Energy 165 (2018) 709–726

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders

Jui-Sheng Chou*, Duc-Son Tran

Department of Civil and Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 15 June 2018
Received in revised form
17 September 2018
Accepted 20 September 2018
Available online 21 September 2018

Keywords:
Energy consumption
Residential buildings
Time-series forecasting
Data mining
Artificial intelligence
Machine learning

ABSTRACT

Energy consumption in buildings is increasing because of social development and urbanization. Forecasting the energy consumption in buildings is essential for improving energy efficiency and sustainable development, and thereby reducing energy costs and environmental impact. This investigation presents a comprehensive review of machine learning (ML) techniques for forecasting energy consumption time series using actual data. Real-time data were collected from a smart grid that was installed in an experimental building and used to evaluate the efficacy and effectiveness of statistical and ML techniques. Well-known artificial intelligence techniques were used to analyze energy consumption in single and ensemble scenarios. An in-depth review and analysis of the 'hybrid model' that combines forecasting and optimization techniques is presented. The comprehensive comparison demonstrates that the hybrid model is more accurate than the single and ensemble models. Both the accuracy of prediction and the suitability for use of these models are considered to support users in planning energy management.

1. Introduction

The rate of development of countries around the world is growing tremendously and is inevitable, leading to an increased demand for energy. Furthermore, the rapid expansion of residential and commercial areas contributes to an increasing in building energy consumption. Forecasting energy consumption has become crucial for estimating energy usage. As well as providing environmental benefits, building energy efficiency can provide great economic benefits. Buildings with efficient energy systems and management strategies have much lower operating costs. Several countries have accelerated the implementation of energy regulations and codes for various types of building. Many computer software packages have also been developed and are being widely used for designing energy-efficient buildings.

Global energy consumption is estimated to increase by 53% from 505 quadrillion Btu in 2008 to 700 quadrillion Btu in 2035 [1]. Numerous countries have accelerated the implementation of energy codes and regulations for various types of buildings. These regulations impose basic requirements to ensure that new buildings designs are energy-efficient with a view to reducing their energy consumption and related CO_2 emissions [2]. Studies reveal that end-use energy efficiency can significantly reduce total global energy consumption [3,4]. Even minor shifts in peak demand can provide major savings for both consumers and utilities [5]. Therefore, improving the energy efficiency of buildings is necessary for decreasing environmental impact and controlling the influence of energy costs.

To optimize the energy performance of buildings, appropriate operational strategies should be applied to energy management systems. Managers must continuously monitor and manage the time series of energy along with factors that influence the energy performance of their buildings. Prediction is an important part of the continuous monitoring and management of energy consumption. Accurately predicting energy consumption is critical for improving building energy efficiency [6]. As a time-series forecasting model learns from previous energy consumption usage patterns, a slightly higher than forecasted energy consumption over a period may cause facility managers and energy systems to be notified of the change.

In the past two decades, various forecasting techniques have been used to predict the energy consumption of buildings [7–14]; they fall into three categories, which are engineering methods, statistical methods and artificial intelligence methods. Of these, the most widely implemented are artificial intelligence (AI) methods,

* Corresponding author.
E-mail addresses: jschou@mail.ntust.edu.tw (J.-S. Chou), son.neutron@gmail.com (D.-S. Tran).

https://doi.org/10.1016/j.energy.2018.09.144
0360-5442/

which include Artificial Neural Networks (ANNs) [9,15], the Classification and Regression Tree (CART) [16], the Support Vector Machine (SVM) [7], the Linear Regression (LR) [17], the autoregressive integrated moving average (ARIMA) [9,18–21], and the seasonal autoregressive integrated moving average (SARIMA) methods [22]. The nonlinearity and volatility of real-time energy usage create the difficulties in predicting energy consumption in residential buildings.

Recent reviews of energy forecasting provide details of the existing forecasting models and their classification. Chirag et al. (2017) has compiled the results of previous researchers [23]. They analyzed the nine most popular forecasting techniques that are based on the machine learning platform. Zhao and Magoules (2012) reviewed recently developed models for predicting energy consumption, including complicated and simplified statistical methods, engineering methods, and artificial intelligence methods [2]. However, previous reviews are based on previous research results but not on the same database. Therefore, the evaluations of models have no standards of comparison.

To fill the aforementioned research gap, the current study presents a comprehensive review of energy consumption prediction using machine learning techniques, based on the energy consumption pattern in a residential building. The goal is to facilitate the efficient 1-day-ahead forecasting of energy consumption for end-users and managers in residential buildings. One-day-ahead forecasts can be applied to increase the productivity of energy usage by controlling electrical equipment in buildings. Energy consumption is predicted using basic AI techniques - ANNs, SVR (Support Vector Regression), CART, LR and SARIMA - using two data analytic platforms (RapidMiner Studio and IBM SPSS Modeler). Additionally, two ensemble models, which are generated by combining these AI techniques, use voting and bagging methods. Two advanced AI models are created by hybridizing machine learning and metaheuristic optimization, which are integrated with the SARIMA.

The objectives of this research are as follows.

- To provide an exhaustive and comprehensive review of techniques for short-term load forecasting time series of energy consumption for buildings.
- To deploy single, ensemble, and hybrid energy consumption forecasting models.
- Both the accuracy of prediction and the suitability for use are considered to evaluate the extent to which the presented models support users in planning energy management.

The rest of this paper is organized as follows. Section 2 reviews the literature on time-series forecasting methods, machine learning methods, the hybrid methods, and existing reviews of methods for predicting the energy consumption of buildings. Section 3 describes methods that are currently used to predict energy consumption. Section 4 elucidates the energy consumption dataset and parameters that are input into the AI models. Section 5 compares the performances of the implemented models. Finally, Section 6 draws conclusions and offers suggestions for end-users of buildings and energy managers.

2. Literature review

2.1. Energy consumption in buildings

The rapidly increasing use of energy around the world has already raised concerns of the depletion energy resources, the exceeding of supply capacities, and severe environmental impacts (global warming, ozone layer depletion, climate change, and others) [24]. The size and location of a residential building are key factors in determining its energy consumption. The amount of energy that is used in residential buildings are related mainly to the weather, architectural design, energy systems and income of the householders. Energy use in the built environment will grow by 34% in the next 20 years, at a mean rate of 1.5% annually. In 2030, dwellings will consume 67% of all energy and the non-domestic sector will consume 33% [24]. Energy consumption represents a large fraction of the total consumption expenditure of society.

Recently, building energy management systems have seen rapid development and end-users have thus been enabled to use electricity in their buildings more effectively [25–35]. For instance, Agheno et al. (2014) presented a case study of the energy and environmental performances of a lighting control system in offices [25]. They evaluated the results of ten offices in Torino (Italy) concerning potential energy savings (from 17% to 32%), taking into account both the monitored annual electric energy consumption and the parasitic energy consumption. Zhou et al. (2014) proposed a method for controlling energy consumption in real time for homes in the United Kingdom [27]. The simulation results indicated that their proposed control approach can optimize the schedule of use of home appliances and battery charging/discharging behaviors, even if the forecasted information is inaccurate.

Moreover, Valor et al. (2001) analyzed the relationship between building electricity load and daily outdoor temperature in Spain, using a population-weighted temperature index [36]. Their study indicated that the sensitivity of electricity load to daily outdoor temperature has increased over time, in a higher degree for summer than for winter, although the sensitivity in the cold season is always more significant than in the warm season. The association between electricity data and outdoor temperature defines the heating and cooling demand functions, which show correlation coefficients of 0.79 and 0.87, and forecast electricity load with standard errors of estimate of 4% and 2%, respectively.

Energy optimization in buildings has been extensively investigated in the last decade. Numerous investigations have sought to address the energy efficiency of buildings [29,37–39] and several research projects have been carried out to identify the best-performing buildings with respect to energy efficiency. Monfet et al. (2014) presented a novel approach to predicting the energy demand of commercial buildings using case-based reasoning. Their method can be used for predicting energy demand and can be implemented in building operation systems [29].

Costa et al. (2012) presented the key factors methodology to support energy managers in determining the optimal building operation strategy with respect to both energy consumption and thermal comfort [40]. Chou and Ngo (2016) presented a framework for smart grid big data analytics and the required components of an energy-saving decision-support system [41]. These studies have focused on developing prediction platforms and methods to enhance the efficiency and reliability of energy management systems.

2.2. AI techniques for predicting time-series energy consumption

Energy consumption analysis involves methods for analyzing time series data to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values [42]. A time series is an ordered sequence of values that are recorded at equal intervals of time. The analysis of a time series comprises two parts [23]. In the first part, the structure and underlying pattern of the observed data are obtained. In the second part, a model is fitted to the data to support future predictions. Time series analysis is used in several fields, including statistics, energy prediction, signal processing, pattern recognition, econometrics, weather forecasting, earthquake prediction, astronomy, and in any domain of applied science and engineering that involves temporal measurements.

Time series analysis is extensively utilized to forecast energy consumption as buildings are increasingly being monitored in real time. Such monitoring can provide historic data that can be used for critical analysis and in the forecasting of the energy consumption of buildings. Many investigations have examined the energy performance of buildings [43–51]. Various time series models have been developed for forecasting energy consumption in buildings [52–58]. This study considers some of the most popular time series forecasting techniques that can be applied to the energy consumption of buildings.

Artificial Neural Networks are typically modeling techniques that are based on a collection of connected units or nodes, called artificial neurons. ANNs models have been extensively used to predict the energy consumption of buildings [59,60]. Adya and Collopy (1998) reviewed the effectiveness of ANNs in forecasting [61]. Ekici and Aksoy (2009) proposed a model to predict the energy requirements of buildings, and their relationship with the orientation of a building, thickness of insulation, and transparency ratio, using ANNs [53]. Their simulation results revealed that ANN provides satisfactory results with deviation of 3.43% and an accuracy prediction rate of 94.8–98.5%. Hamzacebi (2007) developed models using ANNs to predict net electricity consumption in Turkey [62]. The results demonstrated that the ANN technique gives better results than the Model for Analysis of Energy Demand (MAED) technique.

Support Vector Machines are used in many machine learning tasks, such as pattern recognition, object classification, and time series prediction, including especially the forecasting of energy consumption. Dong et al. (2005) applied SVM to predict the energy consumption of buildings in a tropical region [63]. They collected mean monthly utility bills to develop and test models. Their predictions had a coefficients of variance of less than 3% and a percentage error of less than 4%. Bozic et al. (2010) proposed a Least Squares Support Vector Machine (LSSVM) for short-term load forecasting [64]. They used a week of hourly data to make hourly and daily load forecasts. The results yielded mean absolute percentage error values between 0.93% and 3.04% for daily forecasts.

Linear Regression (LR) models have been used to predict the energy consumption of buildings because of their ease of use. Catalina et al. (2008) developed regression models to predict monthly heating demand in the single-family residential sector in temperate climates, to be used by architects or design engineers in the initial stage of their projects to find efficiently energetic solutions with respect to energy performance [65]. Hygh et al. (2012) presented a new modeling method that used a multivariate regression model to quantify the energy performance of buildings in their early design stages [66]. Ease of use is considered to be the main advantage of the LR method because no hyperparameters were tuned. However, the LR method has a major limitation, which is its inapplicability to nonlinear problems [67].

Autoregressive integrated moving average (ARIMA) models are effective statistical tools and provide the most general framework for time series forecasting techniques [18,19,21,22,68]. Such models are widely used linear models of univariate time series. Newsham and Birt (2010) proposed an ARIMA with eXternal input (ARIMAX) model to predict the energy demand for an office building [69]. They used occupancy data as an external predictor to improve the model. Chujai et al. (2013) presented ARIMA and ARMA (autoregressive moving average) models for analyzing time series of household electric consumption data [18]. Their results revealed that the ARIMA model was the best model for finding the most suitable forecasting period. Mohamed et al. (2011) performed short-term load forecasting using a double seasonal ARIMA model [70] and half hourly load demand data for Malaysia over a year. The estimated MAPE was 0.99%.

Ensemble models meta heuristically combine the outputs of many generated individual AI-based models such that the combination outperforms each of the single models [71–74]. Ensemble techniques for forecasting time series are indispensable for many practical data mining applications [75]. Soares et al. (2018) described a variation of the data cloud-based intelligent method known as the typicality and eccentricity-based method for data analysis (TEDA), which uses ensemble models [76]. They developed an ensemble of cloud and fuzzy models as well as aggregation operators to give single-valued and granular predictions of time series. The ensemble provided the best overall accuracy with one out of four datasets, and TEDA provided the best predictions with the remaining datasets.

Hybrid linear and nonlinear models have recently been developed for forecasting time-series data, and especially the energy consumption in buildings [77–80]. Tan et al. (2010) proposed a novel price forecasting method that combined wavelet transformation with ARIMA and generalized autoregressive conditional heteroskedasticity (GARCH) models [20]. They evaluated the accuracy of the proposed hybrid method, which captured the complex features of electricity prices. Chou and Ngo (2016) developed a novel time-series sliding window metaheuristic optimization-based machine learning system for forecasting energy consumption for buildings in real-time, collected using a smart grid [6]. Their prediction system had high and reliable accuracy in making 1-day-ahead predictions of the energy consumption in a building, with a total error rate of 1.181% and a mean absolute error of 0.026 kWh.

Azadeh et al. (2010) presented an integrated fuzzy regression and time series framework to estimate and predict electricity demand seasonal and monthly changes in electricity consumption, especially developing countries such as China and Iran, using non-stationary data [49]. Their results indicate that the proposed algorithm is superior to other available models and algorithms. Lee and Tong (2012) described a novel hybrid dynamic approach that combines a dynamic grey model with genetic programming to forecast energy consumption [78]. They used it to predict energy consumption because of its excellent accuracy, applicability to cases with limited data sets and ease of computation using mathematical software.

This section also summarizes review papers that focus on the AI-based forecasting of energy consumption in buildings [43,46,67,81–83]. For each review paper, the objective of the study, the time series forecasting techniques reviewed, and the models compared, are all provided. Martínez-Álvarez et al. (2015) presented a survey of data mining techniques used in electricity-related time series forecasting [72]. They concluded that data mining outperforms classical techniques. Deb et al. (2017) comprehensively reviewed machine learning techniques for forecasting energy consumption time series [23]. They analyzed the nine most popular forecasting techniques that are based on the machine learning platform. They also provided an in-depth review and analysis of 'hybrid models' that combine at least two forecasting techniques.

Zhao and Magoulès (2012) reviewed recently developed models for predicting the energy consumption of buildings, including elaborate and simplified engineering methods, statistical methods and artificial intelligence methods [2]. They concentrated on the application of these models to new problems of prediction, optimizing the model parameters or input samples to improve performance, and comparing models under certain conditions. Raza and Khosravi (2015) provided a comprehensive and systematic review of the literature of artificial intelligence-based short-term load forecasting techniques for smart grids and buildings [84]. The major objective of their investigation was to review, identify, evaluate and analyze the performance of AI-based load forecasting models and identified gaps in the relevant research.

Ahmad et al. (2014) reviewed AI-based methods, such as the SVM and ANN methods, for forecasting the electrical energy consumption of buildings [7]. Combining models can improve forecasting performance. They asserted that a hybrid model of group method of data handling (GMDH) and LSSVM has great potential for forecasting time series in various fields, such as stock, energy consumption, and others. Wang and Srinivasan (2017) [67] conducted an in-depth review of single AI-based methods such as MLR, ANNs, and SVR, and ensemble prediction methods that combine AI-based prediction models to improve predictive accuracy by a large factor. They elaborated the principles, applications, advantages, and limitations of AI-based prediction methods and discussed the future of research into AI-based methods for predicting the energy use of buildings.

Amasyali and El-Gohary (2018) [85] reviewed works on the development of data-driven models to predict the energy consumption of buildings, with a focus on the scope of prediction, the methods of preprocessing the data and the properties of the data, ML algorithms that are utilized for prediction, and performance metrics. Recent reviews of the forecasting of energy consumption provide details of time-series techniques, forecasting models, and their classification. Aforementioned papers are based on the results of previous studies, but these are not based on the same database. Therefore, no basis for comparison and verification exists. This study conducts a comprehensive review of the forecasting of time series using machine learning techniques that are based on an energy consumption pattern in a residential building.

3. Methodology

3.1. AI-based prediction models

A typical AI-based prediction method consists of four main steps, which are data collection, data preprocessing, model training, and model testing [67]. Many AI-based learning algorithms and prediction models with these steps are used to predict the energy consumption of buildings. AI-based prediction models may be classified into three categories, which are single, ensemble, and hybrid based on their prediction components. Single prediction models use one learning algorithm; ensemble models comprise multiple prediction models, which are integrated in a manner that determines the output data. Hybrid models combine two or more machine learning techniques. These models are more robust than the others as they frequently exhibit the advantages of the incorporated techniques and provide improved forecasting accuracy. Fig. 1 shows the general structure of time series predicting techniques. The following sub-sections introduce these three major categories of model.

3.1.1. Single models

In this study, single prediction methods include all available AI-based prediction techniques. Such methods may be categorized into several types by the phenomena they are used to forecast. This investigation focuses on ML techniques that are used to forecast energy consumption in buildings. Some popular models are presented to provide a broad understanding of single predictive models; these include ANNs, C&R Tree (Classification and Regression Tree), SVR, and LR. Fig. 2 presents the general structure of single prediction models.

3.1.1.1. Artificial neural networks. Artificial neural networks are the most widely used AI models; they simulate the ability of brain neurons to process information. The use of ANNs to solve forecasting problems has recently attracted considerable research attention because they substantially outperform previous implemented techniques for forecasting based on non-linear input variables [15,47,48,53,84]. The function of the synapse, which is the structure that is responsible for storing information in the brain, is modeled as joints between neural units that are located in adjacent layers, each with a modifiable weight.

An ANN based load-forecasting model is developed using a multilayer neural network with an appropriately modified back propagation-learning algorithm. The training is done in a supervised manner that is for every input vector, the desired output is given and weights are adapted to minimize an error function that measures the discrepancy between the desired output and the actual output computed by the network.

3.1.1.2. Support vector machine for classification and regression. The Support Vector Machine (SVM) approach, developed by Vapnik in 1995 [86], is based on statistical learning theory and the principle of structural risk minimization. The Support Vector Regression (SVR) is a conversion of the SVM that is used to solve regression problems. For nonlinear regression problems, the data are transformed using a

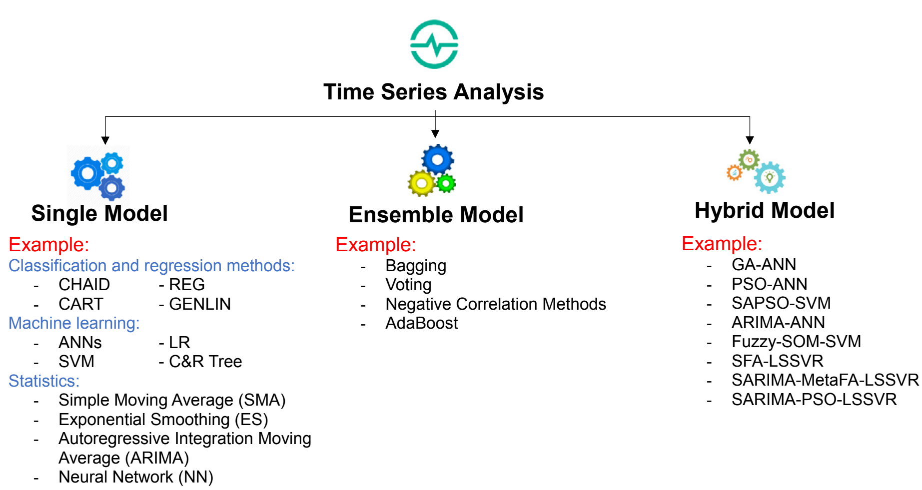

Fig. 1. Stream of time-series predicting techniques.

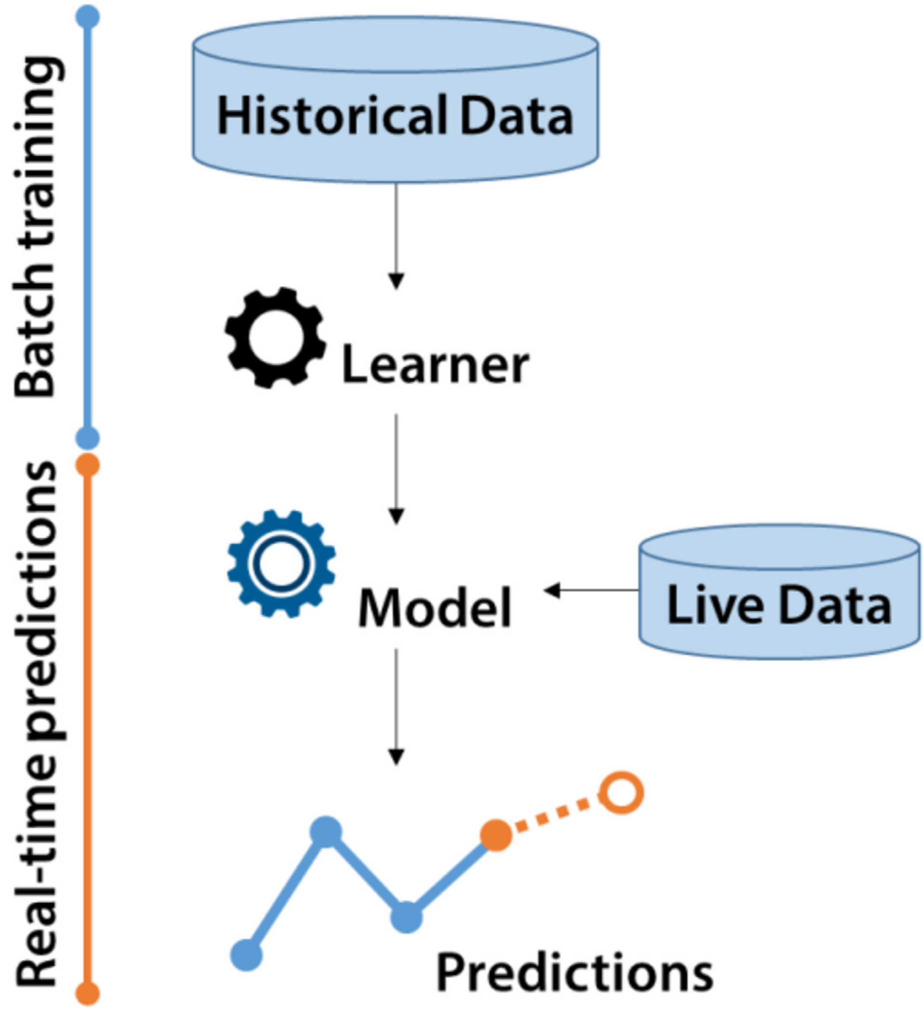

Fig. 2. Framework of single prediction models.

nonlinear kernel function that maps the inputs to a high-dimensional feature space. Accordingly, the general performance of the SVM for classification and regression models depends on the satisfactory selection of the kernel parameters [23]. The basic concept of SVM for function approximation is to map the data into a high-dimensional feature space by a nonlinear kernel mapping and then performing a linear regression in the feature space.

3.1.1.3. Classification and regression tree. Breiman et al. (1984) coined the term "Classification and Regression Trees (CARTs)" to refer to decision tree algorithms that can be used for classification or regression modeling problems [87]. The CART algorithm provides a foundation for important algorithms, such as bagged decision tree, random forest and boosted decision tree algorithms. The representation of a CART model is a binary tree. Each root node represents a single input variable and a split point on that (which is assumed to be numeric). Each of the leaf nodes of the tree contains an output variable, which is used in making predictions.

3.1.1.4. Linear regression. Linear regression (LR) was developed to model the relationship between a scalar dependent variable and one or more independent variables. This investigation used multivariate linear regression (MLR) technique, which is one of multiple versions of LR models. Owing to their ease of use, MLR models have been widely used to predict the energy loads of buildings. This study applies the statistical property of ordinary least squares (OLS) to the linear time series regression model, requiring the specification of the properties of the data and the error term. Several important asymptotic and finite sample results are presented and these results are compared with the statistical properties of the time series regression.

3.1.1.5. Autoregressive integrated moving average. Autoregressive integrated moving average (ARIMA) models are, in theory, the most general class of models for forecasting a time series of data. They are based on the transformation of a time series into a stationary series by a differencing process. A random variable that is a time series is stationary if its statistical properties are all constant over time. Therefore, the ARIMA equation for a time series is a linear equation in which the predictors comprise lags of the dependent variable and of the forecast error.

The problem with using lagged errors as predictors is that the model's predictions are not linear functions of the coefficients, even though they are linear functions of past data. Hence, coefficients in ARIMA models that include lagged errors must be estimated by nonlinear optimization methods rather than by just solving a system of equations. The seasonal autoregressive integrated moving average (SARIMA) model is a simple extension of the ARIMA model. As mentioned, SARIMA models are widely applied for forecasting seasonal time-series data because of their effectiveness in linear prediction.

3.1.2. *Ensemble models*

Ensemble methods have gained substantial attention from the machine learning and soft computing communities in recent years. Ensemble learning methods are extensively used nowadays because of their favorable forecasting predictive performance [73–76,88]. Ensemble learning combines multiple predictions made using multiple methods to get better accuracy than simple prediction and to avoid possible over-fitting. Their combination outperforms each model in baseline cases. Therefore, the following ensemble methods are used herein to obtain improved predicted outputs. Fig. 3 presents the framework of ensemble models.

3.1.2.1. *Voting.* Voting is the simplest ensemble method because it is easy to understand and implement [89]. Voting is a means of meta-classification for combining conceptually similar or dissimilar machine learning classifiers via majority or plurality voting. In this study, the voting method was used to generate 11 ensembles of two, three or four techniques. The two-technique ensembles were ANNs + CART, ANNs + SVR, ANNs + LR, CART + SVR, CART + LR, and SVR + LR. The three-technique ensembles were ANNs + CART + SVR, ANNs + CART + LR, CART + SVR + LR, and ANNs + SVR + LR. The four-technique ensemble was ANNs + CART + SVR + LR.

3.1.2.2. *Bagging.* Bagging is a simple and very powerful ensemble method. Bagging applies the Bootstrap procedure to a high-variance machine learning algorithm, which typically involves decision trees. In the bagging method, bootstrapping is performed to train several models independently and with different training sets. The only parameter in the bagging of decision trees is the number of samples. This number can be set by increasing the number of trees every run until the accuracy stops improving significantly. Very large numbers of models may take a long time to prepare, but will not over-fit the training data. This work develops four uniformity bagging ensemble models using learning techniques; they are an SVR ensemble, an ANN ensemble, a CART ensemble, and an LR ensemble.

3.1.3. *Hybrid models*

Hybrid models combine machine learning techniques with optimization algorithms. They are more powerful than single models as they commonly incorporate the advantages and compensate for the weaknesses of the individual techniques involved, improving forecasting accuracy. Hybrid models can be created with one or more phases, corresponding to different problem-solving goals. The following sections discuss hybrid methods and present two representative hybrid structures for predicting energy consumption.

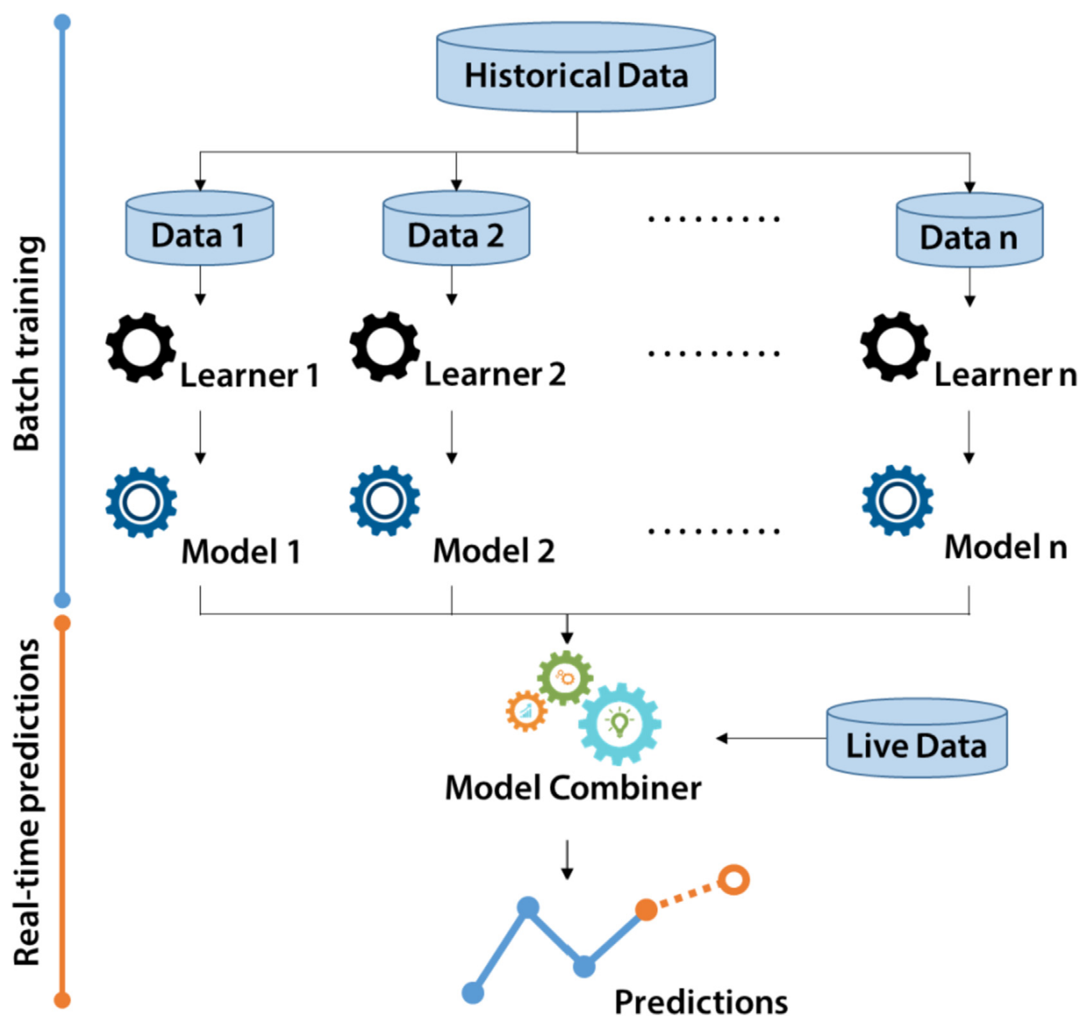

Fig. 3. Framework of ensemble models.

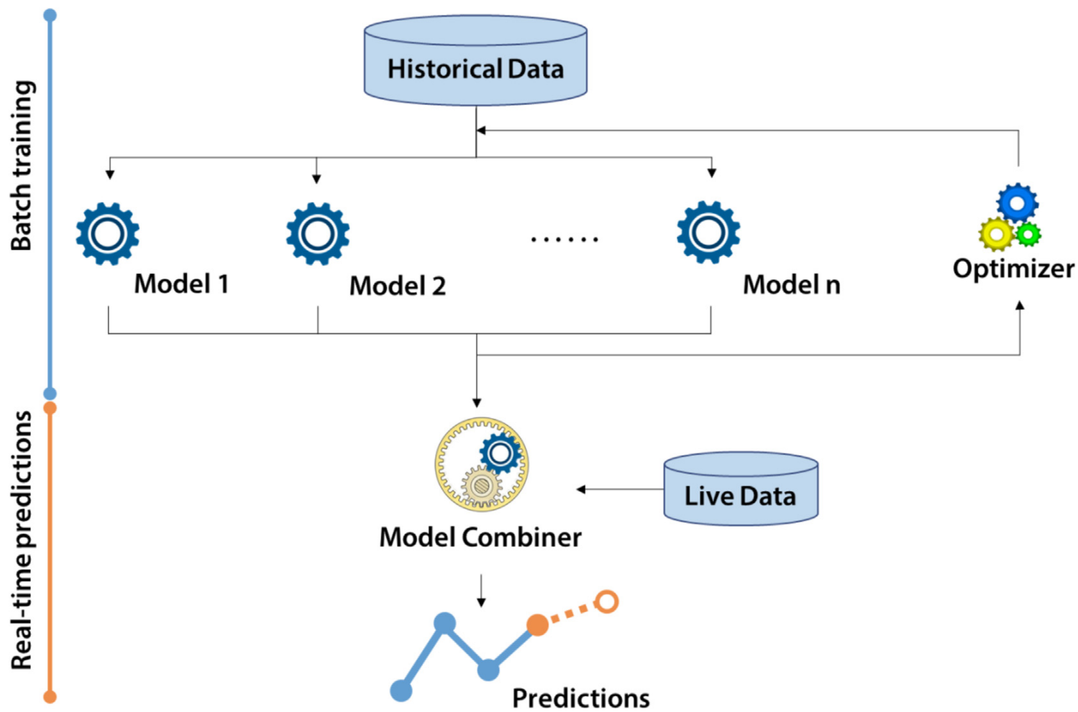

Fig. 4. Stream of single-phase hybrid models.

3.1.3.1. Single-phase models. Hybrid systems have become essential for computational intelligence and soft computing, as they can deal with complex components. The simplest hybrid model is single-phase simulation. A large number of single-phase models have been developed for time-series analysis in recent years [20,90,91]. In the training stage, two or more machine learning techniques are combined, as in the ARIMA-GARCH model. The combined techniques function together to solve all predictive problems. In some cases, these combined models are associated with an optimization algorithm to generate a fine-tuned prediction model, such as the

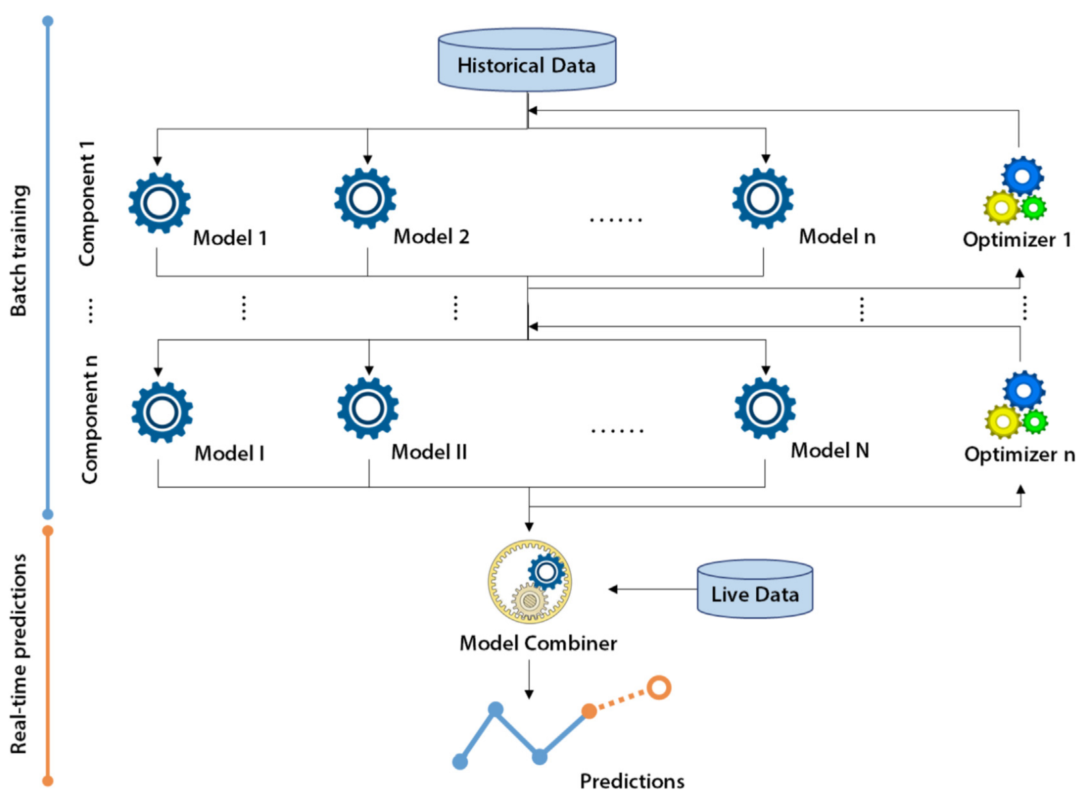

Fig. 5. Stream of multiple-phase hybrid model.

PSO-ANN (particle swarm optimization-artificial neural network) model. Fig. 4 displays the general structure of single-phase hybrid models for forecasting time series.

3.1.3.2. Multiple-phase models. Multiple-phase learning models have been theoretically and empirically demonstrated to significantly outperform single-phase models, especially for high-dimensional, complex regression and classification problems. Numerous academic and industrial applications of multiple-phase models have been developed [6,9,22,23,92]. Multiple-phase models consist of two or more components, each of which implements various machine learning techniques. Each component functions independently of the others to solve a specific problem. For example, the SARIMA-ANN hybrid model has a SARIMA component that solves linear pattern data and an ANN component that solves non-linear pattern data. Combined models will be associated with an optimization algorithm to generate a fine-tuned prediction model. Fig. 5 shows the general structure of multiple-phase hybrid models.

3.1.3.3. Proposed hybrid models. As mentioned above, hybrid models of many kinds exist for forecasting time series. This work presents two metaheuristic optimization-based machine learning models for predicting the energy consumption of a building. They are a SARIMA-PSO-LSSVR (seasonal autoregressive integrated moving average-particle swarm optimization-least squares support vector regression) model and a SARIMA-MetaFA-LSSVR (seasonal autogressive integrated moving average-metaheuristic firefly algorithm-least squares support vector regression) model. These models are multiple-phase hybrid prediction models that have two components – univariate linear modeling and multivariate nonlinear modeling - in the training process. Chou and Ngo (2016) [6] developed a metaheuristic optimization-based machine learning model, SARIMA-MetaFA-LSSVR, to make 1-day-ahead predictions of the energy consumption of a building. This model has outstandingly outperformed single models such as SARIMA, LSSVR and hybrid models such as MetaFA-LSSVR, in predicting energy consumption. This investigation develops the SARIMA-PSO-LSSVR and compares its performance with those of SARIMA-MetaFA-LSSVR and the other models.

SARIMA models are extensively used for forecasting seasonal time-series data owing to their effectiveness in predicting the linear component of data. However, their major limitation is their assumption that time-series values are linearly correlated [6]. Therefore, the SARIMA model cannot properly elucidate nonlinear patterns in a complex data structure, such as that of energy consumption data. The combination of SARIMA model with other nonlinear forecasting models eliminates this limitation. In the

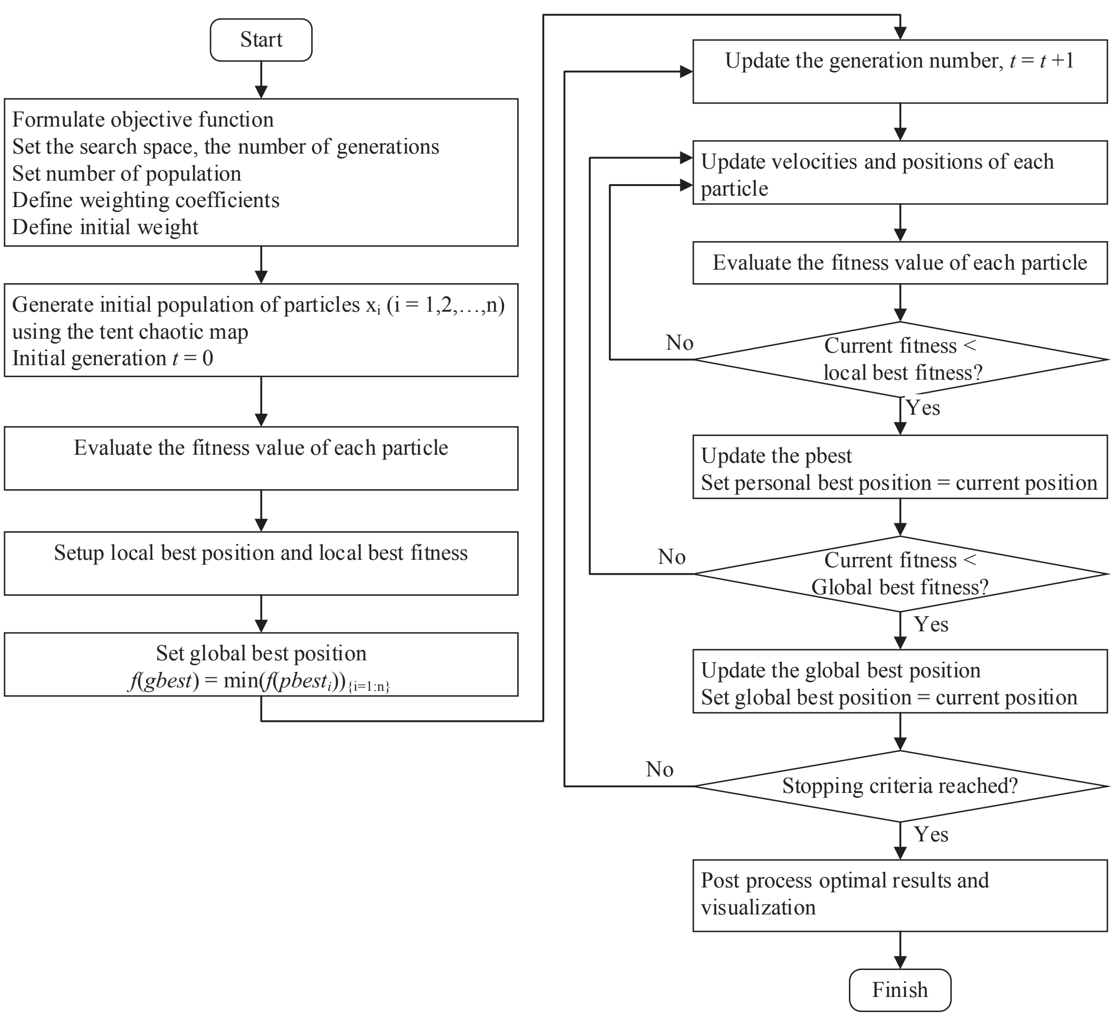

Fig. 6. Flowchart of metaheuristic PSO algorithm.

proposed hybrid models, the outputs of the univariate linear SARIMA model in phase 1 are part of the inputs of multivariate nonlinear models that apply metaheuristic optimization-based machine learning, such as that of MetaFA-LSSVR and PSO-LSSVR, in phase 2.

The SVR model is an ML technique that is based on statistical learning theory. Suykens et al. (2002) presented the LSSVR method to improve upon the computational speed of SVR model [93]. Although the LSSVR model effectively solves prediction problems, its accuracy depends on its hyperparameters, and especially the regularization constant and kernel parameters. To improve the predictive accuracy of the model, automatic optimization that is integrated with LSSVR should involve the regularization parameter (C) and the sigma of the radial basis function (RBF) kernel (σ). Details of the LSSVR can be found elsewhere [6].

Particle swarm optimization (PSO) is a population-based stochastic optimization technique that was developed by Kennedy and Eberhart in 1995 [94], inspired by the social foraging behavior of some animals, including the flocking of birds and the schooling of fish. The PSO algorithm involves a collection of particles that move around a search space, affected by their own best past location and the best past location of any particle in the swarm or a close neighbor. In each iteration, the velocity of every particle is updated.

The PSO algorithm in the proposed model – SARIMA-PSO-LSSVR uses a tent chaotic map. The tent map provides a highly diverse initial population because of its sensitivity to initial conditions. The tent map is piecewise linear, making the tent map easier to analyze than the logistic map. However, although the form of the tent map is simple and the equations are linear, for certain parameter values, the map can yield complex and chaotic behavior. As well as being sensitive to initial conditions, the tent map is also dependent on its parameter value, according to which it exhibits behavior from predictable to chaotic. The tent map can be combined with PSO to create a substantial optimization algorithm. Fig. 6 shows the flowchart of the metaheuristic PSO algorithm.

The metaheuristic firefly algorithm (MetaFA) is an optimization algorithm that was developed by Chou and Ngo in 2016 [6]. The MetaFA has three metaheuristic components, which are chaotic maps, the adaptive inertia weight (AIW), and Lévy flights. The conventional FA is combined with these components to enhance its search and optimization capabilities. A logistic chaotic map is used to generate a highly diverse initial population in the early stage of the MetaFA. Next, a Gauss/mouse chaotic map is used to replace the random parameters that are used in the conventional FA. To determine the convergence, the AIW is dynamically adjusted in the optimization process to control the local and the global exploration capabilities of the FA. Finally, Lévy flights are used to accelerate the local search by generating new optimal neighborhoods: around the derived optimal solution. Details of the MetaFA are provided elsewhere [6].

3.2. Machine learning software

This investigation assesses both the predictive accuracy of models and their ability to meet the demands of practitioners. Therefore, software that provides a user-friendly interface and powerful predictive application is necessary. To obtain accurate predictions from the models, the commonly used data-mining packages RapidMiner Studio, IBM SPSS Modeler, and WEKA are used for easy machine learning. The hybrid model was developed in MATLAB, which has proven powerful for rapid analysis and data processing because of its functions library, toolboxes, and flexibility.

3.3. Performance evaluation

R, RMSE, MAE, MAPE, MaxAE, and SI measures were used to evaluate the predictive accuracy of the proposed models; Eqs. (1)–(6) are their respective formulas.

$$\mathrm{R} = \frac{n\sum y\cdot y' - (\sum y)(\sum y')}{\sqrt{n(\sum y^2) - (\sum y)^2}\sqrt{n(\sum y'^2) - (\sum y')^2}} \tag{1}$$

where y' is the predicted value; y is the actual value, and n is the number of data samples.

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y-y'}{y}\right| \tag{2}$$

Table 1
Default settings of model parameters in RapidMiner Studio.

Model	Parameters	Default value RapidMiner Studio
Artificial Neural Network	Name	Neural Net
	Hidden layer specification	N/A
	Number of hidden layers	1
	Number of hidden nodes	$1+\frac{atts+classes}{2}$
	Number of learning iterations	500
	Momentum	0.2
	Type of normalization	[-1,1]
	Shuffle examples	True
	Learning rate	0.3
Support Vector Machine	Name	SVM
	Kernel type	RBF
	Regularization parameter (C)	0
	ε (epsilon)	0.001
	RBF γ (gamma)	1.0
	Stopping criteria	N/A
Classification and Regression Tree	Name	W-REPTree
	Tree depth	Infinite
	Number of random splits/leaf node	N/A
	Min number of samples/leaf node	2
	Create trainer mode	Single parameter
	Minimum variance for split (V)	1.0E-3
	Number of folds for reduced error pruning (N)	3
	Seed for random data shuffling (S)	1
	Pruning (P)	True
Linear Regression	Name	Linear Regression
	Fitting regression line method	Ordinary Least Squares
	L2 regularization weight	1.0E-8
	Features selection	M5-prime
	Use bias	True
	Eliminate collinear features	True
	Tolerance	0.05

Table 2
Default settings of model parameters in SPSS Modeler.

Model	Parameters	SPSS Modeler
SARIMA	Confident interval	95%
	Consider seasonal	Yes
	Intervals data	15-min
	Periodicity	96
	Maximum number of lags	96

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y - y^{'}\right| \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y^{'} - y\right)^2} \tag{4}$$

$$\text{MaxAE} = \max\left(\left|y - y^{'}\right|\right) \tag{5}$$

$$\text{SI} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{P_i - P_{\min,i}}{P_{\max,i} - P_{\min,i}}\right) \tag{6}$$

where m is the number of performance measures, and $P_i = i^{th}$ performance measure. The range of SI is from 0 to 1; an SI value close to 0 indicates an extremely accurate predictive model.

4. Energy consumption of buildings and development of model thereof

4.1. Data description

Real-time energy consumption data that were collected from a smart grid network installed in a typical three-floor building in Xindian District, New Taipei City, Taiwan, were utilized herein. The data were used to evaluate the applicability of the proposed predictive models in forecasting energy consumption patterns in buildings. The residential building of interest was occupied by a family of five (three children and their parents). The building had a total area of 350 m^2. Details of the building layout and the appliances and electrical equipment in the building can be found elsewhere [6].

The energy consumption data were converted from data with 1min intervals into data with 15-min intervals by executing a conversion procedure in MySQL, based on the actual electricity information demand from electricity companies. Consequently, 96 data points for building energy consumption with 15 min intervals were obtained per day. The database included information on four variables that directly affect the total energy consumption in the building; these are outdoor temperature (^{o}C), day of the week (Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday), and hour of the day (0, 1, 2, …,23). The times at which these values were also recorded.

Seven sets of data, corresponding to the days of the week, were collected. This paper presents short-term load forecasting models using AI techniques. Based on a suggestion of Chou and Telaga (2014) [57], the learning data that were used in this study were collected within a four week long sliding window and the data for the day after learning were the used for testing. For example, the first set of learning data were for June 23 to July 20 and the corresponding testing data pertained to July 21. Details of the data collection and preprocessing are provided elsewhere [6].

4.2. Model development

4.2.1. Single models

RapidMiner Studio software provides user-friendly interfaces for the development of AI techniques in all mentioned single models, but they have various parameter fields with different settings. For efficient operation and ease of use, the default settings of the main parameters were used in each model in the software package. Table 1 shows the comprehensive settings of these individual models in RapidMiner Studio. SPSS Modeler provides a SARIMA model to predict time-series data. Table 2 presents the default settings of single model in SPSS Modeler.

RapidMiner Studio and SPSS Modeler allow users to adjust the structure of a single model manually using integrated modules in the platform. Figs. 7 and 8 display the implementation procedures for single models in RapidMiner Studio and IBM SPSS Modeler.

4.2.2. Ensemble models

As in the single mode scenario, a user can construct ensemble analytic streams by arranging and customizing the modules of AI techniques. The best software for implementing the single scenario was the RapidMiner Studio, which is therefore utilized to increase the predictive accuracy in the ensemble scenario. In RapidMiner Studio, voting and bagging ensemble models were implemented. Figs. 9 and 10 present ensemble streams in RapidMiner Studio.

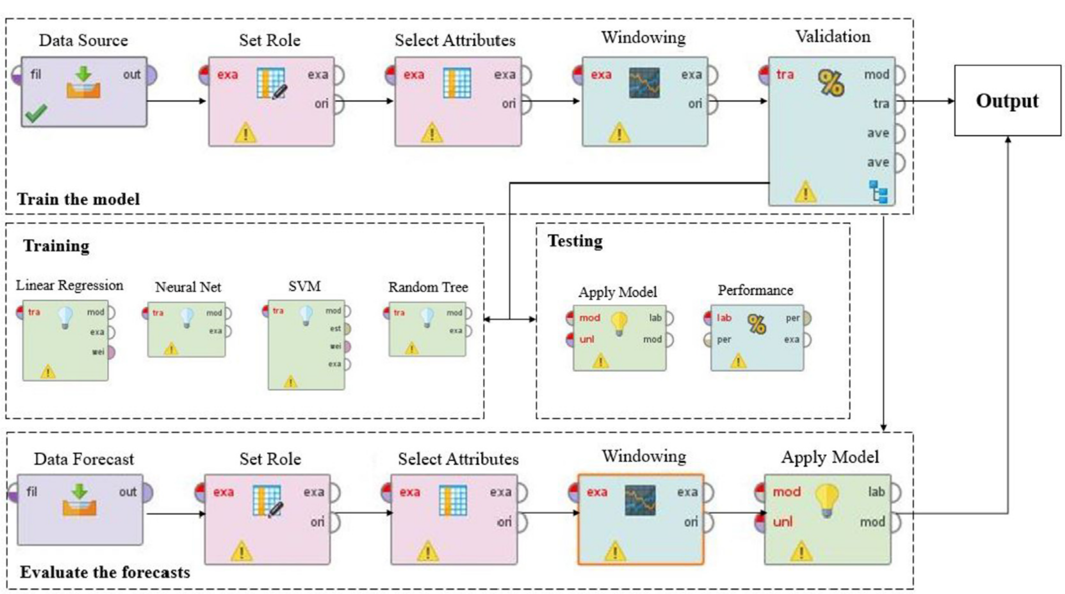

Fig. 7. Stream of single model in RapidMiner Studio.

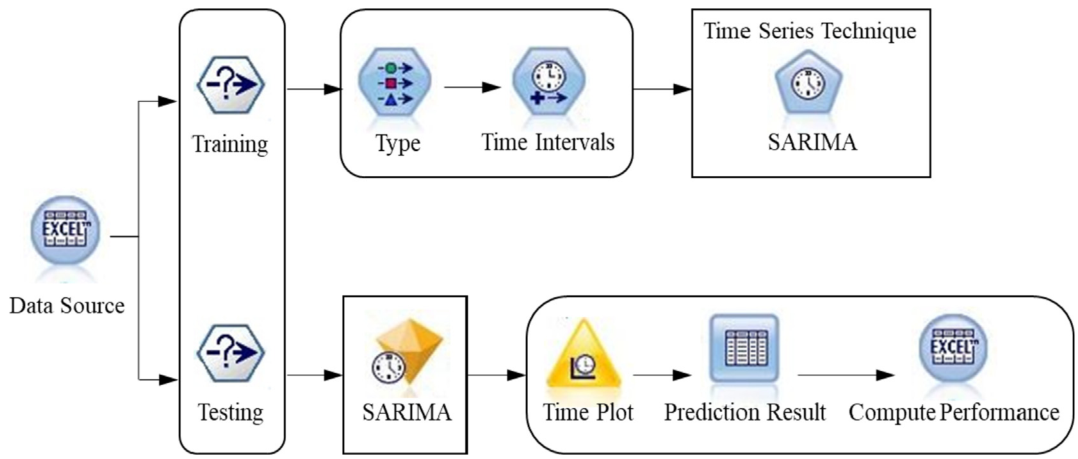

Fig. 8. Stream of single model in IBM SPSS Modeler.

4.2.3. *Hybrid models*

Previous studies have demonstrated that a hybrid model that exhibits both linear and nonlinear modeling abilities is suitable for forecasting energy consumption. The SARIMA and LSSVR models can solve linear and nonlinear components in data, respectively. Since the proposed hybrid models combine the strengths of all the constituent models, they can effectively capture patterns of time-series data. The proposed models (SARIMA-PSO-LSSVR and SARIMA-MetaFA-LSSVR) have two stages; stage 1 involves univariate linear modeling and stage 2 involves multivariate nonlinear modeling. The hybrid metaheuristic optimized models herein were implemented in MATLAB software. Details of prediction model development have been provided elsewhere [6]. Tables 3 and 4 provide the parameter settings for the proposed prediction models. Fig. 11 shows the processes by which the proposed hybrid models are constructed.

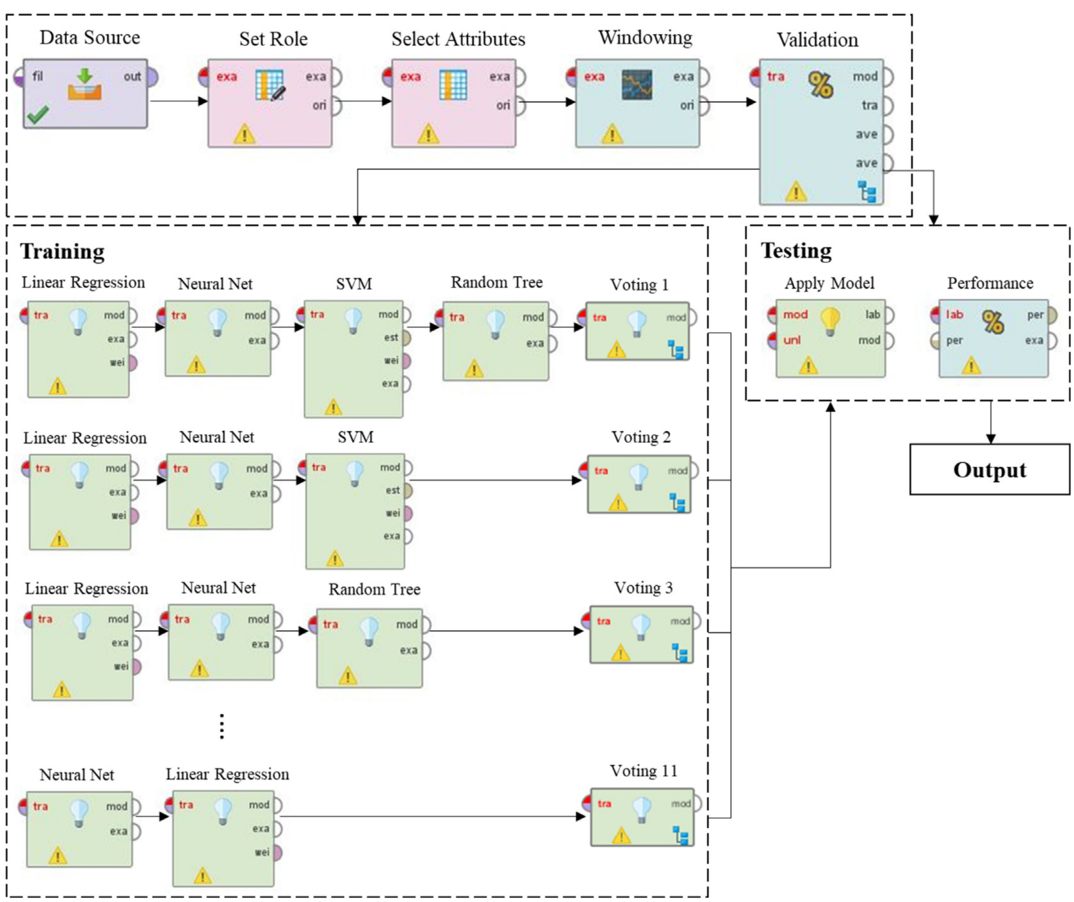

Fig. 9. Stream of voting model in RapidMiner Studio.

5. Experimental results and discussion

5.1. Performance in single scenario

In the first stage of the analytical process, single models were implemented using two software packages - RapidMiner Studio and IBM SPSS Modeler. Their performance measures are used to compare these single models in terms of R, RMSE, MAE, MAPE, MaxAE, and SI. Tables 5 and 6 present the performance of single models in SPSS Modeler and RapidMiner Studio, respectively.

The prediction performance of the single model in SPSS Modeler was compared with those of single models in RapidMiner Studio. Tables 5 and 6 confirm that the ANNs model in RapidMiner Studio is the most effective single model for making predictions using the data. It yielded the lowest values of RMSE (0.092), MAE (0.057), and MAPE (36.834%); the second lowest value of MaxAE (0.353), and the highest R value (0.556).

5.2. Performance in ensemble scenario

Table 7 presents the performance measures of ensemble models in RapidMiner Studio. The bagging ANN ensemble achieved the best results among the four ensemble models in RapidMiner Studio, with the best values of all evaluation measures - RMSE (0.094), MAE (0.049), MAPE (42.31%), MaxAE (0.338) and R (0.607). With respect to voting, the best performance was achieved using the ANNs + LR-based ensemble model, which yielded the best values of R, RMSE, MAE, MAPE, and MaxAE (0.512, 0.164, 0.104, 42.75%, and 0.564), respectively.

5.3. Performance in hybrid scenario

To demonstrate the efficacy and reliability of the proposed hybrid models SARIMA-MetaFA-LSSVR and SARIMA-PSO-LSSVR in MATLAB, their predictive performances were compared as shown in Table 8. The hybrid SARIMA-MetaFA-LSSVR model achieved better results than the hybrid SARIMA-PSO-LSSVR model, with the best values of all evaluation measures - R (0.796), RMSE (0.164), MAE (0.028), MAPE (15.657%), and MaxAE (0.179). The SARIMA-PSO-LSSVR model had a shorter CPU time (0.729s).

5.4. Comprehensive comparison and discussion

This investigation identified the best of the single, ensemble, and hybrid models with data at different intervals. With data at 15 min intervals, the best AI single model was ANNs in RapidMiner

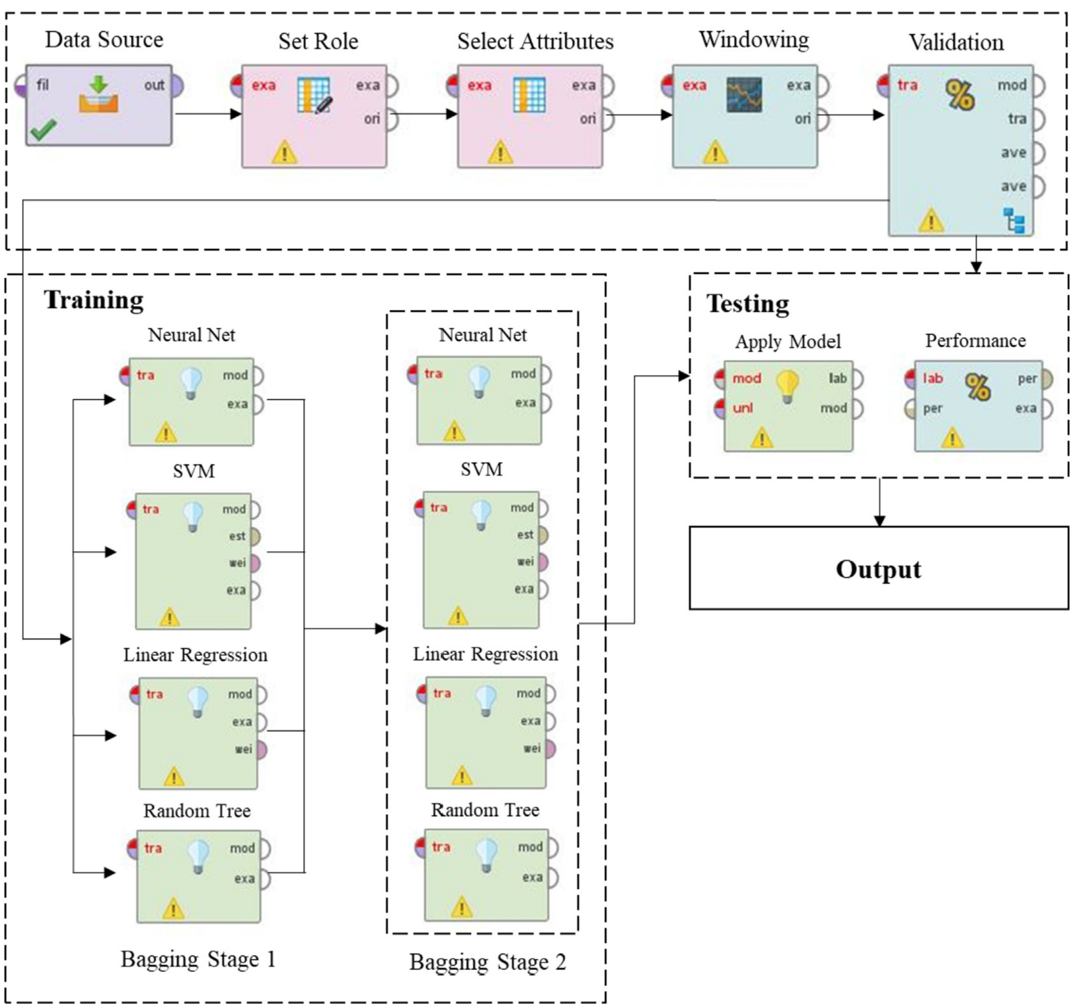

Fig. 10. Stream of bagging model in RapidMiner Studio.

Table 3
Parameter settings for the SARIMA-PSO-LSSVR prediction model.

Parameter	Description	Setting
First stage: SARIMA model		
p	Non-seasonal autoregressive order	1
d	Non-seasonal difference	0
q	Non-seasonal moving average	1
P	Seasonal autoregressive order	96
D	Seasonal difference	0
Q	Seasonal moving average order	96
S	Time span of repeating seasonal pattern	96
Second stage: PSO-LSSVR model		
Range of C	Regularization parameter of LSSVR	$[10^{-3}\ 10^{12}]$
Range of σ	Sigma of RBF kernel function in LSSVR	$[10^{-3}\ 10^{3}]$
No. of particles	Population size of the PSO	50
Stop criteria	Max generation	25
	Three consecutive rates of change in objective function values	$<10^{-8}$
Logistic chaotic map	Generate initial population with high diversity	Random generation
α_o	Initial value of the randomness parameter	0.5
$c_{1,}\ c_2$	Learning parameters	$c_1 = 2.05$ $c_2 = 2.05$

Studio, whereas the best ensemble model was the ANN-based bagging model, whose accuracy was 9% better than that of the single model in the same software package. The better effective hybrid model was SARIMA-MetaFA-LSSVR, whose accuracy was 64% better than that of the single model. Table 9 comprehensively compares these models.

In the existing studies on the ensemble models for energy prediction in buildings and facilities, for instance, Jovanović et al. (2015) presented the ensemble of various neural networks for prediction of heating energy consumption in an university [95]. They achieved high results with daily data to analyze heating energy consumption. However, that is difficult to supervise closely the energy consumption by a daily scale. Based on the analytical results, this study suggests that 15-min intervals data are rather suitable for predicting energy consumption in residential buildings.

Candanedo et al. (2017) presented data-driven prediction models for the energy use of appliances in a low-energy house [96]. The best model - gradient boosting machines (GBM) ensemble model was able to explain 57% of the variance (R^2) and reach 38.29% of MAPE in the testing process when using all the predictors. Compared to their work, this study achieves better results than the GBM model with R^2 (63.4%) and MAPE (15.657%) in the testing process.

Similarly, Wang et al. (2018) applied ensemble bagging trees (EBT) model to predict the energy consumption of an amphitheater in a university [97]. Compared to the conventional single prediction model (i.e., CART), they concluded the homogeneous ensemble model (i.e., EBT) is superior in prediction accuracy, while it requires more computation time and is short of interpretability due to its sophisticated model structure.

The hybrid models are generally more accurate than the single and ensemble models. The single models are the easiest to use because they can be implemented in software packages that integrate AI techniques. Users just have to select models and parameters using the interfaces in the software packages, which are herein RapidMiner Studio and IBM SPSS Modeler. However, their accuracy and efficiency are very low. Therefore, single models are the choice for users who firstly use an AI model to solve the energy consumption problem.

Users who know the basics of machine learning techniques must be able to use models that are more effective than single models. Ensemble models are favorable because they perform

Table 4
Parameter settings for the SARIMA-MetaFA-LSSVR prediction model.

Parameter	Description	Setting
First stage: SARIMA model		
P	Non-seasonal autoregressive order	1
D	Non-seasonal difference	0
Q	Non-seasonal moving average	1
P	Seasonal autoregressive order	96
D	Seasonal difference	0
Q	Seasonal moving average order	96
S	Time span of repeating seasonal pattern	96
Second stage: MetaFA-LSSVR model		
Range of C	Regularization parameter of LSSVR	$[10^{-3}\ 10^{12}]$
Range of σ	Sigma of RBF kernel function in LSSVR	$[10^{-3}\ 10^{3}]$
No. of fireflies	Population size of the MetaFA	50
Stop criteria	Max generation	25
	Three consecutive rates of change in objective function values	$<10^{-8}$
Logistic chaotic map	Generate initial population with high diversity	Random generation
$\beta_{\min}$	Minimum value of attractiveness parameter β	0.1
γ	Absorption coefficient	1
Gauss/mouse chaotic map	Automatically tune the parameter β	Attractive parameter
α	Randomness of firefly movement	$\alpha_o = 0.2$
AIW	Control local and global exploration capabilities of the swarm algorithm	$\theta = 0.9$
Lévy flight	Accelerate local search by generating new solutions around the optimal solution	$\tau = 1.5$

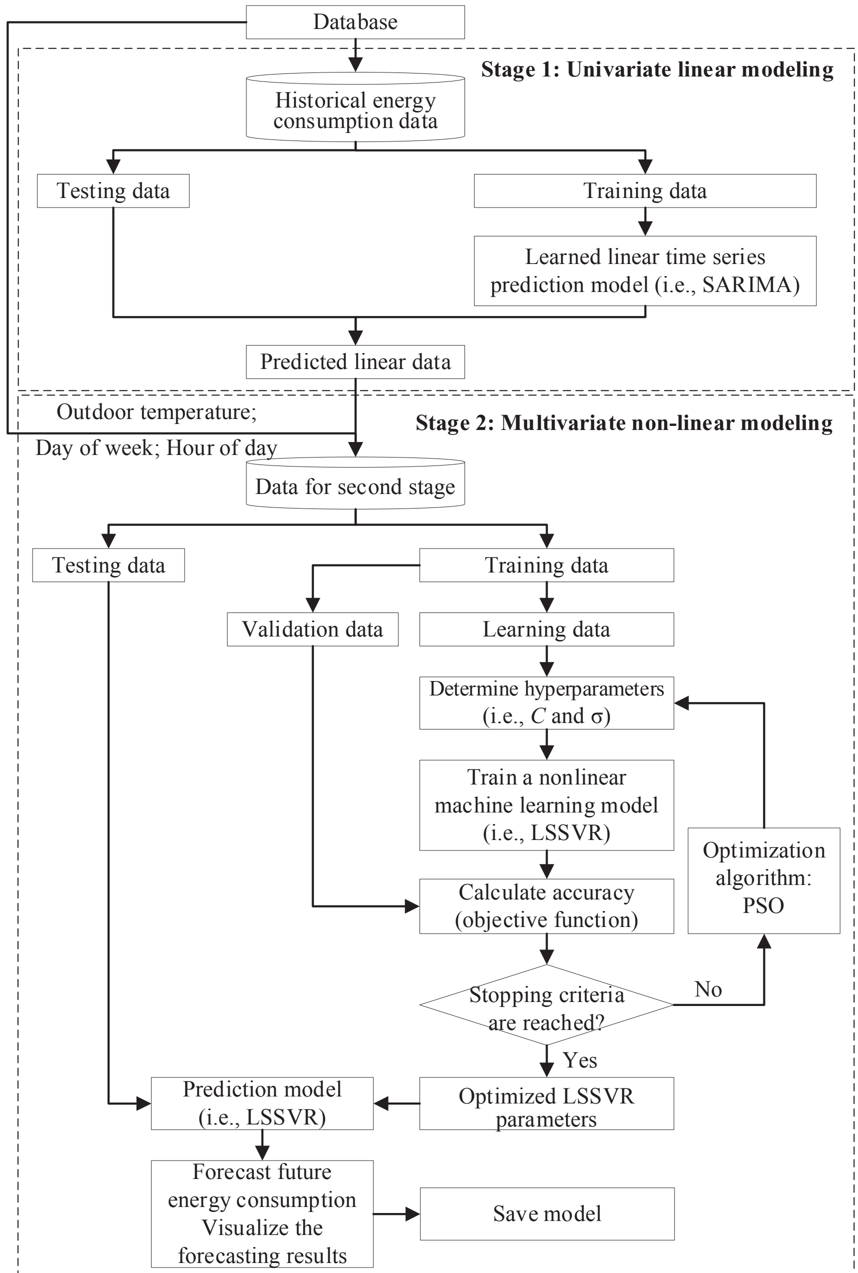

Fig. 11. Metaheuristic optimization-based machine learning models in MATLAB.

satisfactorily in analyzing time series. To build ensemble models, users must know their associated procedures and combine the strengths of the constituent models. Some ensemble models are integrated into software packages, whose users can implement them as easily as implementing single models. An ensemble model is a combination of single models that compensates for errors made using the single models, so they usually outperform baseline cases.

The efficacy and effectiveness of proposed hybrid models were demonstrated by analyzing real-time energy consumption data. The proposed SARIMA-MetaFA-LSSVR and SARIMA-PSO-LSSVR

Table 5
Performance measures of single models in SPSS Modeler.

Model	Measure	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	Ave.	Max	Min
SARIMA	R	0.192	0.519	0.332	0.602	0.733	0.527	0.138	0.435	0.733	0.138
	RMSE (kWh)	0.146	0.081	0.176	0.109	0.074	0.129	0.098	0.116	0.176	0.074
	MAE (kWh)	0.103	0.058	0.161	0.073	0.055	0.102	0.085	0.091	0.161	0.055
	MAPE (%)	49.791	26.226	103.2	47.52	34.01	80.35	82.53	60.518	103.2	26.226
	MaxAE (kWh)	0.18	0.226	0.983	0.265	0.178	0.204	0.213	**0.321**	0.983	0.178

Table 6
Performance measures of single models in RapidMiner Studio.

Model	Measure	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	Ave.	Max	Min
ANNs	R	0.373	0.729	0.518	0.697	0.645	0.651	0.276	**0.556**	0.729	0.276
	RMSE (kWh)	0.097	0.068	0.116	0.105	0.073	0.068	0.116	**0.092**	0.116	0.068
	MAE (kWh)	0.058	0.045	0.068	0.083	0.051	0.052	0.044	**0.057**	0.083	0.044
	MAPE (%)	35.06	19.04	34.95	56.69	29.29	40.82	41.99	**36.834**	56.69	19.04
	MaxAE (kWh)	0.34	0.36	0.529	0.375	0.226	0.177	0.464	0.353	0.529	0.177
SVM	R	0.339	0.371	0.345	0.725	0.535	0.584	0.19	0.441	0.725	0.19
	RMSE (kWh)	0.094	0.109	0.134	0.089	0.077	0.095	0.108	0.101	0.134	0.077
	MAE (kWh)	0.064	0.082	0.085	0.062	0.058	0.065	0.072	0.070	0.085	0.058
	MAPE (%)	34.78	37.25	34.53	36.96	34.76	49.67	74.62	43.224	74.62	34.53
	MaxAE (kWh)	0.308	0.371	0.526	0.335	0.215	0.279	0.276	0.330	0.526	0.215
LR	R	0.296	0.326	0.525	0.593	0.57	0.563	0.239	0.445	0.593	0.239
	RMSE (kWh)	0.099	0.103	0.109	0.111	0.105	0.063	0.119	0.101	0.119	0.063
	MAE (kWh)	0.069	0.068	0.069	0.064	0.076	0.039	0.046	0.062	0.076	0.039
	MAPE (%)	42.62	28.15	37.32	42.89	51.92	22.79	43.93	38.517	51.92	22.79
	MaxAE (kWh)	0.348	0.363	0.39	0.39	0.355	0.264	0.655	0.395	0.655	0.264
C&R Tree	R	0.392	0.519	0.381	0.355	0.425	0.512	0.08	0.381	0.519	0.08
	RMSE (kWh)	0.097	0.103	0.125	0.116	0.099	0.124	0.087	0.107	0.125	0.087
	MAE (kWh)	0.058	0.083	0.068	0.089	0.078	0.085	0.056	0.074	0.089	0.056
	MAPE (%)	35.06	40.41	26.47	65.49	48.96	63.43	57.61	48.204	65.49	26.47
	MaxAE (kWh)	0.409	0.262	0.538	0.347	0.327	0.346	0.305	0.362	0.538	0.262

Note: Bold value denotes the best performance.

Table 7
Performance measures of ensemble models in RapidMiner Studio.

	Model	Measure	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	Ave.	Max	Min
Voting	SVR + LR	R	0.545	0.573	0.528	0.363	0.535	0.658	0.384	0.512	0.658	0.363
		RMSE (kWh)	0.087	0.131	0.126	0.167	0.362	0.101	0.174	0.164	0.362	0.087
		MAE (kWh)	0.048	0.064	0.058	0.056	0.178	0.171	0.153	0.104	0.178	0.048
		MAPE (%)	38.76	45.89	34.65	35.36	47.25	43.58	53.74	42.75	53.74	34.65
		MaxAE (kWh)	0.215	0.391	0.429	0.534	0.641	0.486	1.253	0.564	1.253	0.215
Bagging	**ANNs**	R	0.551	0.616	0.658	0.623	0.734	0.651	0.416	**0.607**	0.734	0.416
		RMSE (kWh)	0.058	0.016	0.072	0.171	0.125	0.149	0.067	**0.094**	0.171	0.016
		MAE (kWh)	0.042	0.068	0.031	0.061	0.034	0.039	0.069	**0.049**	0.069	0.031
		MAPE (%)	34.82	54.95	38.19	35.31	28.07	42.62	62.19	**42.31**	62.19	28.07
		MaxAE (kWh)	0.277	0.549	0.157	0.232	0.451	0.381	0.319	**0.338**	0.549	0.157

Note: Bold value denotes the best performance.

Table 8
Performance comparison of the proposed hybrid models in MATLAB.

Model	Measure	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	Sun.	Ave.	Max	Min
SARIMA-MetaFA-LSSVR	R	0.817	0.855	0.910	0.813	0.867	0.82	0.492	**0.796**	0.91	0.492
	RMSE (kWh)	0.179	0.18	0.18	0.17	0.172	0.168	0.1	**0.164**	0.18	0.1
	MAE (kWh)	0.032	0.032	0.032	0.029	0.03	0.028	0.01	**0.028**	0.032	0.01
	MAPE (%)	16.56	13.98	14.93	16.55	18.42	18.97	10.19	**15.657**	18.97	10.19
	MaxAE (kWh)	0.182	0.213	0.249	0.261	0.151	0.166	0.033	**0.179**	0.261	0.033
	CPU time (s)	0.733	0.815	0.705	0.797	0.772	0.784	0.821	0.775	0.821	0.705
SARIMA- PSO-LSSVR	R	0.731	0.632	0.845	0.908	0.867	0.825	0.493	0.757	0.908	0.493
	RMSE (kWh)	0.24	0.239	0.182	0.18	0.171	0.162	0.1	0.182	0.24	0.1
	MAE (kWh)	0.057	0.057	0.033	0.032	0.029	0.026	0.01	0.035	0.057	0.01
	MAPE (%)	33.85	22.45	14.28	14.95	0.184	17.4	10.2	16.188	33.85	0.184
	MaxAE (kWh)	0.24	0.458	0.234	0.244	0.151	0.182	0.033	0.220	0.458	0.033
	CPU time (s)	0.767	0.724	0.736	0.729	0.724	0.714	0.71	**0.729**	0.767	0.71

Note: Bold value denotes the best performance.

Table 9
Performance comparison among the best models in each scenario.

Scenario	Best model	Performance Measure					
		R	RMSE	MAE	MAPE (%)	SI	Improvement
Single	RapidMiner: ANNs	0.556	**0.092**	0.057	36.834	0.70 (3)	–
Ensemble	RapidMiner: Bagging (ANNs)	0.607	0.094	0.049	42.31	0.64 (2)	9%
Hybrid	**MATLAB: SARIMA-MetaFA-LSSVR**	**0.796**	0.164	**0.028**	**15.657**	**0.25 (1)**	64%

Note: Bold value denotes the best performance; (_) indicates the ranking.

models were highly accurate and reliable in making 1 day-ahead predictions of energy consumption in the building of interest. These models are more powerful than single and ensemble models because they normally incorporate the advantages of, and compensate for the weaknesses of, the individual constituent models, and improve forecasting accuracy by optimizing algorithms. However, the proposed hybrid models are built on platforms (e.g., MATLAB) that are not easy to use. Therefore, their users are required to have a deep knowledge of machine learning and optimization techniques.

6. Conclusion and recommendation

Energy management is essential to reducing the energy consumption of buildings. The importance of saving energy derives from the global need to reduce energy expenditure and environmental impacts and from relevant legislation. This work presents a comprehensive review and analysis of several time-series forecasting techniques for predicting the energy consumption of buildings one-day in advance. Recently, many experiments have been carried out using these techniques with data on the energy consumption of buildings. Each of the techniques mentioned herein has advantages and disadvantages. This study comprehensively presents the forecasting methods and provides suggestions regarding their effective use.

The AI methods that were used in this study are single, ensemble, and hybrid techniques for predicting the energy consumption of buildings. Specifically, five AI techniques (namely ANNs, SVR, CART, LR, and SARIMA) were employed to build single models, whereas their combinations were used to generate two ensemble models, involving voting and bagging. The SARIMA-MetaFA-LSSVR and SARIMA-PSO-LSSVR were both implemented as hybrid models. A comprehensive comparison demonstrated that the hybrid model is more accurate than single and ensemble models. Both the accuracy of prediction and their usability in planning energy management are considered.

Three analytic platforms - RapidMiner Studio, IBM SPSS Modeler, and WEKA - were used to implement the individual AI models. With data at 15 min intervals, the best AI single model was ANNs in RapidMiner Studio, whereas the best ensemble model was the ANN-based bagging model in the same software package. The better effective hybrid model was SARIMA-MetaFA-LSSVR in MATLAB. The ensemble and hybrid models improved overall performance measure (SI) by 9% and 64% as compared to the best single model, respectively.

The novelty of this study is its provision of a comprehensive approach to using AI models for forecasting energy consumption in buildings. This research focuses on finding the most effective model for various users who have to develop predictive models rapidly. The results in this research reveal that the single models are the best choice for users who would like to promptly estimate energy consumption of buildings. Ensemble models are best for users who know the basics of machine learning techniques. Users who have a deep knowledge of AI techniques and require highly accurate and effective models should develop hybrid models by integrating machine learning techniques with optimization algorithms.

This investigation provides a comprehensive review of energy forecasting using AI techniques and some typical models. Users can exploit this study to improve the energy-efficiency of appliances and electrical equipment in their buildings. This study involved a predictive analysis using four variables that were collected from the smart grid infrastructure in a selected building. The influence of factors such as humidity, solar radiation intensity, and times of extreme weather events were not considered in the forecasting of the energy consumption of the building herein. A further study should thus be conducted to collect such data potential to enhance predictive performance. Future work shall develop an intelligent prediction platform with comprehensive auto-analysis and visualizations of useful insights.

Acknowledgement

The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan, for financially supporting this research under grant No. 106-2221-E-011-017-MY3.

References

[1] IEA. World energy outlook 2011 executive summary. International Energy Agency; 2011.
[2] Zhao H-x, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16(6):3586–92.
[3] Lee YM, An L, Liu F, Horesh R, Chae YT, Zhang R. Applying science and mathematics to big data for smarter buildings. Ann N Y Acad Sci 2013;1295(1):18–25.
[4] Reinisch C, Kofler M, Iglesias F, Kastner W. Thinkhome energy efficiency in future smart homes. EURASIP J Embed Syst 2011;2011(1), 104617.
[5] Spees K, Lave L. Impacts of responsive load in PJM - load shifting and real time pricing. Energy J 2008;29:101–22.
[6] Chou J-S, Ngo N-T. Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns. Appl Energy 2016;177:751–70.
[7] Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, et al. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew Sustain Energy Rev 2014;33:102–9.
[8] Biswas MAR, Robinson MD, Fumo N. Prediction of residential building energy consumption: a neural network approach. Energy 2016;117:84–92.
[9] Khandelwal I, Adhikari R, Verma G. Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. Procedia Computer Science 2015;48:173–9.
[10] Zhang W, Liu S, Li N, Xie H, Li X. Development forecast and technology roadmap analysis of renewable energy in buildings in China. Renew Sustain Energy Rev 2015;49:395–402.
[11] Lim Y-D, Kim S-C, Park J-O, Choo Y-G. A predictive demand of the maximum electric power using chaos-fuzzy. In: The 10th IEEE international conference: fuzzy systems; 2001. p. 1511–4.
[12] Lü X, Lu T, Kibert CJ, Viljanen M. Modeling and forecasting energy consumption for heterogeneous buildings using a physical–statistical approach. Appl Energy 2015;144:261–75.
[13] Monfet Arkhipova, Choiniere. Evaluation of a case-based reasoning energy prediction tool for commercial buildings. In: Proceedings of the 13th international conference enhanc build oper; 2013. p. 1–7. Montreal, Quebec, Canada.
[14] Mousavi A, Vyatkin V. Energy Efficient Agent Function Block: a semantic agent approach to IEC 61499 function blocks in energy efficient building automation systems. Autom ConStruct 2015;54:127–42.
[15] Kandananond K. Forecasting electricity demand in Thailand with an artificial neural network approach. Energies 2011;4(12):1246–57.
[16] Hu Y, Peng L, Li X, Yao X, Lin H, Chi T. A novel evolution tree for analyzing the global energy consumption structure. Energy 2018;147:1177–87.
[17] Yip S-C, Wong K, Hew W-P, Gan M-T, Phan RCW, Tan S-W. Detection of energy theft and defective smart meters in smart grids using linear regression. Int J Electr Power Energy Syst 2017;91:230–40.
[18] Chujai P, Kerdprasop N, Kerdprasop K. Time series analysis of household electric consumption with ARIMA and ARMA models. In: International Multiconference of Engineers and Computer Scientists. Hong Kong: IMECS; 2013. 2013.
[19] Khashei M, Bijari M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Appl Soft Comput 2011;11(2): 2664–75.
[20] Tan Z, Zhang J, Wang J, Xu J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Appl Energy 2010;87(11):3606–10.
[21] Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 2003;50:159–75.
[22] Chen K-Y, Wang C-H. A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert Syst Appl 2007;32(1):254–64.
[23] Deb C, Zhang F, Yang J, Lee SE, Shah KW. A review on time series forecasting techniques for building energy consumption. Renew Sustain Energy Rev 2017;74:902–24.
[24] Pérez-Lombard L, Ortiz J, Pout C. A review on buildings energy consumption information. Energy Build 2008;40(3):394–8.
[25] Aghemo C, Blaso L, Pellegrino A. Building automation and control systems: a case study to evaluate the energy and environmental performances of a lighting control system in offices. Autom ConStruct 2014;43:10–22.

[26] Corry E, Pauwels P, Hu S, Keane M, O'Donnell J. A performance assessment ontology for the environmental and energy management of buildings. Autom ConStruct 2015;57:249–59.
[27] Zhou S, Wu Z, Li J, Zhang X-p. Real-time energy control approach for smart home energy management system. Elec Power Compon Syst 2014;42(3–4): 315–26.
[28] Lee Y-S, Tong L-I. Forecasting energy consumption using a grey model improved by incorporating genetic programming. Energy Convers Manag 2011;52(1):147–52.
[29] Monfet D, Corsi M, Choinière D, Arkhipova E. Development of an energy prediction tool for commercial buildings using case-based reasoning. Energy Build 2014;81:152–60.
[30] Allouhi A, El Fouih Y, Kousksou T, Jamil A, Zeraouli Y, Mourad Y. Energy consumption and efficiency in buildings: current status and future trends. J Clean Prod 2015;109:118–30.
[31] Faria P, Vale Z. Demand response in electrical energy supply: an optimal real time pricing approach. Energy 2011;36(8):5374–84.
[32] Guo JJ, Wu JY, Wang RZ. A new approach to energy consumption prediction of domestic heat pump water heater based on grey system theory. Energy Build 2011;43(6):1273–9.
[33] Hsu D. Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data. Appl Energy 2015;160: 153–63.
[34] Radulovic F, Poveda-Villalón M, Vila-Suero D, Rodríguez-Doncel V, García-Castro R, Gómez-Pérez A. Guidelines for Linked Data generation and publication: an example in building energy consumption. Autom ConStruct 2015;57:178–87.
[35] Sandels C, Widén J, Nordström L, Andersson E. Day-ahead predictions of electricity consumption in a Swedish office building from weather, occupancy, and temporal data. Energy Build 2015;108:279–90.
[36] Valor E, Meneu V, Caselles V. Daily air temperature and electricity load in Spain. Appl Meteorol 2001;40(8). 1413-2.
[37] Yang J, Santamouris M, Lee SE, Deb C. Energy performance model development and occupancy number identification of institutional buildings. Energy Build 2016;123:192–204.
[38] Mat Daut MA, Hassan MY, Abdullah H, Rahman HA, Abdullah MP, Hussin F. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: a review. Renew Sustain Energy Rev 2017;70:1108–18.
[39] Fumo N, Rafe Biswas MA. Regression analysis for prediction of residential energy consumption. Renew Sustain Energy Rev 2015;47:332–43.
[40] Costa A, Keane MM, Raftery P, O'Donnell J. Key factors methodology—a novel support to the decision making process of the building energy manager in defining optimal operation strategies. Energy Build 2012;49:158–63.
[41] Chou J-S, Ngo N-T. Smart grid data analytics framework for increasing energy savings in residential buildings. Autom ConStruct 2016;72:247–57.
[42] Brockwell PJ, Davis RA. Introduction time series and forecasting. New York: Springer-Verlag New York; 1996.
[43] De Boeck L, Verbeke S, Audenaert A, De Mesmaeker L. Improving the energy performance of residential buildings: a literature review. Renew Sustain Energy Rev 2015;52:960–75.
[44] Fan C, Xiao F, Wang S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. Appl Energy 2014;127:1–10.
[45] Foucquier A, Robert S, Suard F, Stéphan L, Jay A. State of the art in building modelling and energy performances prediction: a review. Renew Sustain Energy Rev 2013;23:272–88.
[46] Fumo N. A review on the basics of building energy estimation. Renew Sustain Energy Rev 2014;31:53–60.
[47] Li K, Hu C, Liu G, Xue W. Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis. Energy Build 2015;108:106–13.
[48] Mena R, Rodríguez F, Castilla M, Arahal MR. A prediction model based on neural networks for the energy consumption of a bioclimatic building. Energy Build 2014;82:142–55.
[49] Azadeh A, Saberi M, Seraj O. An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: a case study of Iran. Energy 2010;35(6):2351–66.
[50] Azadeh A, Seraja O, Saberib M. A total fuzzy regression algorithm for energy consumption estimation. In: Proceedings of the 6th IEEE international conference ind informatics. Daejeon, Korea; 2008. p. 1562–8.
[51] Naji S, Shamshirband S, Basser H, Keivani A, Alengaram UJ, Jumaat MZ, et al. Application of adaptive neuro-fuzzy methodology for estimating building energy consumption. Renew Sustain Energy Rev 2016;53:1520–8.
[52] Arghira N, Hawarah L, Ploix S, Jacomino M. Prediction of appliances energy use in smart homes. Energy 2012;48(1):128–34.
[53] Ekici BB, Aksoy UT. Prediction of building energy consumption by using artificial neural networks. Adv Eng Software 2009;40(5):356–62.
[54] Hamzacebi C, Es HA. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 2014;70:165–71.
[55] Jung HC, Kim JS, Heo H. Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach. Energy Build 2015;90:76–84.
[56] Yang C, Li H, Rezgui Y, Petri I, Yuce B, Chen B, et al. High throughput computing based distributed genetic algorithm for building energy consumption optimization. Energy Build 2014;76:92–101.
[57] Chou J-S, Telaga AS. Real-time detection of anomalous power consumption. Renew Sustain Energy Rev 2014;33:400–11.
[58] Pereira CM, Almeida NNd, Velloso MLF. Fuzzy modeling to forecast an electric load time series. Procedia Computer Science 2015;55:395–404.
[59] D'Andrea E, Lazzerini B, Rosario SLd. Neural network-based forecasting of energy consumption due to electric lighting in office buildings. Sustainable Internet and ICT for Sustainability (SustainIT) Pisa. Italy: IEEE; 2012.
[60] Mehdi Shahrookh, Hirad, Masih, Singh. Hourly air temperature driven using multi-layer perceptron and radial basis function networks in arid and semi-arid regions. Theor Appl Climatol 2012;109(3–4):519–28.
[61] Adya M, Collopy F. How effective are neural networks at forecasting and prediction? A review and evaluation. Forecasting 1998;17:481–95.
[62] Hamzaçebi C. Forecasting of Turkey's net electricity energy consumption on sectoral bases. Energy Pol 2007;35(3):2009–16.
[63] Dong B, Cao C, Lee SE. Applying support vector machines to predict building energy consumption in tropical region. Energy Build 2005;37(5):545–53.
[64] Božić M, Stojanović M, Stajić Z. Short-term electric load forecasting using least square support vector machines. Automat Contr Robot 2010;9:141–50.
[65] Catalina T, Virgone J, Blanco E. Development and validation of regression models to predict monthly heating demand for residential buildings. Energy Build 2008;40(10):1825–32.
[66] Hygh JS, DeCarolis JF, Hill DB, Ranji Ranjithan S. Multivariate regression as an energy assessment tool in early building design. Build Environ 2012;57: 165–75.
[67] Wang Z, Srinivasan RS. A review of artificial intelligence based building energy use prediction: contrasting the capabilities of single and ensemble prediction models. Renew Sustain Energy Rev 2017;75:796–808.
[68] Zhang P. Time-series analysis with neural networks and ARIMA-neural network hybrids. Neurocomputing 2001;50:159–75.
[69] Newsham GR, Birt BJ. Building-level occupancy data to improve ARIMA-based electricity use forecasts. Zurich, Switzerland: BuildSys; 2010.
[70] Mohamed N, Ahmad M, Ismail Z. Short term load forecasting using double seasonal Arima model. Appl Sci 2011;15:223–31.
[71] Chou J-S, Ho C-C, Hoang H-S. Determining quality of water in reservoir using machine learning. Ecol Inf 2018;44:57–75.
[72] Martínez-Álvarez F, Troncoso A, Asencio-Cortés G, Riquelme J. A survey on data mining techniques applied to electricity-related time series forecasting. Energies 2015;8(12):13162–93.
[73] Chou J-S, Lin C. Predicting disputes in public-private partnership projects: classification and ensemble models. J Comput Civ Eng 2013;27(1):51–60.
[74] Qiu X, Zhang L, Ren Y, Suganthan PN. Ensemble deep learning for regression and time series forecasting. In: IEEE Symposium on Computational Intelligence in Ensemble Learning; 2014. Orlando, FL USA.
[75] Allende H, Valle C. Ensemble Meth Time Series Forecast 2017;349:217–32.
[76] Soares E, Costa P, Costa B, Leite D. Ensemble of evolving data clouds and fuzzy models for weather time series prediction. Appl Soft Comput 2018;64: 445–53.
[77] Sallehuddin R, Shamsuddin M, Abraham A. Forecasting time series data using hybrid grey relational artificial neural network and auto regressive integrated moving average model. Neural Netw World 2007;17(6):573–605.
[78] Lee Y-S, Tong L-I. Forecasting nonlinear time series of energy consumption using a hybrid dynamic model. Appl Energy 2012;94:251–6.
[79] Chen J, Yang J, Zhao J, Xu F, Shen Z, Zhang L. Energy demand forecasting of the greenhouses using nonlinear models based on model optimized prediction method. Neurocomputing 2016;174:1087–100.
[80] Chou J-S, Cheng M-Y, Wu Y-W. Improving classification accuracy of project dispute resolution using hybrid artificial intelligence and support vector machine models. Expert Syst Appl 2013;40(6):2263–74.
[81] Coakley D, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renew Sustain Energy Rev 2014;37: 123–41.
[82] Sapankevych N, Sankar R. Time series prediction using support vector machines: a survey. IEEE Comput Intell Mag 2009;4(2):24–38.
[83] Suganthi L, Samuel AA. Energy models for demand forecasting—a review. Renew Sustain Energy Rev 2012;16(2):1223–40.
[84] Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev 2015;50:1352–72.
[85] Amasyali K, El-Gohary NM. A review of data-driven building energy consumption prediction studies. Renew Sustain Energy Rev 2018;81:1192–205.
[86] Vapnik VN. The nature of statistical learning theory. New York, United States: Springer-Verlag; 1995.
[87] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. New York, United States: Taylor & Francis; 1984.
[88] Lin L, Wang F, Xie X, Zhong S. Random forests-based extreme learning machine ensemble for multi-regime time series prediction. Expert Syst Appl 2017;83:164–76.
[89] Kittler J, Hatef M, Duin RPW, Matas J. On combining classifiers. In: IEEE transactions on pattern analysis and machine intelligence. IEEE; 1998. p. 226–39.
[90] Bashir ZA, El-Hawary ME. Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Trans Power Syst 2009;24(1):20–7.
[91] Pradeepkumar D, Ravi V. Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network. Appl

Soft Comput 2017;58:35–52.
[92] Hansen JV, Nelson RD. Time-series analysis with neural networks and ARIMA-neural network hybrids. Exp Theor Artif Intell 2003;15(3):315–30.
[93] Suykens JAK, Gestel TV, Brabanter JD, Moor BD, Vandewalle J. Least squares support vector machines. Singapore: World Scientific, Singapore; 2002.
[94] Kennedy J, Eberhart R. Particle swarm optimization. WA, Australia: IEEE. Perth; 1995.
[95] Jovanović RŽ, Sretenović AA, Živković BD. Ensemble of various neural networks for prediction of heating energy consumption. Energy Build 2015;94:189–99.
[96] Candanedo LM, Feldheim V, Deramaix D. Data driven prediction models of energy use of appliances in a low-energy house. Energy Build 2017;140:81–97.
[97] Wang Z, Wang Y, Srinivasan RS. A novel ensemble learning approach to support building energy use prediction. Energy Build 2018;159:109–22.