Enhancing Building Energy Efficiency: A Hybrid Meta-Heuristic Approach for Cooling Load Prediction

Chenguang Wang[1]*, Yanjie Zhou[2], Libin Deng[3], Ping Xiong[4], Jiarui Zhang[5], Jiamin Deng[6], Zili Lei[7]
School of Municipal and Geomatics Engineering, Hunan City University, Yiyang, Hunan, 413000, China[1, 5, 6, 7]
Heshan District Garden Landscaping Co., Ltd., Yiyang, Hunan, 413000, China[2]
Hunan Construction Investment Group Co., Ltd., Changsha, Hunan, 410004, China[3]
School of Mechanical and Electrical Engineering, Hunan City University, Yiyang, Hunan, 413000, China[4]

***Abstract*—The research tackles the complex problem of accurately predicting cooling loads in the context of energy efficiency and building management. It presents a novel approach that increases the precision of cooling load forecasts by utilizing machine learning (ML). The main objective is to incorporate a hybridization strategy into Radial Basis Function (RBF) models, a commonly used method for cooling load prediction, to improve their effectiveness. This new method significantly increases accuracy and reliability. The resulting hybrid models, which combine two powerful optimization techniques, outperform the state-of-the-art approaches and mark a major advancement in predictive modelling. The study performs in-depth analyses to compare standalone and hybrid model configurations, guaranteeing an unbiased and thorough performance evaluation. The deliberate choice of incorporating the Self-adaptive Bonobo Optimizer (SABO) and Differential Squirrel Search Algorithm (DSSA) underscores the significance of leveraging the distinctive strengths of each optimizer. The study delves into three variations of the RBF model: RBF, RBDS, and RRBSA. Among these, the RBF model, integrating the SABO optimizer (RBSA), distinguishes itself with an impressive R^2 value of 0.995, denoting an exceptionally close alignment with the data. Furthermore, a low Root Mean Square Error (RMSE) value of 0.700 underscores the model's remarkable precision. The research showcases the effectiveness of fusing ML techniques in the RBSA model for precise cooling load predictions. This hybrid model furnishes more dependable insights for energy conservation and sustainable building operations, thereby contributing to a more environmentally conscious and sustainable future.**

***Keywords*—Building energy; cooling load; machine learning; radial basis function; self-adaptive bonobo optimizer; differential squirrel search algorithm**

I. Introduction

In the contemporary discourse surrounding global energy challenges, there exists an escalating emphasis on imperative energy conservation measures. This urgency is particularly salient within the purview of the global building sector, constituting a substantial fraction of overall energy consumption. Within this context, the imperative to optimize energy efficiency has prompted a meticulous examination of key metrics, where Cooling Load (CL) and Dynamic Air-Conditioning Load assume pivotal roles in the paradigm of building design and operation [1], [2]. Of particular significance is the latter, as it intricately interfaces with the orchestration of heating, ventilation, and air conditioning (HVAC) systems, endowing a spectrum of advantages including expeditious cooling startups, precise management of peak demand, cost optimization, and heightened energy efficiency within cooling storage systems [3]. The Dynamic Air-Conditioning Load, as a multifaceted entity, stands as the linchpin for realizing energy and cost efficiency objectives across a diverse array of HVAC systems. It functions as a discerning orchestrator, facilitating seamless coordination and adjustment of HVAC operations in response to the dynamically evolving thermal requisites of a given structure [4]. This adaptability becomes instrumental in meeting the heterogeneous and evolving demands posed by residential, commercial, and industrial structures alike. Nevertheless, the realization of these efficiency goals is not bereft of challenges. The accurate prediction of building cooling loads remains a formidable task, characterized by the intricate interplay between the optical and thermal properties of the building and meteorological data [5]. The integration of these factors into the cooling load prediction process engenders a complex network that necessitates innovative solutions. The scientific pursuit of building cooling load prediction has endured for several years, embodying an enduring commitment to surmounting these challenges [6].

In the quest for precision, diverse methodological avenues have been explored. Engineering-based feature extraction techniques have surfaced as a valuable approach, probing into the intrinsic thermal properties of diverse building materials. By eliciting domain-specific insights from data, these techniques shed light on the nuanced ways in which distinct materials exert influence on the cooling load [7]. From concrete to glass, each material introduces a unique set of thermal dynamics that contributes to the overall cooling load profile. These insights not only enrich the comprehension of thermal intricacies but also facilitate the development of tailored solutions for augmented energy efficiency. Beyond engineering-based techniques, statistical feature extraction methods have assumed an integral role in the predictive arsenal. Employing mathematical approaches, these methods unearth pertinent information from the data, revealing patterns and relationships that may elude conventional scrutiny [8].

This data-driven approach affords a comprehensive exploration of the intricate factors influencing cooling loads, offering a holistic perspective that transcends traditional analytical confines. Advancing the pursuit of precision, structural feature extraction methods have evolved to scrutinize the underlying data structure itself. These methods endeavour to uncover latent dependencies and patterns that might elude conventional analyses. By delving into the depths of data intricacies, structural feature extraction unveils critical insights that can significantly enhance the precision of cooling load predictions [9].

A multitude of techniques has been devised for optimizing HVAC systems, which can be broadly classified into three categories: artificial intelligence (AI), simulation, and regression analysis. Simulation tools like DOE-2, ESP-r [10], TRNSYS [11], EnergyPlus and ACO [12] are used to estimate CL when complete building data is accessible. Accurately measuring different building parameters, however, presents practical challenges [13]. There are different ways to simplify the process of building models, but it still takes a lot of time and labour to complete. Furthermore, the accuracy of simulated load results depends on the model's precision and the quality of the weather data. In real-time applications like online prediction or optimal operational control, the usefulness of simulation software is constrained [14].

Against this backdrop, recent scholarship has spotlighted the substantial potential of cutting-edge Machine Learning (ML) techniques in transforming the landscape of cooling load predictions [15]. ML, with its adeptness in discerning complex patterns and adapting to evolving data dynamics, emerges as a potent ally in the pursuit of precision and efficiency. The infusion of ML not only refines extant models but also engenders novel approaches capable of navigating the intricate terrain of cooling load prediction with unprecedented precision. Concurrently, feature extraction methods have risen to prominence as indispensable tools within the ML-driven paradigm [16], [17]. These methods assume a pivotal role in harnessing historical data, ensuring that predictions remain not only precise but also computationally manageable. As the volume and intricacy of data burgeon, the symbiosis between ML techniques and feature extraction methods becomes increasingly critical, providing a robust framework to address the evolving challenges of building cooling load prediction [18]. Envisaging the trajectory ahead, these technological strides hold promise in ushering forth a new era of sustainability and energy efficiency in buildings. The amalgamation of ML techniques and feature extraction methods is poised not merely to refine predictive accuracy but also to lay the groundwork for more intelligent and adaptive HVAC systems. By bridging the lacuna between data insights and operational efficiency, these advancements contribute to a future where sustainable practices are not merely aspirational but intrinsic characteristics of contemporary infrastructure [19].

The study introduces a new method for improving cooling load predictions by integrating ML. It uses a hybridization technique to enhance the performance of Radial Basis Function (RBF) models, ensuring greater accuracy and reliability. The hybrid models, which combine two optimization techniques, show superior performance compared to conventional methods. A thorough evaluation was conducted to mitigate potential biases and provide a transparent assessment of the models' performance. The strategic choice of merging Self-adaptive Bonobo Optimizer (SABO) and Differential Squirrel Search Algorithm (DSSA) highlights the importance of each optimizer's strengths. This study demonstrates the effectiveness of ML and the benefits of using tailored hybrid models for improved cooling load predictions. The following section includes the methodology section, which details the study's approach, covering data collection, preprocessing, model development (including machine learning integration), and validation. Results present findings, including performance compared to existing methods, with quantitative data and visualizations. Discussion interprets results within the study's context, evaluating methodology strengths and limitations. The conclusion summarizes key findings, emphasizes research significance, and suggests future research directions.

II. Materials and Methods

A. Data Gathering

Table I provides a comprehensive overview of the statistical properties of the input variables contributing to the cooling load prediction models, as well as the resultant cooling load output. The variables under scrutiny encompass a range of crucial factors influencing building energy dynamics. The Relative Compactness category, denoting the building's surface area-to-volume ratio, exhibits a minimum value of 0.62 and a maximum value of 0.98. The average and standard deviation are reported as 0.76 and 0.105, respectively. Surface area, a pivotal input parameter, ranges from 514.5 to 808.5, with an average of 671.70 and a standard deviation of 88.086. Wall area, roof area, overall height, orientation, glazing area, and glazing area distribution, each with distinct roles in the prediction models, are similarly characterized by their minimum, maximum, average, and standard deviation values. Of particular significance is the Cooling output variable, which represents the anticipated cooling load. The reported statistics for this variable delineate its minimum, maximum, average, and standard deviation as 10.9, 48.03, 24.58, and 9.513, respectively. These values encapsulate the variability in cooling loads expected from the developed prediction models. The statistical summary provides insights into the range and central tendency of each input variable and the variability in cooling load output, which is crucial for understanding data distribution and ensuring the accuracy and reliability of cooling load predictions generated by developed models [20], [21].

TABLE I. STATISTICAL PROPERTIES OF THE INPUT VARIABLE OF COOLING

Variables		Indicators			
	Category	*Min*	*Max*	*Avg*	*St.Dev.*
Relative Compactness	Input	0.62	0.98	0.764	0.105
Surface Area	Input	514.5	808.5	671.708	88.086
Wall Area	Input	245	416.5	318.5	43.626
Roof Area	Input	110.25	220.5	176.60	45.165
Overall Height	Input	3.5	7	5.25	1.751
Orientation	Input	2	5	3.5	1.118
Glazing Area	Input	0	0.4	0.234	0.133
Glazing Area Distribution	Input	0	5	2.812	1.550
Cooling	Output	10.9	48.03	24.58	9.513

B. Radial Basis Function (RBF)

Ojo et al. [22] discusses the challenges associated with traditional path loss prediction models in wireless signal propagation and proposes the use of machine learning algorithms to address these issues. It highlights the trade-off between simplicity and accuracy in deterministic and empirical models. The paper introduces two machine learning-based path loss prediction models, namely the radial basis function neural network (RBFNN) and the multilayer perception neural network (MLPNN), developed using experimental data collected from drive tests in multi-transmitter scenarios. The RBFNN model is found to be more accurate than the MLPNN, exhibiting lower root mean squared errors (RMSEs) when compared to measured path loss. Furthermore, the proposed machine learning-based models are compared against five existing empirical models, with the RBFNN showing the most accurate results. Elansari et al .[23] presents a novel approach called Mixed Radial Basis Function Neural Network (MRBFNN) training using Genetic Algorithm (GA). This study focuses on optimizing the choice of radial basis functions, centers, radius, and weights of the output layer in RBFNNs. They formulate the optimization problem as a mixed-variable problem with linear constraints and employ a genetic algorithm-based approach to solve it. The numerical results validate the theoretical findings and demonstrate improved generalization compared to existing methods. Alitasb et al. [24] discuss the application of RBFNNs in model predictive control for a 4 × 3 Multiple-Input Multiple-Output (MIMO) biomass control system. The study aims to enhance the control performance of a biomass boiler by utilizing RBFNNs to improve the accuracy of the model. They develop a biomass boiler model using system identification techniques and implement the RBFNN model using MATLAB. Simulation results show that the RBFNN-based model predictive controller achieves shorter settling times and tolerable overshoots compared to a state space model-based controller, indicating superior performance in controlling boiler dynamics. Overall, these references demonstrate the versatility and effectiveness of RBFNNs in various applications, including pattern recognition and control systems, showcasing their potential to address complex problems in different domains.

1) Network architecture for RBF: A RBF neural network consists of three distinct layers: the input layer, hidden layer, and output layer, as its fundamental structure. The input vector x in the hidden layer, which is made up of each unique input, $\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3, \dots, \mathrm{x}_\mathrm{n}$ is practical to all neurons. The hidden layer of an RBF network consists of n RBF that have direct connections to each component of the output layer. More information about RBF networks can be found in academic sources, including references. An RBF network's hidden layer nodes increase their output as the input pattern they represent approaches each node's centre. The output of these nodes decreases with increasing distance from the middle if symmetric basis functions are applied. Therefore, neurons with centres near a given input pattern will produce non-zero activation values when that pattern is present, amplifying the input. The neurons' receptive field function, which controls how they react to outside stimuli, is what causes this behavior [23], [25], [26]. The jth covered−up hub's theoretical basis is often represented by a Gaussian exponential function, which can be written as follows:

$$p_j = p(\omega_j) = e\left(-\frac{\omega_j^2}{2\varphi_j^2}\right) \tag{1}$$

The thickness of the jth neuron, signified by φ_j, and ω_j, this is typically computed as the Euclidean distance between the input vector and the neuron center.

$$\omega_j(x) = \|x - c_j\| = \sqrt{\sum_{i=1}^{w}(x_i - c_{i,j})^2} \quad ,i = 1,2,\dots,w \tag{2}$$

Let and $x = [x_1, \dots, x_w]^F$ c_j be the centres of the jth RBF parts, which are vectors having the same dimensions as the jth neuron's input [27]. The network's output, represented by the letter M, is the sum of the weights of all the premise capacities in the hidden layer. The following formula can be used to determine the output node value:

$$M_r = \sum_{j=1}^{u} s_{js} p_j \tag{3}$$

A weighted sum of the output signals from the nodes in the hidden layer yields the output of the r th node in the output layer, which is denoted by M_r as the kth component of M. s_{js} is a representation of the weights connected to the link between the jth neuron in the hidden $layer$ and the sth neuron in the $output$ layer. The algorithm additionally takes into account the output p_j of the jth node in the hidden layer. In essence, the hidden layer nodes calculate a linear combination of basic functions that define the network's output.

2) RBF network training and testing principles neural system RBF operates in two distinct modes: testing and training. The network necessity calculates the ideal number of her RBFs, identifies the centres for each, and constructs the output layer weight matrix during the training phase. Minimizing the total squared error, or (R), is the goal [28] al [28]:

$$R = \frac{1}{2}\sum_{I=1}^{r}\sum_{F}^{n}\{q_f^i - m_f^i(X^i)\}^2 \tag{4}$$

Here q_f^i is the apparent framework control in the event that an input vector X^i is R is the number of cases under preparation that is shown to the organizer. In RBF networks, node centers are positioned during training, and the network's capacity to generalize is greatly influenced by the output of the kernel function. Numerous methods may be used to determine the center of the RBF node in the hidden layer; the most popular method utilized in this research was $p-$ means clustering. The algorithm's objective is to find a set of F cluster centers that minimizes $E,$ which is the sum of the Euclidean distances between each cluster center (yj) and the train points assigned to that cluster.

$$E = \sum_{j=1}^{N}\sum_{I=1}^{N} G_{ij}\left\|y_j - X_i\right\| \tag{5}$$

The membership matrix was given to the network, signified as G_{ij} with dimensions of $m \times N$, where N represents how many training examples there are. There is only one 1 in each column of this binary matrix; all other values are $0s.$ After establishing the RBF unit centre, each RBF unit width is calculated with a parameter named z, which regulates the amount of overlap between adjacent nodes and the $Root - Mean - Square$ distance to the closest RBF node. The $S-$fold cross-validation approach may be used to find the value of $z.$ Eq. (6) provides the breadth of the $j\ -th\ RBF$ unit, which is represented as ς_j.

$$\varsigma_j = \left(\frac{1}{z}\sum_{o=1}^{z}\left\|c_j - c_o\right\|^2\right)^{1/2} \tag{6}$$

After the centers of the nodes near node j are determined as $(c_1, \ldots . c_z)$ the RBF neural network is considered fully determined. Afterwards, Eq. (1) through Eq. (3) can be used to represent a newly provided, easily navigable input vector. The RBF flowchart is displayed in Fig. 1 [29].

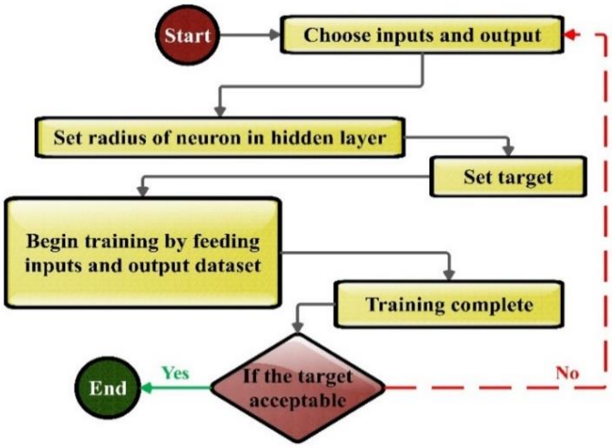


Fig. 1. Flowchart of RBF.

C. Self-adaptive Bonobo Optimizer (SABO)

A detailed argument is presented regarding an improved version of Bonobo Optimizer (BO), acknowledged as SABO with repulsion and a memory-based learning strategy for parameter upgrading. There are four types of mating behaviors found in it: consort ship, extra-group mating, promiscuous mating, and restrictive mating. A few changes have been made to this optimizer to improve performance. In addition, SABO has three additional memorized populations that are distinct from the current population, and the members of these populations mate to give birth to modern bonobos. The SABO's controlling criteria are based on a repulsion-based learning approach [30]. Other than these, p^{th}-bonobo choice in SABO is planned extraordinarily. In this case, the estimate of the subgroup is determined by the calculation taking into account the input obtained during the search procedure. To improve the efficiency of the search process, boundary control technology has also been modified in SABO as opposed to BO. Below is a detailed explanation of each of these key components of the suggested SABO.

1) Memory and its updates: The three populations that make up the proposed SABO are $oldpop$, $worsepop$, and $badpop$, as was previously mentioned. These populations initially match the $SABO's$ the original population. Later, with every iteration, these are revised and updated. Consider that N and d, respectively, are population size and number of decision variables.

a) Oldpop: At any time, i^{th}-new bonobo found superior to the i^{th}-parent bonobo is accepted in the current population position. Furthermore, the i^{th}-parent bonobo is remembered within the i^{th}-position of oldpop.

b) Worsepop: If i^{th}-new bonobo turns out to be worse than i^{th}-parent bonobo in fitness value comparison; it is stored in i^{th}-worse position than the current population.

c) Badpop: The selection process of this N size population is choosing from mixed populace $worsepop$ and $badpop$ in size $2N$. The unique solutions named l, distinguished by objective values, display mixed populations.

If l realized that be less than N, then l number copied to the $badpop$. If l realized that be less than N, then l number copied to the $badpop$. Also, if l found to be greater than or equal to N, $babpop$ is chosen from $l1$ (l1≥N), and it is between N_2 and 1, Eq. (7) calculates N_1.

$$N_2 = ceil(df \times \mathrm{N}) \tag{7}$$

df, diversity factor, is in the range of $[(df_{min}, df_{max}) = (1.2, 1.8)]$.

2) Repulsion-based learning: First, using repulsion-based learning, two controlling variables are identified: phase probability (pp) and sharing co-efficient (sc). Both of this parameter's range are between 0 and 1. If a good solution's number appears superior to predetermined solutions (called N_1) then N_1 solution with the highest change in fitness value will be considered. Eq. (8) calculates the N_1.

$$N_1 = maximum\ of (5, ceil(N \times 0.08)) \tag{8}$$

By taking after a comparative strategy, one can get a worse value that is not upgraded fitness-wise, named pp_{worse}. Considering pp_{worse} and pp_{better} repulsion; their value directions shift in the reverse direction. $pp_{bettter}^{modified}$ determined by Eq. (9).

$$PP_{better}^{modified} = PP_{better} \pm \frac{\sigma \times PP_{better} \times PP_{worse}}{e^{(PP_{better}-PP_{worse})^2}} \tag{9}$$

3) Mating strategies: There are four mating strategies in the proposed $SABO$. There are also conceptions of the positive phase (pp) and negative phase (NP). Phase probabilities (pp) are probabilities that PP or NP use to update the solution at each iteration. Furthermore, in PP, applying a promiscuous or restrictive mating strategy will generate a new bonobo with a probability of 0.5. The probability of mating outside the group is called p_{xgm} which is in the range of $(p_{xgm}^{min}, p_{xgm}^{max})$ and calculates as follows:

$$I_0 = \frac{\sum_{j=1}^{d} \frac{SD_j}{(V_{max_j} - V_{min_j})}}{d} \tag{10}$$

$$I_1 = 0.1 \times I_0 \tag{11}$$

$$P_{xgm}^{min} = minimum\ of\ (I_1, \frac{1}{d}) \tag{12}$$

$$p_{xgm}^{max} = 2 \times I_1 \tag{13}$$

In initial population, SD_j represents the standard deviation of j^{th}-variable; j is in the range of 1 to d, and d is the decision variable number in the problem. The maximum and minimum merits of j^{th}-variable called v_{max_j} and v_{min_j}. $p_{xgm}^{min}\ and\ p_{xgm}^{max}$ are determined by I_0 and I_1. Fig. 2 shows a detailed flowchart of the suggested SABO.

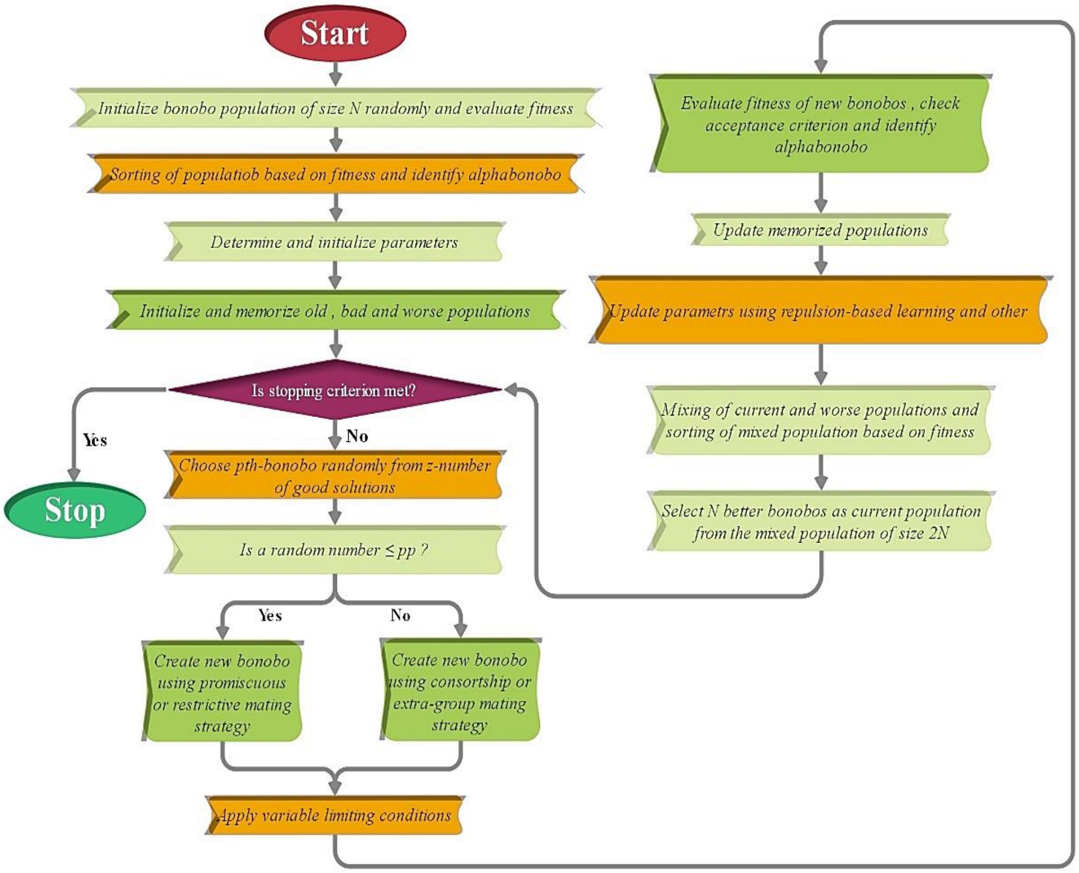

Start
Initialize bonobo population of size N randomly and evaluate fitness
Sorting of populatiob based on fitness and identify alphabonobo
Determine and initialize parameters
Initialize and memorize old , bad and worse populations
Is stopping criterion met?
Yes
Stop
No
Choose pth-bonobo randomly from z-number of good solutions
Is a random number ≤ pp ?
Yes
No
Create new bonobo using promiscuous or restrictive mating strategy
Create new bonobo using consortship or extra-group mating strategy
Apply variable limiting conditions
Evaluate fitness of new bonobos , check acceptance criterion and identify alphabonobo
Update memorized populations
Update parametrs using repulsion-based learning and other
Mixing of current and worse populations and sorting of mixed population based on fitness
Select N better bonobos as current population from the mixed population of size 2N

Fig. 2. Flowchart of suggested SABO.

D. Differential Squirrel Search Algorithm (DSSA)

Manoj et ao et al. introduced an innovative optimization approach referred to as DSSA, which integrates two distinct algorithms, namely the squirrel search algorithm (SSA) and differential evolution algorithm (DE). Squirrels are observed to update their location in the context of the Simulated Squirrel Algorithm (SSA) by closely monitoring the positions of other squirrels living in the hickory or acorn tree. That means that to improve their search strategies, the top squirrels have modified their updating mechanisms. Adding a crossover operation that draws inspiration from Differential Evolution (DE) improves exploration potential. This exposition provides a mathematical model of the various foraging strategies that are included in the Decision Support System for Agriculture (DSSA).

1) Initialization of position and the evaluation of fitness: Initially, the squirrels are arranged at random within the exploratory domain. Following the squirrels' positions, the fitness levels of each are evaluated by inserting the position mentioned earlier into the fitness function. This indicates the nutritional value of the food source that each squirrel obtained within its position. The values of fitness are subsequently subjected to a sorting process to obtain the most optimal squirrel, denoted as PS_{ht}, that currently inhabits the hickory tree. The subsequent 3 optimal function values signify the locations of the squirrels inhabiting the acorn tree PS_{at} (1:3). In addition, they are deemed to be a step towards achieving an optimal location in the next iteration. Thus far, the remaining individuals belonging to the squirrel population are denoted as PS_{nt}^{p1} have yet to discover a viable nourishment source occupying typical locations within the tree habitat.

2) Update of the position: When a squirrel colonizes a tree that provides acorns, it updates its position and moves towards the best source by following the current best path or PS_{ht} in the absence of a predator. By mimicking the movements of the squirrels living in the hickory or acorn tree, the squirrels in the regular tree can track where they are. The observed phenomenon is that when faced with the possibility of predation, squirrels tend to change their foraging route randomly. The following can be used to express the mathematical models that were used to update the squirrel's position. Now, as the others adjust their position, the squirrels in the acorn trees do the following:

$$PS_{at}^{new} = \begin{cases} PS_{at}^{old} + d_g.G_c\left(PS_{ht}^{old} - PS_{at}^{old} - P_{avg}\right), & r_1 \geq P_{dp} \\ random\, position, & otherwise \end{cases} \tag{14}$$

P_{avg} is the mean of whole squirrels' positions in the available population.

The crossover mechanism is incorporated into the DE algorithm to reduce the likelihood of local minima, thereby increasing the diversity of the squirrel population being evaluated. The crossover operation, which was derived from Eq. (15), was applied to the current position and its corresponding new position.

$$PS_{at,i,j}^{cr} = \frac{PS_{at,i,j}^{new},\ if\ \left(rand_j \leq Cr\right)\ or\ j = j_{rand}}{PS_{at,i,j}^{old},\ if\ \left(rand_j > Cr\right)\ or\ j \neq j_{rand}}, \quad j = 1,2,3,\ldots,D \tag{15}$$

$$i\ =\ 1, 2, 3, \ldots, NP$$

$PS_{at,i,j}^{new}$ and $PS_{at,i,j}^{old}$ are new and old squirrels' locations normal, or are acorn trees.

$PS_{at,i,j}^{cr}$ are the squirrels' positions after the operation of the crossover.

NP shows the population size.

Cr indicates the rate of crossover, and it is equal to 0.5.

D shows the dimension of the problem.

$j_{rand}\ \epsilon\ [1, D]$ is a randomly generated index

$rand_j\ \epsilon\ [1, D]$ is the $j - th$ random numbers' assessment uniformly developed in the distinct range.

According to Eq. (16), some of the squirrels that live in typical trees move to new areas by following the squirrels that live in acorn trees.

$$PS_{nt}^{new} = \begin{cases} PS_{nt}^{old} + d_g.G_c(PS_{at}^{old} - PS_{nt}^{old}), & r_2 \geq P_{dp} \\ random\, position, & otherwise \end{cases} \tag{16}$$

r_2 conforms to a uniform distribution across the interval of [0, 1].

The squirrels that are still in traditional trees align with the current optimal location, and their updated locations are expressed as follows:

$$PS_{nt}^{new} = \begin{cases} PS_{nt}^{old} + d_g.G_c\left(PS_{ht}^{old} - PS_{nt}^{old}\right), & r_3 \geq P_{dp} \\ random\, position, & otherwise \end{cases} \tag{17}$$

The application of a crossover operation on squirrels residing in ordinary trees is as follows:

$$PS_{nt,i,j}^{cr} = \frac{PS_{nt,i,j}^{new},\ if\ \left(rand_j \leq Cr\right)\ or\ j = j_{rand}}{PS_{nt,i,j}^{old},\ if\ \left(rand_j > Cr\right)\ or\ j \neq j_{rand}}, \quad j = 1,2,3,\ldots,D \tag{18}$$

One way to speed up the convergence rate is to allow the squirrel in the hickory tree to move about the average of the squirrels' positions inside the tree, as shown by Eq. (19):

$$PS_{ht}^{new} = PS_{ht}^{old} + d_g.G_c\left(PS_{ht}^{old} - PS_{at}^{avg}\right) \tag{19}$$

PS_{at}^{avg} represents the average of whole squirrel's locations in the acorn trees.

To determine who gets to participate in the next development population, the best hybrid positions and their unused positions are compared with the historical positions in the determination handle. Algorithm 1 explains the procedural instructions involved in DSSA.

```
Algorithm 1: DSSA Pseudocode
Input $Itermax, NP, Pdp, sf, Gc, ub, and\ lb$
Initialize the flying squirrels^' location haphazardly using Eq.(14)
Compute the fitness value utilizing the represented fitness function
employing Eq.(15)
While $itr \leq itermax$ do
Sort of all functions of squirrels^' fitness and recognize the
current best, 〖PS〗_ht,positions of squirrels in acorn tree, 〖PS〗
_at (1:3)
and the position of squirrel in the normal tree, 〖PS〗_nt (1:NP-4)
                Develop the new position of squirrels by
comparing the new locations in acorn trees achieved
                 via utilizing Eq.(14) and Eq.(15)
                Develop the new location of squirrels by
                comparing the new positions in normal
                trees achieved via utilizing Eqs.(16) − (18)
Generate the new location of squirrels by
comparing the new positions in
hickory trees achieved via utilizing Eq.(19) and old positions.
Update the population with the best position achieved so far
End While
Return $PS_{ht}$
```

E. Performance Evaluation Methods

The models are evaluated in this article using some metrics, such as the BIAS mentioned earlier, the correlation coefficient (R^2), the Mean Square Error (MSE), the Symmetric Mean Absolute Percentage Error ($SMAPE$), and the root mean square error (RMSE). A high R^2 value indicates that the algorithm performed exceptionally well during the training, validation, and testing phases. Lower RMSE and MAE values, on the other hand, are preferred because they demonstrate less model error. Eq. $(20-24)$ are used to calculate these metrics.

Coefficient of Correlation.

$$R^2 = \left(\frac{\sum_{i=1}^{W}(h_i-\bar{h})(l_i-\bar{l})}{\sqrt{[\sum_{i=1}^{W}(h_i-h)^2][\sum_{i=1}^{W}(l_i-\bar{l})^2]}}\right)^2 \qquad (20)$$

Root Mean Square Error.

$$RMSE = \sqrt{\frac{1}{W}\sum_{i=1}^{W}(l_i-h_i)^2} \qquad (21)$$

Mean Square Error.

$$MSE = \frac{1}{W}\sum_{i=1}^{w}(l_i-h_i)^2 \qquad (22)$$

Symmetric Mean Absolute Percentage Error.

$$SMAPE = \frac{100}{W}\sum_{I}^{W}\frac{2\times|l_i-h_i|}{|l_i|+|h_i|} \qquad (23)$$

BIAS.

$$BIAS = \frac{\bar{l}}{\bar{h}} \qquad (24)$$

- In these equations, h_i and l_i refer to the predicted and experimental values, respectively.
- The mean values of the experimental samples and predicted are represented by $\bar{h}$ and $\bar{l}$.
- Otherwise, W denotes the number of samples being considered.

III. Results

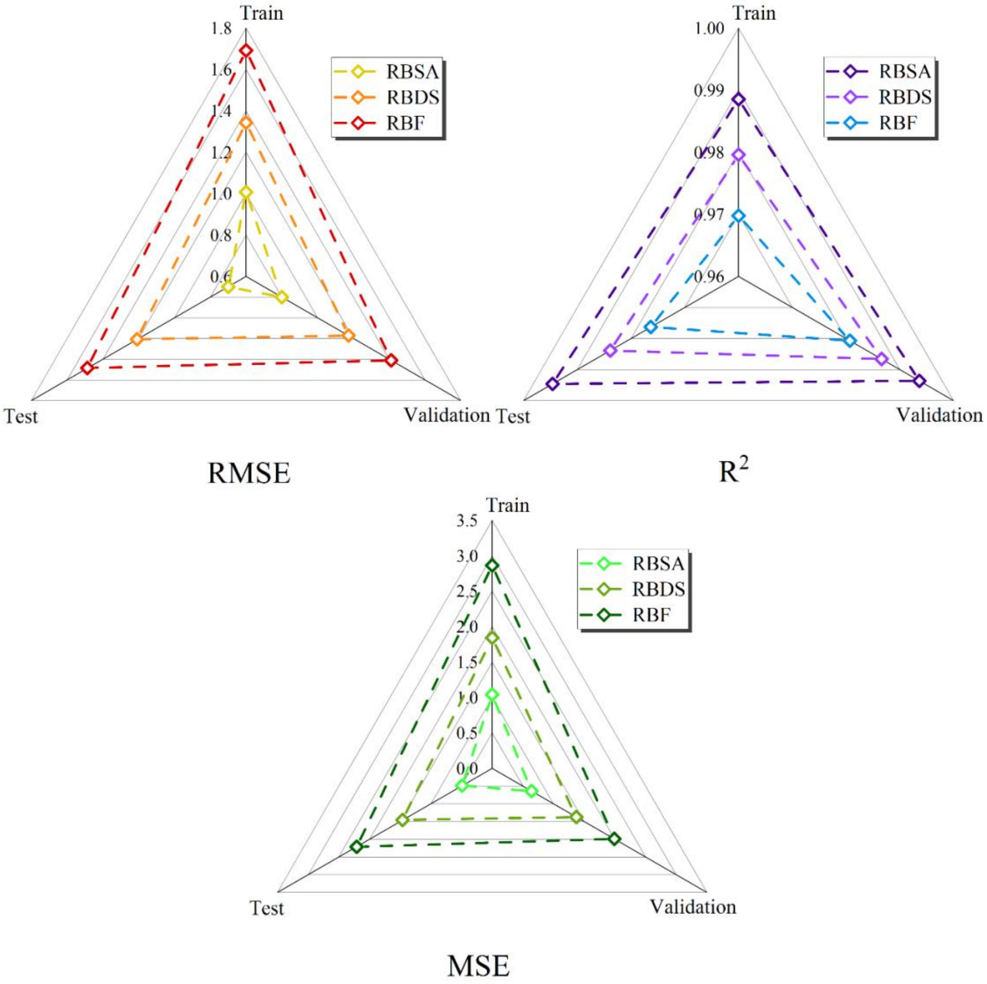
The results of the created RBF models are shown in Table II. It includes an aggregate evaluation that covers all phases and a detailed summary of performance metrics for each of the three phases: training, validation, and testing. Based on RMSE, R^2, MSE, SMAPE, and BIAS, the RBF models RBSA, RBDS, and the generic RBF are assessed. The RBSA model performs well during the training phase, as evidenced by its low RMSE of 1.007, high R^2 of 0.989, and minimized MSE of 1.040. The accuracy of the model in capturing the subtleties of the training dataset is demonstrated by the SMAPE value of 0.00005 and the insignificant BIAS of -0.020. In the same way, RBDS performs admirably during training, showing an RMSE of 1.343, an R^2 of 0.980, and an MSE of 1.844. When the models move on to the validation stage, RBSA continues to perform well, with a lower RMSE (0.800) and a higher R^2 (0.994), indicating that it can generalize far beyond the training set. The RBDS model, on the other hand, shows a little increase in RMSE (1.172) and a lower R^2 (0.987), but it still retains a respectable degree of accuracy during validation. The generic RBF model, on the other hand, shows higher RMSE (1.410) and lower R^2 (0.981) values, indicating a relatively lower fit during the validation phase. When the models are tested, the models' performance is evaluated once more, and RBSA continues to perform well, with a minimum RMSE of 0.700, a high R^2 of 0.995, and an MSE of 0.489. During testing, both the generic RBF model and RBDS perform satisfactorily, with R^2 values of 0.984 and 0.976 and $RMSE$ values of 1.208 and 1.486, respectively. With an overall RMSE of 0.938 and an R^2 of 0.990 when all phases are considered, RBSA performs remarkably well, proving its dependability across a range of datasets. In contrast, the generic RBF model and the RBDS model have higher overall RMSE values of 1.623 and 1.299, respectively. In Fig. 3, the metrics' radar plot is displayed.

These findings show how the RBF models' effectiveness varies depending on the phase, with RBSA constantly showing higher predictive accuracy. The extensive assessment metrics highlight the potential of the RBSA model in producing accurate and trustworthy predictions for cooling loads and offer insightful information about the models' generalization ability.

In Fig. 4, a scatter plot is used to compare the hybrid models' performance over the 3 train, validation, and test phases. R^2 quantifies the degree of agreement between observed and expected values, whereas RMSE shows the prediction error or dispersal. The data points are closely clustered around the centerline of the RBSA model, which shows exceptional accuracy across all three phases. There is little variation between the expected and actual values and a high degree of agreement. The RBF and RBDS models, on the other hand, featured data points that were further from the centerline and comparable performance levels. In comparison to the RBSA model, this broader dispersion predicts a slightly lower precision and a higher inaccuracy.

TABLE III. THE OUTCOME OF THE DEVELOPED *RBF* MODELS

Model	Phase	Index values				
		RMSE	R^2	MSE	SMAPE	BIAS
RBSA	Train	1.007	0.989	1.040	0.00005	-0.020
	Validation	0.800	0.994	0.640	0.00019	0.044
	Test	0.700	0.995	0.489	0.00018	0.038
	All	0.938	0.990	0.896	0.00003	-0.001
RBDS	Train	1.343	0.980	1.844	0.00007	0.002
	Validation	1.172	0.987	1.373	0.00027	0.073
	Test	1.208	0.984	1.460	0.00026	-0.077
	All	1.299	0.981	1.716	0.00004	0.001
RBF	Train	1.692	0.970	2.863	0.00011	-0.093
	Validation	1.410	0.981	1.989	0.00043	0.134
	Test	1.486	0.976	2.209	0.00047	0.174
	All	1.623	0.972	2.634	0.00007	-0.019



RMSE

R^2

MSE

Fig. 3. Radar plot for comparison between the developed models based on metrics.

The study's line plot in Fig. 5 shows the error percentages associated with the models, with the RBSA model being the most prominent due to its low error rate. The majority of error values cluster within the 12.08% range. The RBF and RBDS models show greater variability, with a higher frequency of values exceeding the 23.66% and 17.35% thresholds. Both models display a right-skewed distribution, indicating the presence of specific data points with greater proportions of errors.

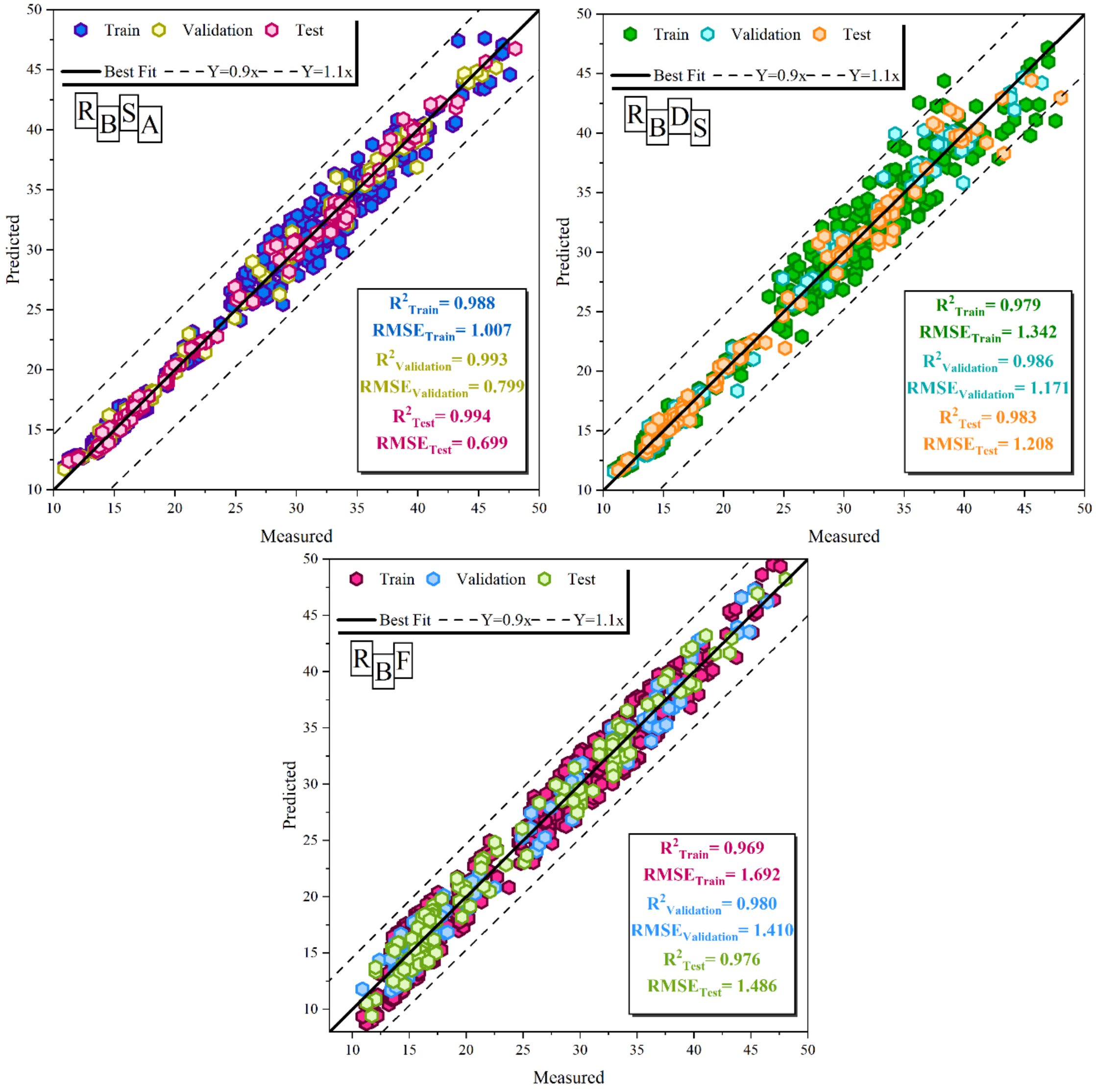


Fig. 4. Scatter plot of the dispersion of evolved hybrid models.

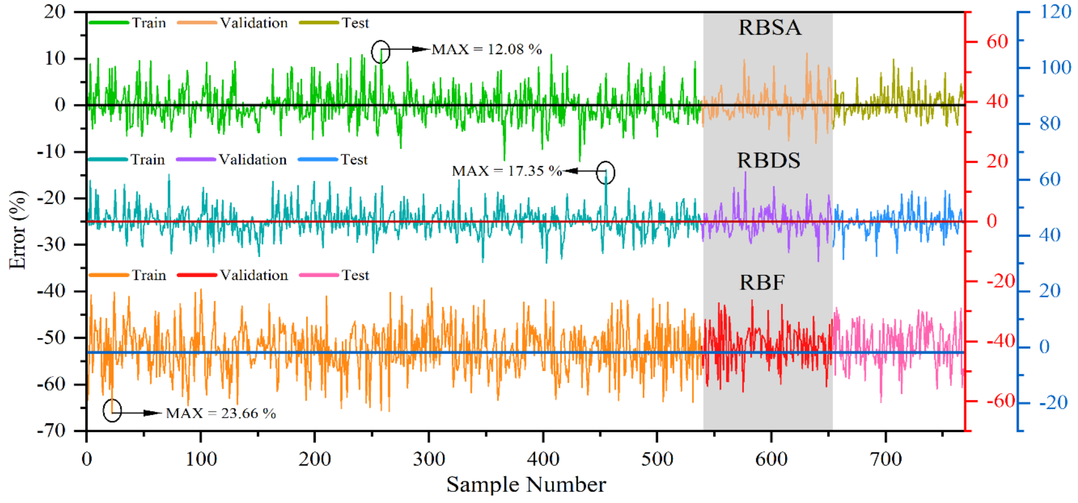


Fig. 5. Error percentage of the models based on the line plot.

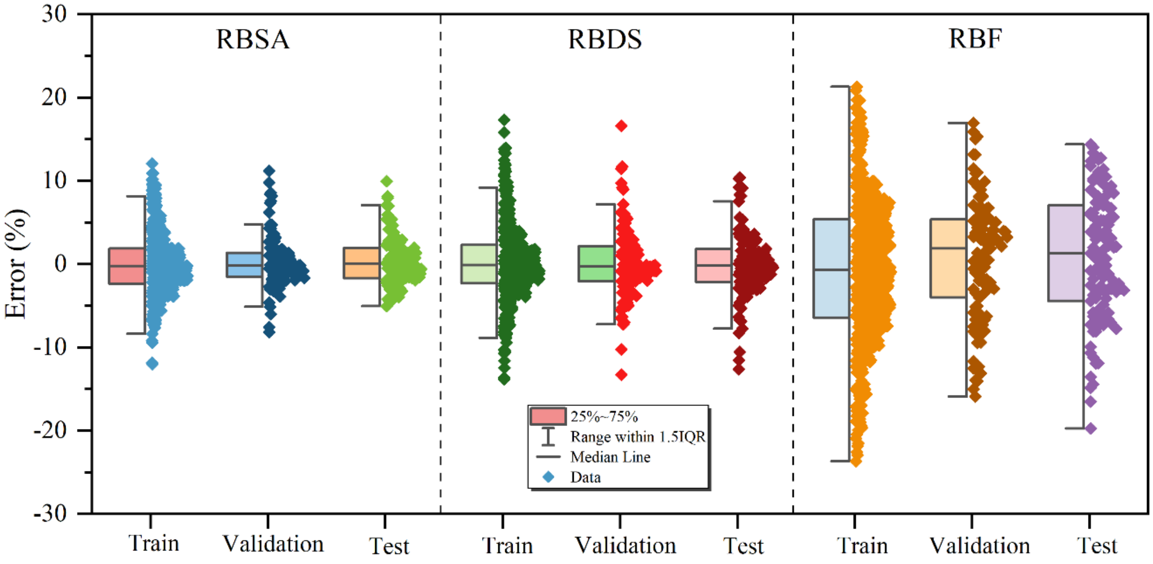


Fig. 6. Half-box plot errors of proposed models.

In Fig. 6, the study displays a half-box plot that shows the three models' respective error percentages. The RBSA model performed remarkably well, with errors kept below 10% and little dispersion. Dispersion was seen in all phases of the RBF model, with a uniform normal distribution and a maximum of 25%. During the assessment stage, the RBDS model exhibited the highest degree of discrepancy among the models, with one outlier data point representing more than 18% of the dataset. The RBF model's Gaussian distribution revealed more dispersion and fewer instances of near-zero frequency.

IV. Discussion

A. Advantages of the Present Study

The novel approach introduced in the study aims to revolutionize the prediction of cooling loads, offering fresh insights and innovative solutions to longstanding challenges in energy efficiency and building management. By harnessing the power of machine learning techniques and hybridization strategies, the proposed methodology endeavors to elevate the accuracy and dependability of cooling load predictions beyond the capabilities of conventional methods. Through rigorous comparative analyses, the research delves into intricate examinations, juxtaposing the proposed methodology with state-of-the-art approaches. These comparisons yield invaluable insights into the efficacy and superiority of the novel methodology, shedding light on its potential to outperform existing techniques in predicting cooling loads with precision and reliability. The practical implications of successfully implementing the study's findings are substantial, encompassing a wide array of benefits ranging from optimized energy consumption in buildings to tangible cost reductions and advancements in sustainability initiatives. Furthermore, the study's contribution to the academic literature is profound, marking a significant advancement in predictive modeling techniques within the realm of cooling load prediction. By pushing the boundaries of knowledge and innovation, the study paves the way for future research endeavors and opens new avenues for exploration in the pursuit of energy-efficient building management practices.

B. Limitations

The effectiveness of the proposed methodology could face constraints due to data availability and quality, which might hinder the reliability of predictions if the data is insufficient or inaccurate. Moreover, the implementation of machine learning algorithms and hybridization strategies adds complexity to the modeling process, potentially necessitating specialized expertise for development and interpretation. Additionally, the study's findings might have limited generalizability to various contexts or building types, as factors like geographical location, building design, and occupancy patterns could influence the applicability of the methodology. Validating the proposed approach poses challenges, particularly in ensuring robustness and reliability across diverse scenarios and conditions. Furthermore, ethical considerations surrounding data privacy, bias, and transparency must be meticulously addressed due to the use of machine learning techniques and optimization algorithms. These ethical concerns underscore the need for responsible and transparent research practices throughout the study.

C. Future Study

For future studies, researchers may explore refining the methodology, analyzing long-term performance, integrating advanced technologies, validating in real-world settings, considering external factors' impact, conducting sensitivity analysis, integrating with building automation systems, and engaging users for feedback. These avenues aim to advance cooling load prediction, enhance energy efficiency, and promote sustainable building management practices.

D. Comparison with Published Papers

Table III shows the comparison between the presented and published papers. From the comparison, it can be observed that the presented study falls within the range of RMSE and R^2 values reported in the published articles. While the RMSE value of the present study (0.938) is higher than that of Roy et al. (0.059) and Gong et al. (0.1929), it is lower than that of Afzal et al. (1.4122). Similarly, the R^2 value of the present study (0.990) is slightly lower than that of Moradzadeh et al. (0.9993) but higher than that of Gong et al. (0.9882) and Afzal et al. (0.9806). Overall, the presented study demonstrates competitive performance in terms of both RMSE and R^2 compared to the published articles, indicating its effectiveness in predicting cooling loads.

TABLE IV. COMPARISON BETWEEN THE PRESENTED AND PUBLISHED ARTICLES

Articles	**Index values**	
	RMSE	**R^2**
Moradzadeh et al. [31]	0.4832	0.9993
Roy et al. [9]	0.059	0.99
Gong et al. [32]	0.1929	0.9882
Afzal et al. [14]	1.4122	0.9806
Present Study	0.938	0.990

V. CONCLUSION

As a result of the significant influence that the global building industry has on total energy consumption, efforts must be made to precisely forecast cooling loads in the context of energy conservation and building operations. The study established a multimodal investigation to improve cooling load prediction models by applying cutting-edge techniques and creative ideas. Initially, the emphasis was on important metrics like Dynamic Air-Conditioning Load (DACL) and Cooling Load (CL), which highlighted the critical role these parameters play in HVAC system optimization. Several different approaches have been adopted to improve prediction accuracy because of the complex interactions that have been identified between the optical and thermal properties of buildings and meteorological data. The process of extracting detailed insights from the data required the application of feature extraction techniques, such as engineering-based, statistical, and structural methods. These techniques, which combined mathematical and domain-specific viewpoints, made it easier to comprehend the thermal dynamics present in different types of construction materials. The rapidly changing field of machine learning (ML) has the potential to transform the precision and effectiveness of cooling load forecasts completely. The combination of machine learning and feature extraction techniques demonstrated the potential to improve current models and introduce new ones that can more accurately predict cooling load with never-before-seen levels of detail. The development and assessment of the RBF models RBSA, RBDS, and generic RBF represented the culmination of these efforts. The thorough statistical characteristics of the input variables served as a fundamental point of reference, allowing for an in-depth analysis of the models' performance during the training, validation, and testing stages. With the Self-adaptive Bonobo Optimizer (SABO) integrated, RBSA proved to be an exceptional performer, exhibiting superior accuracy in every phase. RBSA is a strong and dependable model for cooling load prediction because of its exceptional performance during testing, reduced RMSE and increased R^2, and ability to generalize far beyond the training set. These findings have wider ramifications that go beyond the immediate setting and provide a promising path towards more environmentally friendly and energy-efficient structures. By incorporating advanced modelling techniques, predictive accuracy is improved, and the foundation for intelligent, adaptive HVAC systems is laid. The combination of state-of-the-art machine learning, research methodologies, and feature extraction techniques could lead to the intelligent response of buildings to environmental demands in the future, resulting in a more sustainable and environmentally conscious global infrastructure.

ACKNOWLEDGMENTS

Natural Science Foundation of Hunan Province (2023JJ50348).

Research Foundation of Education Bureau of Hunan Province (22C0514).

REFERENCES

[1] J. Kim, Y. Zhou, S. Schiavon, P. Raftery, and G. Brager, “Personal comfort models: Predicting individuals’ thermal preference using occupant heating and cooling behavior and machine learning,” Build Environ, vol. 129, pp. 96–106, 2018, doi: https://doi.org/10.1016/j.buildenv.2017.12.011.

[2] R. Zhao et al., “Building cooling load prediction based on lightgbm,” IFAC-PapersOnLine, vol. 55, no. 11, pp. 114–119, 2022.

[3] Y. Ding, Q. Zhang, and T. Yuan, “Research on short-term and ultra-short-term cooling load prediction models for office buildings,” Energy Build, vol. 154, pp. 254–267, 2017.

[4] J. Guo et al., “Prediction of heating and cooling loads based on light gradient boosting machine algorithms,” Build Environ, vol. 236, p. 110252, 2023, doi: https://doi.org/10.1016/j.buildenv.2023.110252.

[5] B. S. A. J. khiavi; B. N. E. K. A. R. T. K. hadi Sadaghat;, “The Utilization of a Naïve Bayes Model for Predicting the Energy Consumption of Buildings,” Journal of artificial intelligence and system modelling, vol. 01, no. 01, 2023, doi: 10.22034/JAISM.2023.422292.1003.

[6] C. Lu, S. Li, S. Reddy Penaka, and T. Olofsson, “Automated machine learning-based framework of heating and cooling load prediction for quick residential building design,” Energy, vol. 274, p. 127334, 2023, doi: https://doi.org/10.1016/j.energy.2023.127334.

[7] J. Zhao, X. Yuan, Y. Duan, H. Li, and D. Liu, “An artificial intelligence (AI)-driven method for forecasting cooling and heating loads in office buildings by integrating building thermal load characteristics,” Journal of Building Engineering, vol. 79, p. 107855, 2023, doi: https://doi.org/10.1016/j.jobe.2023.107855.

[8] R. Chaganti et al., “Building heating and cooling load prediction using ensemble machine learning model,” Sensors, vol. 22, no. 19, p. 7692, 2022.

[9] S. S. Roy, P. Samui, I. Nagtode, H. Jain, V. Shivaramakrishnan, and B. Mohammadi-Ivatloo, “Forecasting heating and cooling loads of buildings: A comparative performance analysis,” J Ambient Intell Humaniz Comput, vol. 11, pp. 1253–1264, 2020.

[10] S. T. Kadam, I. Hassan, L. Wang, and M. A. Rahman, “Impact of Urban Microclimate on Air Conditioning Energy Consumption Using Different Convective Heat Transfer Coefficient Correlations Available in Building Energy Simulation Tools,” in Fluids Engineering Division Summer

Meeting, American Society of Mechanical Engineers, 2021, p. V002T03A002.

[11] Q. Si, Y. Peng, Q. Jin, Y. Li, and H. Cai, “Multi-Objective Optimization Research on the Integration of Renewable Energy HVAC Systems Based on TRNSYS,” Buildings, vol. 13, no. 12, p. 3057, 2023.

[12] K. Bamdad, N. Mohammadzadeh, M. Cholette, and S. Perera, “Model Predictive Control for Energy Optimization of HVAC Systems Using EnergyPlus and ACO Algorithm,” Buildings, vol. 13, no. 12, p. 3084, 2023.

[13] X. Li and R. Yao, “A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour,” Energy, vol. 212, p. 118676, 2020, doi: https://doi.org/10.1016/j.energy.2020.118676.

[14] S. Afzal, B. M. Ziapour, A. Shokri, H. Shakibi, and B. Sobhani, “Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms,” Energy, p. 128446, Jul. 2023, doi: 10.1016/j.energy.2023.128446.

[15] G. Bekdaş, Y. Aydın, Ü. Isıkdağ, A. N. Sadeghifam, S. Kim, and Z. W. Geem, “Prediction of Cooling Load of Tropical Buildings with Machine Learning,” Sustainability, vol. 15, no. 11, p. 9061, 2023.

[16] X. Li and R. Yao, “A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour,” Energy, vol. 212, p. 118676, 2020, doi: https://doi.org/10.1016/j.energy.2020.118676.

[17] C. Fan, Y. Liao, G. Zhou, X. Zhou, and Y. Ding, “Improving cooling load prediction reliability for HVAC system using Monte-Carlo simulation to deal with uncertainties in input variables,” Energy Build, vol. 226, p. 110372, 2020.

[18] X. Zhao, F. Li, B. Chen, X. Li, and S. Lub, “Modeling the hardness properties of high-performance concrete via developed RBFNN coupling matheuristic algorithms,” Journal of Intelligent & Fuzzy Systems, no. Preprint, pp. 1–15, 2023.

[19] M. B. Bashir and A. A. Alotaibi, “Smart buildings Cooling and Heating Load Forecasting Models,” IJCSNS, vol. 20, no. 12, p. 79, 2020.

[20] C. Deb, L. S. Eang, J. Yang, and M. Santamouris, “Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks,” Energy Build, vol. 121, pp. 284–297, 2016, doi: https://doi.org/10.1016/j.enbuild.2015.12.050.

[21] O. Probst, “Cooling load of buildings and code compliance,” Appl Energy, vol. 77, no. 2, pp. 171–186, 2004.

[22] S. Ojo, A. Imoize, and D. Alienyi, “Radial basis function neural network path loss prediction model for LTE networks in multitransmitter signal propagation environments,” International Journal of Communication Systems, vol. 34, no. 3, p. e4680, 2021.

[23] T. Elansari, M. Ouanan, and H. Bourray, “Mixed Radial Basis Function Neural Network Training Using Genetic Algorithm,” Neural Process Lett, vol. 55, no. 8, pp. 10569–10587, 2023.

[24] G. K. Alitasb and A. O. Salau, “Multiple-input multiple-output Radial Basis Function Neural Network modeling and model predictive control of a biomass boiler,” Energy Reports, vol. 11, pp. 442–451, 2024.

[25] M. Sahoo et al., “MLP (multi-layer perceptron) and RBF (radial basis function) neural network approach for estimating and optimizing 6-gingerol content in Zingiber officinale Rosc. in different agro-climatic conditions,” Ind Crops Prod, vol. 198, p. 116658, 2023.

[26] W. Bowen, “Research on nonlinear calibration of mine catalytic-combustion-based combustible-gas sensor based on RBF neural network,” Heliyon, vol. 9, no. 3, 2023.

[27] M. Y. Mashor, “Hybrid training algorithm for RBF network,” International Journal of the computer, the Internet and Management, vol. 8, no. 2, pp. 50–65, 2000.

[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] M.-L. Zhang, “M l-rbf: Rbf neural networks for multi-label learning,” Neural Process Lett, vol. 29, pp. 61–74, 2009.

[30] A. K. Das, S. Sahoo, and D. K. Pratihar, “An Improved Design of Knee Orthosis Using Self-Adaptive Bonobo Optimizer (SaBO),” J Intell Robot Syst, vol. 107, no. 1, p. 8, 2023.

[31] 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.

[32] 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.