Enhancing Residential Building Heating Load Prediction With Hybrid MLP Models

Yizhe Guan

School of Architecture and Urban Planing, Yunnan University, Kunming 650000, Yunnan, China

Corresponding author. E-mail: guanyizhe2002@163.com

Received: February 15, 2024; Accepted: May 26, 2024

Efficiently managing the heating load (HL) of residential buildings is a critical aspect of energy conservation and sustainability. This study highlights a novel approach for HL prediction employing Multi-layer Perceptron (MLP) neural networks in conjunction with two innovative optimizers: Flying Fox Optimization (FFO) and Horse Herd Optimizers (HHO). The MLP is a powerful machine learning algorithm known for its ability to capture complex relationships within data. In this study, it is utilized as the core predictive model. FFO and HHO, both inspired by nature, provide optimization techniques that complement the MLP's capabilities. The integration of these optimizers with the MLP model leads to a hybrid approach that leverages the strengths of each component. FFO and HHO are employed to fine-tune the MLP's weights and biases, enhancing its predictive accuracy. The combination of MLP, FFO, and HHO outperforms traditional methods and even standalone MLP models in terms of predictive accuracy and convergence speed. This method improves HL prediction in homes aids energy management, and cuts environmental impact effectively. The hybrid MLFF2 model stands out by showcasing superior accuracy compared to other proposed models with an impressive R^2 value of 0.997 and a remarkably low RMSE of 0.523, MLFF2 has achieved the highest level of performance. Moreover, this research opens doors to exploring the application of nature-inspired optimization techniques in conjunction with neural networks for various other complex problems. The synergistic effect of the MLP and these innovative optimizers showcases the potential of hybrid models in addressing real-world challenges.

Keywords: Heating load; Multi-layer Perceptron; Horse Herd Optimizers; Flying Fox Optimization

http://dx.doi.org/10.6180/jase.202505_28(5).0007

1. Introduction

Lately, there has been a growing emphasis on research endeavors aimed at enhancing the energy efficiency of buildings. This heightened focus could be ascribed to the rising worries concerning energy wastage and its enduring adverse impacts on the ecosystem [1]. Scholars have energetically investigated techniques for improving building efficacy and reducing their environmental footprint, recognizing the substantial influence of buildings on energy consumption and the release of greenhouse gases [2]. Achieving energy conservation in buildings necessitates the development of a range of strategies for efficient energy governance [3]. An essential element in formulating these strategies involves the precise prediction of energy usage, a subject that has gained notable prominence in recent times [4]. Through vigilant tracking of variations in construction power usage, it turns feasible to craft focused and efficient saving power initiatives [5, 6]. They identify prospective chances for conserving energy and rectifying operational inefficiencies, consequently benefiting equally power preservation and the optimal operation of structure facilities [7]. These methods employ predictive data to enhance energy efficiency, reduce consumption, and con-

firm the effective functioning of structure mechanisms [8]. Studies have revealed that equally slight improvements in forecasting structure power ingesting lead to significant decreases in the usage of energy. Precise predictions of energy consumption patterns empower the capacity to make informed choices and proactively manage energy consumption for optimization [9]. These actions may encompass a range of tactics, including regulating settings of heating, ventilation, and air conditioning (HVAC), enhancing illumination agendas, installing efficient energy tackle, and promoting behavioral adjustments to line up with power-conservation objectives. These measures yield advantages not only for building administrators but also for occupants [10, 11].

Accurately predicting the energy consumption of buildings is a fundamental component of energy modeling, yet it frequently fails to replicate real-world outcomes [12]. Several research investigations have emphasized the substantial disparity between projected and actual energy usage, at times exceeding forecasts by a factor of two or even three. Conventional power replicas depend on technical computations grounded in actual principles to gauge structure power utilization, which is appropriate for initial evaluations [13]. However, they possess constraints in precisely representing the difficulties of the real world. To surmount these restrictions, arithmetical imitation methods are utilized for modeling power consumption in constructions. These emulations aid in addressing the hurdles connected within the integration of ML samples into structure power effectiveness through methodically investigative current explore outcomes then constraints [14, 15].

Models of artificial intelligence (AI) offer substantial potential in forecasting and enhancing construction power usage. These representations utilize past information, present device efforts, and artificial intelligence procedures to produce precise forecasts and offer visions aimed at effective management of energy [16–18]. Substantial progress has been completed by the power's domain ingesting forecast over the years. Investigators and experts have devised numerous methods and strategies for consistently predicting power consumption [19, 20].

A neural network was introduced by Kim and Cho [21], which combined CNN and LSTM architectures and was customized for the prediction of household power usage by capturing both temporal and spatial features. The competencies of the LSTM and CNN systems were efficiently utilized by this system to adeptly detention intricate consumption of energy forms. The experimental results indicated the exceptional accuracy of the CNN-LSTM method, especially in the realm of electrical consumption of energy, outperforming conservative prediction methods. Roy et al. [22] presented a tailored Deep Neural Network (DNN) model for forecasting cooling and heating loads in residential buildings. Following this introduction, a comparative evaluation was performed, pitting the DNN model against the (MPMR), (GBM), and (GPR). The outcomes revealed both the GPR and DNN models demonstrated the highest alteration reported intended for (VAF) in forecasting both cooling and heating loads. Moradzadehand the partners [23] utilized (MLP) and (SVR) models for the anticipation of Heating and Cooling Loads. The method of MLP delivered outstanding results, securing the uppermost R-amount of 0.9993 for predicting the Heating Load (HL). Conversely, in the forecast of Cooling Load, the Support vector regression approach demonstrated excellence with the uppermost R-amount of 0.9878. A summary of the different approaches used to estimate energy consumption in the reviewed studies can be found in Table 1.

In this study, an innovative machine learning approach is introduced aimed at achieving optimal and accurate prediction results. The hybridization technique used in this study was carefully designed to improve the performance of the MLP model and ensure that it produced reliable results. By combining two efficient and advanced optimization techniques, the development of these innovative hybrid models goes beyond traditional methods and examines both single and hybrid configurations to fairly evaluate their capabilities. A comprehensive evaluation of these models was conducted. To ensure result reliability, the evaluation of model performance incorporated established metrics like R^2 and RMSE. Moreover, the intentional choice to utilize two distinct optimizers, specifically Flying Fox Optimization (FFO) and Horse Herd Optimizers (HHO), in constructing the hybrid models was driven by the desire to harness the distinctive capabilities of each optimizer, ultimately aiming for improved performance. In general, the study presents an innovative machine-learning approach aimed at precise predictive outcomes. By combining FFO and HHO, hybrid models surpass conventional methods. The careful design of this hybridization technique ensures improved performance in MLP models. The study's contribution lies in advancing predictive modeling through novel methodologies, offering dependable outcomes.

The combination of MLP neural networks with FFO and HHO results in several significant advantages. Firstly, it leads to enhanced predictive accuracy, crucial for efficiently managing the HL of residential buildings. This improvement is pivotal for aiding energy conservation and sustainability efforts. Accurate HL prediction facilitated by the hybrid model enables more effective energy management in

Table 1. A comprehensive review of energy consumption prediction research.

Article	Methodology
[24]	A decision-based model was used to predict building energy consumption and a neural network model was used to evaluate it. Addresses the challenge of predicting the timing of structural power consumption.
[25]	Used SVR for temporal sequence estimation in structure power usage.
[26]	Used extreme ML methods for energy consumption estimation.
[27]	Implemented clustering methods in energy usage estimation.
[28]	Used multivariate autoregressive model for energy consumption estimation.
[29]	Used layered self-encoding structures with intense learning apparatus for structure power usage estimation.
[30]	Utilized gradient boosting regression for dataset withdrawal in structure power usage estimation.
[31]	Used LSTM models for energy consumption estimation.
[32]	Used Gray system models for energy usage estimation.
[33]	Used chaotic timing predicting models for energy consumption estimation.

residential buildings, potentially leading to energy savings and reduced environmental impact. The mechanisms underlying this combination are noteworthy. FFO and HHO optimize the weights and biases of the MLP model, enhancing its predictive accuracy by complementing its ability to capture complex data relationships. Furthermore, this research sets the stage for exploring the broader application of nature-inspired optimization techniques in conjunction with neural networks, offering promising avenues for addressing diverse complex problems beyond HL prediction.

In the following sections, this study provides an in-depth explanation of the dataset in the Materials section (Section 2), followed by an introduction to the model and optimizers. Section 3 focuses on the results of the developed models, evaluating them with specified metrics. Furthermore, the conclusion summarizes the findings and suggests future research directions.

2. Materials and methodology

2.1. Materials

To streamline the training of machine learning models, a dataset is assembled using an extensive body of published literature. Table 2 presents statistical characteristics for the input variables. This dataset encompasses the following parameters:

- Relative Compactness: This parameter quantifies the overall building compactness, directly influencing its heat retention and loss characteristics.
- Surface Area (m^2): The precise measurement of the structure's total surface area is crucial for understanding its exposure to external thermal influences.
- Wall Area (m^2): Accurate recording of the wall area provides insights into the building's thermal insulation properties and its heat transfer capabilities.
- Roof Area (m^2): Including roof area in the dataset is significant due to its role in the building's thermal equilibrium, as a source of heat gain or loss.
- Overall Height (m): The building's vertical dimension is directly linked to its internal volume and, consequently, its HL requirements.
- Orientation: Knowledge of the building's orientation is essential for assessing solar exposure and its impact on heating demands.
- Glazing Area (%): Thoroughly documenting the percentage of glazing or fenestration area within the building envelope is critical for understanding heat transfer dynamics through windows.
- Glazing Area Distribution: The spatial allocation of glazing across different facades of the building plays a pivotal role in shaping interior heat distribution patterns, necessitating meticulous documentation.

This dataset, based on these parameters, facilitates the development of machine learning models for more efficient and accurate predictions in the context of building HL analysis.

Table 3 displays a correlation table that highlights the relationships between the input variables and the output variable (Heating). The table provides a visual representation of the degree of correlation between these variables, with correlation coefficients ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation), and 0 indicating no correlation. RCE (Relative Compactness) exhibits a moderately strong negative correlation (-0.821) with Heating, indicating that as RCE increases, Heating tends to decrease. SA (Surface Area) has a strong positive correlation (0.875) with Heating, suggesting that as SA increases, Heating also tends to increase. OVH (Overall Height) demonstrates a strong negative correlation (-0.819) with Heating, signifying that as OVH increases, Heating

Table 2. Statistic properties of the input variable of heating.

Indicator	Components								
	RCE	*SA*	*WA*	*RA*	OVH	*OR*	*GA*	*GAD*	*HL*
Max.	0.98	808.5	416.5	220.5	7	5	0.4	5	42.96
Min.	0.62	514.5	245	110.25	3.5	2	0	0	6.01
Median	0.75	673.75	318.5	183.75	5.25	3.5	0.25	3	18.95
Avg.	0.764	671.708	318.5	176.604	5.25	3.5	0.234	2.813	22.307
Skew.	0.496	-0.125	0.534	-0.163	$-2.9E-19$	$2.68E-18$	-0.060	-0.089	0.361
St. Dev.	0.106	88.086	43.626	45.166	1.751	1.119	0.133	1.551	10.090

typically decreases. Other variables like WA (Wall Area), RA (Roof Area), OR (Orientation), GA (Glazing Area), and GAD (Glazing Area Distribution) exhibit varying degrees of correlation with Heating, though some of these correlations are relatively weak.

This correlation plot provides valuable insights into how each input variable is related to the Heating output. Researchers and practitioners can use this information to understand which features have a more pronounced impact on the HL of a building, aiding in the development of predictive models and energy-efficient building designs.

2.2. Multi-layer Perceptron (MLP)

A concise overview of the single-layer and one-neuron perceptron is presented to foster comprehension of the MLP [34]. The elementary variant of a neural network, the single-neuron perceptron, comprises a solitary output linked to all inputs. In this neural structure, the weights are symbolized as v_i, with i spanning from 0 to n, where n signifies the overall input count. The outcome y signifies the predictive binary class, whereas the a_i inputs correspond to characteristics or factors [35]. During the evaluation phase, every individual input attribute value gets scaled by its associated coefficient $v_i a_i$. The resulting outcomes are subsequently combined in the following phase $(v_0 a_0 + v_1 a_1 + \ldots + v_n a_n)$. The subsequent phase involves a transformation operation, during which the cumulative value experiences an activation function g (sometimes called a transformation function), leading to the generation of an outcome z, which can be represented as:

$$y = f(z) \text{ and } z = \sum_{i=0}^{n} v_i a_i \tag{1}$$

The result denoted as y is expressed as $a_0 = 1$, where a_0 symbolizes the threshold or offset.

Eq. (1) empowers a perceptron to acquire linearly separable functions exclusively. This function is visualized as a straight line involving two-dimensional input and two characteristics [36]. The representation becomes flat when extending to three dimensions with three attributes. Nevertheless, when the input spans n dimensions, the function is portrayed as a hyperplane. The expression for the hyperplane is articulated as follows:

$$\sum_{i=0}^{n} v_i a_i = 0 \tag{2}$$

The depiction of Eq. (2) is the scalar product involving the weight vector V and the input vector A.

$$V.A = 0 \tag{3}$$

The training or instructional phase refines neural network weights using gradient descent to minimize a cost metric, often a squared error. Convergence is crucial for network functionality. Validation assesses adaptability to new data. Single-layer Perceptrons are used for multiclass scenarios, as illustrated in Fig. 1(A) [37, 38].

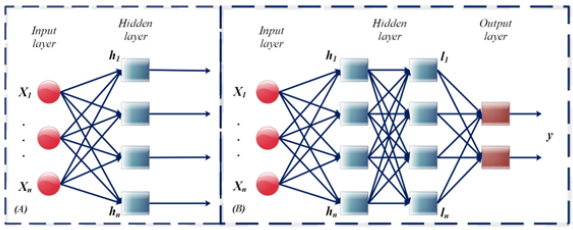

Fig. 1. Displays the layer organization: A) SLP with three inputs and four outputs, and B) MLP with two concealed layers, three inputs, and two outputs.

To address perceptual issues lacking SLP linear separability, the single-layer perceptron's limitations are overcome by constructing an MLP with multiple sequential layers, as shown in Fig. 1(B). The initial and terminal layers are input and output, while intermediate layers are concealed. Data flows unidirectionally through the MLP, each layer using weights and activation functions. The number of hidden layers, units, training data, and hyperparameters like learning rate and momentum significantly impact MLP model performance.

Table 3. The Correlation plot for input and output.

	RCE	SA	WA	RA	OVH	OR	GA	GAD	HL
RCE	1.000								
SA	-0.821	1.000							
WA	-0.109	0.291	1.000						
RA	-0.787	0.875	-0.208	1.000					
OVH	0.823	-0.748	0.308	-0.921	1.000				
OR	0.032	0.051	0.050	0.027	0.021	1.000			
GA	0.019	0.033	0.032	0.019	0.014	0.016	1.000		
GAD	0.019	0.030	0.030	0.017	0.013	0.015	0.223	1.000	
Heating	0.622	-0.574	0.462	-0.819	0.890	0.016	0.280	0.100	1.000

2.3. Flying Foxes Optimization (FFO)

The population-based stochastic algorithm known as various techniques inform FFO, and it is employed by flying foxes to cope with elevated temperatures [39]. A hybrid algorithm structure is utilized by this method, incorporating operators from pre-existing algorithms [40, 41]. The performance of FFO is significantly influenced by the selection of the population size (N), the replacement list (RL), and the attraction constant (b). The steps elaborate in the FFO procedure are delineated as follows.

2.3.1. *Functioning of Flying Foxes Algorithm*

Flying foxes, giant bats without echolocation, return to tree habitats after meals, seeking more astounding trees; overcrowding often results in fatalities [42].

2.4. The Application of FFO Algorithm

An innovative algorithm starts by choosing initial positions based on the movements of flying foxes, represented as a vector (x) in a multi-dimensional space. Then, an objective function evaluates proposed solutions at these positions. The objective is for the flying foxes to find a tree with a lower temperature to survive in extreme heat conditions.

2.4.1. *Movement of Flying Foxes*

Given that flying foxes either trail the paths of their counterparts or seek out the nearest available tree, it is reasonable to infer that when the habitat tree fails to offer the requisite minimal temperature comfort to the flying foxes, they will migrate to an alternative location to evade the overwhelming heat. The description of this relocation can be depicted through the following mathematical expression:

$$x_{i,j}^{t+1} = x_{i,j}^{t} + \text{arand}\left(\text{cool}_j - x_{i,j}^{t}\right) \tag{4}$$

With $x_i^0 \sim U\left(x_{\min}, x_{\max}\right)$, $x_{i,j}^t$ represents the $j-$ th element within $FF(i)$ at iteration t, where a remains a constant, rand $\sim U(0,1)$, and cool signifies the position of the FF situated in the tree with the lowest temperature. Eq. (4) comes into play when $|f(\text{cool}) - f\left(x_i\right)| > \frac{\delta_1}{2}$, Whereby cool designates the vector representing the flying fox positioned at the most optimal location ever identified, signifying the finest solution on record, while the parameter δ_1 corresponds to the maximum conceivable distance between two flying foxes to be considered close. Upon nearing a tree with the most favorable temperature $\left| f(\text{ cool }) - \leq f\left(x_i\right) \right| \frac{\delta_1}{2}$, the flying fox initiates a quest for the closest available space to prevent overcrowding. Subsequent equations offer a more profound elucidation of this occurrence:

$$x_{i,j}^{t+1} = \left\{ \begin{array}{c} nx_{i,j}^{t+1}, \text{ if } j = k \text{ or } r_{nd} \geq pa \\ x_{i,j}^{t}, \text{ otherwise} \end{array} \right\} \tag{5}$$

In this context, where rand $\sim U(0,1)$ signifies random sampling, rnd_j is an arbitrary value ranging from 0 to 1 , and x_{R1}^t and $\mathrm{d}x_{R2}^t$ represent two unspecified individuals within the present population. Additionally, pa denotes a constant probability. Ultimately, k, randomly chosen from the set $\{1, 2, \ldots, m\}$, is a pivotal factor, ensuring that at least one element from $nx_{i,j}^{t+1}$ is selected by $x_{i,j}^{t+1}$, thus preserving the original solution's distinctiveness from the new one. An evaluation of the computed outcomes is conducted.

2.4.2. *Death and Replacement Flying Foxes*

Numerous factors affect flying fox survival, including extreme heat and distance from preferred trees. Replacement List offers optimal solutions.

$$x_{i,j}^{t+1} = \frac{\sum_{k=1}^{n} RL_{k,j}^{t}}{n} \tag{6}$$

In the context of the tth iteration, RL_k^t represents the $k - th(FF)$ on the (RL). Eq. (6) is designed to enhance the likelihood of identifying a suitable area to the maximum extent.

Crowding can lead to flying fox fatalities. Before each iteration, a probability is computed based on foxes in cold regions.

$$p^D = \frac{nc - 1}{\text{population size}} \tag{7}$$

The quantity nc is strongly correlated with the count of Flying Felines (FF) having a congruent objective function as the ideal solution.

2.4.3. *Crossover Process*

Genetic recombination unites two flying foxes. Two diverse parents are chosen randomly, producing two descendants in a subsequent operation.

$$\begin{aligned} \text{offspring 1} &= L \cdot R_1 + (1-1) \cdot R_2 \\ \text{offspring 2} &= L \cdot R_2 + (1-L) \cdot R_2 \end{aligned} \tag{8}$$

R_1 and R_2 represent unique individuals drawn accidentally from the community, while L is an arbitrary value generated at arbitrary, ranging from 0 to 1 .

2.5. Horse herd optimization algorithm

Presented by HOA [43], derived from the documented conduct of equines within their native habitat, it is molded explicitly by six essential behavioral traits: supremacy, communal engagement, grazing, imitation, safeguarding reactions, and locomotion. The Homeowners Association (HOA) is founded on these conducts, guiding the mobility of horses in each rotation, as explained in the Eq. (9):

$$X_m^{\text{Iter},\,AGE} = \vec{V}_m^{\text{Iter},AGE} + X_m^{(Iter-1),AGE}, AGE = \alpha, \beta, \gamma, \delta \tag{9}$$

Within this circumstance, $X_m^{Iter,AGE}$ denotes the place of the $m-th$ equine, AGE signifies the equine's age classification, and Iter correlates with the current cycle. AGE additionally encompasses the age grouping of the horse, while $\vec{V}_m^{\text{Iter,AGE}}$ embodies the velocity vector of the horse. It is significant to mention that horses generally possess a life span of 25 30 years, characterized by a range of accomplishments throughout their existence. These attributes are categorized as $\delta(0-5$ years), $\gamma(5-10$ years), and α (15 years or above) classes. A complete reaction array is utilized to arrange equine ages based on their achievements. The highest tenth constitutes the α group, the subsequent fifth is allocated to the β group, and the remaining three-tenths and four-tenths are assigned to the γ and δ groups, respectively. Gesture pathways for equines at different ages and algorithmic rotations are outlined under these features.

$$\begin{aligned} \vec{V}_m^{\text{Iter},\alpha} &= \vec{G}_m^{\text{Iter},\alpha} + \vec{D}_m^{\text{Iter},\alpha} \\ \vec{V}_m^{\text{Iter},\beta} &= \vec{G}_m^{\text{Iter},\beta} + \vec{H}_m^{\text{Iter},\beta} + \vec{S}_m^{\text{Iter},\beta} + \vec{D}_m^{\text{Iter},\beta} \\ \vec{V}_m^{\text{Iter},\gamma} &= \vec{G}_m^{\text{Iter},\gamma} + \vec{H}_m^{\text{Iter},\gamma} + \vec{S}_m^{\text{Iter},\gamma} + \vec{I}_m^{\text{Iter},\gamma} + \vec{D}_m^{\text{Iter},\gamma} + \vec{R}_m^{\text{Iter},\gamma} \\ \vec{V}_m^{\text{Iter},\delta} &= \vec{G}_m^{\text{Iter},\delta} + \vec{I}_m^{\text{Iter},\delta} + \vec{R}_m^{\text{Iter},\delta} \end{aligned} \tag{10}$$

By employing Eqs. (11) and (12) to elucidate the development of the worldwide matrix, a linkage is formed between locations (X) and their corresponding measurement values ($C(X)$).

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,d} \end{bmatrix}, \quad C(X) = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \tag{11}$$

$$\text{Global Matrix} = \begin{bmatrix} X & C(X) \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \ldots & x_{1,d} & c_1 \\ x_{2,1} & x_{2,2} & \ldots & x_{2,d} & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{m,1} & x_{m,2} & \ldots & x_{m,d} & c_m \end{bmatrix} \tag{12}$$

In the equations, x represents position, $C(x)$ is the associated cost, m is the number of horses, and d is dimension. A global matrix organizes costs and horse ages. Benefits come from a higher probability of success, which favors accuracy and lower speed. When the chance of success is lower, it is better to prioritize speed over precision. The overall speed course is calculated using a specific method. The velocity of horses in the age range of 0 to 5 years:

$$\begin{aligned} \vec{V}_m^{\text{Iter},\delta} &= \left[g_m^{(\text{Iter}-1),\delta} \omega_g (\breve{u} + P\breve{l}) \left[X_m^{(Iter-1)} \right] \right] \\ &+ \left[i_m^{(\text{Iter}-1),\delta} \omega_i \left[\left(\frac{1}{P^N} \sum_{j=1}^{P^N} \hat{X}_j^{\text{Iter}-1} \right) - X^{\text{Iter}-1} \right] \right] \\ &+ \left[r_m^{(Iter-1),\delta} \omega_r P X^{\text{Iter}-1} \right] \end{aligned} \tag{13}$$

The speed of horses within the age range of 5 to 10 years.

$$\begin{aligned} \vec{V}_m^{\text{Iter},\gamma} &= \left[g_m^{(\text{Iter}-1),\gamma} \omega_g (\breve{u} + P\breve{l}) \left[X_m^{(Iter-1)} \right] \right] \\ &+ \left[h_m^{(\text{Iter}-1),\gamma} \omega_h \left[X_*^{(\text{Iter}-1)} - X_m^{(\text{Iter}-1)} \right] \right] \\ &+ \left[S_m^{(Iter-1),\gamma} \omega_S \left[\left(\frac{1}{N} \sum_{j=1}^{N} X_j^{\text{Iter}-1} \right) - X^{\text{Iter}-1} \right] \right] \\ &+ \left[i_m^{(\text{Iter}-1),\gamma} \omega_i \left[\left(\frac{1}{P^N} \sum_{j=1}^{P^N} \hat{X}_j^{\text{Iter}-1} \right) - X^{\text{Iter}-1} \right] \right] \\ &- \left[d_m^{(Iter-1),\gamma} \omega_d \left[\left(\frac{1}{q^N} \sum_{j=1}^{q^N} \hat{X}_j^{\text{Iter}-1} \right) - X^{\text{Iter}-1} \right] \right] \\ &+ \left[r_m^{(Iter-1),AGE} \omega_r P X^{\text{Iter}-1} \right] \end{aligned} \tag{14}$$

The velocity demonstrated by horses in the 10 to 15-year

age group:

$$\vec{V}_m^{\text{Iter},\beta} = \left[g_m^{(\text{Iter}-1),\beta}\omega_g(\check{u}+P\check{l})\left[X_m^{(\text{Iter}-1)}\right]\right] + \left[h_m^{(\text{Iter}-1),\beta}\omega_h\left[X_*^{(\text{Iter}-1)} - X_m^{(\text{Iter}-1)}\right]\right] + \left[S_m^{(Iter-1),\beta}\omega_S\left[\left(\frac{1}{N}\sum_{j=1}^{N} X_j^{\text{Iter}-1}\right) - X^{\text{Iter}-1}\right]\right] - \left[d_m^{(Iter-1),\beta}\omega_d\left[\left(\frac{1}{q^N}\sum_{j=1}^{q^N} \check{X}_j^{Iter-1}\right) - X^{Iter-1}\right]\right] \quad (15)$$

Horses surpassing the age of 15 showcase the following velocity:

$$\vec{V}_m^{\text{Iter},\alpha} = \left[g_m^{(\text{Iter}-1),\alpha}\omega_g(\check{u}+P\check{l})\left[X_m^{(Iter-1)}\right]\right] - \left[d_m^{(Iter-1),\alpha}\omega_d\left[\left(\frac{1}{q^N}\sum_{j=1}^{q^N} \check{X}_j^{\text{Iter}-1}\right) - X^{\text{Iter}-1}\right]\right] \quad (16)$$

2.6. Performance Evaluators

In this section, a collection of metrics has been devised for assessing the developed models. The metrics measure both correlation and error, presenting valuable insights into the models' performance. The formulation for the metrics utilized in this study can be found in Table 4 [44].

As an alternative representation, the variables can be described as follows:

- The measured value is indicated by m_i.
- The means of the predicted and measured values are represented as $\bar{b}$ and $\bar{m}$, correspondingly.
- Predicted values are presented as b_i.
- The sample size is denoted by the n.
- The mean of the predicted values is determined as $\bar{x}$.

3. Result and discussion

3.1. Models' evaluation

Table 5 presents the performance results of the developed models for Multi-layer Perceptron (MLP) in various phases: Training, Validation, Test, and All (combining all phases). The table displays several index values, including RMSE, R^2, MSE, SI, and RAE. These metrics are essential in assessing the accuracy and effectiveness of the models.

MLFF (1) and MLHH (1): In the first set of models, labeled as (1), it is observed that the MLFF model is outperformed by the MLHH model in all phases—Training, Validation, Test, and the combined results. The RMSE values for MLHH are consistently higher than those of MLFF, indicating a higher level of predictive accuracy. Furthermore, the R^2 values for MLHH are notably lower, signifying a less ideal fit to the data. This suggests that, in this context, the hybrid MLP model with Horse Herd Optimization (MLHH) exhibits superior accuracy in predicting HL.

MLFF (2) and MLHH (2): In the second set of models, labeled as (2), there is a noticeable improvement in performance for both MLFF and MLHH. RMSE values decrease significantly, with MLFF achieving the lowest RMSE in the Validation and Test phases. R^2 values are remarkably high, indicating a strong fit to the data. The performance gap between MLFF and MLHH narrows in this set, with MLFF still maintaining a slight advantage.

MLFF (3) and MLHH (3): In the third set of models, labeled as (3), MLFF continues to show strong performance with low RMSE and high R^2 values. However, it's essential to note that in this set, the difference in performance between MLFF and MLHH is less pronounced, particularly in the Test phase. This suggests that the hybrid MLP model with Horse Herd Optimization (MLHH) is showing improved performance in this set of models, but MLFF remains a strong competitor.

In summary, the results indicate that the hybrid MLP models, particularly MLFF, demonstrate significant promise in accurately predicting HL. While MLFF consistently outperforms MLHH in most scenarios, the performance may vary depending on the specific model configuration. These findings provide valuable insights for optimizing MLP models in the context of building HL analysis and emphasize the effectiveness of hybrid optimization techniques in improving predictive accuracy. Further investigations may be needed to fine-tune these models for specific applications and data sets.

In Fig. 2, a scatter plot is provided, offering a visual representation of the performance of the proposed models within the three layers of the Multi-layer Perceptron (MLP). The distribution of predicted data points in relation to the central line, representing the optimal outcome, is highlighted, thereby emphasizing the accuracy of these models. It is evident that a tight clustering of predicted data points around the central lines is consistently demonstrated by MLFF across all layers. Notably, in Layer 2, the predicted data points from MLFF show an exceptionally close alignment with the best-fit distribution, indicating a high level of precision. Conversely, the predicted data points from MLHH exhibit a broader dispersion when compared to those of MLFF. This observation reinforces the superior precision and reliability displayed by MLFF in the prediction of the HL across the various layers of the hybrid

Table 4. Equations of the Performance Evaluators.

Coefficient Correlation (R^2):	$R^2 = \left(\frac{\sum_{i=1}^{n}(b_i-\bar{b})(m_i-\bar{m})}{\sqrt{\left[\sum_{i=1}^{n}(b_i-\bar{b})^2\right]\left[\sum_{i=1}^{n}(m_i-\bar{m})^2\right]}} \right)^2$	(17)
Root Mean Square Error (RMSE):	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i-b_i)^2}$	(18)
Mean Square Error (MSE):	$\mathrm{MSE} = \frac{1}{n}\sum_{j=1}^{n}(m_i-b_i)^2$	(19)
Scatter Index (SI):	$SI = \frac{RMSE}{\mathrm{mean}(m_i)}$	(20)
Relative Absolute Error (RAE):	$RAE = \sum_{j=1}^{n}\frac{b_i}{\lvert m_i-\bar{m}\rvert}$	(21)

Table 5. Equations of the Performance Evaluators.

Model	Phase	Index values				
		RMSE	R^2	MSE	SI	RAE
MLFF (1)	Train	1.676	0.973	1.194	0.00010	0.973
	Validation	1.350	0.983	0.872	0.00035	0.981
	Test	1.409	0.980	1.110	0.00042	0.980
	All	1.593	0.975	1.133	0.00007	0.975
MLHH (1)	Train	2.014	0.962	1.304	0.00011	0.961
	Validation	1.385	0.981	0.985	0.00040	0.980
	Test	1.800	0.970	1.245	0.00047	0.968
MLFF (2)	All	1.901	0.966	1.248	0.00007	0.964
	Train	1.092	0.989	0.790	6.57E − 05	0.988
	Validation	0.523	0.997	0.354	0.00014	0.997
	Test	0.450	0.998	0.348	0.00013	0.998
	All	0.952	0.991	0.659	3.84E − 05	0.991
MLFF (3) (2)	Train	1.541	0.978	1.060	8.82E − 05	0.977
	Validation	0.887	0.992	0.656	0.00027	0.992
	Test	1.332	0.983	0.930	0.00035	0.982
	All	1.431	0.980	0.980	5.72E − 05	0.980
MLHH (3)	Train	1.284	0.984	0.888	7.41E − 05	0.984
	Validation	1.062	0.989	0.720	0.00029	0.988
	Test	0.812	0.993	0.585	0.00022	0.993
	All	1.192	0.986	0.817	4.78E − 05	0.986
	Train	1.724	0.972	1.072	8.93E − 05	0.971
	Validation	1.011	0.990	0.642	0.00026	0.989
	Test	1.219	0.986	0.816	0.00031	0.985
	All	1.568	0.976	0.970	5.66E − 05	0.976

models. The scatter plot visualization is of great value for gaining insights into the distribution patterns and the accuracy of predictive outcomes generated by these models. It underscores the remarkable consistency and precision of the MLFF model across the layers of the MLP, further emphasizing its efficacy in predicting the HL, a critical aspect of this research.

In Fig. 3, a symbol column line plot is indicated to evaluate the correlation between the predicted and measured values in the prediction models. Upon analyzing Fig. 3, it becomes evident that the results of the MLFF models outperform those of the MLHH models in terms of accuracy. Notably, the performance of the MLFF model in Layer 2 is remarkable, displaying exceptional precision in aligning predicted values with measured values, even in cases of potential overfitting. It is clear that the testing phase of the MLFF model in Layer 2 yields the most accurate results compared to the other layers, indicating an enhancement in accuracy relative to its training phase. This observation underscores the effectiveness of the MLFF model in maintaining accuracy across different phases and layers, thus improving its applicability in predicting HL. The symbol column line plot in Fig. 3 visually portrays the degree of correlation between predicted and measured values, offering a clear comprehension of the models' performance. Conversely, the predicted values of the MLHH models exhibit a weaker alignment with the measured values, particularly noticeable in Layer 1. Notable discrepancies between predicted and measured values are evident between sample numbers 25 and 75.

Fig. 4 provides a visual representation of the error rate percentages in the models. In this context, the proximity of

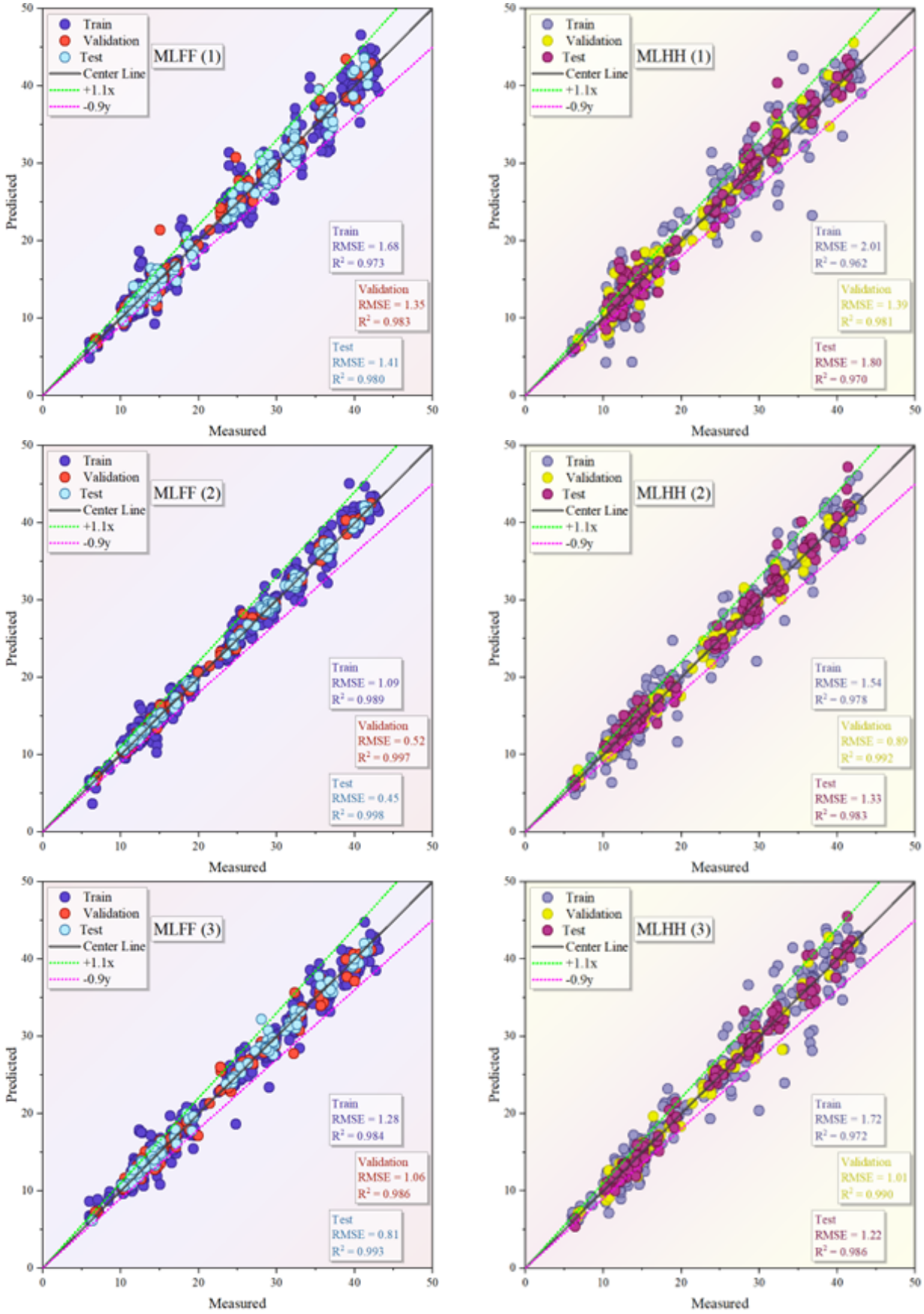

Fig. 2. The scatter plot for developed hybrid models.

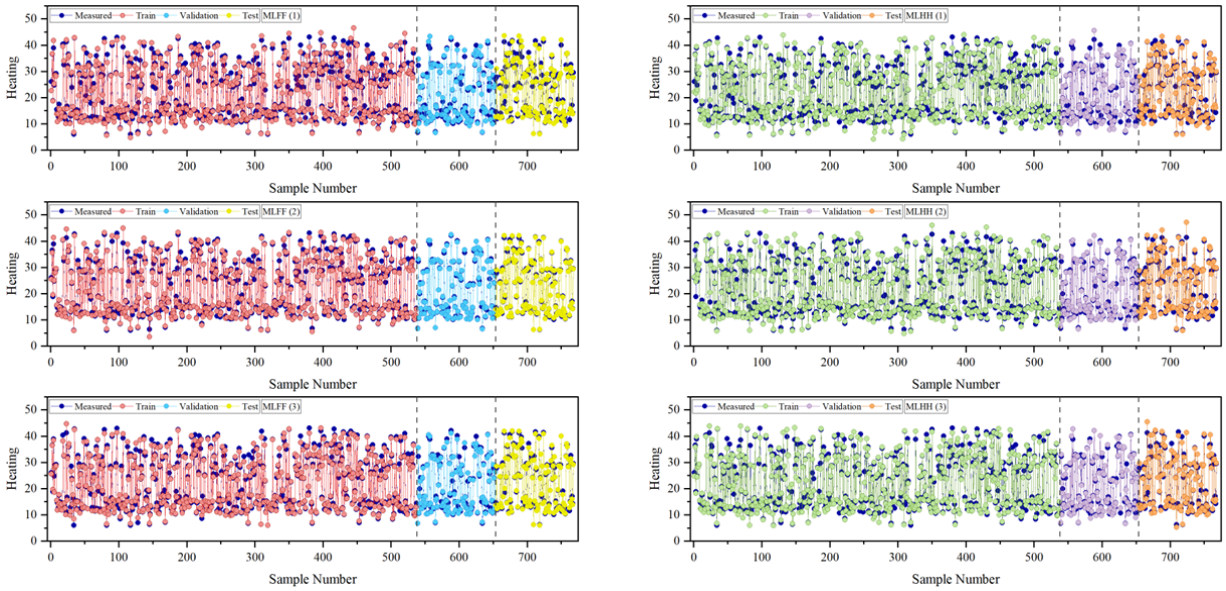

Fig. 3. Comparison of measured and predicted values.

data points to 0 percent error serves as a crucial indicator of the model's accuracy. Notably, the MLFF models exhibit tighter dispersions within the range of 10 to -10 error percentages, underscoring their accuracy. In the first layer, the disparity between the MLFF and MLHH models is relatively comparable. However, progressing to the second and third layers, the MLFF model consistently demonstrates a higher level of accuracy and reliability when compared to the MLHH model. A noteworthy highlight is the exceptional precision and accuracy displayed by the MLFF2 model during the validation phase. These results position it as the most precise and reliable among the model variations.

To perform a comprehensive analysis of performance variation among the previously discussed models, a half-violin plot comparison was employed. This comparison provides a visually intuitive means to comprehend the distribution of predicted values and their associated errors. In addition to the half-violin plots, the utilization of box plots, as illustrated in Fig. 5, offers valuable insights into key aspects of the error distributions within the models. Upon a detailed examination of Fig. 5, it becomes conspicuous that the MLFF model consistently maintains a relatively tight and compact distribution of errors across all layers. This phenomenon strongly suggests a heightened level of consistency and accuracy in its predictions. The clustering of errors close to the central tendency line in each layer underscores the model's capability to deliver reliable outcomes.

In contrast, the MLHH models exhibit a noticeably wider dispersion of errors across all layers, accentuating the inherent variability in their performance. The divergence in error distribution is quite prominent, indicating that the MLHH models, while effective in some cases, may struggle to maintain the same level of precision consistently. Moreover, the box plot comparison doesn't just convey central tendencies and dispersion; it also offers the means to identify potential outliers within the error distributions. This provides a comprehensive insight into the predictive behavior of the models, allowing researchers and practitioners to recognize situations where the models may significantly deviate from the expected performance. The combined use of half-violin plots and box plots in Fig. 5 serves as a powerful visual tool to gain a deeper understanding of the predictive performance of these models. It sheds light on their consistency, variability, and the potential presence of outliers, enabling more informed decisions in their application and fine-tuning.

3.2. Verification and comparison

Table 6 showcases the performance metrics for the PSO-MLP model during the training and testing phases. In the training phase, the model achieved an R^2 value of 0.9126 and an RMSE of 2.9736, indicating a good fit to the training data with relatively low error. During the testing phase, the model's performance improved further, with an R^2 value of 0.9370 and an RMSE of 2.5693. These results suggest that the model generalizes well to unseen data, demonstrating its predictive capability and reliability in estimating HL in residential buildings. Comparing these results

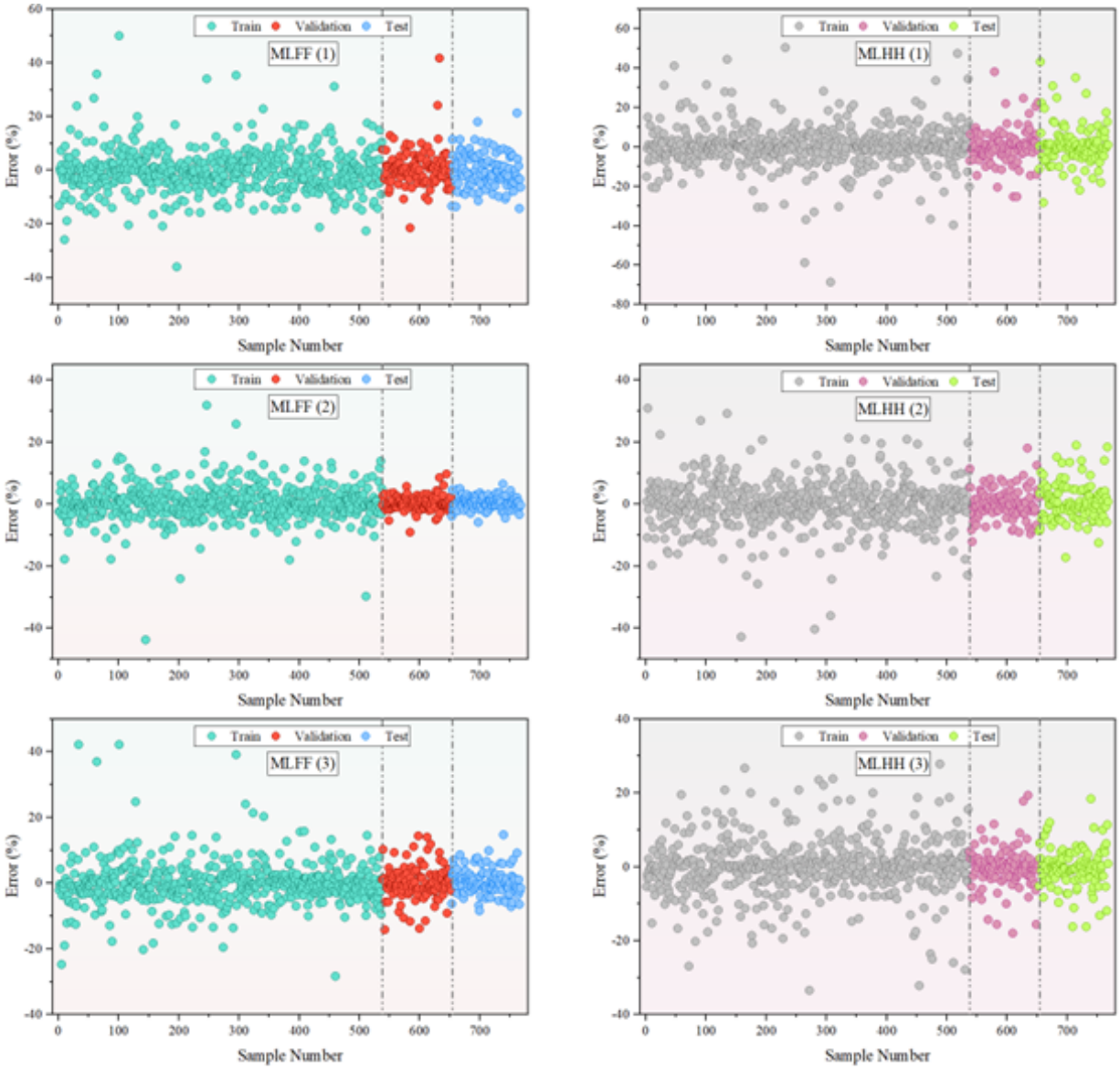

Fig. 4. Error percentage of the hybrid models based on the Scatter plot.

with the developed MLFF2 model mentioned earlier, which achieved an R^2 value of 0.997 and an RMSE of 0.523, highlights the significant improvement and superior accuracy of the hybrid approach integrating MLP with FFO and HHO optimizers. The comprehensive validation and evaluation of the developed model reinforce its effectiveness in accurately predicting HL and its potential for practical application in energy management and sustainability efforts.

Table 7 compares the performance of the best-performing models from the present study with those from related literature articles. Moradzadeh et al. [46] utilized Support Vector Regression (SVR), achieving an RMSE of 0.4832 and an R^2 value of 0.9969 for HL prediction. Sadaghat et al. [47] employed XGCM, yielding an RMSE of 0.8215 and an R^2 of 0.9934. Roy et al. [48] used MPMR techniques, achieving exceptional results with an RMSE of 0.059 and an R^2 of 0.99. Gong et al. [49] utilized Light Gradient Boosting Machines (LGBM), obtaining an RMSE of 0.1929 and an R^2 of 0.9882. In comparison, the present study's MLFF2 model achieved an RMSE of 0.523 and an impressive R^2 value of 0.997 for HL prediction. These results demonstrate the competitive performance of the MLFF2 model against established techniques in the field, showcasing its accuracy and reliability in predicting HL in residential buildings.

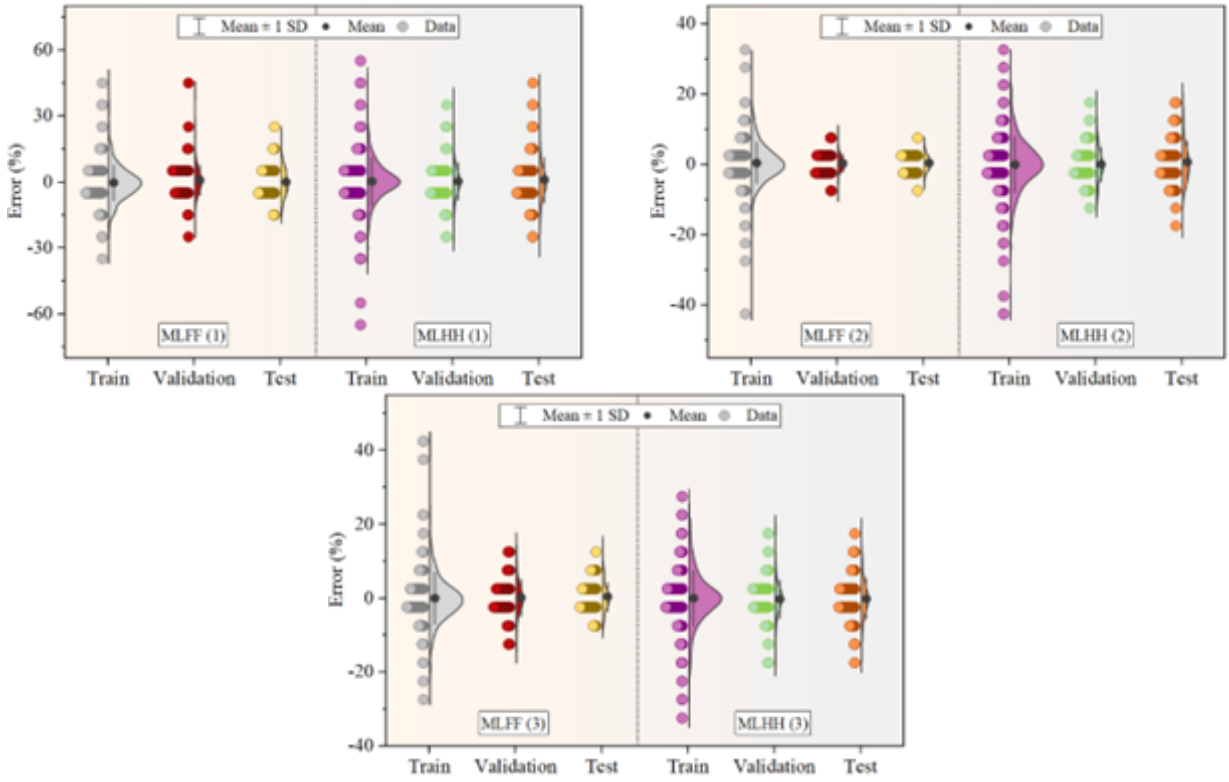

Fig. 5. Half-violin plot for error percentage of developed models.

Table 6. The verification of the developed model.

Article	Model	Phase	Evaluator	
			R^2	RMSE
Zhou et al. [45]	PSO-MLP	Train	0.9126	2.9736
		Test	0.9370	2.5693
Present study	MLFF2	Train	1.092	0.989
		Test	0.523	0.997

Table 7. The comparison of the best-performed model results of the present study with some related literature.

Articles	Index values			
	Models	*Target*	*RMSE*	R^2
Moradzadeh et al. [46]	SVR	HL	0.4832	0.9969
Sadaghat et al. [47]	XGCM	HL	0.8215	0.9934
Roy et al. [48]	MPMR	HL	0.059	0.99
Gong et al. [49]	LGBM	HL	0.1929	0.9882
Present Study	MLFF2	HL	0.523	0.997

4. Conclusion

In the pursuit of enhancing the accuracy and reliability of predicting heating load (HL) in residential buildings, this study has ventured into the innovative modeling and optimization techniques. The goal was to offer valuable insights into the effective management of energy consumption in buildings, an area of growing importance in the face of environmental concerns and energy efficiency initiatives.

This research introduced a novel approach that harnessed the capabilities of Multi-layer Perceptron (MLP) models, coupling them with the power of optimization techniques, namely Flying Fox Optimization (FFO) and Horse Herd Optimization (HHO). The results obtained from these models and techniques have provided valuable contributions to the field of building energy efficiency. The performance of these hybrid models, MLFF and MLHH, was evaluated across various layers of MLP. The results consistently high-

lighted the superiority of the MLFF model in terms of predictive accuracy. This was particularly evident in Layer 2, where MLFF exhibited exceptional precision and alignment with measured values. The testing phase in Layer 2 demonstrated the model's remarkable consistency and improvement in accuracy compared to its training phase, underscoring its efficacy in predicting HL across different phases and layers. In the presented models (MLFF (1) and MLHH (1)), MLHH regularly exceeds MLFF with lower RMSE values, indicating improved predicting accuracy. However, MLHH's lower R^2 values indicate a less accurate data fit. The trade-off between precision and model fit is evident. In the following model sets (MLFF (2) and MLHH (2)), both models improve, with MLFF slightly outperforming in the testing phases. MLFF (3) and MLHH (3) show a decreasing performance disparity, demonstrating developing dynamics between the two techniques. While MLFF excels in accuracy, MLHH shows competitive performance in later model sets. These findings highlight the potential of hybrid MLP models for predicting HL and the significance of customized model setups. Further refinement and application-specific changes are advised for optimal outcomes. This research contributes not only to the body of knowledge in building energy efficiency but also offers practical solutions for a more sustainable and eco-conscious future. Further investigations may focus on refining these models for specific scenarios and expanding their application in real-world settings. In order to address certain scenarios and obstacles, future work in this field may concentrate on improving the hybrid MLP models using FFO and HHO optimization techniques. This could entail adjusting the models to account for differences in building kinds, weather patterns, or energy sources, improving their accuracy and adaptability in a range of environments. Furthermore, investigating the integration of cutting-edge sensors and real-time data may enhance the models' capacity for prediction and their ability to adapt to dynamic variations in heating demand. Moreover, extending the utilization of these models to encompass additional building energy management facets like anticipating cooling load or incorporating renewable energy sources may provide all-encompassing approaches for sustainable and effective building functions. The dataset's probable lack of representation for various residential building HVAC systems is its main drawback. Although the study makes use of cutting-edge methodologies like as MLP models and optimization approaches like FFO and HHO, the dataset may not fully represent the range of HVAC configurations encountered in real-world situations. Because of this, the predictive models that have been built may be biased or have limited generalizability, especially when applied to buildings that have distinct HVAC systems or use different energy sources. Future studies should concentrate on collecting data from a larger range of HVAC systems, taking into account variations in building types, energy sources, climate areas, and operating configurations, in order to address this restriction and improve representativeness.

nomenclature

Acronyms

AI	Artificial intelligent
FFO	Flying Fox Optimization
GA	Glazing Area
GAD	Glazing Area Distribution
HHO	Horse Herd Optimizers
HL	Heating load
HVAC	Heating, ventilation, and air conditioning
ML	Machine learning
MLP	Multi-layer Perceptron
MSE	Mean Square Error
OR	Orientation
OVH	Overall Height
R^2	Coefficient Correlation
RA	Roof Area
RAE	Relative Absolute Error
RCE	Relative Compactness
RMSE	Root Mean Square Error
SA	Surface Area
SI	Scatter Index
WA	Wall Area

References

[1] Y. Lu, Z. Tian, Q. Zhang, R. Zhou, and C. Chu, (2021) *"Data augmentation strategy for short-term heating load prediction model of residential building"* **Energy 235**: 121328. DOI: 10.1016/j.energy.2021.121328.

[2] R. Chaganti, F. Rustam, T. Daghriri, I. de la Torre Díez, J. L. V. Mazón, C. L. Rodríguez, and I. Ashraf, (2022) *"Building heating and cooling load prediction using ensemble machine learning model"* **Sensors 22**: 7692. DOI: 10.3390/s22197692.

[3] C. Wang, J. Yuan, K. Huang, J. Zhang, L. Zheng, Z. Zhou, and Y. Zhang, (2022) *"Research on thermal load prediction of district heating station based on transfer learning"* **Energy 239**: 122309. DOI: 10.1016/j.energy.2021.122309.

[4] M. Protić, S. Shamshirband, M. H. Anisi, D. Petković, D. Mitić, M. Raos, M. Arif, and K. A. Alam, (2015) *"Appraisal of soft computing methods for short term consumers' heat load prediction in district heating systems"* **Energy 82**: 697–704. DOI: 10.1016/j.energy.2015.01.079.

[5] J. Ling, N. Dai, J. Xing, and H. Tong, (2021) *"An improved input variable selection method of the data-driven model for building heating load prediction"* **Journal of**

Building Engineering *44*: 103255. DOI: 10.1016/j.jobe.2021.103255.

[6] F. Dalipi, S. Y. Yayilgan, and A. Gebremedhin, (2016) *"Data-driven machine-learning model in district heating system for heat load prediction: A comparison study"* **Applied Computational Intelligence and Soft Computing *2016***: DOI: 10.1155/2016/3403150.

[7] G. Xue, Y. Pan, T. Lin, J. Song, C. Qi, and Z. Wang, (2019) *"District heating load prediction algorithm based on feature fusion LSTM model"* **Energies *12***: 2122. DOI: 10.3390/en12112122.

[8] K. Kato, M. Sakawa, K. Ishimaru, S. Ushiro, and T. Shibano. "Heat load prediction through recurrent neural network in district heating and cooling systems". In: IEEE, 2008, 1401–1406. DOI: 10.1109/ICSMC.2008.4811482.

[9] J. Yuan, Z. Zhou, H. Tang, C. Wang, S. Lu, Z. Han, J. Zhang, and Y. Sheng, (2020) *"Identification heat user behavior for improving the accuracy of heating load prediction model based on wireless on-off control system"* **Energy *199***: 117454. DOI: 10.1016/j.energy.2020.117454.

[10] Y. Zhang, Z. Zhou, J. Liu, and J. Yuan, (2022) *"Data augmentation for improving heating load prediction of heating substation based on TimeGAN"* **Energy *260***: 124919. DOI: 10.1016/j.energy.2022.124919.

[11] G. Xue, C. Qi, H. Li, X. Kong, and J. Song, (2020) *"Heating load prediction based on attention long short term memory: A case study of Xingtai"* **Energy *203***: 117846. DOI: 10.1016/j.energy.2020.117846.

[12] Q. Zhang, Z. Tian, Z. Ma, G. Li, Y. Lu, and J. Niu, (2020) *"Development of the heating load prediction model for the residential building of district heating based on model calibration"* **Energy *205***: 117949. DOI: 10.1016/j.energy.2020.117949.

[13] E. Guelpa, L. Marincioni, M. Capone, S. Deputato, and V. Verda, (2019) *"Thermal load prediction in district heating systems"* **Energy *176***: 693–703. DOI: 10.1016/j.rser.2015.04.020.

[14] S. Shamshirband, D. Petković, R. Enayatifar, A. H. Abdullah, D. Marković, M. Lee, and R. Ahmad, (2015) *"Heat load prediction in district heating systems with adaptive neuro-fuzzy method"* **Renewable and Sustainable Energy Reviews *48***: 760–767. DOI: 10.1016/j.enbuild.2017.11.002.

[15] Y. Ding, Q. Zhang, T. Yuan, and K. Yang, (2018) *"Model input selection for building heating load prediction: A case study for an office building in Tianjin"* **Energy and Buildings *159***: 254–270. DOI: 10.1016/j.jobe.2019.100950.

[16] M. Gong, Y. Bai, J. Qin, J. Wang, P. Yang, and S. Wang, (2020) *"Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin"* **Journal of Building Engineering *27***: 100950. DOI: 10.1016/j.jobe.2019.100950.

[17] S. S. Roy, P. Samui, I. Nagtode, H. Jain, V. Shivaramakrishnan, and B. Mohammadi-Ivatloo, (2020) *"Forecasting heating and cooling loads of buildings: A comparative performance analysis"* **Journal of Ambient Intelligence and Humanized Computing *11***: 1253–1264. DOI: 10.1007/s12652-019-01317-y.

[18] B. Sadaghat, S. Afzal, and A. J. Khiavi, (2024) *"Residential building energy consumption estimation: A novel ensemble and hybrid machine learning approach"* **Expert Systems with Applications *251***: 123934. DOI: 10.1016/j.eswa.2024.123934.

[19] A. Moradzadeh, A. Mansour-Saatloo, B. Mohammadi-Ivatloo, and A. Anvari-Moghaddam, (2020) *"Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings"* **Applied Sciences *10***: 3829. DOI: 10.3390/app10113829.

[20] G. Zhou, H. Moayedi, M. Bahiraei, and Z. Lyu, (2020) *"Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings"* **Journal of Cleaner Production *254***: 120082. DOI: 10.1016/j.jclepro.2020.120082.

[21] A. Botchkarev, (2018) *"Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology"* **arXiv preprint arXiv:1809.03006**: DOI: 10.48550/arXiv.1809.03006.

[22] F. MiarNaeimi, G. Azizyan, and M. Rashki, (2021) *"Horse herd optimization algorithm: A nature-inspired algorithm for high-dimensional optimization problems"* **Knowledge-Based Systems *213***: 106711. DOI: 10.1016/j.knosys.2020.106711.

[23] S. Kumar, G. G. Tejani, N. Pholdee, and S. Bureerat, (2021) *"Multiobjecitve structural optimization using improved heat transfer search"* **Knowledge-Based Systems *219***: 106811. DOI: 10.1016/j.knosys.2021.106811.

[24] Z. Liu, P. Jiang, J. Wang, and L. Zhang, (2021) *"Ensemble forecasting system for short-term wind speed forecasting based on optimal sub-model selection and multi-objective version of mayfly optimization algorithm"* **Expert Systems with Applications *177***: 114974. DOI: 10.1016/j.eswa.2021.114974.

[25] Y. Li, W. Lu, Z. Pan, Z. Wang, and G. Dong, (2023) *"Simultaneous identification of groundwater contaminant source and hydraulic parameters based on multilayer perceptron and flying foxes optimization"* **Environmental Science and Pollution Research *30***: 78933–78947. DOI: 10.1007/s11356-023-27574-1.

[26] R. Aalloul, A. Elaissaoui, M. Benlattar, and R. Adhiri, (2023) *"Emerging Parameters Extraction Method of PV Modules Based on the Survival Strategies of Flying Foxes Optimization (FFO)"* **Energies *16***: 3531. DOI: 10.3390/en16083531.

[27] A. G. Parlos, K. T. Chong, and A. F. Atiya, (1994) *"Application of the recurrent multilayer perceptron in modeling complex process dynamics"* **IEEE Transactions on Neural Networks *5***: 255–266. DOI: 10.1109/72.279189.

[28] H. Taud and J.-F. Mas, (2018) *"Multilayer perceptron (MLP)"* **Geomatic approaches for modeling land change scenarios**: 451–455. DOI: 10.1007/978-3-319-60801-3_27.

[29] F. Murtagh, (1991) *"Multilayer perceptrons for classification and regression"* **Neurocomputing *2***: 183–197. DOI: 10.1016/0925-2312(91)90023-5.

[30] H. Ramchoun, Y. Ghanou, M. Ettaouil, and M. A. J. Idrissi, (2016) *"Multilayer perceptron: Architecture optimization and training"*: DOI: 10.9781/ijimai.2016.415.

[31] L. Noriega, (2005) *"Multilayer perceptron tutorial"* **School of Computing. Staffordshire University *4***: 444.

[32] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, (2017) *"A review on time series forecasting techniques for building energy consumption"* **Renewable and Sustainable Energy Reviews *74***: 902–924. DOI: 10.1016/j.rser.2017.02.085.

[33] N. Xu, Y. Dang, and Y. Gong, (2017) *"Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China"* **Energy *118***: 473–480. DOI: 10.1016/j.energy.2016.10.003.

[34] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, (2017) *"Building energy consumption prediction: An extreme deep learning approach"* **Energies *10***: 1525. DOI: 0.3390/en10101525.

[35] C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, M. A. Brown, and R. M. Pendyala, (2017) *"Machine learning approaches for estimating commercial building energy consumption"* **Applied energy *208***: 889–904. DOI: 10.1016/j.apenergy.2017.09.060.

[36] J. Zhu, Z. Yang, Y. Chang, Y. Guo, K. Zhu, and J. Zhang. "A novel LSTM based deep learning approach for multi-time scale electric vehicles charging load prediction". In: IEEE, 2019, 3531–3536. DOI: 10.1109/ISGT-Asia.2019.8881655.

[37] N. Mohammadi and J. E. Taylor, (2017) *"Urban energy flux: Spatiotemporal fluctuations of building energy consumption and human mobility-driven prediction"* **Applied energy *195***: 810–818. DOI: 10.1016/j.apenergy.2017.03.044.

[38] D. Hsu, (2015) *"Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data"* **Applied energy *160***: 153–163. DOI: 10.1016/j.apenergy.2015.08.126.

[39] S. Naji, A. Keivani, S. Shamshirband, U. J. Alengaram, M. Z. Jumaat, Z. Mansor, and M. Lee, (2016) *"Estimating building energy consumption using extreme learning machine method"* **Energy *97***: 506–516. DOI: 10.1016/j.energy.2015.11.037.

[40] F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah, (2016) *"Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique"* **Energy and Buildings *126***: 94–103. DOI: 10.1016/j.enbuild.2016.05.028.

[41] M. W. Ahmad, M. Mourshed, and Y. Rezgui, (2017) *"Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption"* **Energy and buildings *147***: 77–89. DOI: 10.1016/j.enbuild.2017.04.038.

[42] A. Moradzadeh, A. Mansour-Saatloo, B. Mohammadi-Ivatloo, and A. Anvari-Moghaddam, (2020) *"Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings"* **Applied Sciences *10***: 3829. DOI: 10.3390/app10113829.

[43] S. S. Roy, P. Samui, I. Nagtode, H. Jain, V. Shivaramakrishnan, and B. Mohammadi-Ivatloo, (2020) *"Forecasting heating and cooling loads of buildings: A comparative performance analysis"* **Journal of Ambient Intelligence and Humanized Computing *11***: 1253–1264. DOI: 10.1007/s12652-019-01317-y.

[44] T.-Y. Kim and S.-B. Cho, (2019) *"Predicting residential energy consumption using CNN-LSTM neural networks"* **Energy *182***: 72–81. DOI: 10.1016/j.energy.2019.05.230.

[45] M. R. Akbarzadeh, H. Ghafourian, A. Anvari, R. Pourhanasa, and M. L. Nehdi, (2023) *"Estimating compressive strength of concrete using neural electromagnetic field optimization"* **Materials *16***: 4200. DOI: 10.3390/ma16114200.

[46] Z. Wang, T. Hong, and M. A. Piette, (2020) *"Building thermal load prediction through shallow machine learning and deep learning"* **Applied Energy *263***: 114683. DOI: 10.1016/j.apenergy.2020.114683.

[47] B. Sedaghat, A. J. Khiavi, B. Naeim, E. Khajavi, and A. R. T. Khanghah, (2023) *"Evaluation of Object-Based and Pixel-Based Technique for Extracting Snow Cover Surface Using Landsat 8 Satellite Images (Case Study Damavand Mountain Range)"* **Advances in Engineering and Intelligence Systems *2***: 87–100. DOI: 10.22034/AEIS.2023.427451.1147.

[48] S. Afzal, B. M. Ziapour, A. Shokri, H. Shakibi, and B. Sobhani, (2023) *"Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms"* **Energy *282***: 128446. DOI: 10.1016/j.energy.2023.128446.

[49] M. Sajjad, S. U. Khan, N. Khan, I. U. Haq, A. Ullah, M. Y. Lee, and S. W. Baik, (2020) *"Towards efficient building designing: Heating and cooling load prediction via multi-output model"* **Sensors *20***: 6419. DOI: 10.3390/s20226419.