

Research article

Chiller power consumption forecasting for commercial building based on hybrid convolution neural networks-long short-term memory model with barnacles mating optimizer



Mohd Herwan Sulaiman^{a,c,*}, Zuriani Mustaffa^b

^a Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Pekan, Pahang, 26600, Malaysia

^b Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Pekan, Pahang, 26600, Malaysia

^c Center for Research in Advanced Fluid and Processes (Fluid Center), Universiti Malaysia Pahang Al-Sultan Abdullah, Gambang Campus, Kuantan, Pahang, 26300, Malaysia

ARTICLE INFO

Keywords:

Barnacles mating optimizer
Chiller power consumption forecasting
Convolution neural networks
Long short-term memory
Metaheuristic algorithms

ABSTRACT

This paper addresses the critical challenge of energy efficiency in commercial buildings, where chillers typically consume 40–50% of total building energy. Accurate forecasting of chiller power consumption is essential for optimizing building energy management systems and reducing operational costs. Despite advances in deep learning, existing forecasting models often struggle with the complex temporal dependencies and non-linear patterns in chiller operation data. This paper presents an innovative approach using a hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) model optimized by the Barnacles Mating Optimizer (BMO). The study compares the proposed CNN-LSTM-BMO against other metaheuristic optimization algorithms, including Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Differential Evolution (DE). The models were evaluated using comprehensive performance metrics and validated through statistical analysis. Results demonstrate that the CNN-LSTM-BMO achieves superior performance with the lowest Root Mean Square Error (RMSE) of 0.5523 and highest R^2 value of 0.9435, showing statistically significant improvements over other optimization methods as confirmed by paired t -tests ($P < 0.05$). Key observations include: (1) the CNN-LSTM-BMO model converges 27% faster than traditional optimization methods; (2) SHapley Additive exPlanations (SHAP) analysis reveals that temperature-related features, particularly saturation temperature, are the most influential predictors across all models; and (3) the proposed model maintains prediction accuracy even under varying operational conditions. The proposed CNN-LSTM-BMO model demonstrates robust convergence characteristics and superior generalization capability, making it particularly suitable for real-world applications in building energy management systems. This research contributes to the advancement of accurate and efficient chiller power consumption forecasting methodologies, offering practical implications for Heating, Ventilation, and Air Conditioning (HVAC) system optimization and energy efficiency improvements in commercial buildings.

1. Introduction

Energy management in buildings and industrial facilities is a critical challenge, particularly in the context of Heating, Ventilation, and Air Conditioning (HVAC) systems, where chillers are key components for maintaining temperature control and ensuring efficient cooling. As chillers often constitute a significant proportion of a building's energy consumption, accurate forecasting of their power usage is essential for

optimizing operational strategies, reducing energy costs, and improving overall efficiency [1,2]. Commercial buildings account for approximately 40% of global energy consumption, with HVAC systems representing the largest end-use category at 40–50% of total building energy usage [3]. Within HVAC systems, the energy distribution varies significantly among components: chillers typically consume 40–50% of HVAC energy, while air handling units (AHUs) account for 40%. The remaining energy is distributed among auxiliary equipment such as

* Corresponding author at: Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Pekan, Pahang, 26600, Malaysia.

E-mail address: herwan@umpsa.edu.my (M.H. Sulaiman).

<https://doi.org/10.1016/j.nxener.2025.100321>

Received 8 February 2025; Received in revised form 11 May 2025; Accepted 13 May 2025

2949-821X/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

pumps and cooling towers. Chillers dominate energy consumption due to 3 primary factors: (1) the thermodynamic energy intensity of the vapor compression cycle required for cooling, (2) continuous operation during occupied hours regardless of partial loading conditions, and (3) the cascading effect where inefficient chiller operation negatively impacts the energy consumption of dependent systems such as pumps and cooling towers [4]. A recent study found that optimizing chiller performance alone can reduce overall building energy consumption by 8–20%, highlighting why focusing on chiller power consumption forecasting offers the greatest potential impact for energy efficiency improvements in commercial buildings [2].

With the advent of the era of big data, buildings have become not only energy-intensive but also data-intensive [5], as the complex nature of chillers shaped by numerous factors such as environmental conditions, operational loads, and system performance presents significant challenges for accurate prediction. Employing data mining technologies to analyze the vast amounts of operational data is essential for enhancing the performance of building energy systems, particularly in optimizing chiller power consumption [6,7]. Traditionally, methods such as regression models and time series analysis have been employed to predict chiller energy consumption. For example, linear regression and autoregressive integrated moving average (ARIMA) models have been applied to HVAC systems, utilizing historical data to forecast energy usage [8–10]. While these methods are straightforward and interpretable, they often fall short when dealing with the highly nonlinear dynamics of chiller systems, particularly under varying operational and environmental conditions [11]. This limitation has driven the exploration of more sophisticated Artificial Intelligence (AI) and machine learning (ML) approaches [12–15].

Recent advances in deep learning (DL) have introduced powerful tools such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for tackling energy forecasting problems [16–20]. CNNs are effective at extracting spatial features from high-dimensional datasets, such as temperature, humidity, and pressure readings, which significantly influence chiller power consumption [21,22]. Meanwhile, LSTM networks, as specialized recurrent neural networks (RNN), are well-suited for modeling temporal dependencies, enabling them to capture trends and seasonal patterns in time-series data [23,24]. These models have demonstrated significant promise in improving forecasting accuracy for HVAC systems, including chillers. For instance, Artificial Neural Networks (ANN) and deep learning models have been successfully applied to predict atmospheric temperatures at both average and extreme levels [25] and daily ozone concentrations [26], while CNN-LSTM hybrid models have shown strong performance in monthly climate prediction tasks [27]. However, they often require extensive fine-tuning of hyperparameters and remain susceptible to issues such as overfitting and suboptimal convergence [28]. Comparative studies across various domains, including urban climate and air quality forecasting, have shown that hybrid deep learning models outperform standalone architectures when properly optimized [29–31]. These findings further motivate the integration of optimization techniques within deep learning frameworks to enhance predictive performance.

To address these challenges, optimization techniques are commonly integrated into DL models. Among these, nature-inspired metaheuristic algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Teaching-Learning Based Optimization (TLBO), Ant Colony Optimization (ACO), Grey Wolves Optimization (GWO) and many more have been widely used to optimize model parameters, enhancing both accuracy and generalization [32–37]. Despite their success, many of these algorithms face limitations in navigating complex, high-dimensional search spaces, often leading to premature convergence and suboptimal solutions [28]. Building on these advancements, the Barnacles Mating Optimizer (BMO), a nature-inspired algorithm, has recently emerged as an effective tool for solving complex optimization problems [38,39]. Inspired

by the mating behavior of barnacles, BMO exhibits strong global search capabilities and robustness against local optima [40–42]. It has been successfully applied in various fields, including structural engineering, machine learning, and energy optimization, making it a promising candidate for optimizing DL models.

This paper presents an innovative hybrid model that combines CNN, LSTM, and BMO to forecast chiller power consumption with unprecedented accuracy. The proposed CNN-LSTM-BMO model incorporates the spatial feature extraction capabilities of CNN, the temporal sequence modeling strength of LSTM, and the optimization power of BMO to enhance predictive performance. By fine-tuning the selected key hyperparameters, BMO addresses the limitations of standalone CNN and LSTM models, leading to reduced prediction errors and improved generalization. Comprehensive comparisons with other metaheuristic algorithms, such as GA, PSO, and Differential Evolution (DE), are conducted to validate the model's superiority. Furthermore, SHapley Additive exPlanations (SHAP) analysis is employed to provide insights into feature importance, highlighting the critical role of temperature-related variables in chiller power consumption forecasting. This study contributes to the growing body of research on energy management by offering a robust and interpretable forecasting framework tailored to chiller systems. The proposed CNN-LSTM-BMO model not only addresses the limitations of existing forecasting methods but also provides practical implications for optimizing HVAC operations, ultimately paving the way for more energy-efficient building management systems.

The structure of the paper is organized as follows: Section 2 introduces the real-world dataset utilized in this research, followed by a concise summary of the CNN-LSTM model in Section 3. Section 4 offers a brief explanation of the BMO, and Section 5 outlines the methodology applied to model chiller power consumption using the CNN-LSTM-BMO approach. The results are thoroughly presented and discussed in Section 6, while Section 7 concludes the paper with key insights and findings.

2. Data set

The data used in this research, obtained from [43,44], includes a variety of temperatures, relative humidity, power, and energy consumptions by the fans. The dataset contains information about the HVAC system's operational settings that manage the indoor comfort of a commercial building in the winter of 2019–2020 and 2020–2021 in Turin, Italy. Fig. 1 illustrates how the chiller power consumption data is distributed and partitioned for model development and evaluation. The dataset consists of 33,888 total instances, with each instance representing measurements taken at 15-min intervals. Following standard DL practices, the data was split into 3 distinct sets: training (70%, 23,722 instances), validation (10%, 3389 instances), and testing (20%, 6777 instances). This partitioning ensures robust model training while maintaining sufficient data for validation and testing phases.

The input features encompass a comprehensive set of HVAC system parameters, including 3 temperature measurements (T_{return} , T_{supply} , $T_{outdoor}$) measured in degrees Celsius ($^{\circ}\text{C}$), 3 relative humidity measurements (RH_{return} , RH_{supply} , $RH_{outdoor}$) expressed as percentages (%), the return air temperature setpoint (SP_{return}) in $^{\circ}\text{C}$, and the saturation temperature in the humidifier ($T_{saturation}$) in $^{\circ}\text{C}$. The output variable shown in Fig. 1 represents the chiller power consumption in kilowatts (kW), which displays the actual power required by the fans in the HVAC system. The output variable is specifically chosen as power consumption in kilowatts (kW), rather than energy, due to the discrete nature of energy measurements in the original dataset.

The time-series visualization reveals distinct patterns in power consumption across all 3 datasets, with values ranging primarily between 0 and approximately 5.3 kW. The training set (shown in blue) exhibits consistent cyclic patterns typical of HVAC system operation, while the validation set (red) and testing set (green) demonstrate similar characteristics, suggesting appropriate data distribution across

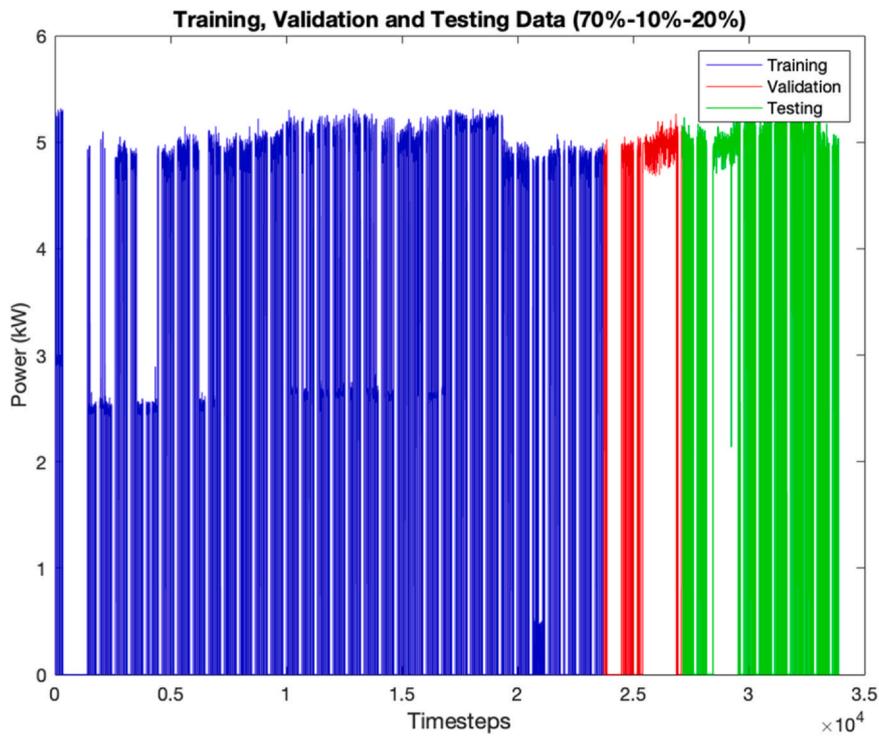


Fig. 1. The split of data into training, validation, and testing for the output of chiller power consumption.

the splits. This consistency is essential for ensuring that the model can effectively learn and generalize from the training data to unseen scenarios. The temporal resolution of 15-min intervals provides sufficient granularity to capture both rapid changes in power consumption and longer-term operational patterns, making it suitable for developing accurate forecasting models. Furthermore, the substantial size of the training dataset (23,722 instances) ensures that the deep learning model has adequate examples to learn the complex relationships between the input features and power consumption patterns.

Fig. 2 presents a visual overview of the proposed chiller power consumption forecasting framework. It illustrates the end-to-end workflow, showing how real-time building operational data flows from input parameters through the hybrid CNN-LSTM network that optimized using the BMO to produce the final prediction of chiller power consumption. The diagram clearly outlines the model’s architecture, with inputs on the left representing real-time building operational data, the core processing units in the center, and the predicted output on the right. This visual representation enhances understanding of the system’s structure and the interactions among its components. The input layer comprises a set of key parameters commonly monitored by Building Management Systems (BMS). These include temperature-related

features such as supply air temperature (T_{supply}), return air temperature (T_{return}), return air setpoint temperature (SP_{return}), saturation temperature ($T_{saturate}$), and outdoor air temperature ($T_{outdoor}$). Additionally, humidity-related features, namely supply air relative humidity (RH_{supply}), return air relative humidity (RH_{return}), and outdoor air relative humidity ($RH_{outdoor}$), are also incorporated as inputs. These parameters provide a comprehensive representation of the building’s thermal environment, enabling the model to capture the complex relationships influencing chiller power consumption.

The central processing unit combines a hybrid CNN-LSTM architecture with the BMO. The CNN layers are designed to extract spatial features and local patterns from the time-series input data, while the LSTM layers are employed to capture temporal dependencies and long-term trends. The BMO algorithm plays a crucial role in optimizing the hyperparameters of the CNN-LSTM model, ensuring optimal performance and generalization capability. The model’s output is the predicted chiller power consumption, measured in kW, providing a valuable tool for building energy management and HVAC system optimization.

3. Convolution neural network – Long short-term memory model

3.1. Fundamental architecture of CNN

A CNN is a specialized deep learning model designed to process data with a grid-like structure, such as images or time-series data. The key components of a CNN are:

Convolutional Layers: These layers apply convolution operations using learnable filters to extract spatial or temporal features. The convolution operation is expressed as:

$$z_{i,j,k} = \sum_{m=1}^M \sum_{n=1}^N x_{i+m-1,j+n-1} \cdot W_{m,n,k} + b_k \tag{1}$$

where:

- $z_{i,j,k}$ is the output feature map,
- x is the input data,

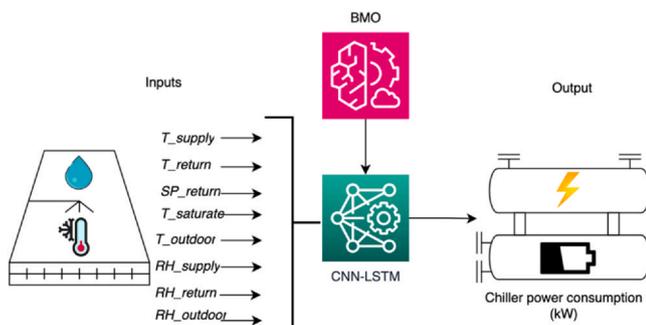


Fig. 2. Proposed CNN-LSTM-BMO for chiller power consumption forecasting. BMO = Barnacles Mating Optimizer; CNNs = Convolutional Neural Networks; LSTM = Long Short-Term Memory.

- $w_{m,n,k}$ is the k -th filter of size $M \times N$,
- b_k is the bias term.

Pooling Layers: These layers down-sample the feature maps, typically using max pooling or average pooling to reduce dimensionality and computation.

Fully Connected Layers: These layers map the extracted features to the output classes or values.

3.2. Fundamental architecture of LSTM

LSTM is a type of RNN capable of learning long-term dependencies. It is particularly effective for sequential data. The key equations governing an LSTM cell are:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{5}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \odot \tanh(C_t) \tag{7}$$

where:

- f_t, i_t, o_t are the forget, input, and output gates, respectively,
- \tilde{C}_t and C_t are the candidate cell state and updated cell state,
- h_t is the hidden state at time step t ,
- σ and \tanh are the sigmoid and hyperbolic tangent activation functions, respectively.

3.3. Hybrid CNN-LSTM architecture

The CNN-LSTM hybrid model is an innovative architecture that uses the strengths of CNN and LSTM networks for sequence modeling. This combination is particularly effective for data with both spatial and temporal characteristics, as it enables the extraction of localized patterns while capturing long-term dependencies. Fig. 3 illustrates the

general architecture of the hybrid CNN-LSTM model, which consists of 4 main components: the Input Layer, Feature Extraction (CNN), Transition Layer, and Sequential Layer (LSTM).

Input Layer: The input layer receives time-series data, such as chiller power consumption features. These features are typically multidimensional, containing both spatial and temporal information. The input data is structured to feed into the subsequent layers for processing.

Feature Extraction (CNN): The CNN component is responsible for extracting spatial or temporal features from the input data. This process begins with convolutional layers, which apply multiple filters to detect patterns such as edges, textures, or specific features in the input data. Following the convolution operations, activation functions (e.g., ReLU) introduce non-linearity to enhance the model's ability to capture complex relationships. Pooling layers, such as max pooling, reduce the spatial dimensions of the feature maps while preserving the most significant information. This step condenses the raw input into a set of high-level feature maps, effectively summarizing the spatial characteristics of the data.

Transition Layer: Once the feature maps are generated, the next step is to reduce their dimensions for further sequential processing. This can be achieved through either flattening the feature maps into a single-dimensional vector or applying global pooling operations, such as global average pooling. These operations transform the feature maps into a compact representation that retains essential information, making them suitable for sequential processing in the LSTM layer. The reduced feature maps are then reshaped and treated as a sequential dataset before being passed to the LSTM layers.

Sequential Layer (LSTM): The LSTM network processes the extracted features as a sequence, allowing it to model temporal dependencies and capture patterns over time. The LSTM's memory cells and gating mechanisms ensure the retention of relevant information while discarding less significant data, making it highly effective for tasks involving time-series data or sequential dependencies. The output of the CNN layers, after being flattened or pooled, serves as the input to the LSTM layers. This integration allows the model to utilize the spatial feature extraction capabilities of CNNs and the sequence modeling strengths of LSTMs. The feature maps extracted by the CNN are interpreted as sequential data by the LSTM, enabling the hybrid architecture to learn both local patterns and their temporal evolution.

Output Layer: The final layer of the model is a dense layer, which produces the predicted output. In the context of chiller power

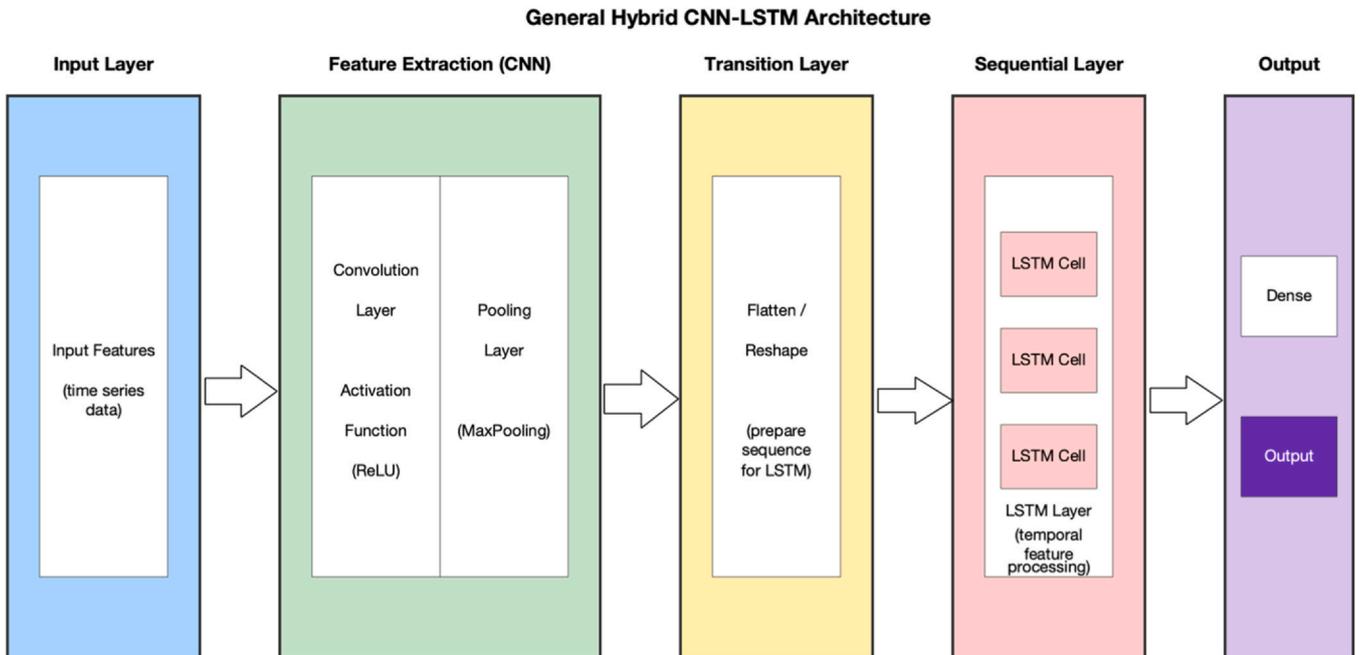


Fig. 3. Hybrid CNN-LSTM architecture. CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory.

consumption forecasting, this output represents the estimated power consumption based on the learned patterns from the input features.

Fig. 3 provides a visual representation of the hybrid CNN-LSTM architecture, highlighting the flow of data from the input layer through the CNN, transition layer, LSTM, and finally to the output layer. This architecture is designed to effectively handle the complexities of time-series data by combining the strengths of CNNs and LSTMs.

3.4. Parameters or variables optimized by BMO

The BMO is employed as an effective strategy to optimize the hyperparameters of the CNN-LSTM hybrid model, ensuring robust and efficient training, where in optimization they are treated as variables to be optimized. The optimization process focuses on minimizing the Mean Squared Error (MSE) as the performance metric, guiding the selection of the most effective hyperparameters. These hyperparameters include the following:

Number of Filters: This determines the depth of the feature maps in the convolutional layers of the CNN. A greater number of filters allows the extraction of more complex and diverse features from the input data. However, excessive filters can lead to overfitting and increased computational costs. As a discrete variable, the number of filters is optimized by BMO to achieve an ideal balance between feature extraction and model efficiency.

Filter Size: This specifies the dimensions of the convolutional kernel in the CNN layers, which influences the receptive field and the scale of patterns the model can detect. An appropriately sized filter ensures the capture of relevant spatial or temporal features without introducing noise or overlooking critical details. The filter size, being a discrete value, is systematically tuned by BMO to match the data characteristics and improve feature extraction.

Number of LSTM Units: This hyperparameter defines the capacity of the LSTM layers to learn and retain temporal dependencies. A sufficient number of units allows the model to capture intricate temporal patterns, while too many units can lead to overfitting and higher computational demands. As a discrete parameter, the number of LSTM units is optimized by BMO to balance temporal learning capability and model efficiency.

Dropout Rate: Regularization is critical to prevent overfitting, and the dropout rate controls the fraction of neurons dropped during training. By randomly disabling neurons in each iteration, the model is encouraged to learn more generalized features. Unlike the first 3 hyperparameters, the dropout rate is a floating-point number, requiring precise optimization to achieve effective regularization while preserving the model's capacity to learn.

Through BMO's biologically inspired optimization mechanisms, these hyperparameters are systematically tuned. The discrete nature of the first 3 variables (number of filters, filter size, and number of LSTM units) and the continuous nature of the dropout rate are handled effectively, ensuring a thorough exploration of the hyperparameter space. This approach ensures that the CNN-LSTM model is well-calibrated for the specific problem, delivering high predictive accuracy while maintaining generalization and efficiency.

The optimization process uses Mean Squared Error (MSE) as the minimization metric, expressed as:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

where:

- y_i is the actual value,
- \hat{y}_i is the predicted value,
- n is the total number of samples.

The BMO iteratively tunes the hyperparameters by minimizing the MSE on the validation set, ensuring optimal performance for the CNN-LSTM model.

4. Barnacles mating optimizer (BMO)

The BMO draws its inspiration from the natural reproductive strategies of barnacles, as described in the literature [38,39,45]. Barnacles employ 2 distinct mating mechanisms: direct copulation and sperm-casting. In direct copulation, physical contact occurs between male and female barnacles, facilitating natural reproduction. Alternatively, sperm-casting serves as a reproductive method for isolated barnacles, where fertilized eggs are released into the surrounding water. This unique reproductive adaptability of barnacles has been translated into an optimization algorithm. The BMO algorithm's structure incorporates the Hardy-Weinberg principle's Punnett square concept [46] for exploitation phases, while its exploration mechanism is modeled after the sperm-casting process, expressed as follow:

$$x_i^{N_new} = p x_{barnacle_d}^N + q x_{barnacle_m}^N \quad (9)$$

$$x_i^{n_new} = rand() \times x_{barnacle_m}^n \quad (10)$$

where p is the normally distributed pseudo random numbers, $q = (1-p)$, $x_{barnacle_d}^N$ and $x_{barnacle_m}^N$ are the variables of *Dad* and *Mum* of barnacles respectively and $rand()$ is the random number between [0,1]. Eq. (9) is used when the selection of parents to be mated are within the pre-set value of parameter pl , which is the only parameter to be set in BMO apart from number of population and maximum iterations. The concept of exploitation and exploration proposed in BMO are adopted from [47].

5. Chiller power consumption forecasting using CNN-LSTM-BMO

The flowchart presented in the Fig. 4 illustrates the comprehensive methodology for implementing the CNN-LSTM-BMO model for chiller power consumption forecasting. The process follows a systematic approach from data preparation through model optimization and evaluation.

The workflow begins with the fundamental data preparation stage, where chiller data is loaded and appropriately segregated into training, validation, and testing sets. This is followed by the crucial initialization of BMO parameters, including the mating factor (pl), population size ($Npop$), maximum iterations, and number of runs, which form the foundation for the optimization process.

The next phase involves generating the initial population, consisting of matrix of [$Npop \times$ the number of variables] to be optimized. This population serves as the starting point for the optimization process. Subsequently, the network architecture is defined by establishing the CNN-LSTM layers and configuring the training parameters, which creates the framework for the deep learning model. The optimization process enters an iterative phase where each population member undergoes evaluation through training and validation. During this process, the MSE is recorded, and the algorithm maintains a record of the best results and corresponding variables. This iteration continues until reaching the maximum number of iterations specified in the initial parameters.

Upon completing the iterations, the model loads the optimized variables and conducts the final training using these optimal parameters. The trained CNN-LSTM-BMO model is then tested using previously unseen testing data to evaluate its generalization capability. This process is repeated until reaching the maximum number of runs, ensuring robust validation of the model's performance. The workflow concludes with the visualization and analysis phase, where the best results are plotted, including comparisons between actual and predicted values, and various performance metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), standard deviation, and R^2 . This comprehensive methodology ensures a systematic approach to developing and validating the hybrid CNN-LSTM-BMO model for accurate chiller power consumption forecasting. The RMSE, MAE, standard deviation, σ and R^2 metrics are expressed as follow:

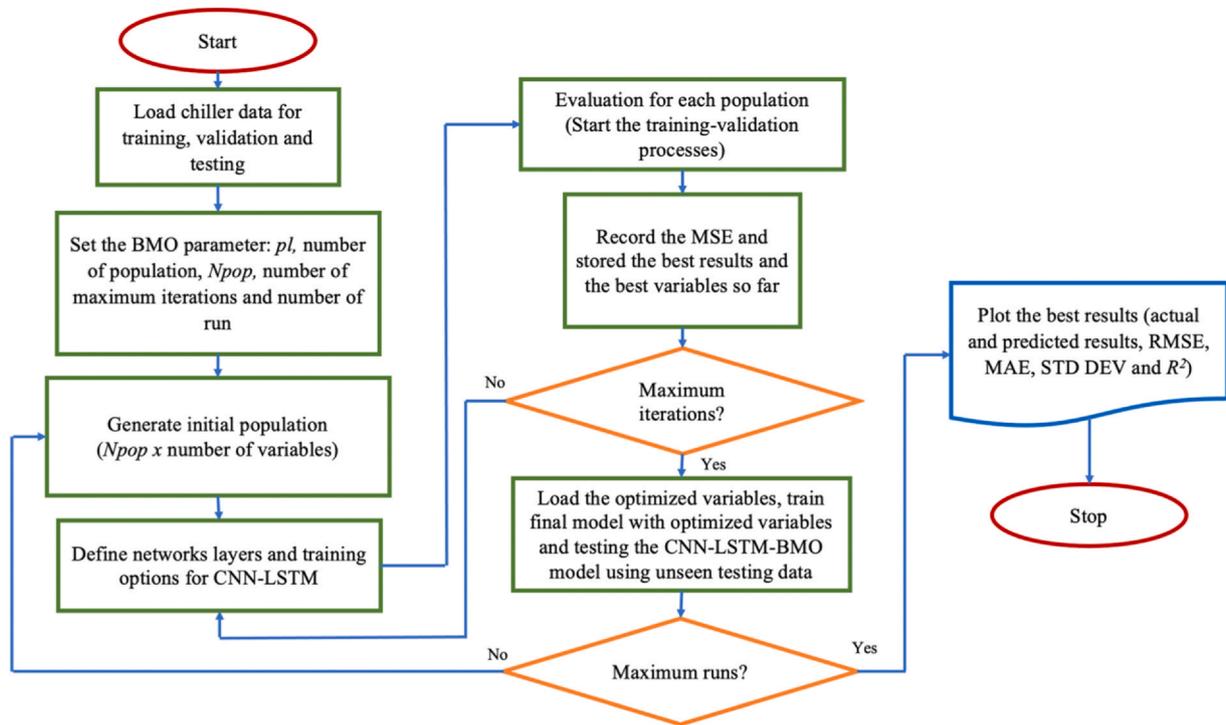


Fig. 4. Chiller power consumption forecasting using CNN-LSTM-BMO. BMO = Barnacles Mating Optimizer; CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory.

$$RMSE = \sqrt{MSE} \tag{11}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \tag{12}$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \tag{13}$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \tag{14}$$

6. Results and discussion

All simulations in this investigation were carried out using MATLAB. Table 1 displays the parameter settings for the chosen algorithms to ensure an equitable comparison. For all algorithms, consistent parameters were established with maximum iterations and population size both set at 10 to facilitate fair comparison. To assess the efficacy of the

Table 1
Parameter setting used for all algorithms

Algorithm	Parameter setting
# All algorithms	Number of populations = 10, maximum iteration = 10, Simulation runs = 5
BMO	Mating factor, $pl = 7/10$
PSO	$w = 1, w_{damp} = 0.99, c_1 = 1.5, c_2 = 2$
GA	Crossover rate = 0.7; Mutation rate = 0.3
ACO	Sample size = 10; Intensification factor, $q = 0.5$; Deviation-Distance Ratio, $\zeta = 1$
DE	Differential weight, $F = 0.8$, Crossover probability, $CR = 0.7$

ACO = Ant Colony Optimization; BMO = Barnacles Mating Optimizer; DE = Differential Evolution; GA = Genetic Algorithm; PSO = Particle Swarm Optimization.

metaheuristic algorithms in optimizing the variables to be optimized viz. number of filters, filter size, number of LSTM units, and dropout rate of CNN-LSTM model, the MSE is employed as the minimization metric. Among these optimization variables, the first 3 (filters, filter size, and LSTM units) are discrete integers, while the dropout rate is a continuous floating-point value. The lower and upper bounds of the variable are set as follows:

$$lb = [8, 2, 32, 0.1] \tag{15}$$

$$ub = [64, 5, 128, 0.5] \tag{16}$$

Table 2 presents a comparative analysis of CNN-LSTM models optimized by different metaheuristic algorithms for chiller power consumption forecasting in commercial buildings. The proposed CNN-LSTM-BMO demonstrates exceptional performance, achieving an RMSE of 0.5523 and R^2 of 0.9435, which translates to highly accurate predictions crucial for real-world building energy management. These performance metrics support the implementation of more effective energy-saving strategies and enable improved planning of chiller operations. The model's low MAE of 0.2966 and standard deviation of 0.5006 further indicate its reliability in making consistent predictions across varying operational conditions, which is essential for maintaining stable building comfort levels while optimizing energy consumption.

The superior performance of CNN-LSTM-BMO can be attributed to the distinctive characteristics of the BMO algorithm combined with its optimal hyperparameter configuration. Operating within defined bounds for number of filters [8–64], filter size [2–5], LSTM units [32–128], and dropout rate [0.1–0.5], BMO discovered an efficient architecture configuration of [8, 4, 81, 0.2889]. This configuration reveals that temporal pattern recognition (evidenced by 81 LSTM units) contributes more significantly to model performance than complex spatial feature extraction (evidenced by only 8 filters), while the moderate dropout rate of 0.2889 provides sufficient regularization to prevent overfitting. The algorithm's balanced exploration and exploitation mechanisms prove particularly effective in navigating this

Table 2
Performance results of CNN-LSTM-metaheuristic models for chiller power consumption forecasting

Models		RMSE	MAE	STD DEV	R^2	Optimized variables
CNN-LSTM-BMO	Best	0.5523	0.2966	0.5006	0.9435	[8, 4, 81, 0.2889]
	Average	0.7345	0.3354	0.7029	0.8951	NA
	Worst	0.9138	0.3655	0.8640	0.8474	[32, 5, 99, 0.4978]
CNN-LSTM-GA	Best	0.6520	0.3241	0.6265	0.9181	[10, 5, 91, 0.2789]
	Average	0.7677	0.3493	0.7228	0.8851	NA
	Worst	1.0721	0.4257	1.0061	0.7930	[50, 4, 76, 0.1422]
CNN-LSTM-PSO	Best	0.6715	0.3219	0.6481	0.9195	[8, 4, 43, 0.3431]
	Average	0.8883	0.4008	0.8395	0.8455	NA
	Worst	1.0897	0.4488	1.0397	0.7767	[64, 5, 44, 0.5000]
CNN-LSTM-ACO	Best	0.6622	0.2665	0.6385	0.9234	[13, 2, 33, 0.1539]
	Average	0.8103	0.3591	0.7812	0.8731	NA
	Worst	0.8951	0.3586	0.8623	0.8522	[64, 4, 74, 0.1120]
CNN-LSTM-DE	Best	0.6334	0.3171	0.5965	0.9255	[15, 3, 84, 0.4913]
	Average	0.8054	0.3784	0.7507	0.8737	NA
	Worst	1.0558	0.5082	0.9488	0.7851	[27, 5, 108, 0.4271]

ACO = Ant Colony Optimization; BMO = Barnacles Mating Optimizer; CNN = Convolutional Neural Network; DE = Differential Evolution; GA = Genetic Algorithm; LSTM = Long Short-Term Memory; MAE = Mean Absolute Error; PSO = Particle Swarm Optimization; RMSE = Root Mean Square Error.

complex hyperparameter landscape, as demonstrated by the narrower range between its best and worst performance metrics (RMSE range: 0.5523–0.9138) compared to alternative methods.

In contrast, the PSO implementation exhibited the weakest performance, with the highest average RMSE (0.8883) and worst-case RMSE (1.0897). Its suboptimal configuration [64, 5, 44, 0.5000] reveals a tendency toward maximum filter complexity but minimal LSTM units, indicating potential overfitting on spatial features while inadequately capturing temporal patterns. This architecture, combined with the maximum dropout rate, likely contributed to its inconsistent predictive performance. Although the GA approach performed better than PSO with its optimal configuration [10, 5, 91, 0.2789], it still demonstrated inferior results compared to BMO. GA's worst-case scenario [50, 4, 76, 0.1422] indicates potential underfitting due to the low dropout rate, highlighting the algorithm's challenge in maintaining a balanced model architecture. DE and ACO demonstrated intermediate performance levels, with DE showing slightly better results (RMSE: 0.6334, R^2 : 0.9255) compared to ACO (RMSE: 0.6622, R^2 : 0.9234). DE's best configuration [15, 3, 84, 0.4913] suggests a more balanced approach to spatial-temporal feature extraction, though the high dropout rate may have limited its potential for further improvement. ACO's optimal configuration [13, 2, 33, 0.1539] utilized minimal filter size and LSTM units, indicating possible underfitting, yet achieved reasonable performance through better generalization with its low dropout rate. However, both methods exhibited higher variability in their performance metrics compared to BMO, suggesting less robust optimization capabilities.

The overall results validate the effectiveness of the proposed CNN-LSTM-BMO approach for chiller power consumption forecasting. The consistent superiority across all evaluation metrics, particularly the high R^2 values (ranging from 0.8474 to 0.9435) and low error rates, indicates that the model successfully captures the complex patterns and relationships in chiller power consumption data. The improved performance of BMO compared to well-established metaheuristic methods demonstrates its unique capability in finding optimal hyperparameter configurations that balance model complexity with generalization ability. These findings have significant implications for building energy management systems, as the accurate power consumption forecasting achieved through this optimized architecture can lead to more efficient chiller operations, reduced energy costs, and improved overall building energy efficiency. The robust performance across different metrics suggests that the CNN-LSTM-BMO model can be reliably deployed in real-world applications, providing facility managers with accurate predictions for more informed decision-making in energy management strategies.

Fig. 5 presents the convergence curves of various CNN-LSTM models combined with different metaheuristic optimization algorithms for

chiller power consumption forecasting. The graph tracks MSE evolution over 10 iterations, revealing the training efficiency and optimization characteristics of each method. Notably, while CNN-LSTM-GA achieved the best convergence performance during the training process with a final MSE of approximately 0.0143, followed closely by CNN-LSTM-BMO, it is significant to highlight that CNN-LSTM-BMO demonstrated superior overall performance in the testing phase as evidenced by the comprehensive metrics shown in Table 2.

The optimization process utilized the parameter settings outlined in Table 1, with consistent algorithm configurations to ensure fair comparison. The simulation was repeated 5 times to ensure result reliability and validate the consistency of the optimization process. The population size and maximum iterations were both set to 10, which proved sufficient for optimizing the 4 critical parameters of the CNN-LSTM model: number of filters, filter size, number of LSTM units, and dropout rate. This relatively modest number of iterations and population size is particularly noteworthy as it demonstrates the efficiency of these optimization approaches in finding optimal hyperparameter combinations without requiring extensive computational resources.

The convergence patterns reveal distinctive optimization behaviors among the algorithms. CNN-LSTM-GA shows remarkable improvement in later iterations, particularly around iteration 9, achieving the lowest training MSE. CNN-LSTM-BMO demonstrates consistent and stable convergence throughout the training process, starting with relatively high MSE but quickly improving and maintaining competitive performance. CNN-LSTM-PSO exhibits the most stable convergence pattern but settles at a higher MSE compared to GA and BMO. CNN-LSTM-ACO shows stepped improvements, particularly around iteration 6, eventually achieving competitive results. CNN-LSTM-DE, despite showing aggressive early optimization, maintains steady but moderate performance throughout the process.

Despite GA's superior convergence in the training phase, it's crucial to emphasize that CNN-LSTM-BMO's overall performance in the testing phase surpassed all other methods, as evidenced by its superior RMSE, MAE, and R^2 (0.9435) values demonstrated in Table 2. This suggests that BMO achieves a better balance between optimization and generalization, making it more suitable for practical applications in chiller power consumption forecasting. The parameter settings, particularly BMO's mating factor namely pI , appear to strike an optimal balance between exploration and exploitation, contributing to its robust performance across both training and testing phases.

Fig. 6 presents a comprehensive visualization of the forecasting results, comparing the observed (actual) and predicted chiller power consumption values over approximately 6777 timesteps, with each timestep representing 15-min intervals. The blue line represents the observed values, while the red line shows the predictions made by the

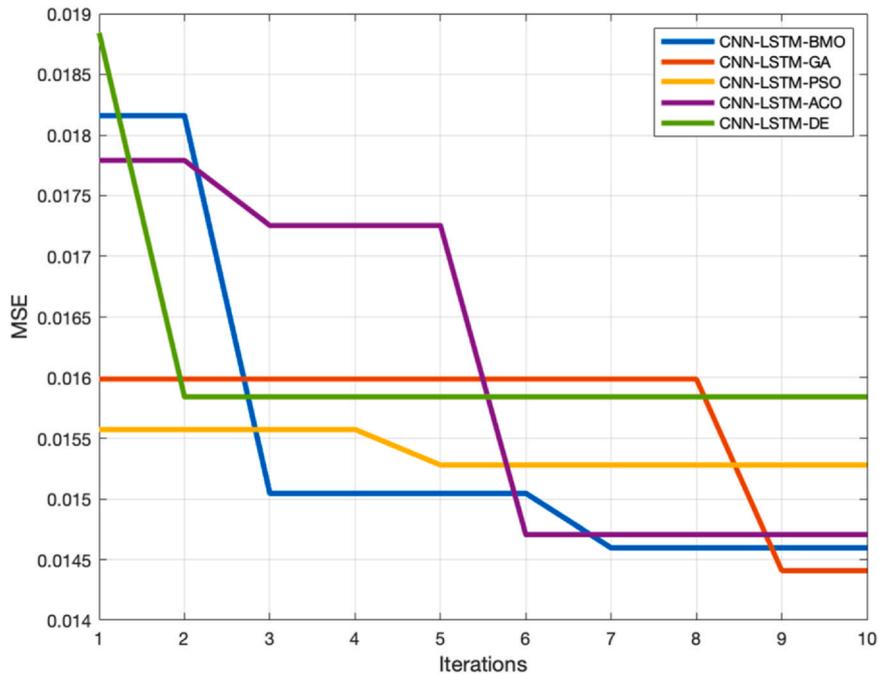


Fig. 5. Convergence curve CNN-LSTM-metaheuristic methods for training process of chiller power consumption forecasting. CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory.

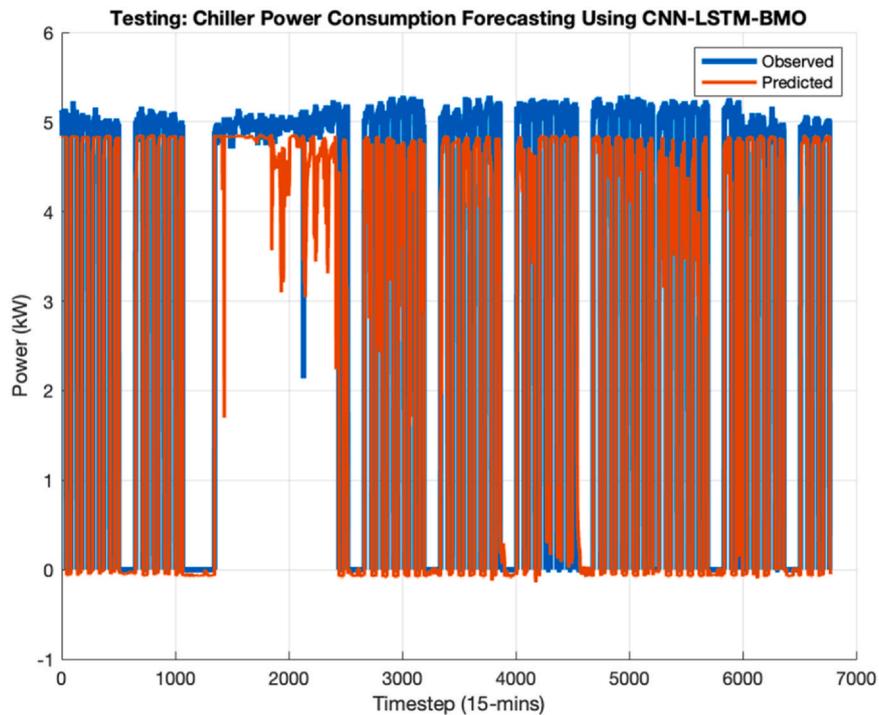


Fig. 6. Forecasting results of testing data of chiller power consumption by CNN-LSTM-BMO. BMO = Barnacles Mating Optimizer; CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory.

CNN-LSTM-BMO model. Several key observations can be drawn from this visualization, firstly, the forecasting performance demonstrates remarkable accuracy in tracking the overall pattern of chiller power consumption. The predicted values (red line) closely follow the observed values (blue line), particularly in capturing the cyclic nature of power consumption patterns. The model successfully predicts both the high consumption periods, where power usage reaches approximately 5 kW, and the low consumption periods, where usage drops to near 0 kW, indicating its ability to capture both peak demands and off-peak periods.

Secondly, a notable aspect of the model's performance is its ability to capture the sharp transitions between operational and non-operational periods of the chiller system. These transitions are particularly evident in the vertical lines dropping from peak usage to zero, representing the chiller's on-off cycles. The CNN-LSTM-BMO model accurately predicts these sudden state changes, which is crucial for practical applications in building energy management systems. Thirdly, looking at the fine details, there are some minor discrepancies between predicted and actual values, particularly during rapid fluctuations in

power consumption. The time-series plot also reveals interesting patterns in chiller operation, such as regular cycling patterns and varying power demand intensities. The CNN-LSTM-BMO model's ability to capture these patterns suggests that it has successfully learned the underlying relationships between temporal features and power consumption, making it a reliable tool for short-term load forecasting in building energy management applications.

It is worth noting that the model's forecasting occasionally underestimates the maximum power consumption peaks, not quite reaching the actual maximum value of 5.316 kW. This phenomenon can be attributed to several factors: (1) the nature of neural network training, which tends to average out extreme values to minimize overall error; (2) the relative rarity of these maximum peaks in the training data, making it challenging for the model to fully capture these exceptional cases; and (3) the model's inherent tendency to make conservative predictions for extreme values to maintain prediction stability and avoid overfitting. Despite this limitation, the model's ability to accurately track the overall consumption patterns and timing of peak usage periods remains robust and reliable for practical applications.

Figs. 7–10 provide a comprehensive visualization of the forecasting results for the hybrid methods CNN-LSTM-GA, CNN-LSTM-PSO, CNN-LSTM-ACO, and CNN-LSTM-DE, respectively. All these hybrid approaches demonstrate generally good forecasting performance, capturing the overall trends and temporal patterns of the chiller power consumption effectively. However, their performance still reveals certain limitations compared to the proposed CNN-LSTM-BMO model. One notable shortcoming is the occasional underestimation of maximum power consumption peaks, where the predicted values fall short of the observed peak value of 5.316 kW. This limitation can be attributed to the inherent challenges of the CNN-LSTM architecture in accurately capturing abrupt changes and extreme variations, particularly during rapid ON/OFF state transitions in chiller operations.

The performance differences between these methods can largely be attributed to the effectiveness of the metaheuristic algorithms in optimizing the 4 critical parameters of the CNN-LSTM model: number of filters, filter size, number of LSTM units, and dropout rate. These parameters play a pivotal role in determining the model's ability to

extract meaningful spatial-temporal features, maintain long-term dependencies, and prevent overfitting. While GA, PSO, ACO, and DE exhibit promising results by improving parameter selection, their optimization capabilities may not be as efficient as BMO. The BMO demonstrates a superior ability to balance exploration and exploitation, enabling it to search the parameter space more effectively and converge toward a globally optimal solution. This advantage is reflected in the superior predictive accuracy of the CNN-LSTM-BMO model, which better handles the complex and dynamic nature of chiller power consumption. While all hybrid models exhibit commendable performance, the occasional underprediction of peak values highlights the challenges of parameter optimization for CNN-LSTM models. This finding highlights the importance of employing robust and efficient metaheuristic algorithms to fine-tune parameters such as the number of filters, filter size, LSTM units, and dropout rate. The BMO's superior performance emphasizes its potential for achieving better convergence and accuracy.

Table 3 shows the snippet results of chiller power consumption forecasting across all approaches, which are part of the outcomes plotted in Figs. 6–10. Observing the actual and predicted values at the given time steps (2000–2030), it is evident that all models closely follow the actual trends. However, variations exist in prediction accuracy, particularly for CNN-LSTM-GA and CNN-LSTM-PSO, where their predicted values exhibit a slightly higher deviation compared to CNN-LSTM-BMO and CNN-LSTM-DE. For instance, at time step 2004, CNN-LSTM-PSO predicts a value of 5.009, which is higher than the actual value 4.920, indicating a noticeable error. On the other hand, CNN-LSTM-BMO and CNN-LSTM-DE maintain better proximity to the actual values, reflecting their superior predictive capability. The CNN-LSTM-BMO approach consistently provides smoother predictions, aligning closer to the actual trends, which is also reflected in its lowest RMSE (0.5523) and highest R^2 (0.9435).

The performance metrics at the bottom of Table 3, particularly RMSE and R^2 , further highlight the strengths and weaknesses of each approach. CNN-LSTM-BMO achieves the best overall accuracy with the lowest RMSE of 0.5523 and the highest R^2 value of 0.9435, demonstrating its ability to generalize well and capture the nonlinear characteristics of chiller power consumption. Conversely, CNN-LSTM-GA

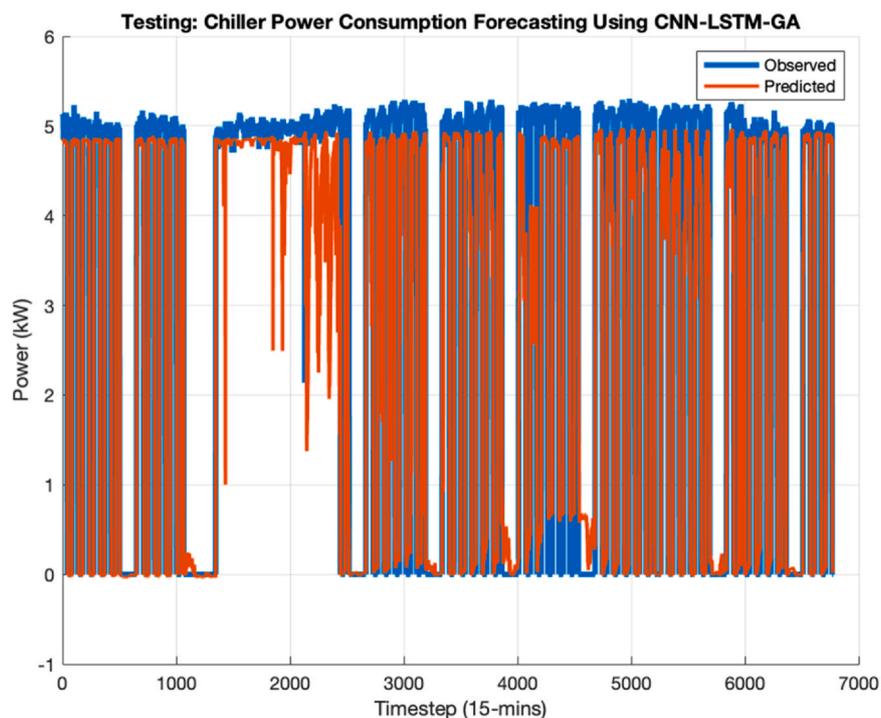


Fig. 7. Forecasting results of chiller power consumption by CNN-LSTM-GA. CNN = Convolutional Neural Network; GA = Genetic Algorithm; LSTM = Long Short-Term Memory.

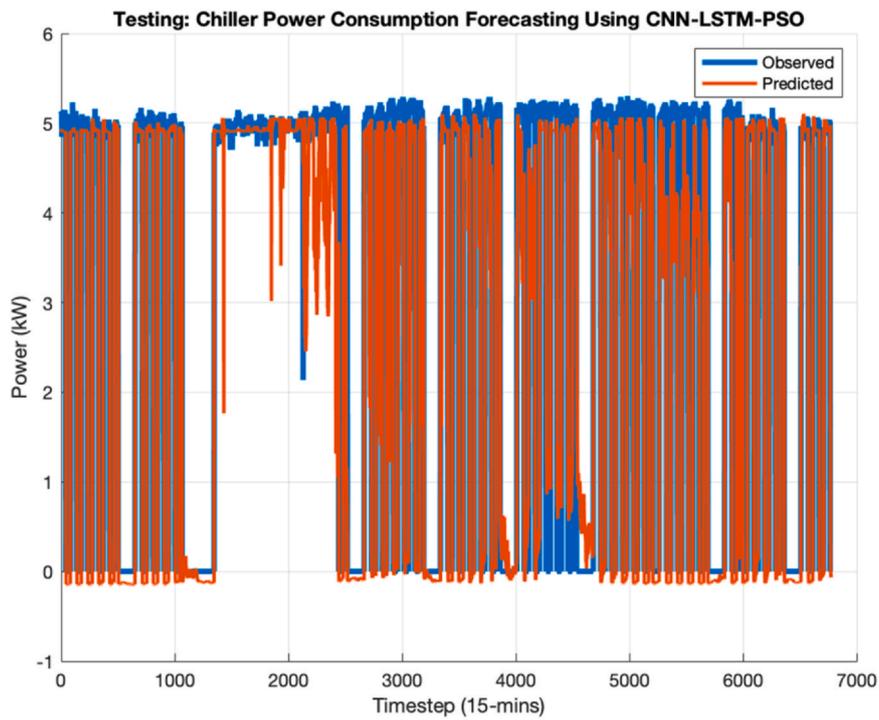


Fig. 8. Forecasting results of chiller power consumption by CNN-LSTM-PSO. CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory; PSO = Particle Swarm Optimization.

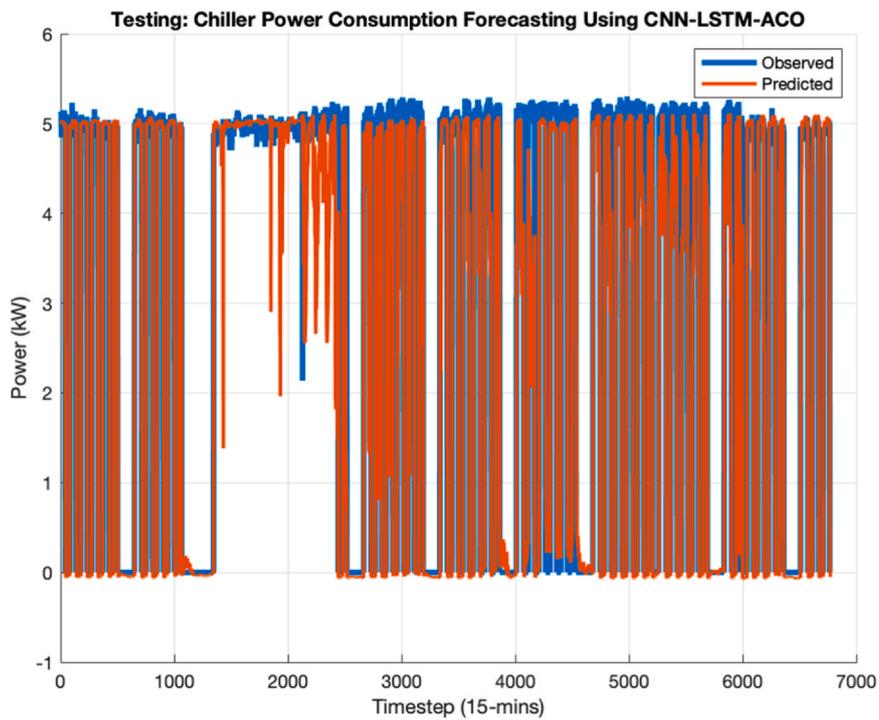


Fig. 9. Forecasting results of chiller power consumption by CNN-LSTM-ACO. ACO = Ant Colony Optimization; CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory.

and CNN-LSTM-PSO show relatively higher RMSE values of 0.6520 and 0.6715, respectively, and slightly lower R^2 values. This indicates that while both GA and PSO-based methods are capable of achieving good predictions, they are more susceptible to overfitting or convergence to local optima compared to the BMO approach. CNN-LSTM-DE and CNN-LSTM-ACO also perform well, with R^2 values of 0.9255 and 0.9234, respectively, confirming their robustness and competitive performance.

Despite the overall high accuracy demonstrated by CNN-LSTM-BMO, CNN-LSTM-DE, and CNN-LSTM-ACO, further improvements can still be achieved. The residual errors, particularly at time steps where sudden variations in actual values occur (e.g., time steps 2004 and 2029), suggest that the models could benefit from incorporating adaptive learning strategies or hybrid optimization techniques to enhance responsiveness to rapid fluctuations. For example, integrating a

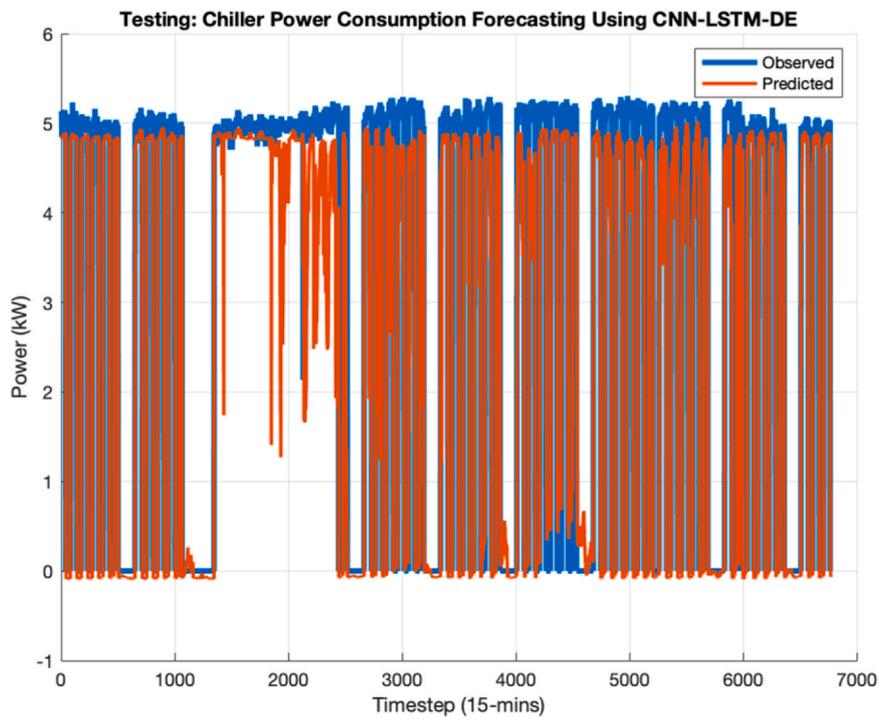


Fig. 10. Forecasting results of chiller power consumption by CNN-LSTM-DE. CNN = Convolutional Neural Network; DE = Differential Evolution; LSTM = Long Short-Term Memory.

Table 3

Snippets results of chiller power consumption forecasting by all approaches

Timesteps	Actual	CNN-LSTM-BMO	CNN-LSTM-GA	CNN-LSTM-PSO	CNN-LSTM-ACO	CNN-LSTM-DE
:	:	:	:	:	:	:
2000	4.908	4.173	4.460	4.895	4.774	4.102
2001	4.968	4.274	4.552	4.962	4.841	4.300
2002	4.980	4.278	4.572	4.977	4.850	4.327
2003	4.956	4.268	4.547	4.962	4.831	4.283
2004	4.920	4.310	4.639	5.009	4.883	4.404
2005	4.908	4.295	4.662	5.005	4.932	4.480
2006	4.932	4.364	4.699	5.029	4.951	4.569
2007	5.004	4.576	4.812	5.054	5.022	4.778
2008	5.040	4.733	4.857	5.012	5.034	4.831
2009	5.016	4.792	4.864	4.978	5.034	4.830
2010	4.968	4.811	4.863	4.954	5.030	4.829
2011	4.992	4.821	4.860	4.941	5.027	4.829
2012	5.040	4.827	4.857	4.932	5.023	4.831
2013	4.992	4.827	4.858	4.933	5.028	4.831
2014	5.076	4.832	4.852	4.923	5.022	4.834
2015	4.980	4.834	4.850	4.919	5.019	4.836
2016	4.992	4.833	4.851	4.919	5.024	4.835
2017	4.920	4.828	4.857	4.929	5.030	4.832
2018	5.028	4.829	4.855	4.927	5.031	4.830
2019	5.004	4.833	4.848	4.916	5.030	4.831
2020	5.040	4.835	4.847	4.915	5.034	4.832
2021	5.028	4.833	4.852	4.920	5.043	4.828
2022	4.980	4.832	4.855	4.926	5.043	4.827
2023	4.980	4.831	4.858	4.931	5.045	4.827
2024	4.908	4.829	4.864	4.929	5.052	4.820
2025	5.052	4.837	4.855	4.912	5.047	4.832
2026	5.028	4.840	4.833	4.904	5.030	4.857
2027	5.040	4.842	4.848	4.906	5.043	4.846
2028	4.944	4.841	4.848	4.904	5.045	4.857
2029	4.956	4.838	4.873	4.914	5.064	4.823
2030	4.992	4.841	4.855	4.905	5.052	4.850
:	:	:	:	:	:	:
RMSE		0.5523	0.652	0.6715	0.6622	0.6334
R ²		0.9435	0.9181	0.9195	0.9234	0.9255

ACO = Ant Colony Optimization; BMO = Barnacles Mating Optimizer; CNN = Convolutional Neural Network; DE = Differential Evolution; GA = Genetic Algorithm; LSTM = Long Short-Term Memory; PSO = Particle Swarm Optimization; RMSE = Root Mean Square Error.

Table 4
Paired *t*-test results: CNN-LSTM-BMO vs. other methods

Methods	<i>P</i> -value	Significant
CNN-LSTM-GA	7.11×10^{-10}	*
CNN-LSTM-PSO	9.66×10^{-6}	*
CNN-LSTM-ACO	0	*
CNN-LSTM-DE	2.57×10^{-2}	*

ACO = Ant Colony Optimization; BMO = Barnacles Mating Optimizer; CNN = Convolutional Neural Network; DE = Differential Evolution; GA = Genetic Algorithm; LSTM = Long Short-Term Memory; PSO = Particle Swarm Optimization.

dynamic regularization approach could further fine-tune the model parameters for improved accuracy. Additionally, increasing the depth of the CNN-LSTM architecture or introducing attention mechanisms may enable the models to better capture subtle temporal patterns and reduce error deviations. Table 3 demonstrates that CNN-LSTM-BMO outperforms other methods in terms of accuracy and reliability for chiller power consumption forecasting. The BMO-based approach's superior performance can be attributed to its effective exploration-exploitation balance, leading to optimal tuning of the CNN-LSTM model parameters. CNN-LSTM-DE and CNN-LSTM-ACO also exhibit promising results, suggesting that DE and ACO remain strong candidates for solving similar forecasting problems.

Results of the paired *t*-test conducted to compare CNN-LSTM-BMO with other algorithms based on raw prediction outputs are tabulated in Table 4. The *P*-values obtained for each comparison indicate whether the performance differences between CNN-LSTM-BMO and the other methods are statistically significant. Notably, all *P*-values are well below the standard significance threshold of 0.05, suggesting that CNN-LSTM-BMO consistently outperforms the alternative approaches with statistically significant results. For example, the *P*-value for CNN-LSTM-BMO vs. CNN-LSTM-GA is 7.11×10^{-10} highlighting an extremely significant difference. This indicates that CNN-LSTM-BMO provides superior predictive performance compared to CNN-LSTM-GA with a very low probability that the observed results occurred by chance.

Similarly, the *P*-value for CNN-LSTM-PSO is 9.6×10^{-6} , which also confirms a significant performance improvement of CNN-LSTM-BMO

over CNN-LSTM-PSO. This result further supports the earlier findings from Table 3, where CNN-LSTM-PSO exhibited higher RMSE and slightly lower R^2 values compared to CNN-LSTM-BMO. The *P*-value for CNN-LSTM-ACO is reported as exactly 0, implying a highly significant difference. This result shows the strength of CNN-LSTM-BMO in achieving more accurate and consistent predictions, as even minor variations in performance are statistically distinguishable. Lastly, for CNN-LSTM-DE, the *P*-value of 2.57×10^{-2} also falls below 0.05, demonstrating statistical significance, albeit less extreme compared to the other methods. This suggests that while CNN-LSTM-DE performs relatively well, it is still outperformed by CNN-LSTM-BMO.

The significant *P*-values across all methods reinforce the reliability and robustness of CNN-LSTM-BMO in chiller power consumption forecasting. The results highlight that the BMO enhances the predictive capability of CNN-LSTM by effectively tuning its parameters and avoiding local optima, which may be encountered by other optimization algorithms like GA and PSO. To further solidify these findings, future studies could employ additional statistical tests, such as Wilcoxon signed-rank tests, to cross-validate the significance of performance differences. Additionally, testing these methods on more extensive and diverse datasets could generalize the conclusions and confirm the robustness of the CNN-LSTM-BMO approach in various forecasting scenarios. Overall, Table 4 provides strong statistical evidence that CNN-LSTM-BMO significantly outperforms CNN-LSTM-GA, CNN-LSTM-PSO, CNN-LSTM-ACO, and CNN-LSTM-DE. The results not only validate the effectiveness of the BMO algorithm in enhancing CNN-LSTM performance but also suggest the potential for further improvements. Incorporating ensemble strategies or hybrid approaches with BMO may lead to even better results, ensuring enhanced accuracy and reliability for real-world forecasting applications.

To further analyze the performance of the developed CNN-LSTM-BMO hybrid model for chiller power consumption prediction, SHAP analysis is conducted to investigate the impact of various input features on the model's predictions. The SHAP analysis provides a detailed interpretation of how each feature contributes to the model's output, revealing both the importance and the directional influence of the features. The results of this analysis help in understanding the sensitivity of the model to each input, offering valuable insights into which variables are most influential for accurate predictions. Fig. 11 presents

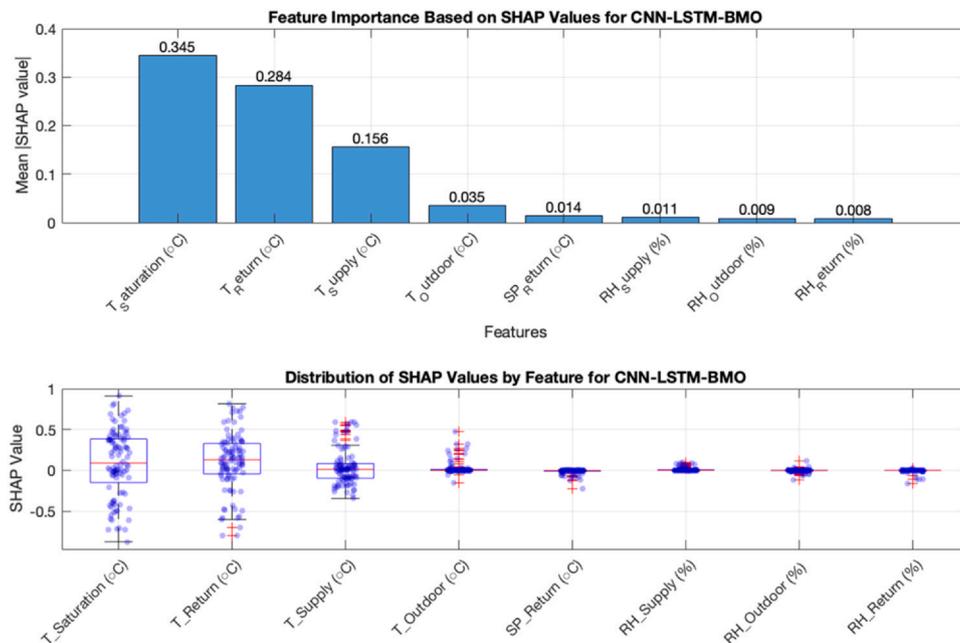


Fig. 11. SHAP analysis for CNN-LSTM-BMO. BMO = Barnacles Mating Optimizer; CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory; SHAP = SHapley Additive explanations.

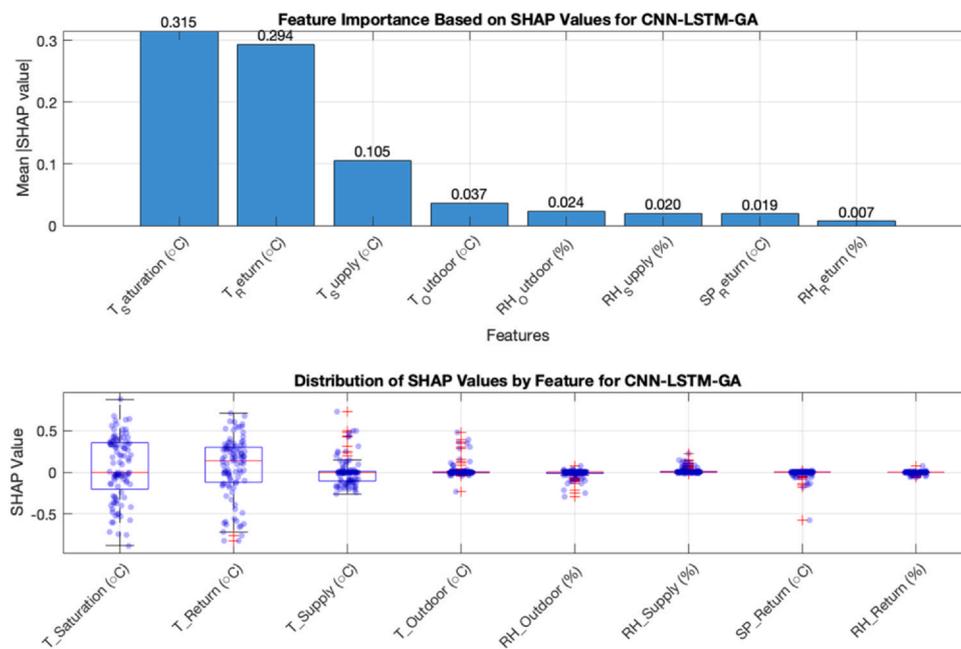


Fig. 12. SHAP analysis for CNN-LSTM-GA. CNN = Convolutional Neural Network; GA = Genetic Algorithm; LSTM = Long Short-Term Memory; SHAP = SHapley Additive exPlanations.

a comprehensive SHAP analysis of the CNN-LSTM-BMO model's feature importance, based on a balanced sample of 100 instances each from the training and test datasets. The analysis is visualized through 2 complementary plots: a bar chart showing mean absolute SHAP values and a distribution plot displaying the range and patterns of SHAP values for each feature. The temperature-related features emerge as the primary drivers of the model's predictions, with $T_{Saturation}$ demonstrating the highest mean absolute SHAP value of 0.345. This substantial influence is further supported by the distribution plot, which shows a wide spread of both positive and negative SHAP values, indicating that $T_{Saturation}$ can significantly increase or decrease the predicted power consumption depending on its value. T_{Return} follows as the second most influential feature with a mean SHAP value of 0.284, while T_{Supply} ranks third with 0.156, both showing similar bidirectional impact patterns in the distribution plot, though with progressively narrower ranges of influence.

The analysis reveals a clear hierarchical structure in feature importance, with a sharp decline in influence after the top 3 temperature-related features. $T_{Outdoor}$ shows a notably lower mean SHAP value of 0.035, followed by SP_{Return} at 0.014, marking a transition point in feature significance. The distribution plot for these features shows more concentrated SHAP values around zero, indicating more consistent but limited impacts on the model's predictions. The humidity-related features (RH_{Supply} , $RH_{Outdoor}$, and RH_{Return}) demonstrate the least influence on the model's predictions, with mean SHAP values ranging from 0.011 to 0.008. This minimal impact is visually confirmed in the distribution plot, where these features show highly concentrated SHAP values near zero with minimal dispersion, suggesting their contributions to power consumption predictions are relatively constant and limited. The box plot reveals few outliers for these features, indicating that even in extreme cases, their impact remains modest.

This detailed SHAP analysis provides valuable insights for both model interpretation and future development. The clear dominance of temperature-related features, particularly $T_{Saturation}$, suggests that these parameters should be prioritized in monitoring and control systems for chiller power consumption optimization. The minimal influence of humidity-related features indicates that while they contribute to the model's overall accuracy, they might be candidates for feature reduction in scenarios where computational efficiency is paramount.

These findings could inform the development of more streamlined models or help focus data collection efforts on the most impactful parameters.

Figs. 12–15 show the SHAP analysis for CNN-LSTM-GA, CNN-LSTM-PSO, CNN-LSTM-ACO, and CNN-LSTM-DE, respectively. The SHAP analysis conducted across CNN-LSTM models with different metaheuristic optimizers reveals consistent patterns in feature importance while highlighting subtle variations introduced by each optimization approach. The CNN-LSTM-BMO model, as illustrated in Fig. 11, demonstrates the highest feature importance values overall, with $T_{Saturation}$ achieving a mean absolute SHAP value of 0.345, surpassing similar values observed in other optimization approaches (GA: 0.3154, PSO: 0.3325, ACO: 0.3478, DE: 0.3186). This superior performance aligns with the model's better prediction accuracy as evidenced in earlier results. The relative importance of temperature-related features maintains a consistent hierarchy across all optimization variants, with $T_{Saturation}$ and T_{Return} consistently emerging as the top 2 influential features. However, the CNN-LSTM-BMO model exhibits a more pronounced differentiation between these features, with T_{Return} showing a mean SHAP value of 0.284, indicating a more refined feature importance distribution. This enhanced feature discrimination likely contributes to the model's superior forecasting performance, as demonstrated in the testing results. A notable distinction in the CNN-LSTM-BMO implementation is its treatment of secondary features such as T_{Supply} and $T_{Outdoor}$. While these features maintain their relative importance ranking across all optimization approaches, the BMO variant shows a more balanced distribution of their influence, with T_{Supply} maintaining a significant contribution (0.156) while avoiding the overshadowing effect seen in some other optimizers. This balanced feature utilization suggests that the BMO algorithm achieves a more optimal weighting of feature contributions.

The handling of humidity-related features (RH_{Supply} , $RH_{Outdoor}$, and RH_{Return}) reveals interesting variations across optimization approaches. The CNN-LSTM-BMO model demonstrates the most consistent treatment of these features, with their SHAP values showing minimal variance and maintaining clear, albeit small, contributions to the prediction process. This contrasts with other optimizers, particularly DE, where RH_{Supply} showed elevated importance levels that might indicate less optimal feature prioritization. The comprehensive

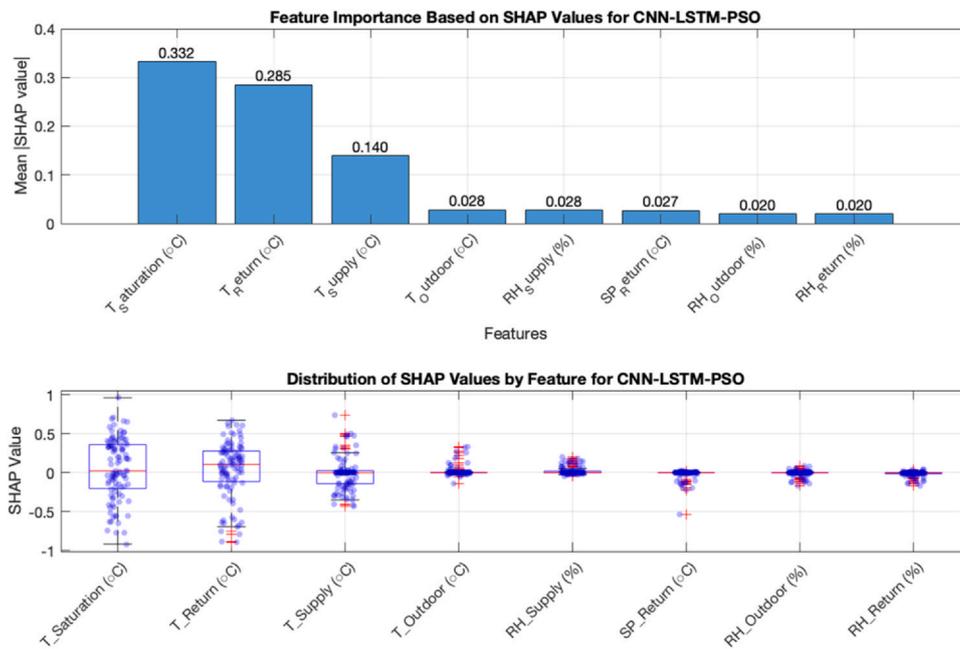


Fig. 13. SHAP analysis for CNN-LSTM-PSO. CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory; PSO = Particle Swarm Optimization; SHAP = SHapley Additive exPlanations.

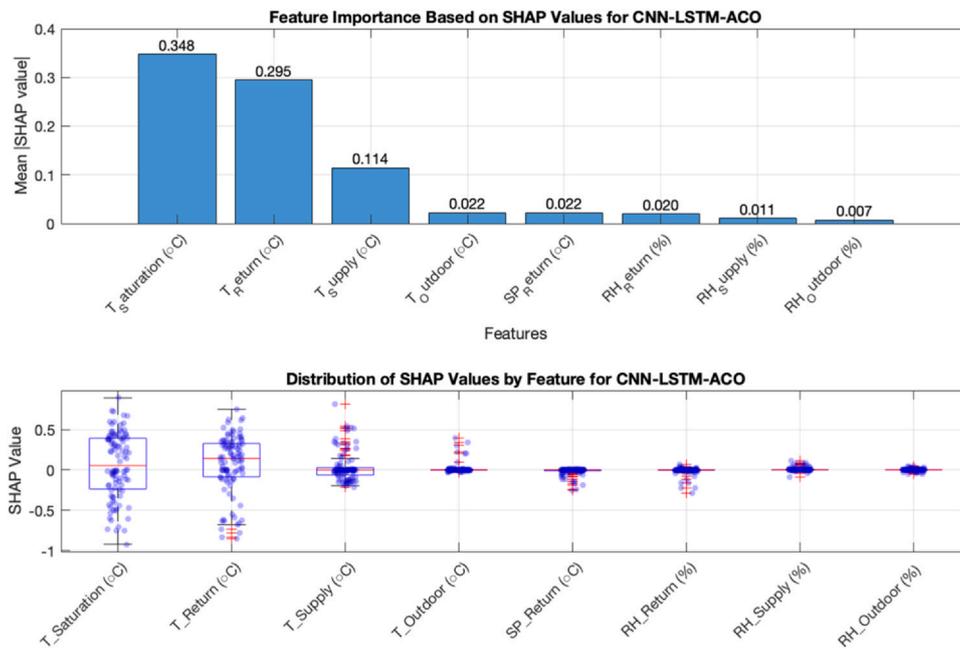


Fig. 14. SHAP analysis for CNN-LSTM-ACO. ACO = Ant Colony Optimization; CNN = Convolutional Neural Network; LSTM = Long Short-Term Memory; SHAP = SHapley Additive exPlanations.

comparison across all optimization variants highlights the CNN-LSTM-BMO model's superior ability to establish a more effective feature hierarchy. This is evidenced not only by the magnitude of SHAP values but also by the clear separation between primary and secondary features, suggesting a more refined optimization process. The model's ability to maintain consistent feature importance patterns while achieving higher overall performance metrics indicates that the BMO algorithm succeeds in finding a more optimal balance in feature utilization compared to traditional optimization approaches. The analysis conclusively demonstrates that while all CNN-LSTM variants identify similar key features, the BMO optimization approach achieves a more nuanced and effective feature importance distribution. This optimized

feature utilization contributes to its superior performance in chiller power consumption forecasting, supporting the choice of BMO as the preferred optimization algorithm for this application.

The experimental results demonstrate the exceptional performance of hybrid CNN-LSTM models optimized with various metaheuristic algorithms for chiller power consumption forecasting. The CNN-LSTM-BMO model achieved superior performance with the lowest RMSE (0.5523) and highest R^2 value (0.9435). This superiority is statistically validated through paired t -tests, which showed significant differences ($P < 0.05$) between CNN-LSTM-BMO and all other optimization approaches. The extremely low p -values, ranging from 0 to 7.11×10^{-10} , provide strong statistical evidence that BMO's superior performance is

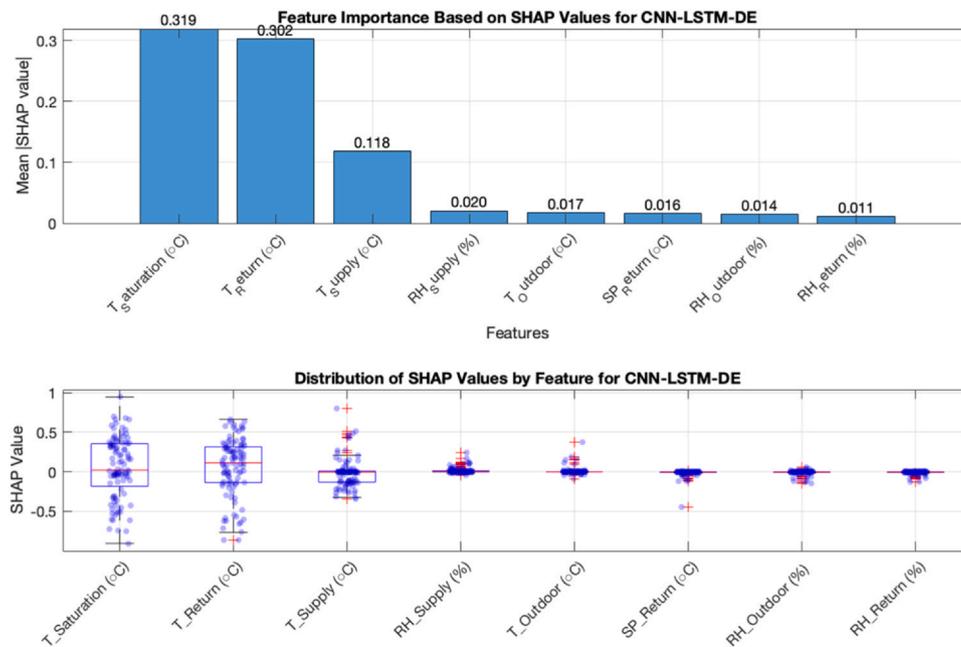


Fig. 15. SHAP analysis for CNN-LSTM-DE. CNN = Convolutional Neural Network; DE = Differential Evolution; LSTM = Long Short-Term Memory; SHAP = SHapley Additive exPlanations.

not due to chance, but rather represents a genuine improvement in forecasting capability. The convergence analysis revealed that while CNN-LSTM-GA showed the best training convergence, CNN-LSTM-BMO demonstrated superior generalization capability in testing. This finding, supported by the statistical significance in the t -tests, suggests that BMO achieves a better balance between optimization and generalization. The SHAP analysis further reinforced this conclusion by showing that BMO achieved more nuanced feature importance distributions, particularly in handling temperature-related parameters, which were consistently identified as the primary drivers of prediction accuracy.

This statistically validated performance has significant implications for HVAC and building management systems. The high prediction accuracy, combined with efficient convergence within 10 iterations, makes these models particularly valuable for real-world applications where both accuracy and computational efficiency are critical. The clear hierarchy of feature importance revealed by SHAP analysis provides practical guidance for system monitoring and sensor deployment strategies. The implementation of metaheuristic optimization in deep learning models presents several advantages, including automated hyperparameter tuning, robust global search capabilities, and adaptability to specific problem characteristics. However, challenges persist, such as parameter sensitivity and computational overhead. The statistically significant superiority of BMO over other methods suggests that newer metaheuristic algorithms may offer solutions to some of these traditional limitations.

Looking forward, several promising research directions emerge:

- Integration of multi-objective optimization approaches that balance prediction accuracy with computational efficiency.
- Development of adaptive parameter tuning mechanisms to enhance robustness.
- Investigation of transfer learning potential for similar building systems.
- Exploration of real-time optimization techniques for dynamic building conditions.
- Research into hybrid optimization approaches that could potentially outperform current methods.

The statistical validation of CNN-LSTM-BMO's superior performance provides a strong foundation for these future developments. The clear

performance advantages demonstrated through both numerical metrics and statistical tests suggest that continued research in this direction could yield further improvements in building energy management systems. This comprehensive analysis, supported by rigorous statistical validation, demonstrates that hybrid CNN-LSTM-metaheuristic models, particularly those optimized with BMO, represent a significant advancement in chiller power consumption forecasting. Their proven capability to deliver accurate predictions while maintaining computational efficiency makes them valuable tools for improving building energy management systems.

7. Conclusion

This research has successfully developed and validated a hybrid CNN-LSTM model optimized by the BMO for accurate chiller power consumption forecasting in commercial building. The comprehensive experimental results demonstrate the superior performance of the proposed CNN-LSTM-BMO model compared to other metaheuristic optimization approaches. The model achieved the lowest RMSE and highest R^2 value, with its superiority confirmed through rigorous statistical validation using paired t -tests. The comparative analysis of different metaheuristic optimizers revealed that while CNN-LSTM-GA showed promising convergence during training, CNN-LSTM-BMO demonstrated better generalization capability and overall performance in testing scenarios. The SHAP analysis provided valuable insights into feature importance, identifying temperature-related parameters as the primary drivers of prediction accuracy, with $T_{s_saturation}$ consistently showing the highest influence across all optimization variants.

The efficient convergence characteristics of the proposed model, achieving optimal performance within 10 iterations, make it particularly suitable for practical applications in building energy management systems. The model's ability to accurately capture both regular patterns and sudden changes in chiller power consumption demonstrates its potential for real-world implementation in HVAC system optimization. However, several challenges and opportunities for future research remain. These include the development of multi-objective optimization approaches, investigation of transfer learning possibilities, and exploration of real-time optimization techniques for dynamic building conditions. The successful implementation and validation of the CNN-LSTM-BMO model provides a strong foundation for these future

developments in building energy management systems. This study contributes significantly to the field by demonstrating the effectiveness of combining deep learning architectures with advanced metaheuristic optimization techniques for energy consumption forecasting. The findings have important implications for improving energy efficiency in commercial buildings and advancing the development of intelligent building management systems.

CRedit authorship contribution statement

Mohd Herwan Sulaiman: Data curation, Conceptualization, Methodology, Formal analysis, Writing – original draft. **Zuriani Mustafa:** Formal analysis, Validation, Writing – review & editing.

Data availability declaration

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) under Distinguished Research Grant (# RDU223003).

References

- [1] M. Vogt, C. Buchholz, S. Thiede, C. Herrmann, Energy efficiency of heating, ventilation and air conditioning systems in production environments through model-predictive control schemes: the case of battery production, 2022/05/20/, *J. Clean. Prod.* 350 (2022) 131354, <https://doi.org/10.1016/j.jclepro.2022.131354>.
- [2] X. Bai, Q. Tang, J. Luo, Y. Mao, C. Liang, X. Zhang, Optimizing energy efficiency in multi-chiller systems: a comprehensive Modelica-based approach, 2024/10/15/, *J. Build. Eng.* 95 (2024) 110087, <https://doi.org/10.1016/j.jobte.2024.110087>.
- [3] H. Shi, Q. Chen, Building energy management decision-making in the real world: a comparative study of HVAC cooling strategies, 2021/01/01/, *J. Build. Eng.* 33 (2021) 101869, <https://doi.org/10.1016/j.jobte.2020.101869>.
- [4] S. Wang, Z. Ma, Supervisory and optimal control of building HVAC systems: a review, 2008/01/01, *HVAC Res.* 14 (1) (2008) 3–32, <https://doi.org/10.1080/10789669.2008.10390991>.
- [5] Y. Zhao, C. Zhang, Y. Zhang, Z. Wang, J. Li, A review of data mining technologies in building energy systems: load prediction, pattern identification, fault detection and diagnosis, 2020/04/01/, *Energy Built Environ.* 1 (2) (2020) 149–164, <https://doi.org/10.1016/j.enbenv.2019.11.003>.
- [6] A.G. Devocioğlu, B. Bilici, V. Oruç, The evaluation and improvement for the energy performance of buildings: a case study, 2024/07/01/, *Energy* 4 (2024) 100126, <https://doi.org/10.1016/j.nxener.2024.100126>.
- [7] K. He, et al., Predictive control optimization of chiller plants based on deep reinforcement learning, 2023/10/01/, *J. Build. Eng.* 76 (2023) 107158, <https://doi.org/10.1016/j.jobte.2023.107158>.
- [8] W.T. Ho, F.W. Yu, Predicting chiller system performance using ARIMA-regression models, 2021/01/01/, *J. Build. Eng.* 33 (2021) 101871, <https://doi.org/10.1016/j.jobte.2020.101871>.
- [9] S.C. Sousa Alcântara, et al., Development of a method for predicting the transient behavior of an absorption chiller using artificial intelligence methods, 2023/08/01/, *Appl. Therm. Eng.* 231 (2023) 120978, <https://doi.org/10.1016/j.applthermaleng.2023.120978>.
- [10] F.-W. Yu, W.-T. Ho, C.-F. Jeff Wong, Improved energy management of chiller system with AI-based regression, 2024/01/01/, *Appl. Soft Comput.* 150 (2024) 111091, <https://doi.org/10.1016/j.asoc.2023.111091>.
- [11] L. Frison, S. Gölzhäuser, M. Bitterling, W. Kramer, Evaluating different artificial neural network forecasting approaches for optimizing district heating network operation, 2024/10/30/, *Energy* 307 (2024) 132745, <https://doi.org/10.1016/j.energy.2024.132745>.
- [12] S. Sinha, Y.M. Lee, Challenges with developing and deploying AI models and applications in industrial systems, 2024/08/16, *Discov. Artif. Intell.* 4 (1) (2024) 55, <https://doi.org/10.1007/s44163-024-00151-2>.
- [13] M. Elahi, S.O. Afolaranmi, J.L. Martinez Lastra, J.A. Perez Garcia, A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipment, 2023/12/07, *Discov. Artif. Intell.* 3 (1) (2023) 43, <https://doi.org/10.1007/s44163-023-00089-x>.
- [14] M.H. Sulaiman, Z. Mustafa, M.S. Saealal, M.M. Saari, A.Z. Ahmad, Utilizing the Kolmogorov-Arnold Networks for chiller energy consumption prediction in commercial building, 2024/11/01/, *J. Build. Eng.* 96 (2024) 110475, <https://doi.org/10.1016/j.jobte.2024.110475>.
- [15] Z. Chen, J. Zhang, F. Xiao, H. Madsen, K. Xu, Probabilistic machine learning for enhanced chiller sequencing: a risk-based control strategy, 2024/03/07/, *Energy Built Environ.* (2024) 1–13, <https://doi.org/10.1016/j.enbenv.2024.03.003>.
- [16] Y. Chen, C. Lin, J. Liu, D. Yu, One-hour-ahead solar irradiance forecast based on real-time K-means++ clustering on the input side and CNN-LSTM, 2024/12/18/, *J. Atmos. Sol. Terr. Phys.* (2024) 106405, <https://doi.org/10.1016/j.jastp.2024.106405>.
- [17] Q. Li, G. Wang, X. Wu, Z. Gao, B. Dan, Arctic short-term wind speed forecasting based on CNN-LSTM model with CEEMDAN, 2024/07/15/, *Energy* 299 (2024) 131448, <https://doi.org/10.1016/j.energy.2024.131448>.
- [18] M. Abdoos, H. Rashidi, P. Esmaili, H. Yousefi, M.H. Jahangir, Forecasting solar energy generation in the mediterranean region up to 2030–2050 using convolutional neural networks (CNN), 2025/06/01/, *Clean. Energy Syst.* 10 (2025) 100167, <https://doi.org/10.1016/j.cles.2024.100167>.
- [19] M.A.A. Al-qaness, A.A. Ewees, A.O. Aseeri, M. Abd Elaziz, Wind power forecasting using optimized LSTM by attraction–repulsion optimization algorithm, 2024/12/01/, *Ain Shams Eng. J.* 15 (12) (2024) 103150, <https://doi.org/10.1016/j.asej.2024.103150>.
- [20] P. Fan, D. Wang, W. Wang, X. Zhang, Y. Sun, A novel multi-energy load forecasting method based on building flexibility feature recognition technology and multi-task learning model integrating LSTM, 2024/11/01/, *Energy* 308 (2024) 132976, <https://doi.org/10.1016/j.energy.2024.132976>.
- [21] G. Li, et al., Interpretation of convolutional neural network-based building HVAC fault diagnosis model using improved layer-wise relevance propagation, 2023/05/01/, *Energy Build.* 286 (2023) 112949, <https://doi.org/10.1016/j.enbuild.2023.112949>.
- [22] J. Hou, Y. Wang, J. Zhou, Q. Tian, Prediction of hourly air temperature based on CNN-LSTM, 2022/12/31, *Geomat. Nat. Hazards Risk* 13 (1) (2022) 1962–1986, <https://doi.org/10.1080/19475705.2022.2102942>.
- [23] J. Cai, H. Yang, C. Song, K. Xu, A novel graph convolutional network-based interpretable method for chiller energy consumption prediction considering the spatiotemporal coupling between variables, 2024/12/15/, *Energy* 312 (2024) 133639, <https://doi.org/10.1016/j.energy.2024.133639>.
- [24] Y. Wang, H. Chen, H. Chen, M. Ye, Y. Ren, C. Yang, A hybrid model based on wavelet decomposition and LSTM for short-term energy consumption prediction of chillers, 2024/12/07/, *J. Build. Eng.* (2024) 111539, <https://doi.org/10.1016/j.jobte.2024.111539>.
- [25] Q. Guo, Z. He, Z. Wang, Prediction of monthly average and extreme atmospheric temperatures in Zhengzhou based on artificial neural network and deep learning models (in English), *Front. For. Glob. Change Orig. Res.* 6 (2023) 1–16, <https://doi.org/10.3389/ffgc.2023.1249300>.
- [26] Q. Guo, Z. He, Z. Wang, Assessing the effectiveness of long short-term memory and artificial neural network in predicting daily ozone concentrations in Liaocheng City, 2025/02/25, *Sci. Rep.* 15 (1) (2025) 6798, <https://doi.org/10.1038/s41598-025-91329-w>.
- [27] Q. Guo, Z. He, Z. Wang, Monthly climate prediction using deep convolutional neural network and long short-term memory, 2024/07/31, *Sci. Rep.* 14 (1) (2024) 17748, <https://doi.org/10.1038/s41598-024-68906-6>.
- [28] M.A.K. Raiaian, et al., A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks, 2024/06/01/, *Decis. Anal. J.* 11 (2024) 100470, <https://doi.org/10.1016/j.dajour.2024.100470>.
- [29] Q. Guo, Z. He, Z. Wang, S. Qiao, J. Zhu, J. Chen, A performance comparison study on climate prediction in weifang city using different deep learning models, *Water* 16 (19) (2024) 2870 <https://www.mdpi.com/2073-4441/16/19/2870>.
- [30] Z. He, Q. Guo, Comparative analysis of multiple deep learning models for forecasting monthly ambient PM2.5 concentrations: a case study in Dezhou City, China, *Atmosphere*, vol. 15, no. 12, p. 1432, 2024. [Online]. Available: <https://www.mdpi.com/2073-4433/15/12/1432>.
- [31] Z. He, Q. Guo, Z. Wang, X. Li, A hybrid wavelet-based deep learning model for accurate prediction of daily surface PM2.5 concentrations in Guangzhou city, *Toxics* 13 (4) (2025) 254 <https://www.mdpi.com/2305-6304/13/4/254>.
- [32] S. Selvarajan, A comprehensive study on modern optimization techniques for engineering applications, 2024/07/04, *Artif. Intell. Rev.* 57 (8) (2024) 194, <https://doi.org/10.1007/s10462-024-10829-9>.
- [33] K.A. Alruwaitee, Predicting stock price movements with combined deep learning models and two-tier metaheuristic optimization algorithm, 2024/12/01/, *J. Radiat. Res. Appl. Sci.* 17 (4) (2024) 101172, <https://doi.org/10.1016/j.jrras.2024.101172>.
- [34] J. Park, J. Cho, K.-Y. Jung, Nature-inspired metaheuristic optimization algorithms for FDTD dispersion modeling, 2024/12/01/, *AEU - Int. J. Electron. Commun.* 187 (2024) 155564, <https://doi.org/10.1016/j.aue.2024.155564>.
- [35] M.H. Sulaiman, Z. Mustafa, Chiller energy prediction in commercial building: a metaheuristic-Enhanced deep learning approach, 2024/06/15/, *Energy* 297 (2024) 131159, <https://doi.org/10.1016/j.energy.2024.131159>.
- [36] H. Dasi, Z. Ying, M.D.F.B. Ashab, Proposing hybrid prediction approaches with the integration of machine learning models and metaheuristic algorithms to forecast the cooling and heating load of buildings, 2024/03/15/, *Energy* 291 (2024) 130297, <https://doi.org/10.1016/j.energy.2024.130297>.
- [37] A. Abdelaziz, V. Santos, M.S. Dias, A.N. Mahmoud, A hybrid model of self-organizing map and deep learning with genetic algorithm for managing energy consumption in public buildings, 2024/01/01/, *J. Clean. Prod.* 434 (2024) 140040, <https://doi.org/10.1016/j.jclepro.2023.140040>.

- [38] M.H. Sulaiman, Z. Mustafa, M.M. Saari, H. Daniyal, Barnacles mating optimizer: a new bio-inspired algorithm for solving engineering optimization problems, 2020/01/01/, *Eng. Appl. Artif. Intell.* 87 (2020) 103330, <https://doi.org/10.1016/j.engappai.2019.103330>.
- [39] M.H. Sulaiman, Z. Mustafa, M.M. Saari, H. Daniyal, I. Musirin, M.R. Daud. Barnacles mating optimizer: an evolutionary algorithm for solving optimization, in: 2018 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), 20-20 Oct. 2018 2018, pp. 99-104, <http://dx.doi.org/10.1109/I2CACIS.2018.8603703>.
- [40] P. Rajesh, F.H. Shajin, N. Vijaya Anand, An efficient estimation model for induction motor using BMO-RBFNN technique, 2021/06/18, *Process Integr. Optim. Sustain.* (2021) 777–792, <https://doi.org/10.1007/s41660-021-00177-4>.
- [41] Z. Mustafa, M.H. Sulaiman, Enhancing battery state of charge estimation through hybrid integration of barnacles mating optimizer with deep learning, 2023/12/01/, *Frankl. Open* 5 (2023) 100053, <https://doi.org/10.1016/j.fraope.2023.100053>.
- [42] M.H. Sulaiman, Z. Mustafa, Optimal chiller loading solution for energy conservation using Barnacles Mating Optimizer algorithm, 2022/03/18/, *Results Control Optim.* (2022) 100109, <https://doi.org/10.1016/j.rico.2022.100109>.
- [43] D. Borda, M. Bergagio, M. Amerio, M.C. Masoero, R. Borchiellini, D. Papurello, Development of anomaly detectors for HVAC systems using machine learning, *Processes* 11 (2) (2023) 535 <<https://www.mdpi.com/2227-9717/11/2/535>>.
- [44] J. Bian, J. Wang, Q. Yece, A novel study on power consumption of an HVAC system using CatBoost and AdaBoost algorithms combined with the metaheuristic algorithms, 2024/09/01/, *Energy* 302 (2024) 131841, <https://doi.org/10.1016/j.energy.2024.131841>.
- [45] M.H. Sulaiman, Barnacles mating optimizer: a bio-inspired algorithm for solving optimization problems, in: 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 27-29 June 2018 2018, pp. 265-270, <http://dx.doi.org/10.1109/SNPD.2018.8441097>.
- [46] J. Lachance, "Hardy-Weinberg Equilibrium and Random Mating, in: R.M. Kliman (Ed.), *in Encyclopedia of Evolutionary Biology*, Academic Press, Oxford, 2016, pp. 208–211.
- [47] G. Gerlach, J. Atema, The Use of Chemical Cues in Habitat Recognition and Settlement, in: C. Brönmark, L.-A. Hansson (Eds.), *in Chemical Ecology in Aquatic Systems*, Oxford University Press, United Kingdom, 2012, p. 0 (sec. Oxford Academic).