Improving Heating Load Prediction With LSSVR: Comparative Analysis Of Optimized Models

Deshen LV[1], Chengquan LIANG[1*], and Xiao LU[2]

[1] *School of Intelligent Manufacturing, Nanning University, Nanning 530200, China*
[2] *School of Information Science and Engineering, Guilin University of Technology, Guilin 541006, China*
[*] *Corresponding author. E-mail: liangchengquan0088@yeah.net*

Received: Jan. 30, 2023; Accepted: May 12, 2024

At present, energy usage is one of the critical components of the global economy and population growth in the construction sector. Buildings play a crucial role in global energy consumption, and it is vital to forecast the heating needs of this sector thoroughly. This necessity is driven by various significant factors, including improving energy efficiency, financial responsibility, promoting environmental health, and developing sustainable and long-lasting solutions. Accurately estimating the heating load of buildings is incredibly important. Machine learning (ML) is one of the most effective techniques among the various methods employed for this purpose. This approach involves the analysis of historical data and an evaluation of the present conditions within the building to deliver precise predictions regarding heating load requirements. This study aims to apply the Least Squares Support Vector Regression (LSSVR) method, a frequently used ML algorithm for predicting continuous numerical values for determining building heating load. The application of the Arithmetic Optimization Algorithm (AOA) and the Ebola Optimization Search Algorithm (EOSA) is geared toward improving accuracy and reducing overall losses in heating load estimation. The research provides significant insights into predicting building heating loads and recommends that employing an LSSVR+EOSA (LSEO) model is the most efficient strategy for optimizing energy consumption. This hybrid model achieved a maximum determination coefficient of 0.985 and a root mean square error of 1.223.

Keywords: Heating Load, Least Squares Support Vector Regression, Arithmetic Optimization Algorithm, Ebola Optimization Search Algorithm.

http://dx.doi.org/10.6180/jase.202505_28(5).0018

Nomenclature

LSSVR	Least Squares Support Vector Regression	EOSA	Ebola Optimization Search Algorithm
AOA	Arithmetic Optimization Algorithm	RMSE	Root Mean Square Error
MAE	Mean Absolute Error	RAE	Relative Absolute Error
RSR	RMSE-observations standard deviation ratio	R^2	Coefficient of Determination
RCE	Relative Compactness	RA	Roof Area
SA	Surface Area	OVH	Overall Height
WA	Wall Area	OR	Orientation
GA	Glazing Area Distribution	GAD	Glazing Area

1. Introduction

According to research, it is foreseen that the world's urban population will reach 5 billion individuals by 2030, representing roughly 60% of the total global populace. This growth in urban numbers will result in an amplified requirement for energy [1]. Typically, final energy consumption is divided into three primary sectors: industry, transportation, and a miscellaneous category that comprises agriculture, the service sector, and residential areas. This classification complicates collecting data on building energy usage [2]. Buildings constitute the primary segment of energy usage within the economic sector, making up more than a third of overall ultimate energy consumption, and bear an equivalent responsibility for roughly one-third of future global carbon dioxide emissions [3]. Employing renewable energy sources represents a critical strategy for achieving low-carbon electricity provision and promoting environmental sustainability. This is particularly pertinent in sectors such as building and transportation, which are associated with elevated carbon emissions and substantial energy consumption [4]. Furthermore, renewable energy-driven supply systems have widespread adoption due to their minimal pollutant emissions and superior energy efficiency [5].

Precise forecasting of building loads, encompassing cooling, Heating, and electricity requirements, is pivotal in optimizing the construction and operational phases [6]. Well-designed and efficiently operated buildings have the potential to yield substantial energy savings, and metrics like Heating Load (HL) and Cooling Load (CL) quantify the energy needed to maintain a comfortable temperature through Heating, Ventilation, and Air Conditioning (HVAC) systems. These accurate load predictions can assist engineers in creating energy-efficient buildings [7]. Extensive research has thoroughly examined buildings' energy consumption throughout their life cycle. These studies have had varying objectives, such as pinpointing the energy usage of individual building components at different levels [8, 9] or assessing energy performance on a national scale [10, 11]. The effective control and enhancement of building energy utilization necessitate a comprehensive comprehension of building performance [12]. This entails the initial identification of a building's energy resources and its primary energy consumers. In the context of a building, energy resources typically encompass the provision of electricity, natural gas, and district heating [13, 14].

In contemporary construction and HVAC design, the precise prediction of heating load is essential for achieving energy efficiency and sustainable building practices. Traditional methodologies often rely on simplified models or manual calculations, potentially overlooking crucial factors and resulting in suboptimal heating system sizing and energy utilization [15]. In pursuing environmental sustainability and energy conservation, the precise control of District Heating Systems (DHSs) is now paramount [16]. Accurately predicting heating loads has gained significant relevance due to its potential to enhance the efficiency of energy-saving measures [17]. Research in this field predominantly falls into 2 main categories based on prediction principles: (a) statistical regression models and (b) ML models [18]. Utilizing data-driven statistical modeling through ML techniques enables the accurate prediction of building heating and cooling loads [19], leveraging existing data, without the need for comprehensive building-specific information [20]. The application of ML encompasses the management of extensive datasets, the identification of intricate patterns, and the capacity to adapt to dynamic conditions [21]. ML has thus far been recognized as an essential resource for building energy prediction, architectural design, and maintenance [22, 23]. However, the advent of ML presents a promising avenue for enhancing the accuracy and reliability of heating load prediction. ML algorithms leverage vast datasets and computational capabilities to uncover intricate relationships between various building parameters and heating requirements [24]. By analyzing historical heating load data alongside factors such as building geometry, insulation materials, weather conditions, and occupancy patterns, ML models can discern complex patterns and nonlinear dependencies that may elude traditional methods [25].

The Least Squares Support Vector Regression (LSSVR) stands as a frequently employed and thoroughly researched ML model in the domain of energy forecasting over an extended duration [26]. Additionally, for managing the interconnections among output variables within a multi-output system, the multi-output LSSVR, which incorporates the consideration of comprehensive fitting errors, demonstrates its reliability and robustness, as substantiated by citations in references [27, 28].

Le et al. [29] developed a custom-tailored LSSVR model for predicting building energy consumption. They integrated the Jellyfish Algorithm for parameter optimization and achieved exceptional forecasting accuracy. The model outperformed alternative regression and time series models and was validated against prior research, confirming its effectiveness in energy consumption prediction.

This study explores the application of ML techniques for predicting heating load in buildings. The objective is to develop models that accurately estimate heating demand while providing insights into the underlying factors influencing energy consumption. Such models have the potential to empower stakeholders, including architects, engineers, and building managers, to make data-informed decisions regarding heating system design, retrofitting strategies, and energy management practices. In addition, 2 optimization algorithms, the Arithmetic Optimization Algorithm (AOA) and the Ebola Optimization Search Algorithm (EOSA), have been introduced to enhance the efficiency of the LSSVR traditional model developing LSEO and LSAO hybrid models. Developed models trained through extracted datasets from previous related works and, after validation, predicted the performance of the models tested. R2, RMSE, MAE, RAE, and RSR evaluation metrics were utilized for a comprehensive comparison between LSSVR and optimized models to introduce the most workable model in heating load prediction and designing more energy-efficient buildings.

2. Materials and methods

2.1. Dataset Description

This research focuses on utilizing experimental data derived from energy consumption patterns within buildings to forecast heating load. To accomplish this objective, the study employs the LSSVR regression ML algorithm and a training process that integrates the Arithmetic Optimization Algorithm (AOA) and the Ebola Optimization Search Algorithm (EOSA). The analysis incorporates various parameters for heating load estimation, with the relevant input and output data encompassing metrics like Minimum, Maximum, Average, and Standard Deviation, as presented in Table 1.

This study's simulation specifically targets residential buildings situated in Athens, Greece. The buildings are designed for a maximum of seven people and are utilized for sedentary activities, generating an estimated 70 (W) of heat per occupant. Various interior design specifications were taken into account, such as a clothing insulation factor of 0.6 , a humidity level of 60%, an air velocity of 0.30 m/s, and an illumination level of 300 Lux. The internal heat gains were defined as sensible at 5 $\left(\mathrm{W/m^2}\right)$ and latent at 2 $\left(\mathrm{W/m^2}\right)$. The infiltration rate, which influences the air change rate, was established at 0.5 with a wind sensitivity factor of 0.25. The material parameters of the facade system are uniform in all 12 buildings in this examination. The U-value, indicating thermal transmittance, is defined for various parts of the building envelope. The study utilized the following U-values:

- Wall:1.78 $\left(\mathrm{W/m^2\,K}\right)$
- Window: 2.26 $\left(\mathrm{W/m^2\,K}\right)$
- Floors: 0.86 $\left(\mathrm{W/m^2\,K}\right)$
- Roofs: 0.5 $\left(\mathrm{W/m^2\,K}\right)$

The U-values offer insights into the thermal conductivity of the specific building components. By using these numbers in the analysis, the thermal efficiency of the structures may be assessed, aiding in precise forecasts of the cooling and heating loads. The buildings' thermal properties were simulated using a mixed-mode technique, with an assumed efficiency of 95%. The dataset utilized to train the intelligent models in this investigation was obtained from a prior study by Tsanas and Xifara [30]. The dataset contains essential information for using the suggested methods and assessing their effectiveness in estimating the cooling and heating loads of buildings.

2.2. Overview of Machine Learning Method and Optimizers

2.2.1. Least Square Support Vector Regression (LSSVR)

The Support Vector Machine (SVM) introduced by Vapnik [31] is a potent supervised learning technique applied in function estimate, nonlinear categorization, and density estimation for nonlinear classification tasks. The Support Vector Machine for Regression (SVR) aims to seek a linear regression model within a multidimensional space to minimize the flatness of the estimated function while maintaining a minor divergence from the training dataset [32]. Under these circumstances, $\{x_i, y_i\}\ i = 1, 2, \ldots, l$ are the training dataset at hand, the regression function can be delineated in the subsequent manner:

$$f(x) = Q^S x + d \quad \text{with} \quad Q, x \in R^n \& d \in R \tag{1}$$

In typical scenarios, the input data is transformed into a higher-dimensional space by the SVR method, primarily because of the nonlinear correlation between the input and output variables. If it is regarded as the mapping function responsible for taking x from the input space and converting it into the feature space. In that case, the linear regression function in the feature space can be consequently represented as demonstrated below:

$$f(x) = Q^S \psi(x) + d \tag{2}$$

To determine $f(x)$ such that the difference between the function value and the actual target for all training data is reduced to less than ε, the epsilon Support Vector Regression ($\varepsilon-\ SVR$) is employed. In this instance, the issue of discovering the regression function is represented by the convex optimization problem with the subsequent constraints:

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

Indicator	Input								Output
	RCE	*SA*	*WA*	*RA*	*OVH*	*OR*	*GA*	*OAD*	*Heating*
Max	0.98	808.5	416.5	220.5	7	5	0.4	5	43.1
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	318.500	176.604	5.250	3.500	0.234	2.813	22.307
Skew	0.496	-0.125	0.533	-0.163	0.000	0.000	-0.060	-0.089	0.360
St. Dev	0.106	88.086	43.626	45.166	1.751	1.119	0.133	1.551	10.090

$$\min \frac{1}{2}\|Q\|^2 \begin{cases} y_i - Q^S\psi(x_i) - d \le \varepsilon \\ Q^S\psi(x_i) + d - y_i \le \varepsilon \end{cases} \tag{3}$$

Considering that achieving convex optimization and a precision ε function may not always be possible, slack variables δ and δ_i^* have been incorporated. Consequently, the reformulated convex optimization problem appears as follows:

$$\min \frac{1}{2}\|Q\|^2 + c\sum_{i=1}^{l}(\delta_i + \delta_i^*)$$
$$\begin{cases} y_i - Q^S\psi(x_i) - d \le \varepsilon + \delta_i \\ Q^S\psi(x_i) + d - y_i \le \varepsilon + \delta_i^* \\ \delta_i, \delta_i^* \ge 0 \end{cases} \tag{4}$$

In this context, the value of c, which is non-negative, acts as the equilibrium point between the smoothness of function f and the permissible deviation level beyond ε. The Least Squares Support Vector Machine (LS-SVM) represents an enhanced iteration of the Support Vector Machine (SVM) algorithm employed for both classification and regression tasks. LSSVM tackles the computational challenges associated with SVM by transforming quadratic optimization issues into linear equations, simplifying the resolution process [33]. More precisely, LS-SVM streamlines the representation by addressing the subsequent optimization problem:

$$\min \frac{1}{2}\|Q\|^2 + \frac{\Omega}{2}\sum_{i=1}^{l} e_i^2 \tag{5}$$

Such that $\quad y_i = Q^S\psi(x_i) + d + e_i,$

Ω acts as a regulation parameter and represents LSSVR strikes a balance between minimizing the training error and promoting the smoothness of the function. This balance is achieved through the use of a Lagrangian function, which is formulated as follows:

$$\mathrm{LF} = \frac{1}{2}\|Q\|^2 + \frac{\Omega}{2}\sum_{i=1}^{l} e_i^2 - \sum_{i=1}^{l}\alpha_i\left(y_i - Q^S\psi(x_i) - d - e_i\right), \tag{6}$$

In this context, α_i represents the Lagrangian multiplier. To resolve the optimization problem, both primal and dual variables must be equated to zero:

$$\frac{\partial LF}{\partial Q} = 0 \to Q = \sum_{i=1}^{l}\alpha_i\psi(x_i), \tag{7}$$

$$\frac{\partial LF}{\partial d} = 0 \to \sum_{i=1}^{l}\alpha_i = 0, \tag{8}$$

$$\frac{\partial LF}{\partial e_i} = 0 \to \alpha_i = \Omega e_i, \tag{9}$$

$$\frac{\partial LF}{\partial \alpha_i} = 0 \to Q^S\psi(x_i) + d + e_i - y_i = 0. \tag{10}$$

Next, the values of Q and e are introduced into the Lagrangian function to construct a linear system as demonstrated below:

$$\begin{bmatrix} 0 & 1^S \\ 1 & \lambda + \Omega^{-1}I \end{bmatrix}\begin{bmatrix} d \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \tag{11}$$

$$\begin{aligned} \alpha &= [\alpha_1 + \alpha_2, \ldots, \alpha_L]^S \\ e &= [e_1 + e_2, \ldots, e_L]^S \\ \lambda &= VV^S \\ V &= [\psi(x_1), \psi(x_2), \ldots, \psi(x_L)]^S, \\ y &= [y_1 + y_2, \ldots, y_L]^S \\ 1 &= [1 + 1, \ldots, 1]^S. \end{aligned} \tag{12}$$

At this point, the kernel manipulation can be employed to construct the matrix λ, as expressed in Eq. (13), which subsequently allows the reconfiguration of the solution to Eq. (14) into the form depicted in Eq. (15):

$$\lambda_{ij} = \psi(x_i)^S\,\psi\left(x_j\right)^S = K\left(x_i, x_j\right) \tag{13}$$

$$c = \frac{1^T\left(\lambda + \Omega^{-1}I\right)^{-1}y}{1^T\left(\lambda + \Omega^{-1}I\right)^{-1}I} \tag{14}$$

$$\alpha = \left(\lambda + \Omega^{-1}I\right)^{-1}(y - d1) \tag{15}$$

The Radial Basis Function (RBF) with parameter v^2 can be chosen as a kernel function, representing an infinitely dimensional function space. Additionally, with the acquired values of d and α, the optimal regression function can be expressed as shown below:

$$f(x) = \sum_{i=1}^{l} \alpha_i K(x_i, x) + d, \tag{16}$$

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2v^2}\right) \tag{17}$$

Wherever, v^2 represents the squared bandwidth of the kernel.

In prior studies, diverse optimal parameters have been recommended for Ω and v^2. For instance, Hsu and colleagues [34] recommended 1 and $1/K$ (where K represents the count of input patterns) for v^2 and Ω, correspondingly. Aiyer and colleagues [35] proposed an experimental and iterative approach, which can be quite time-intensive when it comes to estimating Ω and v^2 since they can assume any positive real values. Consequently, optimization techniques can be employed to enhance these parameters for LSSVR.

2.2.2. *Arithmetic optimization algorithm (AOA)*

Arithmetic optimization stems from number theory, a fundamental field within modern mathematics. This strategy leverages ordinary math operations such as multiplication, subtraction, division, and addition to examine numerical values. It uses them to distinguish the ideal component from a collection of arrangements according to predefined criteria [36]. To tackle mathematical problems, the Arithmetic Optimization Algorithm is inspired by using multiplication operations. Resembling traditional population-based optimization methods, AOA comprises 2 fundamental phases: exploration and exploitation. In the exploration stage, the specified search space is comprehensively examined to reveal potential solutions. In contrast, the exploitation stage focuses on refining the obtained solution's accuracy [37]. The AOA algorithm includes three primary stages, as elaborated in the following sections.

Initialization phase

In the first phase of the Arithmetic Optimization Algorithm, a set of potential solutions (denoted as X) is generated randomly. At each cycle, the algorithm hypothesizes that the best potential solution is either optimal or lies within its local vicinity.

$$X = \begin{bmatrix} x_{1,1} & \cdots & \cdots & x_{1,j} & x_{1,n-1} & x_{1,n} \\ x_{2,1} & \cdots & \cdots & x_{2,j} & \cdots & x_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N-1,1} & \cdots & \cdots & x_{N-1,j} & \cdots & x_{N-1,n} \\ x_{N,1} & \cdots & \cdots & x_{N,j} & x_{N,n-1} & x_{N,n} \end{bmatrix} \tag{18}$$

Before commencing the AOA procedure, a decision must be taken regarding whether to start with the exploration or exploitation phases. Subsequently, the Math Optimizer Accelerated (MOA) function, representing the function value at the ith iteration, is computed using Eq. (19).

$$MOA(B_{Iter}) = Min + B\_Iter \times \left(\frac{Max - Min}{M\_Iter}\right) \tag{19}$$

M_Iter Represents the maximum allowable number of iterations, while B_Iter spans from 1 to M_{Iter} -1 . The designations 'Max' and 'Min' delineate the uppermost and lowest values of the accelerated function. These serve as benchmarks for adjusting the step size of arithmetic operators during the global search phase.

Exploration phase

The exploratory search process integrates mathematical calculations employing either Division (DO) or Multiplication (MO) operators, resulting in widely dispersed values or options. Hence, they experience challenges directly attaining the intended goal, often necessitating several iterations to discover a solution that approaches optimality and can be utilized to support the exploitation phase. Eq. (20) encapsulates the development of 2 central search strategies during the exploration phase, outlining the mathematical expressions for updating positions.

$$x_{i,j}(B_{\text{Iter}} + 1) = \begin{cases} \text{best}(x_j) \div (MOP + \varepsilon) \times \left(\left(UB_j - LB_j\right) \times \lambda + LB_j\right) & , r2 < 0.5 \\ \text{best}(x_j) \times MOP \times \left(\left(UB_j - LB_j\right) \times \lambda + LB_j\right) & , \text{otherwise} \end{cases} \tag{20}$$

In this framework, best (x_i) represents the j th location discovered in the most advantageous solution identified to date, while $x_{i,j}(B\_\text{Iter})$ signifies the j th position of the ith solution at a particular iteration(B_{Iter}). ε represents a small integer value, while LB_j and UB_j denote the lower and upper boundary values for the jth location. λ Functions as a control parameter. The expression for the function value of $MOP(B\_Iter)$ is articulated as follows:

$$\operatorname{MOP}\left(B_{Iter}\right)=1-\frac{B\_Iter^{\frac{1}{\beta}}}{M\_Iter^{\frac{1}{\beta}}} \tag{21}$$

Here, β serves as a pivotal parameter that defines the precision of the exploitation process across the iterations.

Exploitation phase

In mathematical computations, the Subtraction (SO) and Addition (AO) operators produce densely concentrated results. Therefore, during the exploitation phase, these operators can effectively approach the target through numerous iterations. Eq. (22) delineates the 2 primary search strategies in the exploitation phase and presents the equations for updating positions. The exploitation stage of the Arithmetic Optimization Algorithm employs the Subtraction (SO) and Addition (AO) operators to avoid the algorithm's confinement within local search regions and to aid associated search methodologies in locating the optimal solution.

$$x_{i,j}\left(B_{\text{Iter}}+1\right) = \begin{cases} \text{best}\left(x_j\right)-MOP\times\left(\left(UB_j-LB_j\right)\times\lambda+LB_j\right) & ,r3<0.5 \\ \text{best}\left(x_j\right)+MOP\times\left(\left(UB_j-LB_j\right)\times\lambda+LB_j\right) & ,\text{ otherwise}\end{cases} \tag{22}$$

2.3. Ebola Optimization Search Algorithm (EOSA)

A metaheuristic methodology known as the Ebola Optimization Search algorithm (EOSA) is employed, utilizing insights from the transmission process of the infectious Ebola virus disease to enhance. Search processes through a bio-inspired approach [38]. The SEIR-HVQD model, perceived as an enhanced iteration of the frequently employed susceptible-infectionrecovery (SIR) model, was utilized in formulating the EOS algorithm. Ebola disease dissemination among individuals residing in impacted regions is assessed via the SEIR-HVQD framework, which considers both direct and indirect broadcast means. The current investigation delineates the procedure for subsequently formulating the EOS algorithm:

The revision of the positions of all susceptible public institutions is executed through Eq. (23):

$$xL_u^{i+1}=xL_u^i+\tau X(L) \tag{23}$$

- τ indicates the factor for adjusting the scale.
- $X(L)$ symbolizes the pace at which individuals move.
- xL_u^i serves to indicate the current longitudinal position of the discrete.
- xL_u^{i+1} is employed for denoting the updated three-dimensional location of the separate.

The rates of mobility for uninfected entities and those who have encountered the contagious ailment are conveyed via Eqs. (24) and (25), correspondingly.

$$X(L)=\delta_{\text{rate}}\times R(0,1)+L\left(\text{ Best }_{\text{individual}}\right) \tag{24}$$

$$X(\delta)=\varepsilon_{\text{rate}}\times R(0,1)+L\left(\text{ Best }_{\text{individual}}\right) \tag{25}$$

The minimal and maximal levels of the rate of mobility are denoted by δ_{rate} and ε_{rate}, respectively. An area restriction modulates these 2 parameters, thereby controlling their standards. Upon reaching or surpassing a value of 0.5 for the parameter mentioned above, an increase in the movement duration for infected individuals occurs, resulting in an elevated risk of transmission. Conversely, when the value falls below 0.5 , the probability of infection propagation is reduced.

The population consists of individuals created randomly, with their initial positions set to zero. The mathematical representation of this population initialization process can be denoted by Eq. (26):

$$P_u=lb_u+R(0,1)\times\left(ub_u+lb_u\right) \tag{26}$$

- ub_u, lb_u allude to the higher and inferior limitations, correspondingly.
- R is a random variable that follows a uniform distribution over the interval $[0,1]$.
- Populace size $=\{1,2,3,\ldots,n\}$.

At time T, the best solution is selected from the group of infected individuals, and simultaneously, the highest-performing global solution is identified through Eq. (27).

$$Soll_{best}=\begin{cases} G_{best}, C_{best}^{fitness}<G_{best}^{fitness} \\ C_{best}, C_{best}^{fitness}\geq G_{best}^{fitness}\end{cases} \tag{27}$$

The labels assigned to the top-tier solution, the most up-to-date optimal outcome, and the supreme global resolution are denoted by, C_{best} , G_{best} , and Sol_{best} , respectively. G_{best} and C_{best} are instrumental in differentiating individuals into spreader and superspreader groups. Within this framework, "fitness" relates to the quantitative representation of the objective function in an optimization scenario. The cases of Exposed (E), Susceptible (S), Recovered (R), Infected (I), Vaccinated (V), Hospitalized (H), Dead (D), and Quarantined (Q) can be updated through the use of ordinary differential equations. In this field, differential calculus gathers information about changes in infection-related quantities at a specific time, denoted as T. Differential calculus is a crucial mathematical tool employed to determine the rate of change of a given quantity in this context. The subsequent calculations demonstrate the modifications in the amounts of parameters S, I, R, H, V, Q, and D.

$$\frac{\partial S(T)}{\partial T} = \pi - (\alpha_1 I + \alpha_2 (IA)\eta + \alpha_3 D + \alpha_4 R)\, S - (vS + \tau I) \tag{28}$$

$$\frac{\partial I(T)}{\partial T} = (\alpha_1 I + \alpha_2 (IA)\gamma + \alpha_3 D + \alpha_4 R)\, S - (\tau + \lambda) I - vS \tag{29}$$

$$\frac{\partial R(T)}{\partial T} = \lambda l - \tau R \tag{30}$$

$$\frac{\partial H(T)}{\partial T} = \beta l - (\lambda + \omega) H \tag{31}$$

$$\frac{\partial V(T)}{\partial T} = \lambda l - (\kappa + \varpi) V \tag{32}$$

$$\frac{\partial Q(T)}{\partial T} = (\pi l - (\lambda R + \tau D)) - \zeta Q \tag{33}$$

$$\frac{\partial D(T)}{\partial T} = (vS + \tau I) - \mu D \tag{34}$$

$\alpha_1, \alpha_2, \alpha_3$ and α_4 indicate the transmission frequency for contagious agents, pathogens, those with infections, and individuals who have recuperated, respectively. π represents the scalar function.

- ω refers to the rate of response to treatment.
- IA indicates individuals spreading the infection.
- μ mentions to the interment degree.
- η represents the deterioration amount.
- κ exhibits the immunization effectiveness rate.
- ϖ represents the percentage of individuals who have been vaccinated.
- ζ functions to showcase the isolation percentage of people.
- v represents the inherent mortality ratio.
- τ reveals the mortality rate attributed to the disease.
- λ denotes the retrieval degree.

Algorithm 1. EOSA Optimization.

Consequence: Finest answer
Contribution: objfunc, lb, ub, epoch, psize, evdincub
Production: solution, gbest
$S, D, I, H, R, V, Q, Sol_{best} \leftarrow \emptyset$
$S \leftarrow createSusceptibleIndvd\ (p\ size,\ S)$ using $Eq.$ (26);
$Sol_{best} \leftarrow generatedIndexCase\ (S)$;
$G_{best},\ C_{best} \leftarrow Sol_{best}$;
whereas $e \leq epoch \wedge \text{len}(I) > 0$ **do**
$Q \leftarrow \text{rand}(0, I)$;
$fracI = I - Q$;
for $i \leftarrow 1$ **to** $len(fracI)$ **do**
$pos_i \leftarrow movrate()\ using\ Eq.$ (23);
$d_i \leftarrow \text{rand}()$;
if $d_i > evdincub$ **then**
$neighborhood \leftarrow \text{prob}\,(\text{pos}_i)$;
if $neighborhood < 0.5$ **then**
$tmp \leftarrow \text{rand}\ (0,\ Eq.\ (32) \times I \times \delta_{rate})$;
end
else
tmp $\leftarrow \text{rand}\ (0,\ Eq.\ (32) \times I \times \varepsilon_{rate})$;
end
$new\ I+ \leftarrow tmp$;
end
$I + newI$
end
$h \leftarrow \text{rand}\ (0,\ Eq.\ (31) \times I),\ HC\ h$;
$r \leftarrow \text{rand}(0,\ Eq.\ (32) \times I),\ RC\ r$;
$v \leftarrow \text{rand}(0,\ Eq.\ (33) \times h), VC\ v$;
$d \leftarrow \text{rand}(0,\ Eq.\ (34) \times I), DC\ d$;
$I+ \leftarrow I - \text{add}(r;\ d)$;
$S+ \leftarrow r$
$S- \leftarrow d$;
$C_{best} = fitness\,(objfunc,\ I)$;
if $C_{best} > G_{best}$ **then**
$G_{best} = C_{best}$;
$Sol_{best} \leftarrow G_{best}$;
end
end
return $G_{best},\ Sol_{best}$;
end while

Through evaluating the variation in quantity rate relative to the vector's present magnitude, one can ascertain the count of individuals vulnerable to infection at period T. The first algorithm furnishes a brief depiction of the EOSA algorithm through the medium of pseudocode [39]. Algorithm 1 represents the EOSA optimization.

2.4. Performance Evaluation Metrics

In the assessment of a regression model's performance, it is standard practice to utilize the following metrics:

1. **Coefficient of Determination** $(\mathbf{R}^2)$ **:**

   Frequently represented as $(\mathbf{R}^2)$, this metric quantifies the proportion of variance in the dependent variable explained by independent variables in a statistical model, illustrated by the following formula:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (t_i - \bar{w})(v_i - \bar{v})}{\sqrt{\left[\sum_{i=1}^{n} (v_i - w)^2\right] \left[\sum_{i=1}^{n} (v_i - \bar{v})^2\right]}} \right)^2 \tag{35}$$

2. **Error evaluation metrics (RMSE, MAE, and RAE):**

   RMSE (Root Mean Square Error) and MAE (Mean Absolute Error) are statistical measures employed to assess the mean magnitude and accuracy of disparities between predicted and observed values within a model. RMSE, specifically, highlights the square root of the squared differences. These metrics are formally depicted in Eqs. (36) and (37). Relative Absolute Error (RAE) is a statistical measure employed to assess the effectiveness and precision of predictive models, specifically focusing on their application in regression analysis and forecasting. RAE quantifies the degree to which a model's predictions correspond to the observed values, serving as a valuable gauge of model performance. RAE is denoted by Eq. (38).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (v_i - w_i)^2} \tag{36}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} (v_i - w_i)^2 \tag{37}$$

$$RAE = \frac{\left[\sum_{i=1}^{n} (v_{i-} w_i)^2\right]^{1/2}}{\left[\sum_{i=1}^{n} w_i^2\right]^{1/2}} \tag{38}$$

3. **RMSE-observations standard deviation ratio (RSR):**

   The RMSE-observations standard deviation ratio (RSR) is a model evaluation metric that standardizes RMSE using the standard deviation of observed values. RSR ranges from 0 to infinity, effectively normalizing RMSE based on observed data variability. Eq. (39) indicates RSR:

$$RSR = \frac{\left[\sqrt{\sum_{i=1}^{n} (w_i - v_i)^2}\right]}{\left[\sqrt{\sum_{i=1}^{n} (w_i - \bar{w})^2}\right]} \tag{39}$$

   In all equations:

- n : quantity of samples,
- v_i : denotes the individual predicted value,
- $\bar{v}$: indicates the mean of the predicted values,
- w_i : stands for the experimentally measured value,
- $\bar{w}$: represents the average of the experimentally measured values.

3. Results

In this study, the LSSVR model was utilized to forecast heating load, and the performance of this model was improved through the application of 2 distinct optimization algorithms: AOA and EOSA. The prediction is applied in three phases including train, validation, and test. The 70% of samples related to train phases, 15% for validation, and 15% for testing.

The data presented in Table 2 demonstrates that the R^2 falls within an acceptable range. The lowest R^2 value is 0.955 , which corresponds to the LSSVR model, while the highest R^2 value is 0.988 , associated with the LSEO model.

RMSE and MAE are metrics used to measure the extent of error based on their respective definitions. When these measures have lower values, it signifies improved model performance. The RMSE and MAE values for the LSSVR model are 1.990 and 1.736, respectively, while the lowest of these values is for the LSEO model, equal to 1.118 and 0.863. The LSAO model is in second place for model performance with values of 1.545 and 1.282. furthermore, RAE is a measure used to evaluate the accuracy of a predictive model. Therefore, it can be confirmed that a decreased error value demonstrates enhanced model performance. Table 2 shows that the LSEO model performs significantly better, scoring 126.45 , while the LSSVR model lags at 208.83. The LSEO model, with an improvement of about 39%, surpasses the LSSVR model significantly.

RSR is a metric used to assess the precision of a predictive model in comparison to real values within a scale from zero to one. When a parameter gradually approaches zero, it signifies a strong level of agreement. In this situation, it is evident that LSEO demonstrates the highest performance with a value of 0.121 , the LSAO model follows in second place with a value of 0.153 , and the LSSVR model ranks last with a value of 0.189 .

The scatter plot in Fig. 1 visually illustrates the agreement between observed values and their predictions, employing the RMSE and R^2 metrics. RMSE manages data dispersion, with lower values indicating denser, more concentrated data. In contrast, R^2 optimizes the positioning of data points toward the central axis to improve the fit. The simple LSSVR model consistently has lower R^2 values in all three training, validation, and testing stages. Based on the plots, it can be concluded that the LSEO model outperforms the other 2 models, with LSAO and LSSVR taking the second and third positions. For instance, in the R^2 values of the training phase, the models have the following values:

LSSVR at 0.968 , LSEO at 0.988 , and LSAO at 0.980. The performance of the LSEO model is nearly 2% better than the LSSVR model. In the training phase, LSEO achieved an RMSE value of 1.12, representing a roughly 38% reduction compared to LSSVR's RMSE.

Fig. 2 visually compares the RMSE, R^2, and MAE parameters across models during three distinct phases. Each parameter is portrayed as a percentage within its designated section of the pie chart. Regarding RMSE, the LSEO model demonstrates superior performance in the training, validation, and testing phases, with percentages of 25.41%, 27.69%, and 27.18% on each pie chart. This level of performance surpasses that of the LSSVR and LSAO models. Based on the definition of R^2, a model occupying a larger percentage of the pie chart signifies superior performance among the models. Examining the percentages for this parameter reveals that all three models fall within an acceptable range. However, the LSEO model stands out with the best performance, scoring 33.72% $(R^2 = 0.98)$ for the test section. LSAO follows closely with 33.25% $(R^2 = 0.97)$, and LSSVR secures the third position with 32.97%. As mentioned earlier, a reduced error suggests that the predictions closely align with the actual values. Consequently, LSEO consistently performs exceptionally, with MAE error measuring 24.2%, 25.51%, and 25.25% (equivalent to 0.86, 1.05, 1.05) for the training, validation, and test phases.

Fig. 3 presents a histogram that compares the error rate percentages for the models developed in the study, specifically focusing on the frequency of values near zero per cent. The error range is similar for LSEO and LSAO models but larger for LSSVR. LSEO has a higher occurrence of errors, measuring less than 0.07 , and a denser distribution of errors extending beyond the zero interval. On the contrary, LSAO has a value below 0.06 , while LSSVR records a value below 0.05 .

Fig. 4 presents a multi-line diagram visualizing the error percentages for the models. The horizontal axis represents the sample count, and the vertical axis is divided into three sections: the blue section depicts the error rate of model LSSVR. In contrast, the red and green sections correspond to models LSEO and LSAO, respectively. The error percentage range for LSSVR extends approximately from -50 to 50 throughout the training, validation, and testing phases. On the other hand, the range of values for LSEO extends from slightly below -20 to just above 20, whereas the LSAO model's range spans from approximately -30 to 40.

4. Discussion

4.1. Findings

The findings of this study indicate that the application of ML approaches, specifically LSSVR models, significantly contributes to promoting energy efficiency and adopting sustainable building practices. Unlike traditional methods, which often rely on simplified models, this study focuses on regression tasks aimed at predicting continuous numerical values, such as heating loads in buildings.

The study evaluates three variations of the LSSVR model: a single model, an enhanced version using the LSAO, and an improved version employing the LSEO. These findings highlight the LSEO model's capacity to accurately predict energy consumption in buildings, facilitating anticipatory modeling for energy usage and advocating sustainability and ecological conservation practices. The study's significance lies in its contribution to energy efficiency and sustainability, its utilization of ML for regression tasks, the identification of a top-performing model, and its potential to inform proactive energy management and sustainable building practices.

4.2. Implications for Energy Efficiency and Sustainability

- Capacity of LSEO Model in Predicting Energy Consumption:

  The study underscores the remarkable accuracy of the LSEO model in forecasting heating load in buildings. This precision holds immense potential for optimizing energy consumption by ensuring heating systems operate at their most efficient levels.

- Advancements in Anticipatory Modeling for Energy Usage:

  By leveraging the LSEO model's predictive prowess, proactive energy management strategies can be employed. Anticipatory adjustments to heating systems based on predicted loads can lead to significant energy savings and contribute to sustainable energy practices.

- Advocacy for Sustainability and Ecological Conservation:

  Accurate heating load prediction facilitated by the LSEO model aligns with broader sustainability objectives. Through optimized heating system operation and reduced energy consumption, there's potential for substantial reductions in carbon emissions and environmental impact.

4.3. Significance of Study

- Contribution to Energy Efficiency and Sustainability:

Table 2. The result of developed models for LSVVR.

Model	Phase	Index values				
		RMSE	R	MAE	RSR	RAE
LSSVR	Train	1.829	0.968	1.475	0.180	135.02
	Validation	1.990	0.961	1.621	0.202	539.33
	Test	2.129	0.955	1.736	0.213	38.74
	All	1.901	0.965	1.536	0.189	208.83
LSEO	Train	1.118	0.988	0.863	0.110	83.36
	Validation	1.451	0.980	1.046	0.147	113.38
	Test	1.427	0.980	1.050	0.143	24.25
	All	1.223	0.985	0.918	0.121	126.45
LSAO	Train	1.452	0.980	1.228	0.143	101.21
	Validation	1.800	0.967	1.435	0.183	377.49
	Test	1.695	0.971	1.373	0.170	26.18
	All	1.547	0.977	1.281	0.153	159.90

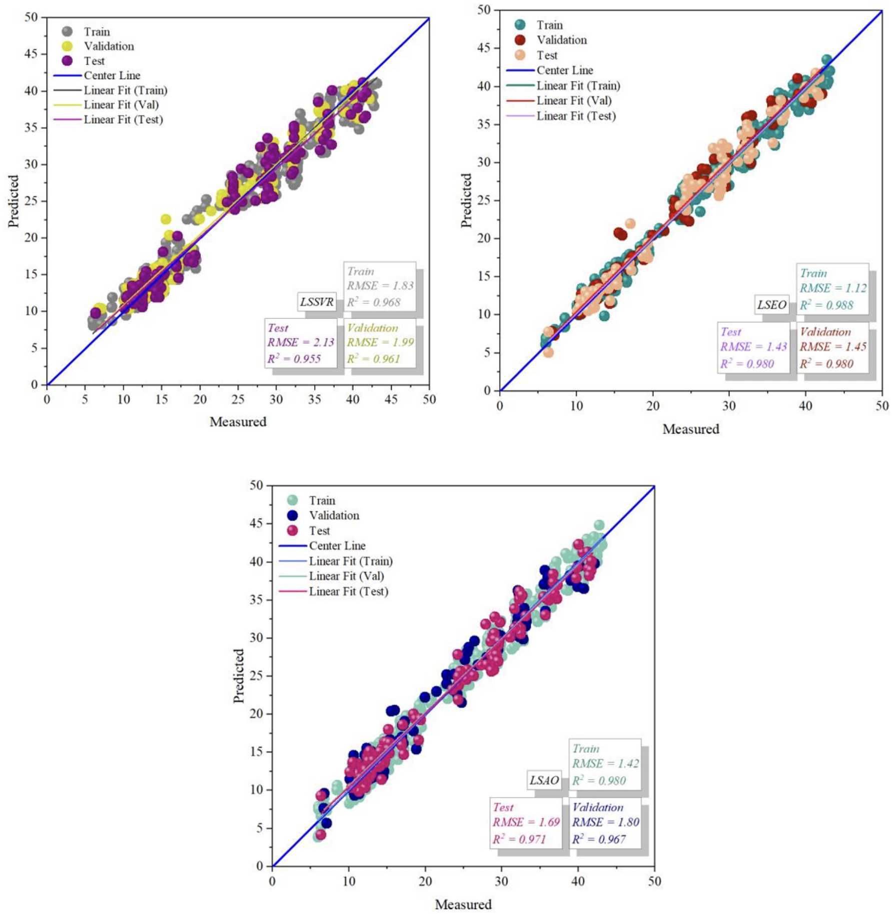

Fig. 1. The scatter plot for developed hybrid models.

This study makes a significant stride towards enhancing energy efficiency and sustainability within the building sector. Utilizing ML techniques for heating load prediction aligns with global efforts to curb energy consumption and combat climate change.

- Utilization of ML for Regression Tasks:

The study highlights the importance of employing ML approaches, particularly in regression tasks, for accurate heating load prediction. This departure from traditional methods underscores ML's ability to handle complex relationships and improve prediction accuracy.

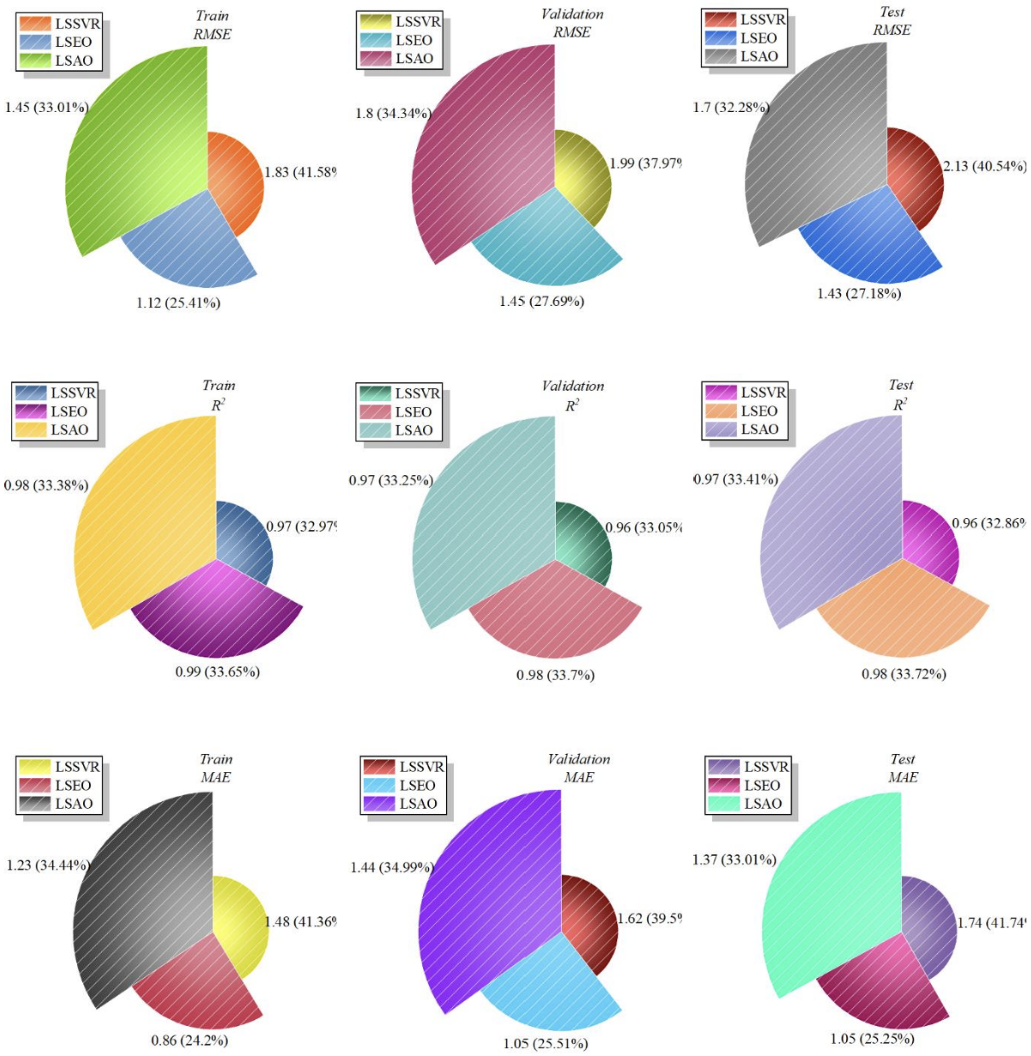

Fig. 2. Comparison between models RMSE, R^2, MAE.

- Potential for Proactive Energy Management and Sustainable Building Practices:

  The findings of this study offer practical implications for stakeholders in the building industry. By employing the LSEO model, decision-makers can inform strategies related to heating system design and energy management, promoting sustainable building practices and reducing environmental impact.

4.4. Comparison

The comparison analysis between the findings of the present study and those reported in prior papers is presented in Table 3. The articles provided encompass the works of Moradzadeh et al. [40], Roy et al. [41], Gong et al. [42], and Afzal et al. [43]. The comparison is conducted using 2 primary index values, namely the RMSE and the R^2. Moradzadeh et al. obtained an RMSE of 0.4832 and an R^2 value of 0.9993 . Roy et al. exhibit a remarkable RMSE of 0.059 , accompanied by an R^2 value of 0.99 . The authors Gong et al. documented an RMSE value of 0.1929 and an R^2 value of 0.9882 . The study conducted by Afzal et al. demonstrated an RMSE of 1.4122 and an R^2 of 0.9806 . In contrast, the present investigation produced an RMSE of 1.223 and an R^2 value of 0.985 . The findings of this study illustrate the competitive performance of the current research, specifically about RMSE. The RMSE falls within the range observed in the published publications, but the R^2 value is greater, suggesting a high level of predictive accuracy.

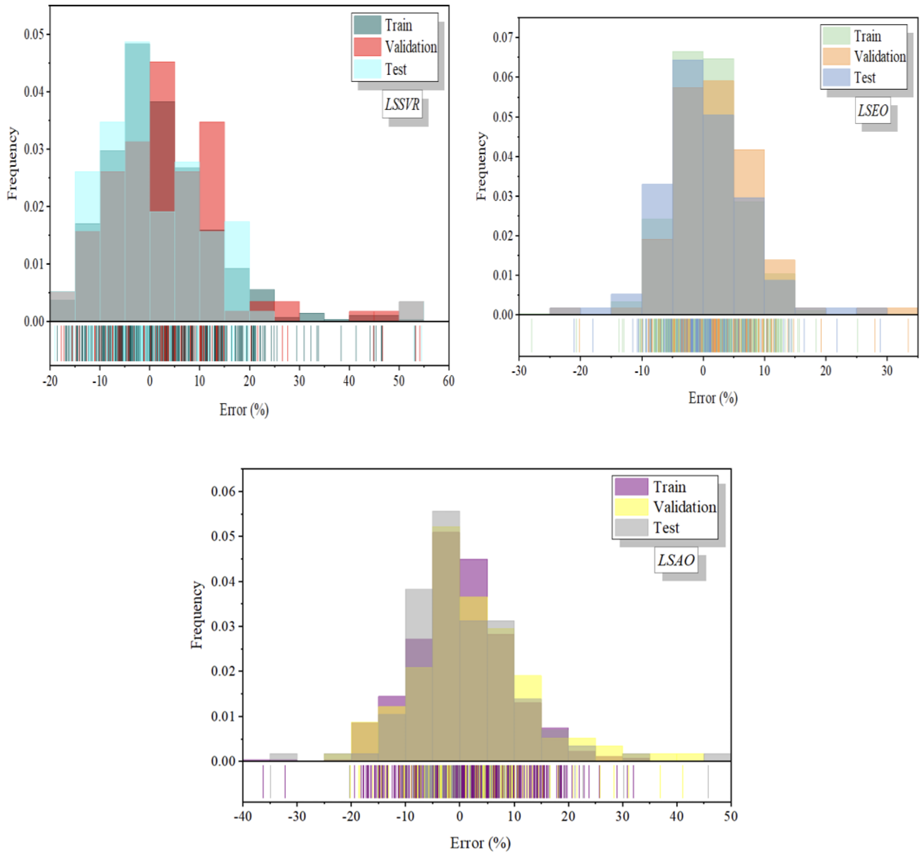

Fig. 3. The error percentage for the hybrid models is based on the Histogram plot.

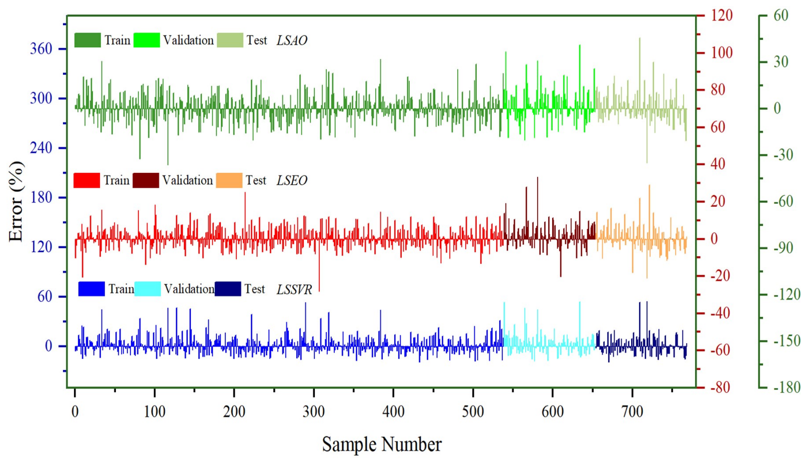

Fig. 4. The Column of errors among the developed models.

5. Conclusion

In conclusion, this study represents a significant advancement in promoting energy efficiency and sustainable building practices through the utilization of ML techniques for regression tasks in predicting continuous numerical values. Specifically, the estimation of heating loads in buildings was conducted using Least Squares Support Vector Regression (LSSVR) models, including a single model, an enhanced version with the Arithmetic Optimization Algorithm (LSAO), and an improved iteration employing the Ebola Optimization Search Algorithm (LSEO). Through rigorous analysis of input variables and comprehensive per-

Table 3. Comparison between the published and presented articles.

Articles	Index values	
	RMSE	R^2
Moradzadeh et al. [40]	0.4832	0.9993
Roy et al. [41]	0.059	0.99
Gong et al. [42]	0.1929	0.9882
Afzal et al. [43]	1.4122	0.9806
Present Study	1.223	0.985

formance evaluations, the LSEO model emerged as the top performer, boasting an impressive Root Mean Square Error (RMSE) of 1.223. This outperformed the RMSE values of 1.901 and 1.547 obtained by the LSSVR and LSAO models, respectively. Moreover, the superiority of the LSEO model was further confirmed by metric values, notably achieving an R-squared $\left(R^2\right)$ value of 0.985 . This represented a substantial increase of 2.07% and 0.8% compared to the R^2 values of 0.965 and 0.977 for the LSSVR and LSAO models, respectively. These simulation results underscore the LSEO model's remarkable capacity to accurately predict energy consumption in buildings. Furthermore, its implementation facilitates anticipatory modeling for energy usage, thereby advocating sustainability and ecological conservation by optimizing heating system operations and reducing energy consumption. The study's contributions are manifold, including its significant impact on energy efficiency and sustainability, the innovative use of ML for regression tasks, the identification of a top-performing model, and its potential to inform proactive energy management and sustainable building practices. These findings render this study a valuable resource for professionals engaged in the field of green construction and energy conservation, offering insights and methodologies to drive positive change toward a more sustainable built environment.

References

[1] I. Dadon, (2019) *"Planning the Second Generation of Smart Cities: Technology to handle the pressures of urbanization"* **IEEE Electrification Magazine 7**(3): 6–15. DOI: 10.1109/MELE.2019.2925729.

[2] B. Sadaghat, A. Javadzade Khiavi, B. Naeim, E. Khajavi, H. Sadaghat, and A. R. Taghavi Khanghah, (2023) *"The utilization of a naıve bayes model for predicting the energy consumption of buildings"* **Journal of Artificial Intelligence and System Modelling *1***(01): 73–91. DOI: 10.22034/JAISM.2023.422292.1003.

[3] R. Jing, M. Wang, R. Zhang, N. Li, and Y. Zhao, (2017) *"A study on energy performance of 30 commercial office buildings in Hong Kong"* **Energy and Buildings *144***: 117–128. DOI: 10.1016/j.enbuild.2017.03.042.

[4] G. Yang and X. Zhai, (2019) *"Optimal design and performance analysis of solar hybrid CCHP system considering influence of building type and climate condition"* **Energy *174***: 647–663. DOI: 10.1016/j.energy.2019.03.001.

[5] L. Zhang, J. Wen, Y. Li, J. Chen, Y. Ye, Y. Fu, and W. Livingood, (2021) *"A review of machine learning in building load prediction"* **Applied Energy *285***: 116452. DOI: 10.1016/j.apenergy.2021.116452.

[6] X. Zhang, R. Yan, R. Zeng, R. Zhu, X. Kong, Y. He, and H. Li, (2022) *"Integrated performance optimization of a biomass-based hybrid hydrogen/thermal energy storage system for building and hydrogen vehicles"* **Renewable Energy *187***: 801–818. DOI: 10.1016/j.renene.2022.01.050.

[7] X.-N. Bui, H. Moayedi, and A. S. A. Rashid, (2020) *"Developing a predictive method based on optimized M5Rules–GA predicting heating load of an energy-efficient building system"* **Engineering with Computers *36***: 931–940. DOI: 10.1007/s00366-019-00739-8.

[8] X. Li and R. Yao, (2020) *"A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour"* **Energy *212***: 118676. DOI: 10.1016/j.energy.2020.118676.

[9] J. Guo, S. Yun, Y. Meng, N. He, D. Ye, Z. Zhao, L. Jia, and L. Yang, (2023) *"Prediction of heating and cooling loads based on light gradient boosting machine algorithms"* **Building and Environment *236***: 110252. DOI: 10.1016/j.buildenv.2023.110252.

[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] J. Song, L. Zhang, G. Xue, Y. Ma, S. Gao, and Q. Jiang, (2021) *"Predicting hourly heating load in a district heating system based on a hybrid CNN-LSTM model"* **Energy and Buildings *243***: 110998. DOI: 10.1016/j.enbuild.2021.110998.

[12] R. Chaganti, F. Rustam, T. Daghriri, I. d. l. T. 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**(19): 7692. DOI: 10.3390/s22197692.

[13] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, and X. Zhao, (2018) *"A review of data-driven approaches for prediction and classification of building energy consumption"* **Renewable and Sustainable Energy Reviews *82***: 1027–1047. DOI: 10.1016/j.rser.2017.09.108.

[14] M. Rana, S. Sethuvenkatraman, and M. Goldsworthy, (2022) *"A data-driven approach based on quantile regression forest to forecast cooling load for commercial buildings"* **Sustainable Cities and Society *76***: 103511. DOI: 10.1016/j.scs.2021.103511.

[15] 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.rser.2015.04.020.

[16] 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.

[17] Y. Feng, Q. Duan, X. Chen, S. S. Yakkali, and J. Wang, (2021) *"Space cooling energy usage prediction based on utility data for residential buildings using machine learning methods"* **Applied energy *291***: 116814. DOI: 10.1016/j.apenergy.2021.116814.

[18] 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.

[19] J. Cai, W. Yu, B. Li, R. Yao, T. Zhang, M. Guo, H. Wang, Z. Cheng, J. Xiong, Q. Meng, et al., (2019) *"Particle removal efficiency of a household portable air cleaner in real-world residences: A single-blind cross-over field study"* **Energy and Buildings *203***: 109464. DOI: 10.1016/j.enbuild.2019.109464.

[20] X. Zhou, W. Lin, R. Kumar, P. Cui, and Z. Ma, (2022) *"A data-driven strategy using long short term memory models and reinforcement learning to predict building electricity consumption"* **Applied Energy *306***: 118078. DOI: 10.1016/j.apenergy.2021.118078.

[21] G. Baasch, P. Westermann, and R. Evins, (2021) *"Identifying whole-building heat loss coefficient from heterogeneous sensor data: An empirical survey of gray and black box approaches"* **Energy and Buildings *241***: 110889. DOI: 10.1016/j.enbuild.2021.110889.

[22] H. Liu, C. Chen, Z. Guo, Y. Xia, X. Yu, and S. Li, (2021) *"Overall grouting compactness detection of bridge prestressed bellows based on RF feature selection and the GA-SVM model"* **Construction and Building Materials *301***: 124323. DOI: 10.1016/j.conbuildmat.2021.124323.

[23] Y. Liu, H. Chen, L. Zhang, and Z. Feng, (2021) *"Enhancing building energy efficiency using a random forest model: A hybrid prediction approach"* **Energy Reports *7***: 5003–5012. DOI: 10.1016/j.egyr.2021.07.135.

[24] D. Tien Bui, H. Moayedi, D. Anastasios, and L. Kok Foong, (2019) *"Predicting heating and cooling loads in energy-efficient buildings using two hybrid intelligent models"* **Applied Sciences *9***(17): 3543. DOI: 10.3390/app9173543.

[25] V. V. Mokeev. "Prediction of heating load and cooling load of buildings using neural network". In: *2019 International Ural Conference on Electrical Power Engineering (UralCon)*. IEEE. 2019, 417–421. DOI: 10.1109/URALCON.2019.8877655.

[26] P. Dai, P. Zhou, Y. Liang, et al., (2019) *"Adaptive nonlinear predictive control based on multi output least squares support vector regression modeling and its application"* **Control theory and application *36***(1): 43–52.

[27] P. Dai, P. Zhou, Y. Liang, et al., (2019) *"Adaptive nonlinear predictive control based on multi output least squares support vector regression modeling and its application"* **Control theory and application *36***(1): 43–52.

[28] G. G. Tejani, B. Sadaghat, and S. Kumar, (2023) *"Predict the maximum dry density of soil based on individual and hybrid methods of machine learning"* **Advances in engineering and intelligence systems *2***(03): 98–109. DOI: 10.22034/AEIS.2023.414188.1129.

[29] T.-L. Le, D.-N. Truong, and D. S. Nguyen. "Nature inspired based optimize combination of time series and machine learning model for predict energy consumption of a residence". In: *AIP Conference Proceedings*. ***2560***. 1. AIP Publishing. 2023. DOI: 10.1063/5.0124782.

[30] A. Tsanas and A. Xifara, (2012) *"Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools"* **Energy and buildings *49***: 560–567. DOI: 10.1016/j.enbuild.2012.03.003.

[31] V. N. Vapnik, (1998) *"The nature of statistical learning"*:

[32] M. Ayubi Rad and M. S. Ayubirad, (2017) *"Comparison of artificial neural network and coupled simulated annealing based least square support vector regression models for prediction of compressive strength of high-performance concrete"* **Scientia Iranica *24***(2): 487–496.

[33] Q. B. Pham, T.-C. Yang, C.-M. Kuo, H.-W. Tseng, and P.-S. Yu, (2019) *"Combing random forest and least square support vector regression for improving extreme rainfall downscaling"* **Water *11***(3): 451. DOI: 10.3390/w11030451.

[34] C.-W. Hsu, C.-C. Chang, C.-J. Lin, et al. *A practical guide to support vector classification*. 2003.

[35] B. G. Aiyer, D. Kim, N. Karingattikkal, P. Samui, and P. R. Rao, (2014) *"Prediction of compressive strength of self-compacting concrete using least square support vector machine and relevance vector machine"* **KSCE Journal of Civil Engineering *18***: 1753–1758. DOI: 10.1007/s12205-014-0524-0.

[36] M. K. Habib and A. K. Cherri, (1998) *"Parallel quaternary signed-digit arithmetic operations: addition, subtraction, multiplication and division"* **Optics & Laser Technology *30***(8): 515–525. DOI: 10.1016/S0030-3992(99)00004-3.

[37] F. A. Hashim, K. Hussain, E. H. Houssein, M. S. Mabrouk, and W. Al-Atabany, (2021) *"Archimedes optimization algorithm: a new metaheuristic algorithm for solving optimization problems"* **Applied intelligence *51***: 1531–1551. DOI: 10.1007/s10489-020-01893-z.

[38] O. N. Oyelade, A. E.-S. Ezugwu, T. I. Mohamed, and L. Abualigah, (2022) *"Ebola optimization search algorithm: A new nature-inspired metaheuristic optimization algorithm"* **IEEE Access *10***: 16150–16177. DOI: 10.1109/ACCESS.2022.3147821.

[39] O. N. Oyelade and A. E. Ezugwu, (2021) *"Ebola Optimization Search Algorithm (EOSA): A new metaheuristic algorithm based on the propagation model of Ebola virus disease"* **arXiv preprint arXiv:2106.01416**: DOI: 10.48550/arXiv.2106.01416.

[40] 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***(11): 3829. DOI: 10.3390/app10113829.

[41] 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.

[42] 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.

[43] 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.