited lower MSE values. To be more specific, the NBSC model achieved the most favorable MSE value, registering at 1.359, while the NB model yielded the least favorable result with the highest MSE value of 20.263 among all the models. This difference in MSE values underscores the superior predictive precision of the NBSC model when compared to the alternatives.

- **Nash-Sutcliffe Efficiency (NSE)**

Considering the NSE values, it becomes clear that the NBSC model demonstrated superior performance compared to the NB and NBAV models. The NSE values spanned from a minimum of 0.797 for the NB model to a maximum of 0.987 for the NBSC model among all the examined models. This fluctuation in NSE values highlights that the NB model exhibited a broader spectrum of uncertainty in its predictions in contrast to the NBSC and NBAV models, with the NBSC model being the most accurate among the three.

- **Weighted Absolute Percentage Error (WAPE):**

Upon reviewing the WAPE values, it becomes evident that, during the training phase, the NBSC model attained the most favorable WAPE value, registering at 0.043. This positions it as the top-performing model among all the models under consideration. Conversely, the NB model was identified as the least effective model, displaying the highest WAPE value of 0.156. This difference in WAPE values emphasizes the superior predictive precision of the NBSC model when compared to the relatively poorer performance of the NB model.

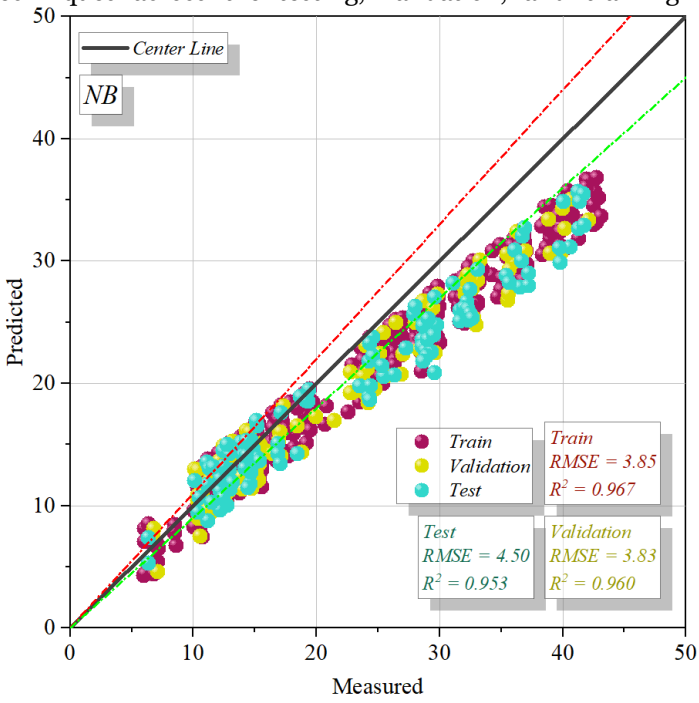
To sum it up, when assessing all three models for estimating HL values and taking into account both evaluator ratings and error metrics, it becomes clear that the most favorable model is the one that combines NB with the SCSO algorithm, referred to as NBSC. This model stands out for its outstanding performance, boasting the lowest error value at 1.166 during the training phase. Moreover, it also achieved the highest $R^2$ value of 0.987 during the same training phase, surpassing the performance of all three individual components. Consequently, the NBSC model emerges as the optimal choice among the three models for estimating HL values.

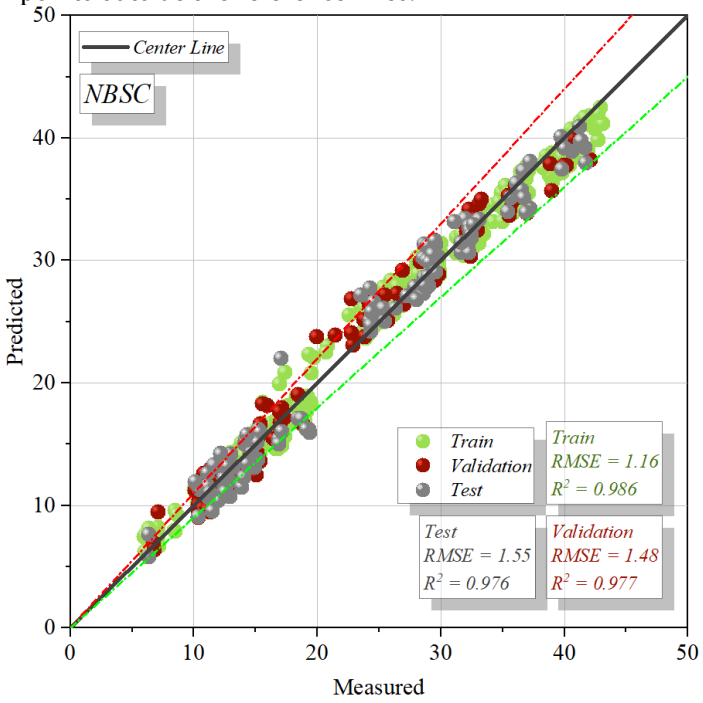
**Table 2.** The result of developed models for NB

| Model | Phase | Index values | | | | |
|---|---|---|---|---|---|---|
| | | RMSE | $R^2$ | MSE | WAPE | NSE |
| NB | Train | 3.859 | 0.967 | 14.894 | 0.139 | 0.855 |
| | Validation | 3.834 | 0.960 | 14.699 | 0.145 | 0.848 |
| | Test | 4.501 | 0.954 | 20.263 | 0.156 | 0.797 |
| | All | 3.958 | 0.964 | 15.669 | 0.142 | 0.846 |
| NBSC | Train | 1.166 | 0.987 | 1.359 | 0.043 | 0.987 |
| | Validation | 1.480 | 0.977 | 2.192 | 0.056 | 0.977 |
| | Test | 1.552 | 0.976 | 2.409 | 0.054 | 0.976 |
| | All | 1.281 | 0.984 | 1.641 | 0.046 | 0.984 |
| NBAV | Train | 1.510 | 0.978 | 2.280 | 0.056 | 0.978 |

| | | | | | |
|---|---|---|---|---|---|
| Validation | 1.925 | 0.962 | 3.705 | 0.072 | 0.962 |
| Test | 1.850 | 0.966 | 3.421 | 0.067 | 0.966 |
| All | 1.632 | 0.974 | 2.664 | 0.060 | 0.974 |

Fig. 6 employs scatter plots to illustrate the correlation between predicted and observed values for HL, explicitly focusing on assessing RMSE and $R^2$ metrics. RMSE, indicative of data point concentration, diminishes as accuracy improves, whereas $R^2$ draws data points nearer to the central axis. Three models (NB, NBSC, and NBAV) were devised by integrating the NB model with two optimization techniques across the testing, validation, and training phases. Fig. 6 serves as a visual summary of the outcomes, vividly demonstrating the superior performance of the NBSC hybrid model, a fusion of the NB approach with the SCSO optimizer. This superiority is evidenced by the tight clustering of data points aligned with the central line. Conversely, the figures reveal that the NB single model performed the worst, as indicated by the abundance of data points outside the reference lines.



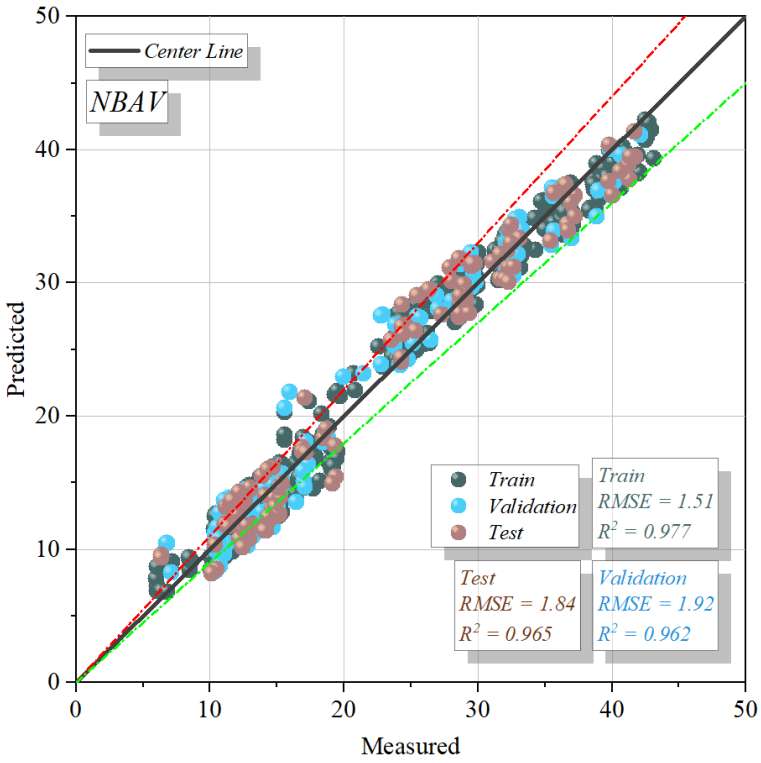
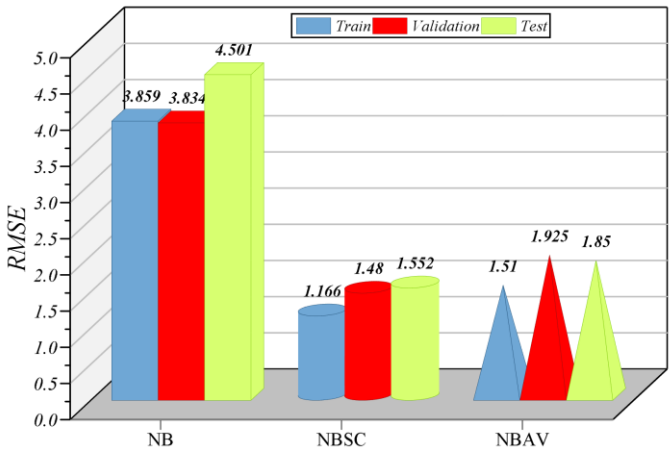
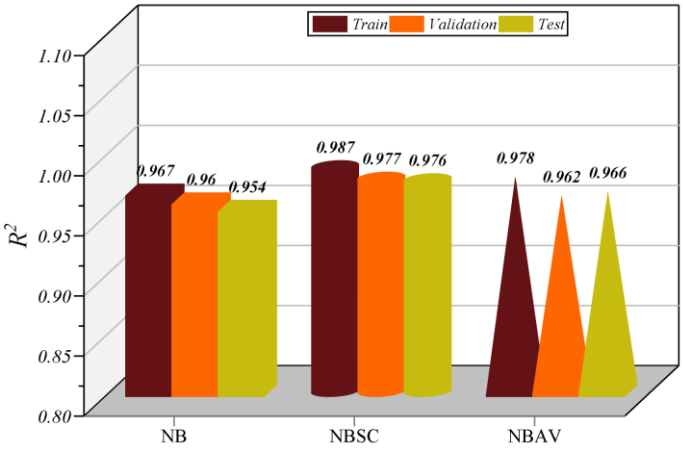



**Fig. 6.** The scatter plot for developed hybrid models

Fig. 7 illustrates a comparison among three machine learning models regarding their performance in predicting HL for HVAC systems: NB, NBSC, and NBAV. Their performance is assessed using $R^2$, RMSE, and NSE metrics. $R^2$ gauges the extent to which the model explains variance in HL; a score of 1 signifies a perfect explanation, while 0 suggests no explanation. RMSE measures the square root of the average squared difference between predicted and actual loads, with lower values indicating better fit. NSE compares prediction accuracy against a baseline model, with 1 indicating perfect prediction and 0 indicating no improvement over the baseline. The NBSC model surpasses the others across all metrics, boasting an $R^2$ of 0.987, RMSE of 1.166, and NSE of 0.954. This underscores NBSC's status as the most accurate and dependable model for HVAC heating load prediction.



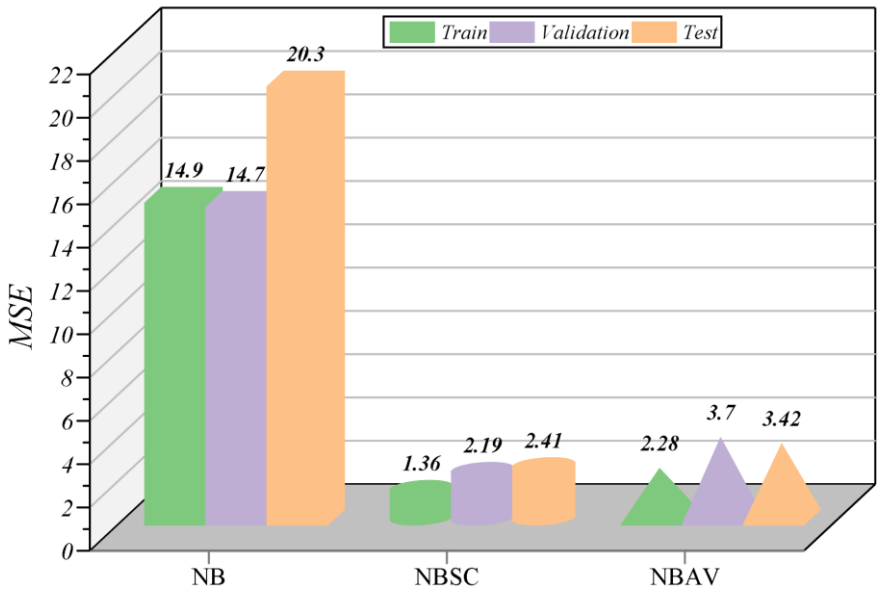
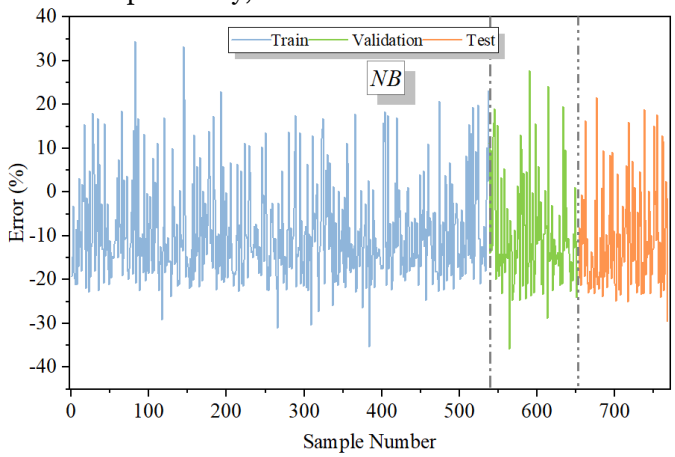
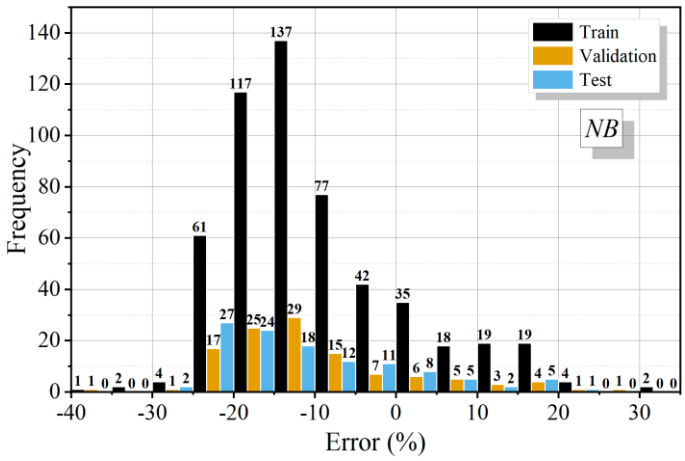



**Fig. 7.** Comparison between models RMSE, $R^2$, and NSE

The error percentages for hybrid machine learning models are determined through analysis of Line and histogram plots in Fig. 8, showcasing a comparison among these models used in HVAC system heating load prediction. The hybrids considered are NBSC, NBAV, and NB, with error percentages calculated across training, validation, and test datasets. Fig. 8 depicts that the NBSC model exhibits the lowest error percentages across all three datasets. Specifically, it reaches a maximum error percentage of 19% for training, 17% for validation, and 15% for the test dataset. Meanwhile, the NBAV model demonstrates commendable performance, displaying maximum error percentages of 24% for training, 22% for validation, and 20% for the test dataset. In contrast, the NB model records the highest error percentages among the trio, peaking at 35% for training, 33% for validation, and 31% for the test dataset.



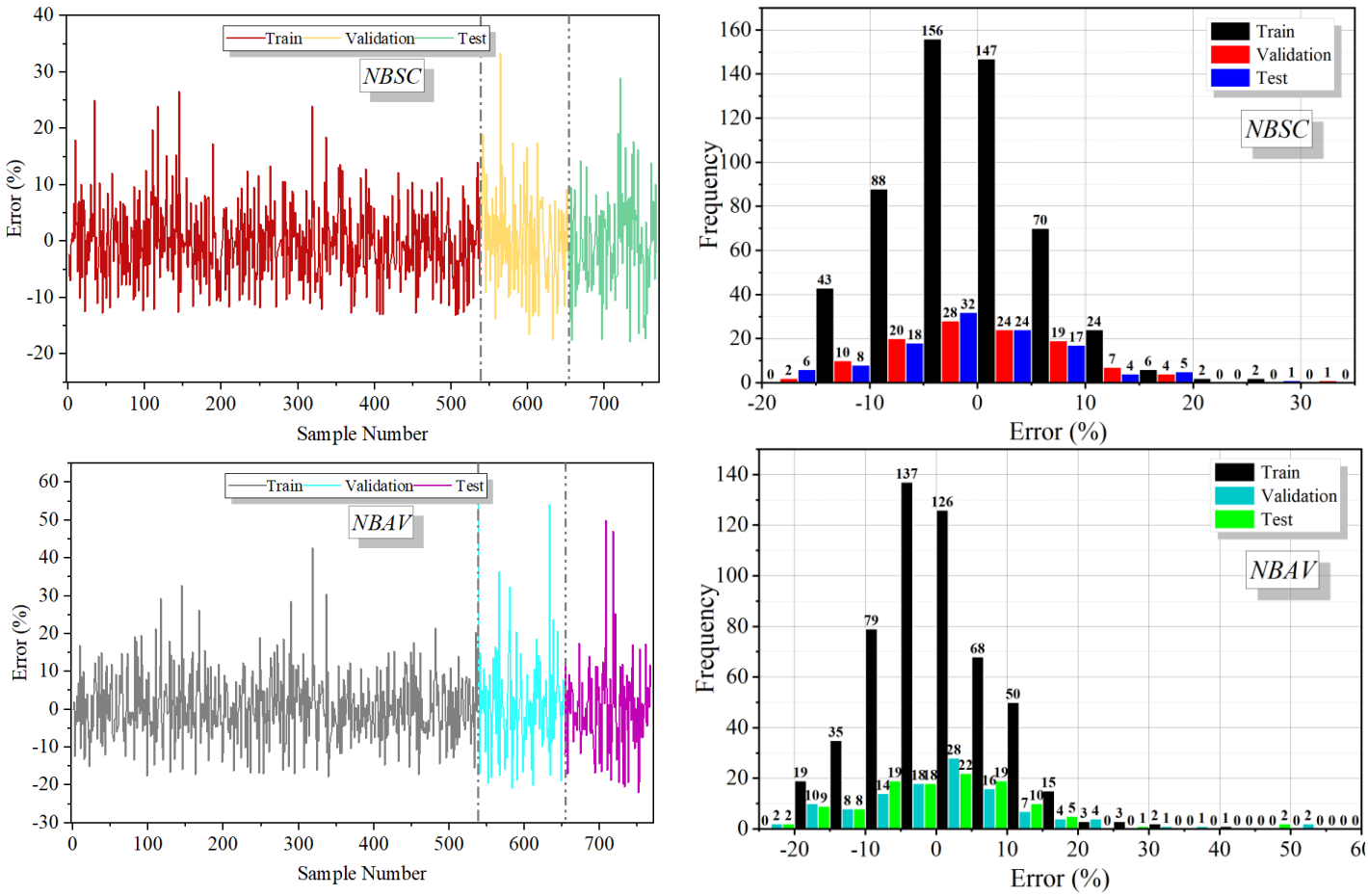



**Fig. 8.** The error percentage for the hybrid models is based on Line and histogram plots

Fig. 9 provides a detailed graphic representation of the error distribution during the training, validation, and testing stages when predicting HL values using three different models. Importantly, during the training phase, the NB model experienced the highest errors, whereas the NBSC model experienced significantly fewer errors. After close examination, a recurring trend that favored the NBSC hybrid model in every stage was discovered. More specifically, in the training stage, the NB model showed errors in the interval of -30 to 25, demonstrating a wider range of deviations. On the other hand, the NBSC model, which was hailed as the best performer, demonstrated exceptional accuracy with errors mostly confined to the smaller range of -16 to 17. The NBSC model stands out among the three due to its concentrated error range, which highlights its superior predictive accuracy.

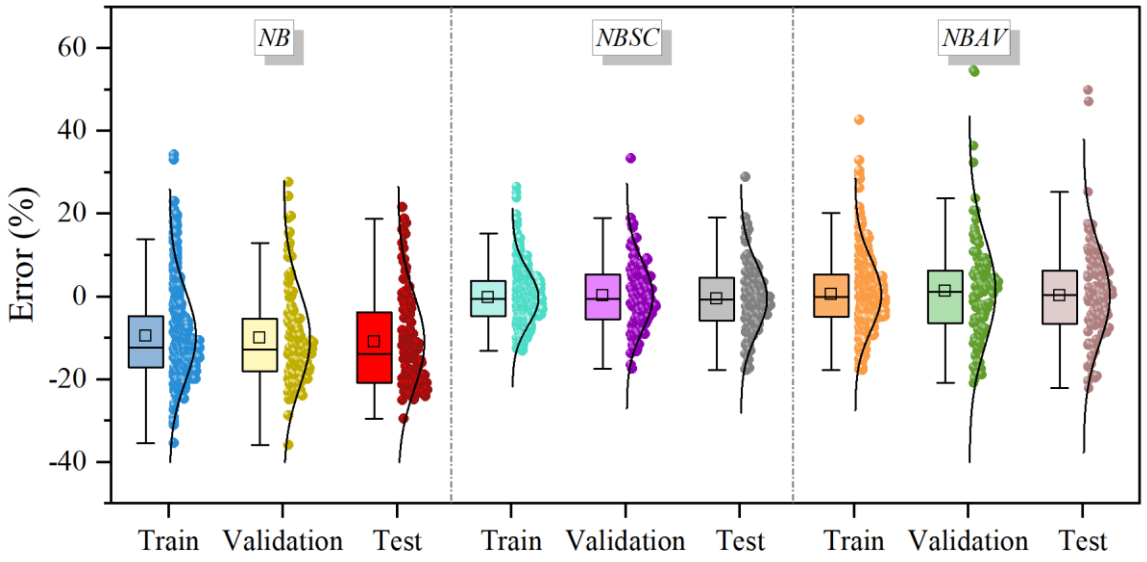



**Fig. 9.** The box normal of errors among the developed models

Fig. 10, depicted as a Taylor diagram, offers a comparative analysis of three distinct machine learning models employed in predicting heating loads of residential buildings: NB, NBSC, and NBAV. This diagram serves as a visual representation of these models' performance based on three pivotal metrics: correlation coefficient, standard deviation, and centered root mean squared error (RMSE).

The correlation coefficient gauges the strength of the linear relationship between predicted and actual heating loads, while the standard deviation denotes the variability in predicted heating loads. Meanwhile, the centered RMSE provides insight into the average error between predicted and actual heating loads. Remarkably, the NBSC model emerges as the closest to the reference point within the Taylor diagram, signifying its superior overall performance. With a correlation coefficient of 0.987, a standard deviation of 1.166, and a centered RMSE of 0.954, the NBSC model showcases exceptional accuracy. Following closely, the NBAV model also exhibits favorable performance, closely trailing the reference point. It holds a correlation coefficient of 0.976, a standard deviation of 1.241, and a centered RMSE of 0.932.

Conversely, the NB model significantly lags, positioned farthest from the reference point within the diagram, denoting its comparatively poorer overall performance. Evidenced by a correlation coefficient of 0.962, a standard deviation of 1.36, and a centered RMSE of 0.907, the NB model showcases less accuracy compared to its counterparts.

These results notably advocate for the NBSC model's position as the most accurate and reliable model for predicting heating loads within HVAC systems, reaffirming its superiority among the evaluated models.

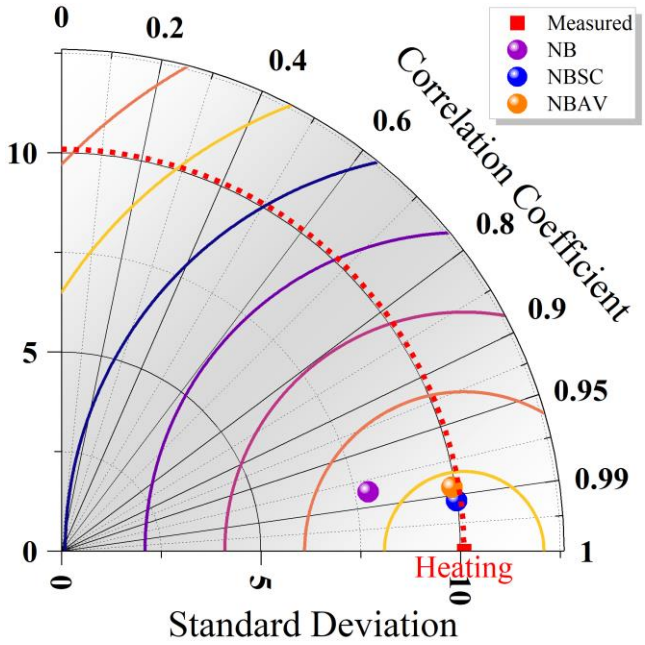



**Fig. 10.** The Taylor diagram of related models

## 5. Conclusion

The pursuit of robust and precise predictive models holds significant promise for efficiency gains and cost reduction in the case of Heating Load (HL) prediction. This research rested upon the Naïve Bayes (NB) foundation as the fundamental framework for constructing these predictive models. Two optimization algorithms, specifically the African Vultures Optimization Algorithm (AVOA) and Sand Cat Swarm Optimization (SCSO), were seamlessly integrated to enhance the model's precision and effectiveness. The research findings underscore the effective utilization of both optimization techniques in

developing predictive models for HL value estimation. The outcomes unambiguously demonstrate the NBSC model's outstanding precision compared to its counterparts. This is notably manifested in its achievement of the RMSE value ($RMSE_{train} = 1.166$) and the highest coefficient of determination value ($R^2{}_{train} = 0.987$). These results highlight the model's exceptional ability to estimate HL values accurately. Incorporating AVOA and SCSO optimization techniques within the unified NB model has yielded substantial enhancements in R-squared values, with increments of 1.13% and 2.07%, respectively. It is imperative to emphasize that the NB model exhibited the least favorable performance among all the models examined. This was evident due to its recording of the highest error value, specifically $RMSE_{test} = 4.501$, and the lowest Coefficient of Determination ($R^2{}_{test} = 0.954$). In summary, these metrics collectively indicate a diminished effectiveness of the NB model in accurately predicting HL values. Consequently, the NBSC model, an amalgamation of the NB model with the SCSO optimizer, emerges as an exceptional performer, showcasing remarkable predictive capabilities in stark contrast to the NB model's relatively suboptimal performance. While the study excels in accurately predicting heating load within HVAC systems using innovative algorithms, it has limitations. These include a narrow focus primarily on predictive accuracy, potential constraints in applying the models to diverse scenarios, reliance on specific algorithms without exploring alternatives, lack of extensive real-world validation, limited consideration of long-term model performance, and reliance on specific datasets. Future research could address these limitations for broader applicability and a more comprehensive understanding of energy-efficient building management.

## REFERENCES


[1] H. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.

[2] W. Jin *et al.*, "A novel building energy consumption prediction method using deep reinforcement learning with consideration of fluctuation points," *Journal of Building Engineering*, vol. 63, p. 105458, 2023.

[3] A. N. Sharif, S. K. Saleh, S. Afzal, N. S. Razavi, M. F. Nasab, and S. Kadaei, "Evaluating and Identifying Climatic Design Features in Traditional Iranian Architecture for Energy Saving," 2022.

[4] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, "Vector field-based support vector regression for building energy consumption prediction," *Applied Energy*, vol. 242, pp. 403–414, 2019.

[5] Y. Zhao, C. Zhang, Y. Zhang, Z. Wang, and J. Li, "A review of data mining technologies in building energy systems: Load prediction, pattern identification, fault detection and diagnosis," *Energy and Built Environment*, vol. 1, no. 2, pp. 149–164, 2020.

[6] D. Darwazeh, J. Duquette, B. Gunay, I. Wilton, and S. Shillinglaw, "Review of peak load management strategies in commercial buildings," *Sustainable Cities and Society*, vol. 77, p. 103493, 2022.

[7] J. Wang, J. Hou, J. Chen, Q. Fu, and G. Huang, "Data mining approach for improving the optimal control of HVAC systems: An event-driven strategy," *Journal of Building Engineering*, vol. 39, p. 102246, 2021.

[8] A. Gellert, U. Fiore, A. Florea, R. Chis, and F. Palmieri, "Forecasting electricity consumption and production in smart homes through statistical methods," *Sustainable Cities and Society*, vol. 76, p. 103426, 2022.

[9] W. Zhang, Q. Chen, J. Yan, S. Zhang, and J. Xu, "A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting," *Energy*, vol. 236, p. 121492, 2021.

[10] J. Runge and R. Zmeureanu, "Forecasting energy use in buildings using artificial neural networks: A review," *Energies*, vol. 12, no. 17, p. 3254, 2019.

[11] B. B. Ekici and U. T. Aksoy, "Prediction of building energy consumption by using artificial neural networks," *Advances in Engineering Software*, vol. 40, no. 5, pp. 356–362, 2009.

[12] M. R. Akbarzadeh, H. Ghafourian, A. Anvari, R. Pourhanasa, and M. L. Nehdi, "Estimating Compressive Strength of Concrete Using Neural Electromagnetic Field Optimization," *Materials*, vol. 16, no. 11, p. 4200, 2023.

[13] T. Čegovnik, A. Dobrovoljc, J. Povh, M. Rogar, and P. Tomšič, "Electricity consumption prediction using artificial intelligence," *Central European Journal of Operations Research*, pp. 1–19, 2023.

[14] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, and A. Amira, "Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives," *Applied Energy*, vol. 287, p. 116601, 2021.

[15] S. Bourhnane, M. R. Abid, R. Lghoul, K. Zine-Dine, N. Elkamoun, and D. Benhaddou, "Machine learning for energy consumption prediction and scheduling in smart buildings," *SN Applied Sciences*, vol. 2, pp. 1–10, 2020.

[16] F. Masoumi, S. Najjar-Ghabel, A. Safarzadeh, and B. Sadaghat, "Automatic calibration of the groundwater simulation model with high parameter dimensionality using sequential uncertainty fitting approach," *Water Supply*, vol. 20, no. 8, pp. 3487–3501, Dec. 2020, doi: 10.2166/ws.2020.241.

[17] S. Shamshirband *et al.*, "Heat load prediction in district heating systems with adaptive neuro-fuzzy method," *Renewable and Sustainable Energy Reviews*, vol. 48, pp. 760–767, 2015.

[18] G. Zhou, H. Moayedi, M. Bahiraei, and Z. Lyu, "Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of

heating and cooling loads of residential buildings," *Journal of Cleaner Production*, vol. 254, p. 120082, 2020.
[19] J. Song, G. Xue, X. Pan, Y. Ma, and H. Li, "Hourly heat load prediction model based on temporal convolutional neural network," *IEEE Access*, vol. 8, pp. 16726–16741, 2020.
[20] S. Sajjadi *et al.*, "Extreme learning machine for prediction of heat load in district heating systems," *Energy and Buildings*, vol. 122, pp. 222–227, 2016.
[21] Y. Liu, X. Hu, X. Luo, Y. Zhou, D. Wang, and S. Farah, "Identifying the most significant input parameters for predicting district heating load using an association rule algorithm," *Journal of Cleaner Production*, vol. 275, p. 122984, 2020.
[22] M. Gong, Y. Bai, J. Qin, J. Wang, P. Yang, and S. Wang, "Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin," *Journal of Building Engineering*, vol. 27, p. 100950, 2020.
[23] A. Moradzadeh, A. Mansour-Saatloo, B. Mohammadi-Ivatloo, and A. Anvari-Moghaddam, "Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings," *Applied Sciences*, vol. 10, no. 11, p. 3829, 2020.
[24] V. V Mokeev, "Prediction of heating load and cooling load of buildings using neural network," in *2019 International Ural Conference on Electrical Power Engineering (UralCon)*, IEEE, 2019, pp. 417–421.
[25] I. Karijadi and S.-Y. Chou, "A hybrid RF-LSTM based on CEEMDAN for improving the accuracy of building energy consumption prediction," *Energy and Buildings*, vol. 259, p. 111908, 2022.
[26] À. Nebot and F. Mugica, "Energy performance forecasting of residential buildings using fuzzy approaches," *Applied Sciences*, vol. 10, no. 2, p. 720, 2020.
[27] M. Adegoke, A. Hafiz, S. Ajayi, and R. Olu-Ajayi, "Application of Multilayer Extreme Learning Machine for Efficient Building Energy Prediction," *Energies*, vol. 15, no. 24, p. 9512, 2022.
[28] Q. He *et al.*, "Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF Classifier, and RBF Network machine learning algorithms," *Science of the total environment*, vol. 663, pp. 1–15, 2019.
[29] K. Vembandasamy, R. Sasipriya, and E. Deepa, "Heart diseases detection using Naive Bayes algorithm," *International Journal of Innovative Science, Engineering & Technology*, vol. 2, no. 9, pp. 441–444, 2015.
[30] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol*, vol. 2, no. 1, p. 53, 2017.
[31] H. Zhang and D. Li, "Naïve Bayes text classifier," in *2007 IEEE International Conference on Granular Computing (GRC 2007)*, IEEE, 2007, p. 708.
[32] B. Abdollahzadeh, F. S. Gharehchopogh, and S. Mirjalili, "African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems," *Computers & Industrial Engineering*, vol. 158, p. 107408, 2021.
[33] A. Seyyedabbasi and F. Kiani, "Sand Cat swarm optimization: A nature-inspired algorithm to solve global optimization problems," *Engineering with Computers*, vol. 39, no. 4, pp. 2627–2651, 2023.
[34] Y. Li and G. Wang, "Sand cat swarm optimization based on stochastic variation with elite collaboration," *IEEE Access*, vol. 10, pp. 89989–90003, 2022.