Aalborg Universitet

# Heating and Cooling Loads Forecasting for Residential Buildings Based on Hybrid Machine Learning Applications

*A Comprehensive Review and Comparative Analysis*

Moradzadeh, Arash ; Mohammadi-Ivatloo, Behnam ; Abapour, Mehdi; Anvari-Moghaddam, Amjad; Roy, Sanjiban Sekhar

*Published in:*
IEEE Access

*DOI (link to publication from Publisher):*
10.1109/ACCESS.2021.3136091

*Publication date:*
2022

*Document Version*
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

*Citation for published version (APA):*
Moradzadeh, A., Mohammadi-Ivatloo, B., Abapour, M., Anvari-Moghaddam, A., & Roy, S. S. (2022). Heating and Cooling Loads Forecasting for Residential Buildings Based on Hybrid Machine Learning Applications: A Comprehensive Review and Comparative Analysis. *IEEE Access*, *10*, 2196-2215. https://doi.org/10.1109/ACCESS.2021.3136091

**General rights**
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

**Take down policy**
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: December 04, 2025

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2017.Doi Number*

# Heating and Cooling Loads Forecasting for Residential Buildings Based on Hybrid Machine Learning Applications: A Comprehensive Review and Comparative Analysis

**Arash Moradzadeh[1], Student Member, IEEE, Behnam Mohammadi-Ivatloo[1,2,*], Senior Member, IEEE, Mehdi Abapour[1], Member, IEEE, Amjad Anvari-Moghaddam[2], Senior Member, IEEE, Sanjiban Sekhar Roy[3], Member, IEEE**

[1]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
[2]Integrated Energy Systems Laboratory, Department of Energy (AAU Energy), Aalborg University, 9220 Aalborg, Denmark
[3]School of Computer Science and Engineering, Vellore Institute of Technology

Corresponding author: Behnam Mohammadi-Ivatloo (e-mail: bmohammadi@tabrizu.ac.ir; bmo@et.aau.dk).

This research is part of the “HeatReFlex-Green and Flexible Heating/Cooling” project (www.heatreflex.et.aau.dk) funded by Danida Fellowship Centre and the Ministry of Foreign Affairs of Denmark under the grant no. 18-M06-AAU.

**ABSTRACT** Prediction of building energy consumption plays an important role in energy conservation, management, and planning. Continuously improving and enhancing the performance of forecasting models is the key to ensuring the performance sustainability of energy systems. In this connection, the current paper presented a new improved hybrid model of machine learning application for forecasting the cooling load (CL) and the heating load (HL) of residential buildings after studying and analyzing various types of CL and HL forecasting models. The proposed hybrid model, called group support vector regression (GSVR), is a combination of group method of data handling (GMDH) and support vector regression (SVR) models. To forecast CL and HL, this study also made use of base methods such as back-propagation neural network (BPNN), elastic-net regression (ENR), general regression neural network (GRNN), k-nearest neighbors (kNN), partial least squares regression (PLSR), GMDH, and SVR. The technical parameters of the building were utilized as input variables of the forecasting models, and the CL and HL were adopted as the output variables of each network. All models were saved in the form of black box after training and initial testing. Finally, comparative analysis was performed to assess the predictive performance of the suggested model and the well-known basic models. Based on the results, the proposed hybrid method with high correlation coefficient (R) and minimal statistical error values provided the most optimal prediction performance.

**INDEX TERMS** Heating load (HL), cooling load (CL), forecasting, machine learning, artificial neural network (ANN), regression

## ABBREVIATIONS

| | | | |
|---|---|---|---|
| HVAC | Heating ventilation and air conditioning | GPR | Gaussian process regression |
| EBP | Performance of buildings | MPMR | Minimax probability machine regression |
| HL | Heating load | MLR | Multiple linear regression |
| CL | Cooling load | ELN | Elastic net |
| WSHP | Water source heat pump | RF | Random forests |
| GSVR | Group support vector regression | XGB | Extreme gradient boosting trees |
| GMDH | Group method of data handling | RNN | Recurrent neural networks |
| SVR | Support vector regression | GRU | Gated recurrent units |
| BPNN | Back propagation neural network | LSTM | Long short-term memory |



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

| | | | |
|---|---|---|---|
| ENR | Elastic-net regression | SLFF-ANN | Single layered feed-forward artificial neural network |
| GRNN | General regression neural network | CART | Classification and regression tree |
| kNN | k-nearest neighbor | CHAID | Chi-squared automatic interaction detector |
| PLSR | Partial least squares regression | GLR | General linear regression |
| ELM | Extreme learning machine | AR | Auto regressive |
| SVM | Support vector machine | GA-SVR | Genetic algorithm support vector regression |
| DT | Decision tree | GA-WD-SVR | Genetic algorithm wavelet decomposition support vector regression |
| MILP | Mixed-integer linear program | EDL | Extreme deep learning |
| OSELM | Online sequential extreme learning machine | LR | Linear regression |
| MARS | Multivariate adaptive regression splines | ETR | Extra tree regressor |
| ANN | Artificial neural network | RC | Resistance Capacitance |
| LS-SVM | Least square support vector machine | GBDT | Gradient Boosted Decision Tree |
| DNN | Deep neural network | XGBoost | Extreme Gradient Boosting |
| FFNN | Feed forward neural network | TB | Tree bagger |
| RBFNN | Radial basis function neural network | BoostedT | Boosted tree |
| MLP | Multilayer perceptron | BaggedT | Bagged tree |
| ANFIS | Adaptive neuro-fuzzy interference system | VAF | Variance accounted for |
| WDSVR | Wavelet decomposition support vector regression | CV | Coefficient of variation |
| PCA | Principal component analysis | $R$ | Correlation coefficient |
| KPCA | Kernel principal component analysis | $R^2$ | Accuracy |
| ARX | Autoregressive with exogenous | MSE | Mean square error |
| MNR | Multiple nonlinear regression | RMSE | Root mean square error |
| PENN | Probabilistic entropy-based neural | MAE | Mean absolute error |
| LSVM | Least square support vector machine | MAPE | Mean absolute percentage error |
| GBM | Gradient boosted machine | MRE | Mean relative error |
| NN | Neural network | APNN | Multi-layer hybrid |

## I. INTRODUCTION

Today, the need for energy has increased in many folds, and the significant energy consumption worldwide is mainly related to residential and commercial sectors. Therefore, managing industries such as transportation and construction and saving energy can be challenging tasks [1], [2]. Recent research has indicated that residential buildings have a significant share of energy consumption because of the increased population [3], [4]. It is highly necessary to have complete information concerning the performance of buildings in order to manage and optimize their energy consumption. Initially, this requires identifying the energy sources and end-uses of the building [1]. District heating supply, electricity, and natural gas are the important energy resources in a building; however, the main end-use applications include heating ventilation and air conditioning (HVAC), lighting, elevators, domestic hot water, and kitchen equipment. Of note, among the aforementioned energy resources and major end-uses of buildings, HVAC operation scheduling and indoor and outdoor conditions are two effective factors in building performance assessment [5], [6]. Considered as a fundamental infrastructure in a building, HVAC plays a major role by either adding or removing the cooling load (CL) and heating load (HL) to/from the indoor climate of residential buildings. A major concern is that these systems consume approximately 40% of the total energy, particularly in office buildings [7], [8]. Nonetheless, very few fundamental solutions have been developed to improve the performance and management of HVAC systems. Due to the significant impact of meteorological factors on HL and CL, the performance of HVAC systems cannot be adapted to external climate change; on the other hand, improper performance of these systems can augment the energy consumption and reduce convenience in terms of cooling and heating [9], [10]. Improving the energy efficiency of buildings through sustainable construction management in urban environments and accurate dynamic load forecasting can be key steps to ameliorating the performance of HVAC systems and saving energy in residential buildings [11], [12].

Improper design and structure of buildings lead to the overuse of these technologies, high energy consumption, and an increase in the emission of carbon dioxide by about 40 %. Accordingly, widespread concerns regarding energy loss and its negative impacts on the environment have made the energy performance of buildings (EPB) the subject of more recent research around the globe [13]–[15]. Designing energy-efficient buildings with enhanced energy conservation properties are among the most important approaches to energy management in order to decrease energy demand and save energy. This can be achieved by initial forecasts of HL and CL in residential buildings. To forecast the required cooling and heating capacity, construction designers need information concerning building specifications and weather conditions in the area [16], [17]. Temperature, solar radiation, wind speed,



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



humidity, and pressure are among the most influential climatic factors in forecasting the CL and HL of buildings. When considering the CL and HL of a building, the following should be taken into account: the relative compactness of buildings, the roof size, the wall surface, the glazing area, height of the roof, the number of walls, and area [18], [19].

Currently, building energy simulation tools are widely utilized in various fields to facilitate an efficient design, optimal performance of energy-efficient buildings, and comparison of buildings with identical scales, where the effect of a single modified parameter is observed over a range of values. The simulation results regarding the design and comparison in some work have often been able to accurately represent the actual measurements [20], [21].

Generally, using building energy simulation software can be a good solution for analyzing the impact of building design indicators; however, this method is time-consuming and requires expert users to perform the simulations, and sometimes there is inconsistency in the accuracy of the estimated results in various building simulation software packages [20]. Hence, in some studies, modern methods such as statistical analysis, artificial neural networks, and machine learning are implemented to forecast the CL and HL of buildings and analyze the effect of various building parameters [15, 16]. The advantage of these methods is that after sufficiently training the model, an exact and reliable response can be obtained even by changing certain design parameters of the building. Furthermore, methods such as statistical analysis reinforce our understanding concerning the impact of quantities that are differently focused by designers or architects.

Various data mining techniques such as principal component analysis (PCA) [24], extreme learning machine (ELM) [10], [25], support vector machine (SVM) applications [26]–[32], k-means [33], deep learning techniques [31], [32], [34]–[40], decision tree (DT) [13], diverse regression approaches and artificial neural networks [16], [20], [29], [41]–[64], and hybrid methods based on regression models [32], [65]–[67] have been used in EPB and forecasting the required energy of residential buildings. A valuable study [68], examines facade retrofitting measures to minimize the demand for heating and cooling in residential and commercial buildings.

A mixed-integer linear program (MILP) in [69] has also been applied to an integrated model prediction for optimal operation of the HVAC in large-scale buildings. The relationship between the design and structural features of a building with HL and CL using new ELM-based methods and its variants in online sequential ELM (OSELM) was studied to forecast the HL and CL in [18]. The CL and HL of a residential building in [25] were forecasted via multivariate adaptive regression splines (MARS), ELM, and hybrid model, where the structural parameters of the building were considered as network inputs. In another study [70], to design an energy-efficient building, ELM method was employed to forecast the CL and HL of residential buildings. In [26], the energy consumption of a building in a tropical region was predicted by accounting for three samples of weather data as input characteristics for the support vector machine. In [27], the CL of an office building was forecasted using artificial neural network (ANN) and SVM methods; their results were ultimately compared to specify the impact of input variables on the prediction accuracy. CL was predicted to optimize an HVAC system [28] by considering the meteorological information for the least square support vector machine (LS-SVM). In [37], a deep neural network (DNN) was employed along with the structural features of the building as inputs to predict CL and HL. A short-term CL prediction in an educational building in Hong Kong was performed in [34] by the deep neural network. In their work, time variables and temperature data were considered as input variables. Another study in Hong Kong [34] carried out a short-term CL prediction of an educational building via deep recurrent neural network-based strategies; they considered two types of time variables and indoor/outdoor ambient temperature data as network input variables. In [36], a deep learning application, called long short-term memory (LSTM) network, was implemented to predict the lighting loads, miscellaneous electrical loads, number of occupants, and internal heat gains in two office buildings. In [13], the energy demand forecasting model of a residential building was expanded using the DT method. [58] employed a feed-forward neural network (FFNN) for thermal comfort and saved 36.5% of the energy in the building. [59] predicted the HL of a building for above-normal energy consumption detection in a university campus with ANN techniques such as FFNN, radial basis function network (RBFN), and adaptive neuro-fuzzy interference system (ANFIS); in this regard, the heating consumption of the previous day, temperature data, and day of the week were selected as input variables for networks. CL and HL prediction of a building in [12, 44] was performed through the use of ANN in order to manage the HVAC system; in these studies, 11 air-handling units and meteorological data were utilized as input variables, respectively. Using back propagation neural network (BPNN), [61]forecasted the cooling demand in a building for operational planning of stable energy supply and energy savings. This work considered air temperature and relative humidity as additional inputs to the network. The CL of a building was predicted in [62] so as to determine the daily operation of the building, and it was optimized using ANN method. In their approach, the network input variables were water temperature of in indoor equipment, outdoor humidity/temperature, chilled water prepared in ice tanks or elsewhere, electric current used by chiller, on/off status of compressors, and date data. [63] forecasted the HL of a building by use of the MLP method and considering the meteorological data as network input variables. In another work [64], the MLP method was employed to forecast the CL and HL of a building in order to design an energy-efficient building, in which the



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access


meteorological and date data were considered as input variables. In [41], prediction of the CL of a building was done using an artificial neural network model called multilayer perceptron, and climatic data inputs were considered as parameters. In [16], the energy performance of a building was examined by forecasting the CL and HL via machine learning applications such as general linear regression, ANN, DT, support vector regression (SVR), and ensemble inference model. [42] forecasted the CL and HL of a building to discover the peak load of water source heat pump (WSHP) in the building environment and supply the heating and cooling demand; in this connection, six regression models of data mining methods were used, in which the meteorological data, previous load consumption, and time and date information were taken as the input variables of the network. For increasing the building energy efficiency, four regression applications, namely chaos-SVR, wavelet decomposition (WD)-SVR, SVR, and BP were utilized to forecast the CL of the building in [71]. In some studies [28, 29], CL prediction was performed in a non-residential building to enhance the operation of the HVAC system using a combination of PCA, kernel (k)-PCA, and SVR methods. In [57], autoregressive with exogenous (ARX) method was employed to predict the CL and HL of a building so as to manage supply and demand of energy; the meteorological data were considered as network input variables. The CL of a building in [45] was forecasted via general regression neural network (GRNN), in which the network input variable was the external hourly temperature for a 24 h period obtained from the Kuwait Institute for Scientific Research over the five years of data collection. In [46], to predict the hourly CL of a building, four methods (traditional BPNN, radial basis function neural network (RBFNN), GRNN and SVM) were used; finally, after comparing the results, SVM and GRNN were proven to have a better performance accuracy. [47] investigated the influence of the structural conditions of the building and its interior design on the cooling load of residential and office buildings. By developing different regression models, [48] predicted the HVAC system energy demand from the CL and HL demand of building. [34, 35] forecasted the CL and electric demand of commercial buildings in short-term and ultra-short-term models to manage the energy demand and improve the energy efficiency of HVAC systems via hybrid SVR. In [51], using the SVR method, the CL of a large office building in a coastal town of China was forecasted , and a new model of vector-based SVR was presented to increase the robustness and forecasting accuracy. Via a multiple nonlinear regression (MNR) model, [52] predicted the optimal operation of HVAC systems and short-term cooling loads; in their study, the network inputs were the physical characteristics of the building, meteorological information, and HVAC operational variables . [53] forecasted the CL of a building using a probabilistic entropy-based neural (PENN) model. In [54], the HL of a building was predicted through the use of a multiple regression method and considering certain climatic conditions such as sol-air temperature, thermal resistance, and surface-to-volume ratio as network inputs. In [55], CL and HL were predicted using neural networks and extraction of a black box model, which was the trained network. In their method, climate data include temperature, humidity, wind speed, and sunlight were used as input data. The prediction of HL and CL associated with residential building has been performed based on existing historical data and using different machine learning techniques called SVM, RF, FFNN, Gradient Boosted Regression Trees, XGBoost in [72]. In another study, indoor temperature was predicted using a network (called nonlinear autoregressive) based on back-propagation neural networks; the considered network inputs were meteorological parameters and HVAC equipment parameters [56]. In [65], a hybrid regression model of ANN and SVM applications was used to forecast the thermal load of 10 residential buildings located in Rottne, Sweden, in order to control district cooling and heating systems. In [73], a multi-layer hybrid model (APNN) has been proposed to forecast HL and CL of a residential building by considering the technical specifications of the building and climatic data as the input of the network. In this paper, the suggested hybrid technique is based on combination an ARX and a particle swarm optimization neural network. [66] presented a hybrid solution comprised of a gray box and a black box containing certain machine learning methods based on regression using internal heat gain and weather data as inputs in order to predict the mean indoor air temperatures. In another study [67], for energy planning, management, and conservation, a novel enhanced integration model (stacking model) was proposed to forecast the HL of two educational buildings in Tianjin, China. The heating demand estimating of buildings in [74] has been done by considering the weather conditions and presenting a Global Forecast System sflux model. In a valuable study [75], considering the two energy flexibility indicators called the efficiency of active demand response events and the available heat storage capacity, the thermal inertia potential of the three apartments have been analyzed to correct their heat load pattern. In this study, user behavior, the effects of building envelopes, and weather conditions are selected as three effective indicators in evaluations.

Based on the above literature, the CL and HL of buildings were mainly predicted by implementing various methods on different data. These methods are categorized according to the type of datasets and utilized them to predict the CL and HL of buildings. In the end, the accuracy of the presented methods was assessed for similar data. Table 1 classifies the studies performed based on various data. In this classification, the datasets were divided into four clusters, namely real measurements, simulated (Ecotect), simulated (Energy plus), and simulated (DeST). Real-world measurement data were specific to each study.



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



**TABLE. 1.** **Reviewing and categorizing the conducted research based on the various datasets**

| Type of dataset | References | Types of input variables in dataset | Building type | Purpose of prediction |
|---|---|---|---|---|
| Simulated (Ecotect) | [16], [18], [25], [37] | Relative compactness, surface area, wall area, roof area, orientation, glazing area, overall height, and glazing area distribution | Residential | Estimating energy consumption and energy-efficient building design |
| Simulated (Energy plus) | [48], [57], [64], [69], [70] | Daily average dry-bulb and wet-bulb temperatures, solar radiation and aperture, daylight aperture, overhang, side-fins projections, daily average clearness index, and date data | Residential and non-residential | Energy-efficient building design |
| Simulated (DeST) | [28] | Dry-bulb temperature, solar radiation, humidity | Non-residential | HVAC system optimization |
| Real measurement datasets | [13], [20], [26], [27], [33]–[35], [40], [40]–[45], [48], [50]–[52], [57]–[62], [64]–[66], [70], [72] | Different in various datasets | Residential and non-residential | Different in various studies |

As mentioned in the literature above, different studies have used different algorithms of ANNs, machine learning, deep learning, ELM, and a variety of hybrid solutions to forecast the CL and HL of buildings. In order to easily access the analysis provided in each of these studies, they were reviewed and categorized in Table 2 based on the type of algorithm used.

**TABLE. 2.** **Reviewing and categorizing the conducted research based on the employed algorithm**

| Method | Algorithm | References |
|---|---|---|
| ANN | MLP | [16], [18], [20], [25], [27], [37], [41], [42], [48], [59], [60], [62]–[65], [72] |
| | BPNN | [28], [46], [51], [57], [61], [71] |
| | RBFNN | [46], [59] |
| | GRNN | [45], [46] |
| | ANFIS | [59] |
| | PENN | [53] |
| Machine learning | SVM | [26], [27], [43], [44], [46], [65], [72] |
| | SVR | [16], [34], [48], [51], [67], [71] |
| | GPR | [37], [42] |
| | GPM | [37] |
| | MLR | [34], [42], [51], [52], [57] |
| | XGB | [34], [42], [67] |
| | GBM | [34] |
| | DT | [13] |
| | AR | [52], [57] |
| | ARX | [52], [57], [73] |
| | MNR | [52] |
| | KNN | [67] |
| Deep learning | DNN | [34], [37] |
| | RNN | [28], [35] |
| | LSTM | [31], [35], [36] |
| ELM | - | [18], [25], [70] |
| MILP | - | [69] |
| Hybrid models | - | [16], [18], [25], [43], [44], [49], [54], [58], [66], [67], [71], [73] |

As observed in Table 2, most studies typically used MLP, SVM, and SVR algorithms to forecast the CL and HL of buildings. However, a review of recent studies shows that researchers are increasingly turning to the idea of combining models to improve the prediction of CL and HL. The methods used to predict CL and HL have their own unique accuracy coefficient and error. In different studies, various types of statistical performance metrics have been utilized to evaluate the performance of each algorithm. Table 3 categorizes and reviews the conducted research according to the statistical performance metrics of each algorithm.

This paper reviewed most of the studies performed to predict the CL and HL of buildings based on different criteria. Table 4 depicts a comparative approach to assessing the effectiveness of each of the methods employed in the studies. Given that the comparison of results should be made for the same data used for each algorithm, these comparisons are made for the same datasets used in several studies.

**TABLE. 3.** **Reviewing and categorizing the conducted research based on the statistical performance metrics**

| Performance | References |
|---|---|
| MAE | [18], [34], [66], [73] |
| MAPE | [25], [42], [59], [65], [71], [73] |
| MSE | [26], [63] |
| RMSE | [16], [27], [28], [35], [43], [44], [46], [64], [70], [72], [73] |
| MRE | [60] |
| $R^2$ | [45], [48], [49], [52], [54], [58], [67] |
| R | [41], [51], [53] |
| CV | [57], [62] |
| VAF | [37] |



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



**TABLE. 4. Performance comparison of various algorithms implemented for the same data in predicting the CL and HL of buildings**

| Simulated (Energy plus) | | | | |
|---|---|---|---|---|
| Ref. | Method | Performance | Type of energy consumption forecasted | Building type |
| [70] | ELM | RMSE=74.02 kWh | Cooling and Heating | Residential |
| [64] | MLP | RMSE=752-1410 kWh | Cooling | Residential |
| | | RMSE=607-785 kWh | Heating | |
| | | RMSE=607-785 kWh | Lighting | |
| | | RMSE=607-785 kWh | Overall | |
| [57] | MLR | CV=173.2% - 249.8% | Heating | Non-residential |
| | | CV=47.5%-48.5% | Cooling | |
| | AR | CV=124.6%-185.9% | Heating | |
| | | CV=32.4%-37.6% | Cooling | |
| | ARX | CV=49.6%-178.8% | Heating | |
| | | CV=8.8%-34% | Cooling | |
| | BPNN | CV=121.2%-174.9% | Heating | |
| | | CV=25.3%-30.6% | Cooling | |
| | MLR | CV=58.5%-70.7% | Heating | Residential |
| | | CV=17.9%-25.9% | Cooling | |
| | AR | CV=14.7%-23.7% | Heating | |
| | | CV=8.3%-8.6% | Cooling | |
| | ARX | CV=10.8%-22% | Heating | |
| | | CV=5%-6.9% | Cooling | |
| | BPNN | CV=23%-23.3% | Heating | |
| | | CV=6.4%-7.1% | Cooling | |
| [48] | Regression (Linear model) | $R^2$=96.31% | Cooling (VAV) | Non-residential |
| | | $R^2$=95.18% | Cooling (CAV) | |
| | | $R^2$=99.76% | Cooling (FC) | |
| | | $R^2$=97.08% | Cooling (EMB) | |
| | | $R^2$=96.70% | Cooling (ALU) | |
| | | $R^2$=97.59% | Heating (VAV) | |
| | | $R^2$=97.48% | Heating (CAV) | |
| | | $R^2$=99.02% | Heating (FC) | |
| | | $R^2$=98.21% | Heating (EMB) | |
| | | $R^2$=98.13% | Heating (ALU) | |
| | Regression (Multivariate model) | $R^2$=99.39% | Cooling (VAV) | |

| Simulated (Ecotect) | | | |
|---|---|---|---|
| Ref. | Method | Performance | Type of energy consumption forecasted |
| [37] | DNN | VAF=99.76% | Cooling |
| | GBM | VAF=98.53% | |
| | GPR | VAF=99.13% | |
| | MPMR | VAF=89.55% | |
| | DNN | VAF=98.05% | Heating |
| | GBM | VAF=96.54% | |
| | GPR | VAF=99.84% | |
| | MPMR | VAF=88.022% | |
| [16] | ANN | RMSE=1.678 kW | Cooling |
| | SVR | RMSE=1.647 kW | |
| | CART | RMSE=1.841 kW | |
| | CHAID | RMSE=1.859 kW | |
| | GLR | RMSE=1.74 kW | |
| | SVR+ANN | RMSE=1.566 kW | |
| | ANN | RMSE=0.61 kW | Heating |
| | SVR | RMSE=0.346 kW | |
| | CART | RMSE=0.8 kW | |
| | CHAID | RMSE=0.909 kW | |
| | GLR | RMSE=1.039 kW | |
| | SVR+ANN | RMSE=0.428 kW | |
| [18] | ELM(RBF) | MAE=0.0454 kW | Cooling |
| | ELM(Hardlim) | MAE=0.037 kW | |
| | ELM(Sig) | MAE=0.0348 kW | |
| | OSELM(RBF) | MAE=0.369 kW | |
| | OSELM(Hardlim) | MAE=0.4008 kW | |
| | OSELM (Sig) | MAE=0.3229 kW | |
| | ELM(RBF) | MAE=0.0596 kW | Heating |
| | ELM(Hardlim) | MAE=0.0566 kW | |
| | ELM(Sig) | MAE=0.0389 kW | |
| | OSELM(RBF) | MAE=0.313 kW | |
| | OSELM(Hardlim) | MAE=0.3964 kW | |
| | OSELM (Sig) | MAE=0.2755 kW | |



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



| | | | | |
|---|---|---|---|---|
| | | $R^2$=98.39% | Cooling (CAV) | |
| | | $R^2$=99.89% | Cooling (FC) | |
| | | $R^2$=98.44% | Cooling (EMB) | |
| | | $R^2$=98.23% | Cooling (ALU) | |
| | | $R^2$=98.32% | Heating (VAV) | |
| | | $R^2$=99.69% | Heating (CAV) | |
| | | $R^2$=99.05% | Heating (FC) | |
| | | $R^2$=98.38% | Heating (EMB) | |
| | | $R^2$=98.34% | Heating (ALU) | |
| [69] | MILP | Overall cost = 136.64 € | Cooling and heating | Residential and non-residential |

| | | | |
|---|---|---|---|
| [25] | MARS | MAPE=19.2319 | Cooling |
| | ELM | MAPE=27.7105 | |
| | Hybrid Model | MAPE =19.5096 | |
| | MARS | MAPE =2.2043 | Heating |
| | ELM | MAPE =17.739 | |
| | Hybrid Model | MAPE =2.2835 | |
| [73] | Hybrid Model | MAPE=8.5600 RMSE=0.367 MAE=0.284 | Heating |
| | | MAPE=11.1490 RMSE=0.7640 MAE=0.611 | Cooling |

Most of the above-mentioned methods were studied and considered as a conventional method for forecasting and controlling the CL and HL of residential buildings. However, some of the work showed that building specifications could have a significant impact on the CL and HL of a building [37], [76]. To forecast the CL and HL of buildings with high accuracy coefficient and least prediction error by regression-based methods, it is necessary to establish a linear mapping between the input variables (building characteristics) and output variables (CL and HL). In this paper, a hybrid model called group support vector regression (GSVR) is proposed by combining the group method of data handling (GMDH) and SVR models in order to estimate the relationship between input and output data and improve the CL and HL prediction results. GMDH and SVM have already been combined in several areas with significant improvements in predictive accuracy over other neural networks, machine learning, and other statistical models; these studies have also shown that combining these two models can improve the efficiency and robustness in the prediction operation [55, 56]. To present a comparative analysis approach to evaluating the results of CL and HL forecasting, in addition to the proposed hybrid model, other artificial neural network and machine learning applications, including BPNN, elastic-net regression (ENR), GRNN, k-nearest neighbors (kNN), partial least squares regression (PLSR), GMDH, and SVR are implemented. Technical specifications and structural characteristics of a building influence its energy; therefore, structural characteristics are used such as relative compactness, surface area, overall height, roof area, orientation, wall area, glazing area distribution, and glazing area as inputs for the proposed methods to determine the output variables: CL and HL of residential buildings. CL and HL were considered as the output variables of the networks in this study. The energy forecasting framework is shown in Fig. 1.

The remainder of this paper is organized as follows: Section 2 introduces the case study. The proposed methods are

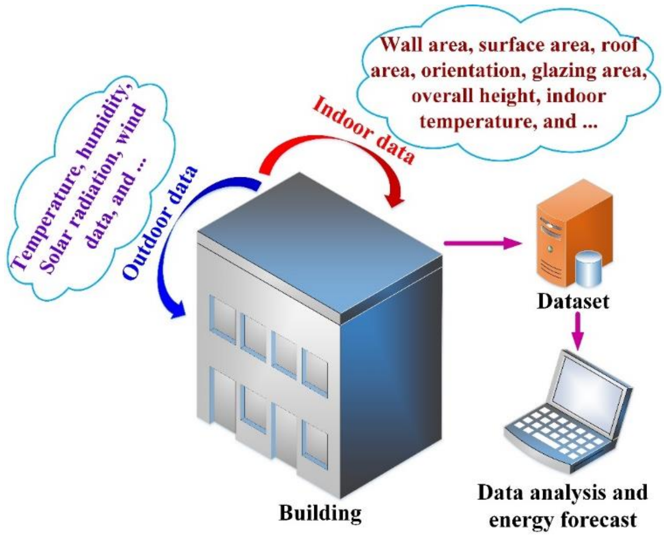



**Figure 1.** Energy forecasting framework

described in Section 3. Section 4 presents the CL and HL forecasting results using the suggested methods. In Section 5, the results and performance analysis of the suggested solutions are compared. Finally, Section 6 concludes the paper.

## II. CASE STUDY

A dataset comprising 12 different building shapes is used, which is simulated by Tsanas and Xifara (2012) in Ecotect software [76]. Regarding the simulations, it was assumed that the buildings were located in Greece, Athens. The shapes of the buildings were created through considering the initial cubes (3.5 × 3.5 × 3.5), each with 18 elements. All buildings were constructed with the same materials and a volume of 771.75 m3. The simulated buildings differed in orientation, glaze area, interior dimensions, glaze area distribution, and certain other parameters. The envelope U-values of these buildings were considered as the ceiling (0.500), walls (1.780), window (2.260), and floor (0.860). U-values indicated the



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



overall performance in retaining heat and preventing it from escaping to the outside. The interior design of a building can have a major impact on its energy. In the simulation, the values of air speed: 0.30 m/s, clothing: 0.6, humidity: 60%, lighting level: 300 Lux were considered as interior design conditions. The internal gains were set to rational: 5 and latent: 2 W/m2 while the infiltration rate was set to 0.5 for the rate of air change with the wind sensitivity being 0.25 air changer per hour. 95% of the complex state was assumed for thermal properties, and a temperature range of 19-24° C was considered for the thermostat range. 15-20 hours of weekdays and 10-20 hours of weekends were set for thermostat operation hours [1, 11].

The dataset consisted of 768 simulated buildings, each with eight characteristic properties (considered as input variables and represented in the data by $X$) and two actual valuable responses (CL and HL intended as output variables and marked with $Y$). Table 5 lists the input and output variables with the mathematical delegation for each variable and shows the number of conceivable values for these data [76].

The data used in this study were obtained by simulating and modeling different aspects of building energy and can be contrasted with real-world data; nevertheless, according to the literature cited by the producers of this dataset, recent works have shown that no inconsistency in real-world data and simulated data can affect the results of the used methods.

## III. METHODOLOGIES

This section briefly describes artificial neural network algorithms and regression techniques used to analyze the data for predicting the CL and HL of buildings and finally evaluate the results.

### A. Artificial neural networks and regression techniques

Artificial neural networks are a suitable means of modeling and estimation purposes. They are capable of learning from data and show generalized behavior after the learning process. Alongside artificial neural networks, regression applications can be used as statistical tools for analyzing or understanding a binary or multivariate relationship [29], [79].

**TABLE 5. The input and output variables with mathematical symbols for each variable**

| Mathematical symbols | Variable names | Number of values |
|---|---|---|
| *X1* | Relative compactness | 12 |
| *X2* | Surface area | 12 |
| *X3* | Wall area | 7 |
| *X4* | Roof area | 4 |
| *X5* | Overall height | 2 |
| *X6* | Orientation | 4 |
| *X7* | Glazing area | 4 |
| *X8* | Glazing area distribution | 6 |
| *Y1* | Heating load | 586 |
| *Y2* | Cooling load | 636 |

Today, regression is used as a powerful tool in the scientific, commercial, and industrial fields for forecasting, modeling, and optimization [80]–[82]. In the remainder of this section, different types of neural networks and regression methods are discussed.

#### 1) BACK-PROPAGATION NEURAL NETWORK (BPNN)

Developed by Parker in 1972, BPNN is one of the well-known and widely used neural network algorithms in engineering applications. This algorithm is similar to a FFNN in every aspect except that it adds a back propagation to its structure. It is estimated that today, this algorithm is utilized in some 80% of the structure of other neural networks [58, 61]. As shown in Fig. 2, BPNN comprises an input layer, a hidden layer, and an output layer in its structure [84]. The number of nodes in the input layer, hidden layer, and the output layer are represented by letters $m$, $T$, and $n$, respectively. $w_{ij}, w_{jk}$ reflects the weights of connections in the layers. Network input values and network prediction values are expressed by $X$ and $Y$, respectively. Input signals with weight are received by the input layer and transmitted to the hidden layer after processing. As an internal processing layer, the hidden layer processes the signals and transmits them to the output layer [85]. The neural network training process requires an activation function; $f(X)$ is the sigmoid activation function used in this paper [79]:

$$f(x) = \frac{1}{(1+e^{-x})} \tag{1}$$

The output of each neuron in the input layer is calculated as follows [86]:

$$O_{ik} = f\left(\sum_{j=0}^{l} W_{kj} O_{ij}\right), k = 1, 2, \ldots, n \tag{2}$$

where $W_{kj}$ is the connection weight from the $j$th input neuron to the $k$th hidden neuron. The output of each neuron in the output layer providing the network prediction is calculated as follows [86]:

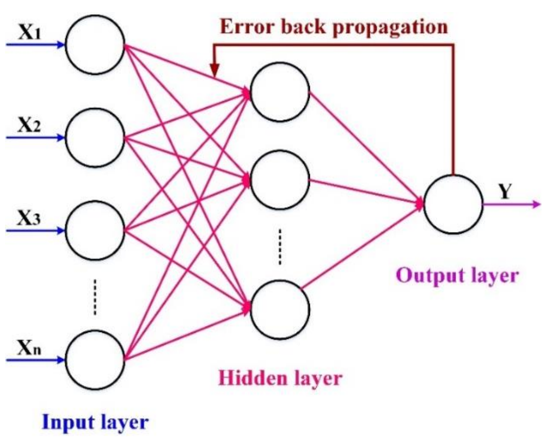



Figure 2. The structure of BPNN



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access


$$Y_i = f\left(\sum_{k=0}^{n} W_k O_{ik}\right), k = 1, 2, \dots, n \tag{3}$$

where $n$ is the number of hidden neurons, and $W_k$ is the connection weight from $k$th hidden neuron to the output neuron.

### 2) GENERAL REGRESSION NEURAL NETWORKS (GRNN)

The general regression neural network (GRNN) is a one-pass learning algorithm suggested by Specht as an alternative to the back-error propagation training algorithm for the FFNN [87]. In fact, GRNN is an RBFN that can be used to predict and assess continuous variables, estimate nonlinear relationships between the output variable and a set of independent variables, and converge to the underlying regression level. Similar to other probabilistic neural networks, GRNN requires training samples to process. In this network, predictability and processing are also dependent on the data type [88]. Among the important advantages of this neural network, mention can be made of rapid learning even with large numbers of samples, establishing a relationship between the input variables and target variables of time series data, and fast convergence to a desirable regression level. As can be seen in Fig. 3, in the structure of GRNN, there are four layers, the input $(X)$, pattern $(P)$, summation $(S)$, and output $(Y)$, which are completely interconnected [89].

The input layer transfers the input variables to the pattern layer without any processing. The number of neurons in the training layer is determined by the training data, which functions as follows [90], [91]:

$$p_i = exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right], i = 1, 2, \dots, n \tag{5}$$

The outcome of the neuron $i$ is expressed as follows:

$$D_i^2 = (X - X_i)^T (X - X_i) \tag{6}$$

where $D_i^2$ is the square of the Euclidean distance between the input variable and the sample corresponding to $X$, $X$ is the network input variable, $T$ is the transpose symbol, and $X_i$ is the learning sample corresponding to the neuron $i$.

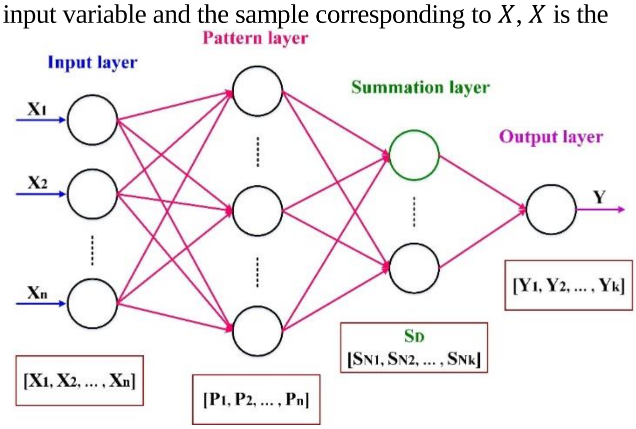



**Figure 3.** The four layers of GRNN

The extracted patterns are transferred to the summation layer as input, where two types of summation are calculated. The first summation is derived from equation (7), in which the arithmetic sum of all neurons is computed and the output is demonstrated as equation (8). The second summation is derived from equation (9), in which the weight coefficients are added to the neurons and the output is expressed as equation (10) [89], [91]:

$$\sum_{i=1}^{n} exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \tag{7}$$

$$S_D = \sum_{i=1}^{n} p_i \tag{8}$$

$$\sum_{i=1}^{n} Y_i exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \tag{9}$$

where $Y_i$ defines the $i$-th connection weight.

$$S_{Nj} = \sum_{i=1}^{n} Y_{ij} P_i \quad , \quad j = 1, 2, \dots, k. \tag{10}$$

After calculating both weighted and arithmetic summations, their results are transferred to the output layer. Finally, through dividing the two summation types, the final output of the network is calculated as follows [92]:

$$Y_j = \frac{S_{Nj}}{S_D} \quad , \quad j = 1, 2, \dots, k. \tag{11}$$

### 3) K-NEAREST NEIGHBORS (KNN)

kNN is an efficient technique for regression problems, classification, and non-parametric feature extraction. The main objective behind kNN is to provide an appropriate estimate based on Euclidean distance or, in other words, the mean of the nearest neighbors in feature space for the training set. These neighbors are selected from a set of training points. kNN is a precision algorithm with a unique property as it delays and continues the calculation operation until it obtains the best results [93]. As in other artificial neural networks and machine learning algorithms, datasets and, in particular, the number of k neighbors selected, have a great impact on the performance of kNN. Choosing small quantities for k leads to over-fitting, while large quantities entail a poor performance. A comparison of different subsets of training data is the best approach to selecting the optimal value of $k$. If the feature space is multi-dimensional, the data must be normalized


This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



beforehand due to the distance comparison to selecting the nearest neighbor [94]. In this paper, kNN is used for regression. In this case, the response of testing point $X_t$ is estimated as the weighted mean of the reactions of $k$ training points such as $X_1, X_2, \dots, X_k$. A kernel function was utilized to calculate the weight of each neighbor. If $X$ is a dataset of $X_1, X_2, \dots, X_k$ members, each with $N$ features, equation (12) calculates the proximity of each training point $X_i$ to the testing point $X_t$ using the weighted Euclidean distance [95], [96].

$$d(X_t, X_i) = \sqrt{\sum_{n=1}^{N} w_n (x_{t,n} - x_{i,n})^2} \tag{12}$$

The following equation shows how to derive the arrangement statistics for distances $d(X_t, X_i), i = 1, 2, \dots, M$.

$$0 \le d\big(X_t, X_{(1)}\big) \le d\big(X_t, X_{(2)}\big) \le \dots \le d(X_t, X_{(M)}) \tag{13}$$

Then, $k$ nearest training points $X_{(1)}, X_{(2)}, \dots, X_{(k)}$ as the $k$-nearest neighbors of $X_t$ are calcualted.

Fig. 4 depicts a transparent example of a two-dimensional feature space $(N = 2)$ of the k-nearest neighbors. In this example, $k = 5$ is assumed. As observed, the five training points are in the nearest Euclidean distance, with one test point selected as the nearest neighbor [96]. A kernel function can be used for the flexibility of the kNN regression model. In this paper, the Gaussian radial basis function (RBF) kernel was used and defined as follows:

$$\phi\big(X_t, X_{(i)}\big) = e^{-d(X_t, X_{(i)})/\beta} \tag{14}$$

where $\phi(X_t, X_{(i)})$ is a kernel function centered at the $i$th training point $X_{(i)}$ and $\beta$ is a Gaussian rottenness factor [97]. Ultimately, by applying the kernel regression, the following estimate of the response $X_t$ is obtained [96]:

$$\hat{f}(X_t) = \frac{\sum_{i=1}^{k} \phi\big(X_t, X_{(i)}\big) f(X_{(i)})}{\sum_{i=1}^{k} \phi\big(X_t, X_{(i)}\big)} \tag{15}$$

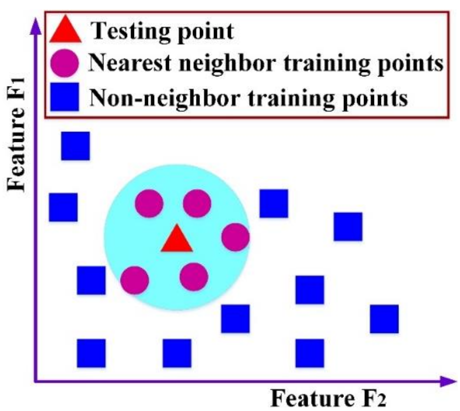



**Figure 4.** **kNN of a testing point in a two-dimensional feature space**

where $k$ defines the number of the closest neighbors used for regression and $f(X_{(i)})$ is the known response of $X_{(i)}$.

### 4) GROUP METHOD OF DATA HANDLING (GMDH)

GMDH was first introduced in 1968 by Prof. Alexey Grigorevich Ivakhnenko as one of the data mining applications for performing nonlinear regression, prediction, function approximation, mathematical modeling, and pattern recognition. GMDH is further considered as a polynomial neural network. Continuous change during the training process is the major difference between GMDH and other neural networks [98], [99]. This network attempts to generate a performance in a feedback network based on a quadratic transmission function. Fig. 5 illustrates the GMDH neural network structure, which includes the input layers, the hidden layers, and the output layers. This network automatically specifies the effective input variables, the structure of an optimal model, the number of layers, and the number of neurons in the hidden layers [100]. High prediction accuracy, self-organized learning process, the accurate linear mapping between input and output variables, and precise identification of complex nonlinear systems are among the significant advantages of GMDH over other neural networks [101].

As a brief introduction for GMDH formulation description, a nonlinear system with n number of input variables $(X_n)$ and output variables $(Y)$connected by the function $f$ is considered as follows [100], [101]:

$$Y = f(X_1, X_2, \dots, X_n) \tag{16}$$

The relationship between the input and output variables in the above system can be rewritten based on the Kolmogorov–Gabor polynomial as below [102]:

$$Y = a_0 + \sum_{i=1}^{m} a_i X_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} X_i X_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk} X_i X_j X_k + \cdots \tag{17}$$

where $X$ stands for inputs, $m$ is the number of inputs, $a$ shows the weights or coefficients, and $Y$ is the model estimation of output. Each layer in the neural network functions based on its neurons. In a GMD neural network, each neuron must have two inputs per output. The following equation describes the number of neurons in the first layer for the $n$ number of inputs [101]:

$$m = \frac{n^2 - n}{2} \tag{18}$$



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



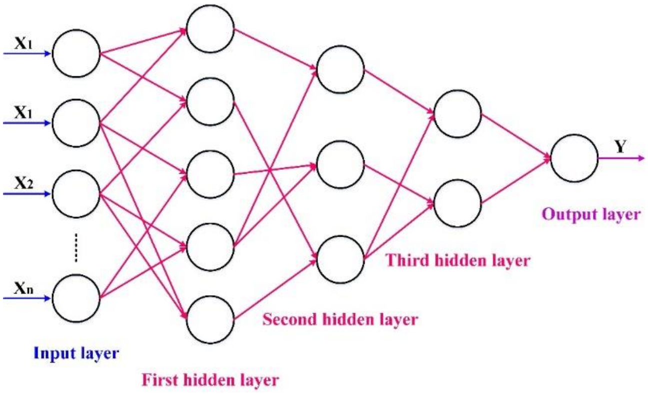



**Figure 5. The principle structure of GMDH**

In order for the GMDH algorithm to identify unknown $\alpha_i$coefficients in the Kolmogorov–Gabor polynomial relation, it can be analyzed as follows [102], [103]:

$$G(X_i, X_j) = a_0 + a_1 X_i + a_2 X_j + a_3 X_i^2 + a_1 X_j^2 + a_5 X_i X_j \tag{19}$$

The regression methods for each pair of input variables $X_i$ and $X_j$ can optimally solve the coefficients $\alpha_i$, according to which, the function $G$ is given by considering the least squares error (20) principle as equation (21) [100], [102].

$$E = \frac{1}{M} \sum_{i=1}^{M} (Y_i - G_i O)^2 \tag{20}$$

where E is the least squares error, M shows the number of samples, and $G_i O$ depicts the desired output or target of the network.

$$Y_i = f(X_{i1}, X_{i2}, X_{i3}, \dots, X_{im}), i = 1, 2, 3, \dots, m \tag{21}$$

### 5) ELASTIC-NET REGRESSION (ENR)

In machine learning applications and the methods related to regression, classification, and prediction of different models, such factors as variance and bias affect the accuracy coefficient of the model. Until now, different methods have been proposed for regression or linear modeling in machine learning. The linear regression (LR) is one of the most common methods for linear modeling. Nevertheless, this method is sometimes observed to have poor performance in forecasting and analyzing. Over the recent years, researchers have proposed various suitable techniques for improving LR performance such as lasso, ridge regression, and ENR [104], [105].

ENR combines ridge regression and lasso, which was primarily introduced by Zou and Hastie in 2003 as a powerful technique for prediction and variable selection. Therefore, this method is utilized to generate regression and linear mapping between input and output variables for practical solutions requiring high analytical accuracy [105], [106]. For an observed vector/signal $(Y \epsilon R_n)$, a certain measurement matrix $(X \epsilon R_{n \times p})$, an unknown correct vector/signal to be recovered $(\beta^* \epsilon R_p)$, and the Gaussian noise term $(z)$, the linear regression in the standard model is generally expressed as follows [107]:

$$Y = X\beta^* + z \tag{22}$$

ENR is feasible by solving the following minimization [107], [108]:

$$min\, \lambda_1 \|\beta\|_1 + \frac{\lambda_2}{2} \|\beta\|_2^2 + \frac{1}{2} \|X\beta - Y\|_2^2 \tag{23}$$

where $\lambda_1, \lambda_2 \geq 0$ are both regularization parameters. In the above equation, two particular cases of elastic regression are further included: ridge regression in [109] with $\lambda_1 = 0$ and lasso in [110] with $\lambda_2 = 0$.

### 6) SUPPORT VECTOR REGRESSION (SVR)

As one of the SVM applications, SVR is used for regression, producing a linear mapping between input and target variables, and function estimation. Solving regression problems is the main function of SVR [111], [112]. Among SVR models, the classical model (ε-SVR) is a version of SVR basically considered in engineering and also used in this work. Fig. 6 shows the principle structure of SVR. This method presents the principle idea by minimizing the rate of error and singularizing the hyperplane, which gives the maximum margin [46]. In the SVR, the risk function (24) is minimized by training the function (25) and vectors contained in (26) [113]:

$$R = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{l} (y1, \langle w, x \rangle) \tag{24}$$

$$f(x) =< w, x > +b \tag{25}$$

$$\{(x1, y1), \dots, (xi, yi), i = 1, \dots, l\} \tag{26}$$

where $w$ adjusts the model smoothness, $\langle w, x \rangle$ is a function of the input space fitting to the feature space, b shows the bias, and $(y1, \langle w, x \rangle)$ depicts the selected loss function. Solving this model requires solving an optimization problem as follow: [113], [114]:

$$min \left( \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \right) \tag{27}$$

Subject to:

$$\begin{cases} y_i - \langle W, X_i \rangle - b & \leq \varepsilon + \xi_i \\ \langle W, X_i \rangle + b - y_i & \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* & \geq 0 \end{cases} \tag{28}$$



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access

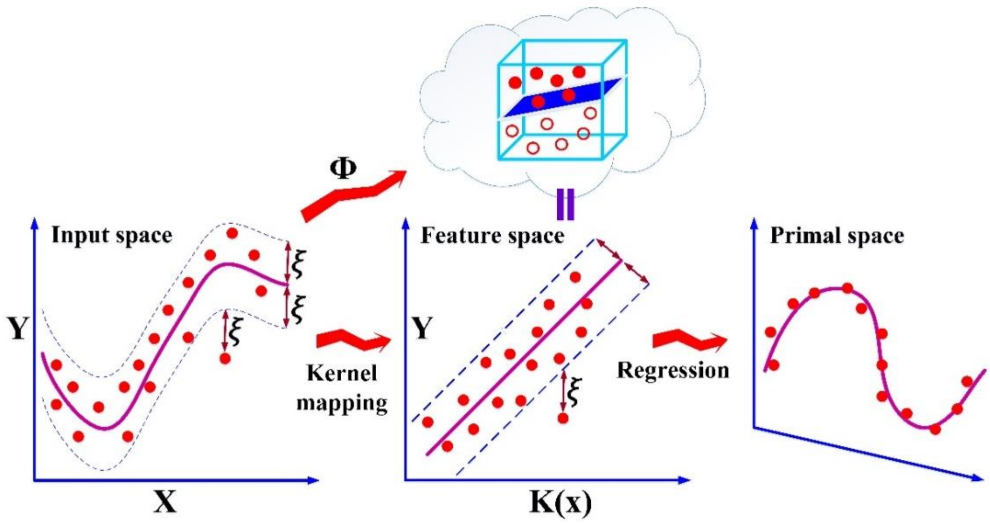




**Figure 6.** **The principle approach of SVR**

where $\xi_i$ and $\xi_i^*$ apply the functional constraints. The following expression is obtained after solving equation (27) [115], [116]:

$$f(X, \alpha_i, \alpha_i^*) = \sum_{i=1}^{N_{sv}} (\alpha_i - \alpha_i^*) \left(K\left(X_i, X_j\right)\right) + b \qquad (29)$$

where $\alpha_i$ and $\alpha_i^*$ are Lagrange coefficients. According to (29), new predictions are made using new training data [115].

7) PARTIAL LEAST SQUARES REGRESSION (PLSR)

PLSR is known as data-driven modeling and multivariate regression. This method is able to create a linear regression model and analyze a wide range of data via transmitting the input variables (input matrix $X$) and output variables (output matrix $Y$) to a new space. PLSR basically uses the covariance between the input and output variables instead of analyzing the hyperplanes with the least variance between the dependent and independent variables [117], [118]. This method interrelates input and output variables in a database using a linear multivariate model. The analysis of $X$ and $Y$ for $m$ number of pulse components is as follows [117], [119]:

$$X = \sum_{h=1}^{m} t_h p_h^T + E = TP^T + E \qquad (30)$$

$$Y = \sum_{h=1}^{m} u_h q_h^T + F = UQ^T + F \qquad (31)$$

where $X$ shows $n \times K$ matrix while $K$ stands for the number of input variables, $Y$ depicts $n \times M$ matrix while $M$ represents the number of output variables, $n$ displays the number of observations, $E$ shows the input residual, while $p$ is the input loading vector, and $F$ depicts the output residual, while $q$ is the output loading vector [117]. PLS is comprised of inner and outer parts [120]; in the inner part, the relationship between the $X$ and $Y$ matrices is established indirectly through their scores via the internal model which is a function of $(t)$ on $(u)$. In the outer part, $T$ and $P$ are the score and loading matrices for the $X$ data set, $U$ and $Q$ are the score and loading matrices for the $Y$ data set.

8) GROUP SUPPORT VECTOR REGRESSION (GSVR)

GSVM is a hybrid model of GMDH and SVR. The input data of the proposed hybrid model are selected based on the decision made by GMDH and utilized as the input data of SVR to forecast the output [56, 100]. In each layer of GSVR, all combinations of the two input variables $(X_i, X_j)$ are generated, for each of which the regression is constructed through forming the polynomial function which approximates the output y in Eq. (19). The GMDH output data, which generates the least error, is used together with the input variables as the SVR inputs. In order for the output data to have the least error, the GSVR model must have 3-5 iterations [56, 101]. Finally, the GSVR model with the least error is chosen as the output model. The GSVR structure is shown in Fig. 7 [121].

### *B. EVALUATION OF MODEL PERFORMANCE*

Different methods can be used to ensure the accuracy of results and their evaluation. The correlation coefficient of determination (R), mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) are statistical performance metrics used in this paper to assess the performance of the proposed methods [123], [124].

Each of the indicators has various emphases. R indicates the correlation coefficient between the forecasted value and the actual value of the designed model. MAE shows the mean distance between the forecasted value and the actual value. MSE is the mean of the squares of the errors, meaning the mean squared difference between the forecasted values and the actual value of the designed model. RMSE is utilized to



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



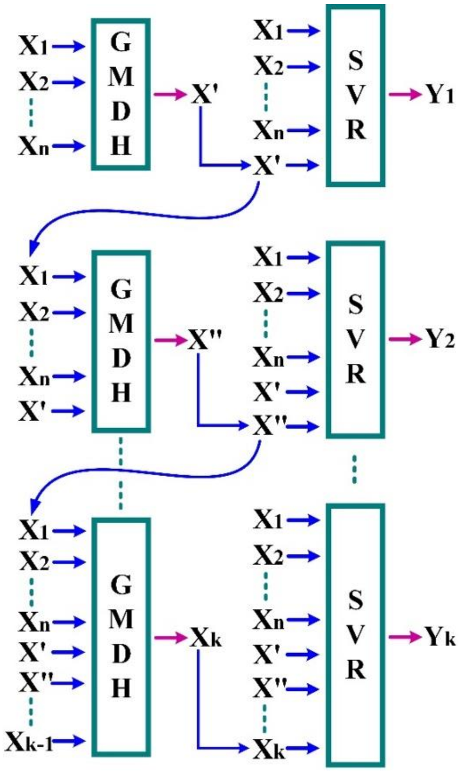


Figure 7. **The principle structure of GSVR**

recognize great errors and assessment of the variation in method response regarding variance and. The larger the value for R indicator is, and the smaller the values of MAE, MSE, and RMSE are, the better the model performance will be. The above-mentioned statistical performance metrics for $N$ number of inputs were calculated using the following formulas [32], [123]:

$$R = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2 \sum_{i=1}^{N}(Y_i - \bar{Y})^2}} \quad (32)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(X_i - Y_i)| \quad (33)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(X_i - Y_i)^2 \quad (34)$$

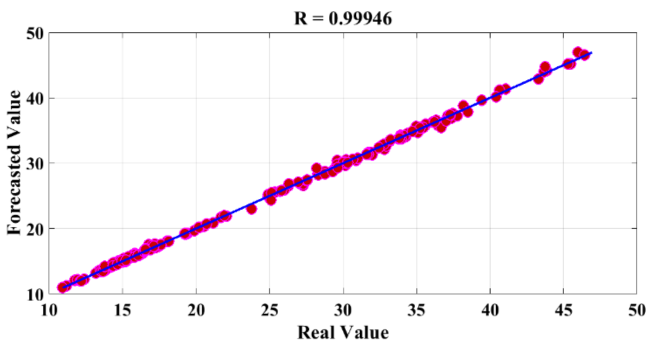
$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - Y_i)^2} \quad (35)$$

where $X_i$, $\bar{X}$, $Y_i$, and $\bar{Y}$ represent the actual value, mean of actual values, estimated value, and mean of actual values, respectively.

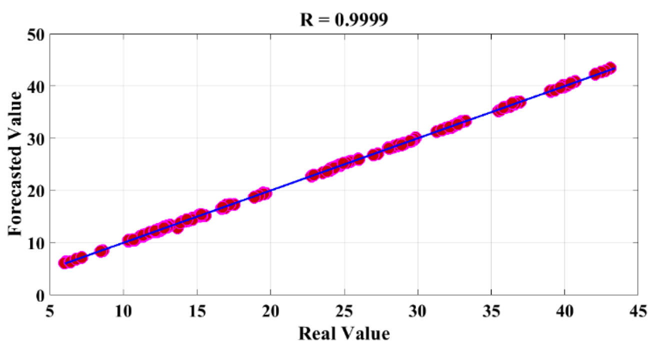
## IV. SIMULATION RESULTS

The proposed methods in this paper require initial design, structure formation, and selection of the number of neurons and the related coefficients in the hidden layers and structures. After the design, any neural network or machine learning application requires a dataset as input to enter the training phase. In the operational data of this study, as described in the relevant section, there exist two outputs (CL and HL) for each sample. Therefore, each network must be trained twice to perform forecasts, one for CL forecasting and the other for HL forecasting. Out of the total available datasets of 768 samples, 90% (688 samples) are selected for initial training and testing of each network; moreover, the remaining 10% (80 samples) are kept as unknown data to test the trained networks and forecast the required energy of buildings. For each network in each case of CL and HL, 70% and 30% of the intended dataset are considered as training and testing data, respectively. The neural networks and machine learning algorithms to forecast CL and HL models were applied in the MATLAB environment (Ver. 2018b).

The proposed networks were trained with relevant data and the results of each are described below. Figure 8 demonstrates an acceptable correlation and excellent overlap between the target data and the output of the designed GSVR for CL and HL training data. After training the network via CL and HL training data, the CL and HL were predicted as the testing data (Fig. 9). Figs. 10 and 11 depict the GSVR testing error for CL and HL prediction in the forms of MSE, RMSE, and histogram.




(a)




(b)

Figure 8. **Fit regression between input variables and outputs: (a) for CL; (b) for HL**



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



The regression and types of errors were shown only for the GSVR network. These results are expressed numerically for the rest of the methods in the following tables. Tables 6 and 7 show the CL and HL prediction results for training and testing data, respectively.

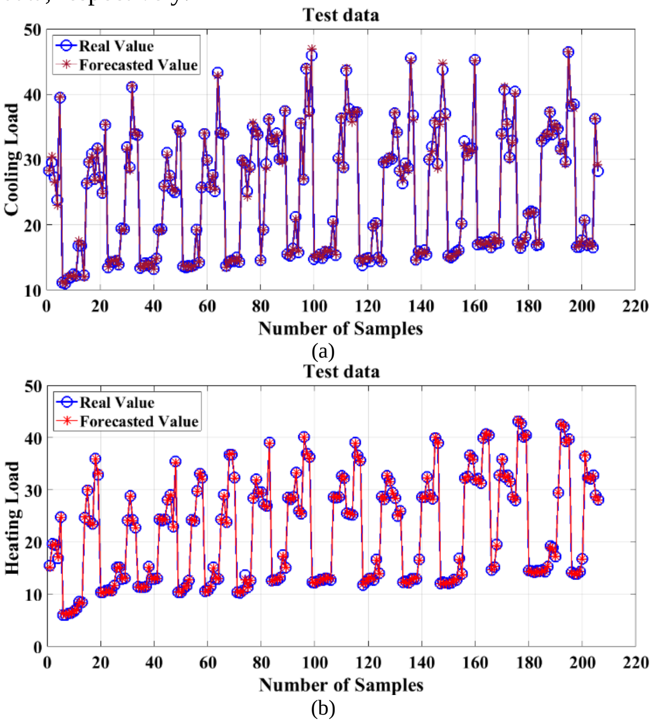
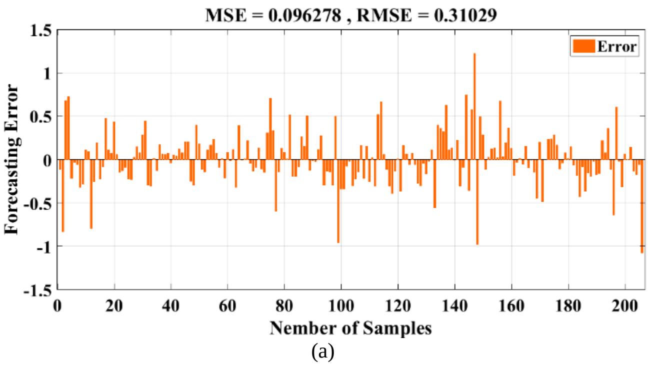



**Figure 9.** **Prediction of CL and HL for testing data via GSVR: (a) for CL; (b) for HL**



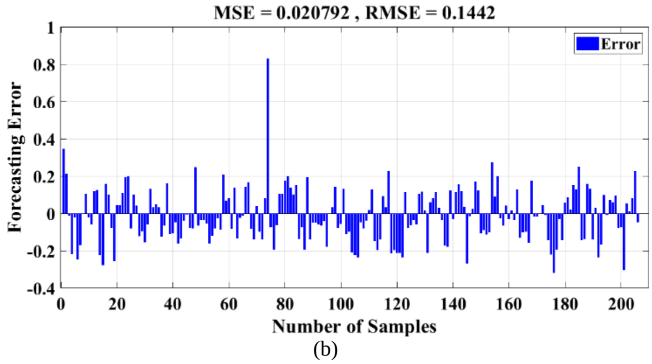




**Figure 10.** **Testing error for CL and HL prediction in the form of MSE and RMSE; (a) for CL; (b) for HL**

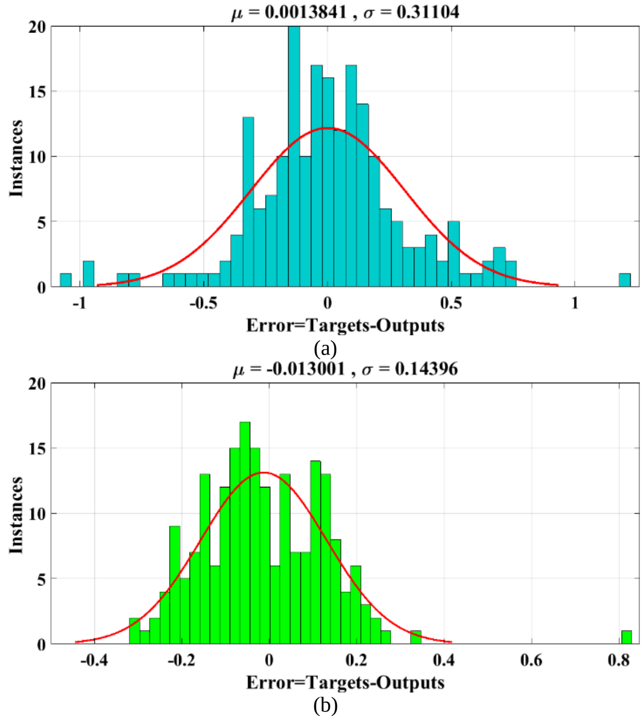



**Figure 11.** **Testing error for CL and HL prediction in the form of histogram; (a) for CL; (b) for HL**

**TABLE 6.** **Results of accuracy and error for CL prediction via all proposed methods**

| Methods | Train | | | | Test | | | |
|---|---|---|---|---|---|---|---|---|
| | R | MSE | RMSE | MAE | R | MSE | RMSE | MAE |
| BPNN | 0.9977 | 0.4679 | 0.6840 | 0.5808 | 0.9899 | 1.9802 | 1.4071 | 0.8448 |
| GRNN | 0.9987 | 0.1458 | 0.3818 | 0.2634 | 0.998 | 0.3831 | 0.6189 | 0.4810 |
| kNN | 0.9795 | 3.0158 | 1.7366 | 1.2471 | 0.9708 | 3.6891 | 1.9207 | 1.5457 |
| GMDH | 0.9991 | 0.0701 | 0.2647 | 0.1647 | 0.9988 | 0.2104 | 0.4586 | 0.3154 |
| ENR | 0.9808 | 2.4857 | 1.5734 | 1.1258 | 0.9741 | 3.2584 | 1.8051 | 1.3064 |
| SVR | 0.9877 | 1.638 | 1.2798 | 0.6409 | 0.9874 | 2.1897 | 1.4798 | 0.8809 |
| PLSR | 0.9854 | 1.985 | 1.4089 | 0.9185 | 0.9792 | 3.0189 | 1.7374 | 1.2651 |
| GSVR | 0.9994 | 0.0428 | 0.2069 | 0.1349 | 0.999 | 0.0962 | 0.3102 | 0.2260 |



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

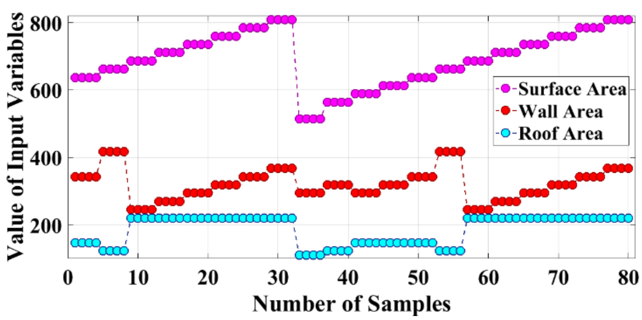
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



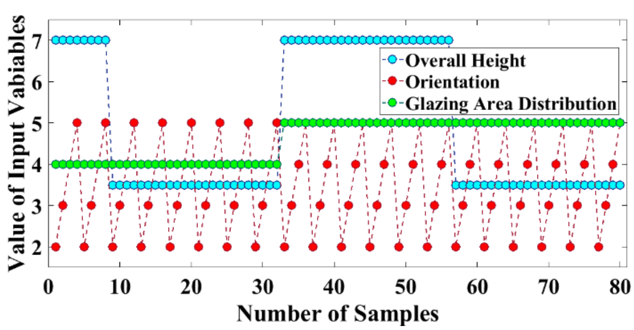
**TABLE 7.** Results of accuracy and error for HL prediction via all proposed methods

| Methods | Train | | | | Test | | | |
|---|---|---|---|---|---|---|---|---|
| | R | MSE | RMSE | MAE | R | MSE | RMSE | MAE |
| BPNN | 0.9996 | 0.0682 | 0.2611 | 0.1346 | 0.9990 | 0.1706 | 0.4130 | 0.2305 |
| GRNN | 0.9997 | 0.0589 | 0.2426 | 0.1371 | 0.9991 | 0.1672 | 0.4089 | 0.2193 |
| kNN | 0.9887 | 0.9485 | 0.9739 | 0.6154 | 0.9871 | 1.3785 | 1.174 | 0.7362 |
| GMDH | 0.9997 | 0.0562 | 0.237 | 0.1362 | 0.9991 | 0.1528 | 0.3908 | 0.209 |
| ENR | 0.9814 | 2.3646 | 1.5377 | 1.0028 | 0.9801 | 2.4504 | 1.5653 | 1.1183 |
| SVR | 0.9985 | 0.17782 | 0.42169 | 0.28775 | 0.9980 | 0.27164 | 0.52119 | 0.3803 |
| PLSR | 0.9938 | 0.5015 | 0.7081 | 0.4125 | 0.9896 | 0.8458 | 0.9196 | 0.5476 |
| GSVR | 0.9999 | 0.01663 | 0.12899 | 0.10134 | 0.9993 | 0.02079 | 0.1442 | 0.11465 |

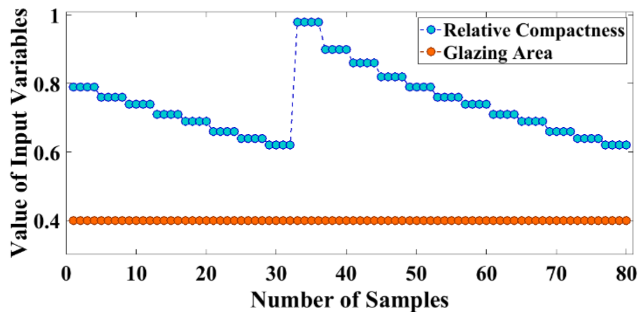
The above tables present the results of accuracy coefficients and error of training and testing for each of the networks. The results showed that the proposed methods were capable of predicting the test data after training. Studies of data mining applications showed that training and prediction operations were highly dependent on the type of data in regard to a variety of neural networks and machine learning algorithms. After training, all of the above networks were saved in the form of black box and utilized as a toolbox. This toolbox contained the features and detected patterns extracted from data or input variables. The CL and HL for new and unknown data were extracted using these black boxes. To do this, 10% (80 samples) of the main database (kept as new and unknown data) were utilized as inputs for each trained and saved network. Fig. 12 shows that each of the input variables (building characteristics) and output variables (CL and HL) were used as new and unknown data.



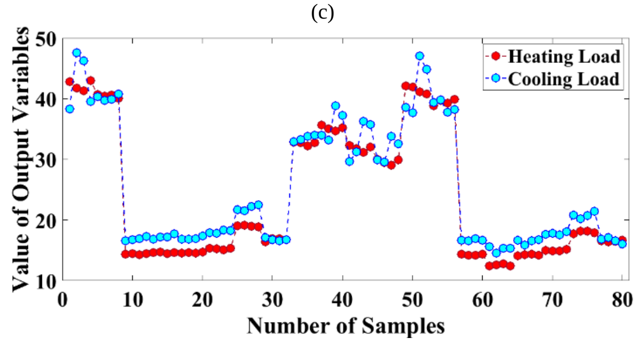

(a)




(b)




(c)




(d)

**Figure 12.** The characteristics and the energy load data of buildings: (a) Surface area, wall area, and roof area; (b) Overall height, orientation, and glazing area distribution; (c) Relative compactness and glazing area; (d) HL and CL

The key step in performing operations such as load forecasting and home energy forecasting is to perform forecasting operations for new and unknown data based on prior training. The CL and HL of residential buildings were predicted through the networks saved for the new data (Fig. 13).

Based on the results of Fig. 13, it can be deduced that each of the proposed methods was able to forecast the CL and HL for the new data based on the recognized patterns of the input variables via the training operation. The forecasted results were very close to the actual values, indicating that the proposed methods can be used as powerful tools for managing and saving the energy of residential and office buildings. Table 8 presents the precision and error rates for each of the proposed methods. The results of the proposed solutions are compared in the next section.



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

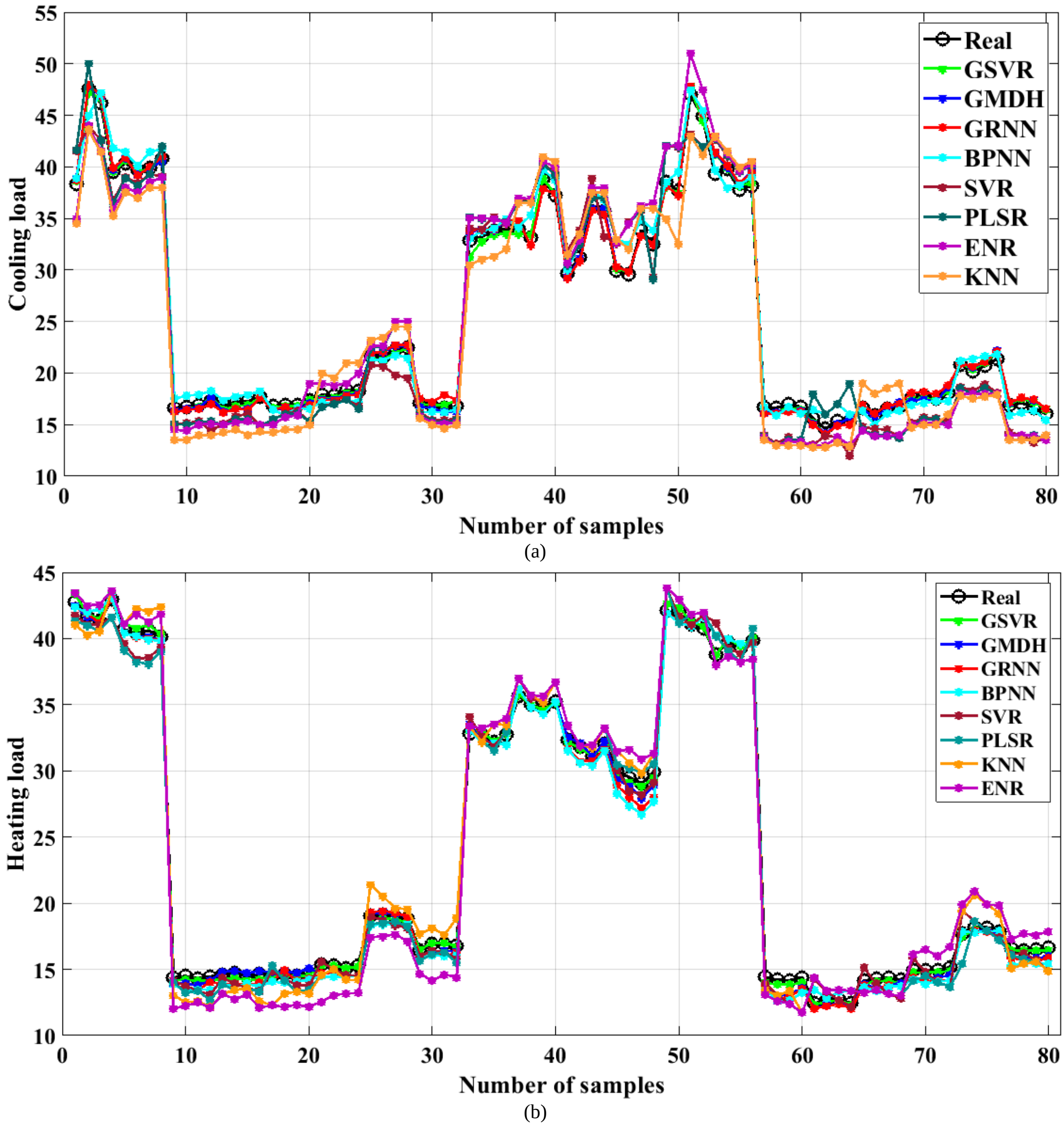



**Figure 13.** Forecasting the CL and HL of residential buildings for new data: (a) CL forecasting; (b) HL forecasting

**Table 8.** Various errors in forecasting CL and HL

| Methods | HL | | | | CL | | | |
|---|---|---|---|---|---|---|---|---|
| | R | MSE | RMSE | MAE | R | MSE | RMSE | MAE |
| BPNN | 0.9983 | 0.7544 | 0.8685 | 0.7304 | 0.9958 | 1.1175 | 1.0571 | 0.8671 |
| GRNN | 0.9989 | 0.488 | 0.6985 | 0.5827 | 0.9986 | 0.3330 | 0.5770 | 0.4961 |
| kNN | 0.9933 | 1.8311 | 1.3531 | 1.244 | 0.9704 | 8.0046 | 2.8292 | 2.72 |
| GMDH | 0.9992 | 0.2709 | 0.5204 | 0.435 | 0.9989 | 0.2678 | 0.5174 | 0.4305 |
| ENR | 0.9924 | 2.6267 | 1.6207 | 1.4933 | 0.9815 | 6.4678 | 2.5431 | 2.3666 |
| SVR | 0.9965 | 0.9837 | 0.9918 | 0.8064 | 0.9843 | 5.6663 | 2.3803 | 2.1621 |
| PLSR | 0.9961 | 1.2018 | 1.0432 | 0.9717 | 0.982 | 5.6418 | 2.3752 | 2.1731 |
| GSVR | 0.9999 | 0.063 | 0.2509 | 0.2078 | 0.9992 | 0.1783 | 0.4222 | 0.313 |



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

## V. ANALYSIS AND COMPARISON OF RESULTS

The effectiveness of each method was determined by comparing the results with other similar solutions. The parameters R, MSE, RMSE, MAE obtained from the results of the proposed methods were compared with each other and with the results obtained in other papers. All of the comparison methods were implemented in the same dataset. The division of the database for training, testing, and validation operations was different among these studies, hence the fact that the direct comparison should be drawn with caution. Table 5 summarizes the results obtained after testing the saved networks with the new data. As seen, the GSVR model with the best R value obtained for HL on new test data (0.9999) was in the range of best forecasting models. For HL prediction, GSVR had minimal errors in terms of MSE (0.0630), MAE (0.2078), and RMSE (0.2509). ENR with the lowest R value (0.9924) had the maximal errors in the form of MSE (2.6267), MAE (1.4933), and RMSE (1.6207). After the GSVR model, the GMDH and GRNN models were the most suitable methods for forecasting the HL with the R values of 0.9992 and 0.9989, respectively. Regarding CL prediction, GSVR had the most optimal R value (0.9992), and the minimal errors in terms of MSE (0.1783), MAE (0.3130), and RMSE (0.4222). In this forecasting, after GVR, GMDH and GRNN were the most appropriate models with R values of 0.9989 and 0.9986, respectively. However, the kNN method with the lowest R value (0.9704) and the highest error values in the forms of MSE (8.0046), MAE (2.7200), and RMSE (2.8292) provided the weakest forecast for CL. By comparing the results presented in Table 5 with the information presented in Table 1, the energy prediction results associated with buildings using the proposed methods can be compared with other models used in other studies for similar data.

## VI. CONCLUSION

The importance of forecasting the energy consumption of buildings, especially CL and HL in order to estimate, manage, and save energy, has raised many challenges. Today, most researchers are looking for an improved model with high predictive performance. This paper introduced a variety of energy forecasting models for residential and non-residential buildings and assessed the performance of each. Through the use of hybrid models, this paper proposed a new (hybrid) model of machine learning application called GSVR to forecast the CL and HL of a residential building. The principle goal of the proposed model was to combine the GMDH and SVR models. In addition to the proposed model, several other models of neural network and regression applications, including BPNN, ENR, GRNN, kNN, PLSR, GMDH, and SVR were used in this paper to forecast CL and HL. Eight technical parameters of the building (X1, X2, ..., X8) were employed as input variables, and CL and HL were selected as the output variables of the proposed models. After training and saving the trained networks, the CL and HL were forecasted using new data. The presented results indicated that the suggested hybrid model was the best model for predicting building energy as it had the highest R value (0.9999) and the lowest error values in the forms of MSE (0.063), RMSE (0.2509), and MAE (0.2078) for forecasting HL and the most optimal R value (0.9992) and minimal errors in terms of MSE (0.1783), RMSE (0.4222), and MAE (0.313) in forecasting CL. Noteworthy, the proposed method is also applicable to real-world data as well as all available energies, such as water and gas.

## REFERENCES

[1] J. Leitao, P. Gil, B. Ribeiro, and A. Cardoso, "A Survey on Home Energy Management," *IEEE Access*, vol. 8, pp. 5699–5722, 2020, doi: 10.1109/ACCESS.2019.2963502.

[2] H. Gong, V. Rallabandi, M. L. McIntyre, E. Hossain, and D. M. Ionel, "Peak Reduction and Long Term Load Forecasting for Large Residential Communities Including Smart Homes With Energy Storage," *IEEE Access*, vol. 9, pp. 19345–19355, 2021, doi: 10.1109/ACCESS.2021.3052994.

[3] M. A. Hannan *et al.*, "A Review of Internet of Energy Based Building Energy Management Systems: Issues and Recommendations," *IEEE Access*, vol. 6, pp. 38997–39014, 2018, doi: 10.1109/ACCESS.2018.2852811.

[4] O. Sadeghian *et al.*, "A comprehensive review on energy saving options and saving potential in low voltage electricity distribution networks: Building and public lighting," *Sustainable Cities and Society*, vol. 72, p. 103064, Sep. 2021, doi: 10.1016/j.scs.2021.103064.

[5] H. Shareef, M. S. Ahmed, A. Mohamed, and E. Al Hassan, "Review on Home Energy Management System Considering Demand Responses, Smart Technologies, and Intelligent Controllers," *IEEE Access*, vol. 6, pp. 24498–24509, 2018, doi: 10.1109/ACCESS.2018.2831917.

[6] A. M. Al-Ghaili, H. Kasim, N. M. Al-Hada, B. N. Jorgensen, M. Othman, and J. Wang, "Energy Management Systems and Strategies in Buildings Sector: A Scoping Review," *IEEE Access*, vol. 9, pp. 63790–63813, 2021, doi: 10.1109/ACCESS.2021.3075485.

[7] O. Sadeghian, A. Moradzadeh, B. Mohammadi-Ivatloo, B. Mohammadi-Ivatloo, M. Abapour, and F. P. G. Marquez, "Generation units maintenance in combined heat and power integrated systems using the mixed integer quadratic programming approach," *Energies*, vol. 13, no. 11, p. 2840, Jun. 2020, doi: 10.3390/en13112840.

[8] H. Nami, A. Anvari-Moghaddam, and A. Arabkoohsar, "Application of CCHPs in a centralized domestic heating, cooling and power network—Thermodynamic and economic implications," *Sustainable Cities and Society*, vol. 60, p. 102151, Sep. 2020, doi: 10.1016/j.scs.2020.102151.

[9] T. Lin, Y. Pan, G. Xue, J. Song, and C. Qi, "A novel hybrid spatial-temporal attention-LSTM model for heat load prediction," *IEEE Access*, vol. 8, pp. 159182–159195, 2020, doi: 10.1109/ACCESS.2020.3017516.

[10] J. Zhao and X. Liu, "A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis," *Energy and Buildings*, vol. 174, pp. 293–308, 2018, doi: 10.1016/j.enbuild.2018.06.050.

[11] Q. Chen, M. Xia, T. Lu, X. Jiang, W. Liu, and Q. Sun, "Short-Term Load Forecasting Based on Deep Learning for End-User Transformer Subject to Volatile Electric Heating Loads," *IEEE Access*, vol. 7, pp. 162697–162707, 2019, doi: 10.1109/ACCESS.2019.2949726.

[12] T. Hong, C. Koo, J. Kim, M. Lee, and K. Jeong, "A review on sustainable construction management strategies for monitoring, diagnosing, and retrofitting the building's dynamic energy performance: Focused on the operation and maintenance phase," *Applied Energy*, vol. 155. pp. 671–707, 2015, doi: 10.1016/j.apenergy.2015.06.043.

[13] Z. Yu, F. Haghighat, B. C. M. Fung, and H. Yoshino, "A decision



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



tree method for building energy demand modeling," *Energy and Buildings*, vol. 42, no. 10, pp. 1637–1646, 2010, doi: 10.1016/j.enbuild.2010.04.006.

[14] G. Platt, J. Li, R. Li, G. Poulton, G. James, and J. Wall, "Adaptive HVAC zone modeling for sustainable buildings," *Energy and Buildings*, vol. 42, no. 4, pp. 412–421, 2010, doi: 10.1016/j.enbuild.2009.10.009.

[15] R. Yao, B. Li, and K. Steemers, "Energy policy and standard for built environment in China," *Renewable Energy*, vol. 30, no. 13, pp. 1973–1988, 2005, doi: 10.1016/j.renene.2005.01.013.

[16] J. S. Chou and D. K. Bui, "Modeling heating and cooling loads by artificial intelligence for energy-efficient building design," *Energy and Buildings*, vol. 82, pp. 437–446, 2014, doi: 10.1016/j.enbuild.2014.07.036.

[17] A. Moradzadeh, H. Moayyed, S. Zakeri, B. Mohammadi-Ivatloo, and A. P. Aguiar, "Deep Learning-Assisted Short-Term Load Forecasting for Sustainable Management of Energy in Microgrid," *Inventions*, vol. 6, no. 1, p. 15, Feb. 2021, doi: 10.3390/inventions6010015.

[18] S. Kumar, S. K. Pal, and R. P. Singh, "A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes," *Energy and Buildings*, vol. 176, pp. 275–286, 2018, doi: 10.1016/j.enbuild.2018.06.056.

[19] S. Chen *et al.*, "An energy planning oriented method for analyzing spatial-temporal characteristics of electric loads for heating/cooling in district buildings with a case study of one university campus," *Sustainable Cities and Society*, vol. 51, p. 101629, Nov. 2019, doi: 10.1016/j.scs.2019.101629.

[20] A. Yezioro, B. Dong, and F. Leite, "An applied artificial intelligence approach towards assessing building performance simulation tools," *Energy and Buildings*, vol. 40, no. 4, pp. 612–620, 2008, doi: 10.1016/j.enbuild.2007.04.014.

[21] A. Tsanas *et al.*, "The Windkessel model revisited: A qualitative analysis of the circulatory system," *Medical Engineering and Physics*, vol. 31, no. 5, pp. 581–588, 2009, doi: 10.1016/j.medengphy.2008.11.010.

[22] Z. Wang and R. S. Srinivasan, "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models," *Renewable and Sustainable Energy Reviews*, vol. 75. pp. 796–808, 2017, doi: 10.1016/j.rser.2016.10.079.

[23] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81. pp. 1192–1205, 2018, doi: 10.1016/j.rser.2017.04.095.

[24] A. Moradzadeh, O. Sadeghian, K. Pourhossein, B. Mohammadi-Ivatloo, and A. Anvari-Moghaddam, "Improving residential load disaggregation for sustainable development of energy via principal component analysis," *Sustainability (Switzerland)*, vol. 12, no. 8, p. 3158, Apr. 2020, doi: 10.3390/SU12083158.

[25] S. Sekhar Roy, R. Roy, and V. E. Balas, "Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM," *Renewable and Sustainable Energy Reviews*, vol. 82. pp. 4256–4268, 2018, doi: 10.1016/j.rser.2017.05.249.

[26] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, no. 5, pp. 545–553, 2005, doi: 10.1016/j.enbuild.2004.09.009.

[27] Y. Ding, Q. Zhang, T. Yuan, and F. Yang, "Effect of input variables on cooling load prediction accuracy of an office building," *Applied Thermal Engineering*, vol. 128, pp. 225–234, 2018, doi: 10.1016/j.applthermaleng.2017.09.007.

[28] X. Li, J. H. Lǔ, L. Ding, G. Xu, and J. Li, "Building cooling load forecasting model based on LS-SVM," in *Proceedings - 2009 Asia-Pacific Conference on Information Processing, APCIP 2009*, 2009, vol. 1, pp. 55–58, doi: 10.1109/APCIP.2009.22.

[29] A. Moradzadeh, A. Mansour-Saatloo, B. Mohammadi-Ivatloo, and A. Anvari-Moghaddam, "Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings," *Applied Sciences (Switzerland)*, vol. 10, no. 11, p. 3829, May 2020, doi: 10.3390/app10113829.

[30] A. Moradzadeh, S. Zeinal-Kheiri, B. Mohammadi-Ivatloo, M. Abapour, and A. Anvari-Moghaddam, "Support Vector Machine-Assisted Improvement Residential Load Disaggregation," in *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, 2020, pp. 1–6, doi: 10.1109/ICEE50131.2020.9260869.

[31] X. J. Luo, L. O. Oyedele, A. O. Ajayi, and O. O. Akinade, "Comparative study of machine learning-based multi-objective prediction framework for multiple building energy loads," *Sustainable Cities and Society*, vol. 61, p. 102283, Oct. 2020, doi: 10.1016/j.scs.2020.102283.

[32] A. Moradzadeh, S. Zakeri, M. Shoaran, B. Mohammadi-Ivatloo, and F. Mohammadi, "Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms," *Sustainability (Switzerland)*, vol. 12, no. 17, p. 7076, Aug. 2020, doi: 10.3390/su12177076.

[33] Y. Ding, H. Su, X. Kong, and Z. Zhang, "Ultra-Short-Term Building Cooling Load Prediction Model Based on Feature Set Construction and Ensemble Machine Learning," *IEEE Access*, vol. 8, pp. 178733–178745, 2020, doi: 10.1109/ACCESS.2020.3027061.

[34] C. Fan, F. Xiao, and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Applied Energy*, vol. 195, pp. 222–233, 2017, doi: 10.1016/j.apenergy.2017.03.064.

[35] C. Fan, J. Wang, W. Gang, and S. Li, "Assessment of deep recurrent neural network-based strategies for short-term building energy predictions," *Applied Energy*, vol. 236, pp. 700–710, 2019, doi: 10.1016/j.apenergy.2018.12.004.

[36] Z. Wang, T. Hong, and M. A. Piette, "Data fusion in predicting internal heat gains for office buildings through a deep learning approach," *Applied Energy*, vol. 240, pp. 386–398, 2019, doi: 10.1016/j.apenergy.2019.02.066.

[37] S. S. Roy, P. Samui, I. Nagtode, H. Jain, V. Shivaramakrishnan, and B. Mohammadi-ivatloo, "Forecasting heating and cooling loads of buildings: a comparative performance analysis," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 3, pp. 1253–1264, 2020, doi: 10.1007/s12652-019-01317-y.

[38] A. Moradzadeh, B. Mohammadi-Ivatloo, M. Abapour, A. Anvari-Moghaddam, S. Gholami Farkoush, and S.-B. Rhee, "A practical solution based on convolutional neural network for non-intrusive load monitoring," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 10, pp. 9775–9789, Oct. 2021, doi: 10.1007/s12652-020-02720-6.

[39] J. Song, G. Xue, X. Pan, Y. Ma, and H. Li, "Hourly heat load prediction model based on temporal convolutional neural network," *IEEE Access*, vol. 8, pp. 16726–16741, 2020, doi: 10.1109/ACCESS.2020.2968536.

[40] J. Liu, X. Wang, Y. Zhao, B. Dong, K. Lu, and R. Wang, "Heating Load Forecasting for Combined Heat and Power Plants Via Strand-Based LSTM," *IEEE Access*, vol. 8, pp. 33360–33369, 2020, doi: 10.1109/ACCESS.2020.2972303.

[41] S. S. K. Kwok, R. K. K. Yuen, and E. W. M. Lee, "An intelligent approach to assessing the effect of building occupancy on building cooling load prediction," *Building and Environment*, vol. 46, no. 8, pp. 1681–1690, 2011, doi: 10.1016/j.buildenv.2011.02.008.

[42] T. Ahmad and H. Chen, "Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches," *Energy and Buildings*, vol. 166, pp. 460–476, 2018, doi: 10.1016/j.enbuild.2018.01.066.

[43] X. Li, L. Ding, J. Lǔ, G. Xu, and J. Li, "A novel hybrid approach of KPCA and SVM for building cooling load prediction," in *3rd International Conference on Knowledge Discovery and Data Mining, WKDD 2010*, 2010, pp. 522–526, doi: 10.1109/WKDD.2010.137.

[44] L. Jinhu, L. Xuemei, D. Lixing, and J. Liangzhong, "Applying principal component analysis and weighted support vector machine in building cooling load forecasting," in *CCTAE 2010 - 2010 International Conference on Computer and Communication Technologies in Agriculture Engineering*, 2010, vol. 1, pp. 434–437, doi: 10.1109/CCTAE.2010.5543476.

[45] A. E. Ben-Nakhi and M. A. Mahmoud, "Cooling load prediction



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

for buildings using general regression neural networks," *Energy Conversion and Management*, vol. 45, no. 13–14, pp. 2127–2141, 2004, doi: 10.1016/j.enconman.2003.10.009.

[46] Q. Li, Q. Meng, J. Cai, H. Yoshino, and A. Mochida, "Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks," *Energy Conversion and Management*, vol. 50, no. 1, pp. 90–96, 2009, doi: 10.1016/j.enconman.2008.08.033.

[47] S. Schiavon, K. H. Lee, F. Bauman, and T. Webster, "Influence of raised floor on zone design cooling load in commercial buildings," *Energy and Buildings*, vol. 42, no. 8, pp. 1182–1191, 2010, doi: 10.1016/j.enbuild.2010.02.009.

[48] I. Korolija, Y. Zhang, L. Marjanovic-Halburd, and V. I. Hanby, "Regression models for predicting UK office building energy consumption from heating and cooling demands," *Energy and Buildings*, vol. 59, pp. 214–227, 2013, doi: 10.1016/j.enbuild.2012.12.005.

[49] Y. Ding, Q. Zhang, and T. Yuan, "Research on short-term and ultra-short-term cooling load prediction models for office buildings," *Energy and Buildings*, vol. 154, pp. 254–267, 2017, doi: 10.1016/j.enbuild.2017.08.077.

[50] Y. Chen and H. Tan, "Short-term prediction of electric demand in building sector via hybrid support vector regression," *Applied Energy*, vol. 204, pp. 1363–1374, 2017, doi: 10.1016/j.apenergy.2017.03.070.

[51] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, "Vector field-based support vector regression for building energy consumption prediction," *Applied Energy*, vol. 242, pp. 403–414, 2019, doi: 10.1016/j.apenergy.2019.03.078.

[52] C. Fan and Y. Ding, "Cooling load prediction and optimal operation of HVAC systems using a multiple nonlinear regression model," *Energy and Buildings*, vol. 197, pp. 7–17, 2019, doi: 10.1016/j.enbuild.2019.05.043.

[53] S. S. K. Kwok, "A study of the importance of occupancy to building cooling load in prediction by intelligent approach," *Energy Conversion and Management*, vol. 52, no. 7, pp. 2555–2564, 2011, doi: 10.1016/j.enconman.2011.02.002.

[54] T. Catalina, V. Iordache, and B. Caracaleanu, "Multiple regression model for fast prediction of the heating energy demand," *Energy and Buildings*, vol. 57, pp. 302–312, 2013, doi: 10.1016/j.enbuild.2012.11.010.

[55] B. Gunay, W. Shen, and G. Newsham, "Inverse blackbox modeling of the heating and cooling load in office buildings," *Energy and Buildings*, vol. 142, pp. 200–210, 2017, doi: 10.1016/j.enbuild.2017.02.064.

[56] Z. Afroz, T. Urmee, G. M. Shafiullah, and G. Higgins, "Real-time prediction model for indoor temperature in a commercial building," *Applied Energy*, vol. 231, pp. 29–53, 2018, doi: 10.1016/j.apenergy.2018.09.052.

[57] K. Yun, R. Luck, P. J. Mago, and H. Cho, "Building hourly thermal load prediction using an indexed ARX model," *Energy and Buildings*, vol. 54, pp. 225–233, 2012, doi: 10.1016/j.enbuild.2012.08.007.

[58] T. Chaudhuri, Y. C. Soh, H. Li, and L. Xie, "A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings," *Applied Energy*, vol. 248, pp. 44–53, 2019, doi: 10.1016/j.apenergy.2019.04.065.

[59] R. Jovanović, A. A. Sretenović, and B. D. Živković, "Ensemble of various neural networks for prediction of heating energy consumption," *Energy and Buildings*, vol. 94, pp. 189–199, 2015, doi: 10.1016/j.enbuild.2015.02.052.

[60] Z. Hou, Z. Lian, Y. Yao, and X. Yuan, "Cooling-load prediction by the combination of rough set theory and an artificial neural-network based on data-fusion technique," *Applied Energy*, vol. 83, no. 9, pp. 1033–1046, 2006, doi: 10.1016/j.apenergy.2005.08.006.

[61] R. Yokoyama, T. Wakui, and R. Satake, "Prediction of energy demands using neural network with model identification by global optimization," *Energy Conversion and Management*, vol. 50, no. 2, pp. 319–327, 2009, doi: 10.1016/j.enconman.2008.09.017.

[62] J. Yang, H. Rivard, and R. Zmeureanu, "On-line building energy prediction using adaptive artificial neural networks," *Energy and Buildings*, vol. 37, no. 12, pp. 1250–1259, 2005, doi: 10.1016/j.enbuild.2005.02.005.

[63] S. Paudel, M. Elmtiri, W. L. Kling, O. Le Corre, and B. Lacarrière, "Pseudo dynamic transitional modeling of building heating energy demand using artificial neural network," *Energy and Buildings*, vol. 70, pp. 81–93, 2014, doi: 10.1016/j.enbuild.2013.11.051.

[64] S. L. Wong, K. K. W. Wan, and T. N. T. Lam, "Artificial neural networks for energy analysis of office buildings with daylighting," *Applied Energy*, vol. 87, no. 2, pp. 551–557, 2010, doi: 10.1016/j.apenergy.2009.06.028.

[65] D. Geysen, O. De Somer, C. Johansson, J. Brage, and D. Vanhoudt, "Operational thermal load forecasting in district heating networks using machine learning and expert advice," *Energy and Buildings*, vol. 162, pp. 144–153, 2018, doi: 10.1016/j.enbuild.2017.12.042.

[66] B. Cui *et al.*, "A hybrid building thermal modeling approach for predicting temperatures in typical, detached, two-story houses," *Applied Energy*, vol. 236, pp. 101–116, 2019, doi: 10.1016/j.apenergy.2018.11.077.

[67] R. Wang, S. Lu, and W. Feng, "A novel improved model for building energy consumption prediction based on model integration," *Applied Energy*, vol. 262, p. 114561, Mar. 2020, doi: 10.1016/j.apenergy.2020.114561.

[68] S. Sarihi, F. Mehdizadeh Saradj, and M. Faizi, "A Critical Review of Façade Retrofit Measures for Minimizing Heating and Cooling Demand in Existing Buildings," *Sustainable Cities and Society*, vol. 64, p. 102525, Jan. 2021, doi: 10.1016/j.scs.2020.102525.

[69] G. Bianchini, M. Casini, D. Pepe, A. Vicino, and G. G. Zanvettor, "An integrated model predictive control approach for optimal HVAC and energy storage operation in large-scale buildings," *Applied Energy*, vol. 240, pp. 327–340, 2019, doi: 10.1016/j.apenergy.2019.01.187.

[70] S. Naji *et al.*, "Estimating building energy consumption using extreme learning machine method," *Energy*, vol. 97, pp. 506–516, 2016, doi: 10.1016/j.energy.2015.11.037.

[71] Z. Xuan, Z. Xuehui, L. Liequan, F. Zubing, Y. Junwei, and P. Dongmei, "Forecasting performance comparison of two hybrid machine learning models for cooling load of a large-scale commercial building," *Journal of Building Engineering*, vol. 21, pp. 64–73, 2019, doi: 10.1016/j.jobe.2018.10.006.

[72] S. Seyedzadeh, F. Pour Rahimian, P. Rastogi, and I. Glesk, "Tuning machine learning models for prediction of building energy loads," *Sustainable Cities and Society*, vol. 47, p. 101484, May 2019, doi: 10.1016/j.scs.2019.101484.

[73] K. Liu, T.-Z. Liu, P. Jian, and Y. Lin, "The re-optimization strategy of multi-layer hybrid building's cooling and heating load soft sensing technology research based on temperature interval and hierarchical modeling techniques," *Sustainable Cities and Society*, vol. 38, pp. 42–54, Apr. 2018, doi: 10.1016/j.scs.2017.11.034.

[74] J. López Gómez, F. Troncoso Pastoriza, E. A. Fariña, P. Eguía Oller, and E. Granada Álvarez, "Use of a numerical weather prediction model as a meteorological source for the estimation of heating demand in building thermal simulations," *Sustainable Cities and Society*, vol. 62, p. 102403, Nov. 2020, doi: 10.1016/j.scs.2020.102403.

[75] J. Vivian, U. Chiodarelli, G. Emmi, and A. Zarrella, "A sensitivity analysis on the heating and cooling energy flexibility of residential buildings," *Sustainable Cities and Society*, vol. 52, p. 101815, Jan. 2020, doi: 10.1016/j.scs.2019.101815.

[76] A. Tsanas and A. Xifara, "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools," *Energy and Buildings*, vol. 49, pp. 560–567, 2012, doi: 10.1016/j.enbuild.2012.03.003.

[77] M. Malvoni, M. G. De Giorgi, and P. M. Congedo, "Forecasting of PV Power Generation using weather input data-preprocessing techniques," in *Energy Procedia*, 2017, vol. 126, pp. 651–658, doi: 10.1016/j.egypro.2017.08.293.

[78] A. S. Ahmad *et al.*, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renewable and Sustainable Energy Reviews*, vol. 33. pp. 102–



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



109, 2014, doi: 10.1016/j.rser.2014.01.069.

[79] I. H. Witten, E. Frank, and Mark A. Hall, *Data Mining: Practical Machine learning*. 2011.

[80] T. Y. Kwok and D. Y. Yeung, “Constructive algorithms for structure learning in feedforward neural networks for regression problems,” *IEEE Transactions on Neural Networks*, vol. 8, no. 3, pp. 630–645, 1997, doi: 10.1109/72.572102.

[81] A. Moradzadeh and K. Pourhossein, “Early Detection of Turn-to-Turn Faults in Power Transformer Winding: An Experimental Study,” in *Proceedings 2019 International Aegean Conference on Electrical Machines and Power Electronics, ACEMP 2019 and 2019 International Conference on Optimization of Electrical and Electronic Equipment, OPTIM 2019*, 2019, pp. 199–204, doi: 10.1109/ACEMP-OPTIM44294.2019.9007169.

[82] D. Bienvenido-Huertas, F. Farinha, M. J. Oliveira, E. M. J. Silva, and R. Lança, “Comparison of artificial intelligence algorithms to estimate sustainability indicators,” *Sustainable Cities and Society*, vol. 63, p. 102430, Dec. 2020, doi: 10.1016/j.scs.2020.102430.

[83] N. Bisoyi, H. Gupta, N. P. Padhy, and G. J. Chakrapani, “Prediction of daily sediment discharge using a back propagation neural network training algorithm: A case study of the Narmada River, India,” *International Journal of Sediment Research*, vol. 34, no. 2, pp. 125–135, 2019, doi: 10.1016/j.ijsrc.2018.10.010.

[84] A. Moradzaeh and K. Khaffafi, “Comparison and Evaluation of the Performance of Various Types of Neural Networks for Planning Issues Related to Optimal Management of Charging and Discharging Electric Cars in Intelligent Power Grids,” *Emerging Science Journal*, vol. 1, no. 4, Jan. 2018, doi: 10.28991/ijse-01123.

[85] B. Guo and J. Zhen, “Coherent Signal Direction Finding With Sensor Array Based on Back Propagation Neural Network,” *IEEE Access*, vol. 7, pp. 172709–172717, 2019, doi: 10.1109/ACCESS.2019.2956555.

[86] S. Lu, Q. Li, L. Bai, and R. Wang, “Performance predictions of ground source heat pump system based on random forest and back propagation neural network models,” *Energy Conversion and Management*, vol. 197, 2019, doi: 10.1016/j.enconman.2019.111864.

[87] D. Antanasijević, V. Pocajt, M. Ristić, and A. Perić-Grujić, “Modeling of energy consumption and related GHG (greenhouse gas) intensity and emissions in Europe using general regression neural networks,” *Energy*, vol. 84, pp. 816–824, 2015, doi: 10.1016/j.energy.2015.03.060.

[88] C. M. Hong, F. S. Cheng, and C. H. Chen, “Optimal control for variable-speed wind generation systems using General Regression Neural Network,” *International Journal of Electrical Power and Energy Systems*, vol. 60, pp. 14–23, 2014, doi: 10.1016/j.ijepes.2014.02.015.

[89] A. Mansour-Saatloo, A. Moradzadeh, B. Mohammadi-Ivatloo, A. Ahmadian, and A. Elkamel, “Machine learning based PEVs load extraction and analysis,” *Electronics (Switzerland)*, vol. 9, no. 7, pp. 1–15, Jul. 2020, doi: 10.3390/electronics9071150.

[90] Y. W. Huang, M. Q. Chen, Y. Li, and J. Guo, “Modeling of chemical exergy of agricultural biomass using improved general regression neural network,” *Energy*, vol. 114, pp. 1164–1175, 2016, doi: 10.1016/j.energy.2016.08.090.

[91] E. E. Elattar, S. K. Elsayed, and T. A. Farrag, “Hybrid Local General Regression Neural Network and Harmony Search Algorithm for Electricity Price Forecasting,” *IEEE Access*, vol. 9, pp. 2044–2054, 2021, doi: 10.1109/ACCESS.2020.3048519.

[92] A. Shinypradeepa and C. Vaithilingam, “Voltage Profile Assessment in Power Distribution System Using Generalized Regression Neural Network,” in *Energy Procedia*, 2017, vol. 117, pp. 209–215, doi: 10.1016/j.egypro.2017.05.124.

[93] R. Becker and D. Thrän, “Completion of wind turbine data sets for wind integration studies applying random forests and k-nearest neighbors,” *Applied Energy*, vol. 208, pp. 252–262, 2017, doi: 10.1016/j.apenergy.2017.10.044.

[94] Y. Hamed, A. Ibrahim Alzahrani, A. Shafie, Z. Mustaffa, M. Che Ismail, and K. Kok Eng, “Two steps hybrid calibration algorithm of support vector regression and K-nearest neighbors,” *Alexandria Engineering Journal*, 2020, doi: 10.1016/j.aej.2020.01.033.

[95] M. Yesilbudak, S. Sagiroglu, and I. Colak, “A novel implementation of kNN classifier based on multi-tupled meteorological input data for wind power prediction,” *Energy Conversion and Management*, vol. 135, pp. 434–444, 2017, doi: 10.1016/j.enconman.2016.12.094.

[96] C. Hu, G. Jain, P. Zhang, C. Schmidt, P. Gomadam, and T. Gorka, “Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery,” *Applied Energy*, vol. 129, pp. 49–55, 2014, doi: 10.1016/j.apenergy.2014.04.077.

[97] A. Navot, L. Shpigelman, N. Tishby, and E. Vaadia, “Nearest neighbor based feature selection for regression and its application to neural activity,” in *Advances in Neural Information Processing Systems*, 2005, pp. 995–1002.

[98] H. Jafarian, H. Sayyaadi, and F. Torabi, “Modeling and optimization of dew-point evaporative coolers based on a developed GMDH-type neural network,” *Energy Conversion and Management*, vol. 143, pp. 49–65, 2017, doi: 10.1016/j.enconman.2017.03.015.

[99] A. G. Ivakhnenko, “Polynomial Theory of Complex Systems,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 1, no. 4, pp. 364–378, 1971, doi: 10.1109/TSMC.1971.4308320.

[100] M. G. De Giorgi, M. Malvoni, and P. M. Congedo, “Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine,” *Energy*, vol. 107, pp. 360–373, 2016, doi: 10.1016/j.energy.2016.04.020.

[101] T. N. Nguyen, S. Lee, H. Nguyen-Xuan, and J. Lee, “A novel analysis-prediction approach for geometrically nonlinear problems using group method of data handling,” *Computer Methods in Applied Mechanics and Engineering*, vol. 354, pp. 506–526, 2019, doi: 10.1016/j.cma.2019.05.052.

[102] I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, and F. Khoshbin, “GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs,” *Engineering Science and Technology, an International Journal*, vol. 18, no. 4, pp. 746–757, 2015, doi: 10.1016/j.jestch.2015.04.012.

[103] S. Jeddi and S. Sharifian, “A hybrid wavelet decomposer and GMDH-ELM ensemble model for Network function virtualization workload forecasting in cloud computing,” *Applied Soft Computing Journal*, vol. 88, 2020, doi: 10.1016/j.asoc.2019.105940.

[104] Z. Zhang, Z. Lai, Y. Xu, L. Shao, J. Wu, and G. Sen Xie, “Discriminative Elastic-Net Regularized Linear Regression,” *IEEE Transactions on Image Processing*, vol. 26, no. 3, pp. 1466–1481, 2017, doi: 10.1109/TIP.2017.2651396.

[105] Q. Liu, Y. Sun, C. Wang, T. Liu, and D. Tao, “Elastic net hypergraph learning for image clustering and semi-supervised classification,” *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 452–463, 2017, doi: 10.1109/TIP.2016.2621671.

[106] J. K. Pant and S. Krishnan, “Group sparse structure and elastic-net regularization for compressive sensing of pulse-type physiological signals,” *Biomedical Signal Processing and Control*, vol. 57, 2020, doi: 10.1016/j.bspc.2019.101685.

[107] Y. Shen, B. Han, and E. Braverman, “Stability of the elastic net estimator,” *Journal of Complexity*, vol. 32, no. 1, pp. 20–39, 2016, doi: 10.1016/j.jco.2015.07.002.

[108] G. Pei, Y. Wang, Y. Cheng, and L. Zhang, “Joint Label-Density-Margin Space and Extreme Elastic Net for Label-Specific Features,” *IEEE Access*, vol. 7, pp. 112304–112317, 2019, doi: 10.1109/access.2019.2934742.

[109] A. E. Hoerl and R. W. Kennard, “Ridge Regression: Applications to Nonorthogonal Problems,” *Technometrics*, vol. 12, no. 1, pp. 69–82, 1970, doi: 10.1080/00401706.1970.10488635.

[110] R. Tibshirani, “Regression Shrinkage and Selection Via the Lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.

[111] Z. Tang, H. Yin, C. Yang, J. Yu, and H. Guo, “Predicting the electricity consumption of urban rail transit based on binary



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



nonlinear fitting regression and support vector regression," *Sustainable Cities and Society*, vol. 66, p. 102690, Mar. 2021, doi: 10.1016/j.scs.2020.102690.

[112] J. Zhao and K. P. Lam, "Influential factors analysis on LEED building markets in U.S. East Coast cities by using Support Vector Regression," *Sustainable Cities and Society*, vol. 5, no. 1, pp. 37–43, Dec. 2012, doi: 10.1016/j.scs.2012.05.009.

[113] H. Taghavifar and A. Mardani, "A comparative trend in forecasting ability of artificial neural networks and regressive support vector machine methodologies for energy dissipation modeling of off-road vehicles," *Energy*, vol. 66, pp. 569–576, 2014, doi: 10.1016/j.energy.2014.01.022.

[114] K. S. Ni and T. Q. Nguyen, "Image Superresolution Using Support Vector Regression," *IEEE Transactions on Image Processing*, vol. 16, no. 6, pp. 1596–1610, Jun. 2007, doi: 10.1109/TIP.2007.896644.

[115] J. Antonanzas, R. Urraca, F. J. Martinez-De-Pison, and F. Antonanzas-Torres, "Solar irradiation mapping with exogenous data from support vector regression machines estimations," *Energy Conversion and Management*, vol. 100, pp. 380–390, 2015, doi: 10.1016/j.enconman.2015.05.028.

[116] A. Moradzadeh and K. Pourhossein, "Application of Support Vector Machines to Locate Minor Short Circuits in Transformer Windings," in *2019 54th International Universities Power Engineering Conference, UPEC 2019 - Proceedings*, 2019, pp. 1–6, doi: 10.1109/UPEC.2019.8893542.

[117] S. Hosseinpour, M. Aghbashlo, M. Tabatabaei, and E. Khalife, "Exact estimation of biodiesel cetane number (CN) from its fatty acid methyl esters (FAMEs) profile using partial least square (PLS) adapted by artificial neural network (ANN)," *Energy Conversion and Management*, vol. 124, pp. 389–398, 2016, doi: 10.1016/j.enconman.2016.07.027.

[118] A. A. Shawul, S. Chakma, and A. M. Melesse, "The response of water balance components to land cover change based on hydrologic modeling and partial least squares regression (PLSR) analysis in the Upper Awash Basin," *Journal of Hydrology: Regional Studies*, vol. 26, 2019, doi: 10.1016/j.ejrh.2019.100640.

[119] M. A. Mohd Noor *et al.*, "Quality Prediction of Refined Bleached Deodorized Palm Oil (RBDPO) Using Partial Least Square Regression Technique," in *Energy Procedia*, 2017, vol. 142, pp. 3002–3007, doi: 10.1016/j.egypro.2017.12.391.

[120] A. I. Abdel-Rahman and G. J. Lim, "A nonlinear partial least squares algorithm using quadratic fuzzy inference system," *Journal of Chemometrics*, vol. 23, no. 10, pp. 530–537, 2009, doi: 10.1002/cem.1249.

[121] R. Samsudin, P. Saad, and A. Shabri, "A hybrid least squares support vector machines and GMDH approach for river flow forecasting," *Hydrology and Earth System Sciences Discussions*, vol. 7, no. 3, pp. 3691–3731, 2010, doi: 10.5194/hessd-7-3691-2010.

[122] R. Samsudin, P. Saad, and A. Shabri, "A hybrid GMDH and least squares support vector machines in time series forecasting," *Neural Network World*, vol. 21, no. 3, pp. 251–268, 2011, doi: 10.14311/NNW.2011.21.015.

[123] B. Choubin, S. Khalighi-Sigaroodi, A. Malekian, and Ö. Kişi, "Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals," *Hydrological Sciences Journal*, vol. 61, no. 6, pp. 1001–1009, 2016, doi: 10.1080/02626667.2014.966721.

[124] A. Moradzadeh, H. Teimourzadeh, B. Mohammadi-Ivatloo, and K. Pourhossein, "Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults," *International Journal of Electrical Power & Energy Systems*, vol. 135, p. 107563, Feb. 2022, doi: 10.1016/j.ijepes.2021.107563.

**ARASH MORADZADEH** (Student Member IEEE) was born in Tabriz, Iran, in 1991. He received the B.S. degree in electrical power engineering and the M.S. degree in power electronics and electrical machines from Islamic Azad University of Tabriz, Tabriz, Iran, respectively in 2016 and 2019. He is currently working toward the Ph.D. degree in power electrical engineering with the University of Tabriz, Tabriz, Iran.

His current research interests include power and energy systems, cyber-physical systems, energy management, intelligent energy systems, transient in power systems, diagnostics and condition monitoring of power transformers, frequency response analysis, and application of data mining methods to design, optimize and analysis of energy systems.

**BEHNAM MOHAMMADI-IVATLOO** (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from University of Tabriz, Tabriz, Iran, in 2006, and the M.Sc. and Ph.D. degree from Sharif University of Technology, Tehran, Iran, in 2008, all with honors. He is currently a Professor with Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran. His main area of research is economics, operation, and planning of intelligent energy systems in a competitive market environment.

He as an Associate Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE ACCESS, IET SMART GRID, and the SUSTAINABILITY.

**MEHDI ABAPOUR** (Member IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from The University of Tabriz, Tabriz, Iran, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering from The Tarbiat Modares University, Tehran, Iran, in 2013. Currently, he is an Associate Professor with the School of Electrical and Computer Engineering, University of Tabriz. His research interests include Reliability, Energy management, demand response and Power electronics.

**AMJAD ANVARI-MOGHADDAM** (Senior Member, IEEE) received the Ph.D. degree (Hons.) in power systems engineering from the University of Tehran, in 2015. Currently, he is an Associate Professor and Leader of Intelligent Energy Systems and Flexible Markets (iGRIDS) Research Group at the Department of Energy (AAU Energy), Aalborg University where he is also acting as the Vice-Leader of Power Electronic Control, Reliability and System Optimization (PESYS) and the coordinator of Integrated Energy Systems Laboratory (IES-Lab). He has coauthored more than 250 technical articles, five books, and nine book chapters, in the field. His research interests include planning, control, and operation management of microgrids, renewable/hybrid power systems, and integrated energy systems with appropriate market mechanisms. He as an



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3136091, IEEE Access



Associate Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE ACCESS, the IEEE SYSTEMS JOURNAL, the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, and the IEEE POWER ENGINEERING LETTERS. He is the Vice-Chair of IEEE Power and Energy Society (PES) Danish Chapter and serves as a Technical Committee Member for several IEEE PES/IES/PEL and CIGRE working groups. He was a recipient of 2020 DUO– India Fellowship Award, DANIDA Research Fellowship grant from the Ministry of Foreign Affairs of Denmark in 2018, IEEE-CS Outstanding Leadership Award 2018 (Halifax, Nova Scotia, Canada), and the 2017 IEEE-CS Outstanding Service Award (Exeter-UK).

**SANJIBAN SEKHAR ROY** (Member IEEE) is an Associate Professor in the School of Computer Science and Engineering, Vellore Institute of Technology. He uses Deep Learning and machine learning techniques to solve many complex engineering problems, especially those are related to imagery. Dr Roy has vast experience in research specially in the field of advanced machine learning and deep learning. He is specialized in ML & deep convolutional neural networks towards solving various image related complex problem and various engineering problems such as computational biology, civil, energy inspired problems. Dr. Roy also has edited special issues for journals and many books with reputed international publishers such as Elsevier springer and IGI Global. Dr Roy has been an associatet researcher to Ton Duc Thang University, Veitnam from 2019-2020. Very recently, Ministry of National Education, Romania in collaboration with "Aurel Vlaicu" University Arad Faculty of Engineers, Romania has awarded him with "Diploma of Excellence" as a sign of appreciation for the special achievements obtained in the scientific research activity in 2019.



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/