Renewable and Sustainable Energy Reviews 74 (2017) 902–924

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

A review on time series forecasting techniques for building energy consumption

Chirag Deb[a], Fan Zhang[b], Junjing Yang[a,*], Siew Eang Lee[a], Kwok Wei Shah[a]

[a] Department of Building, School of Design and Environment, National University of Singapore, Singapore 117566, Singapore
[b] Institute of Systems Science, National University of Singapore, Singapore 119615, Singapore

ARTICLE INFO

Keywords:
Building energy
Time series forecasting
Energy forecasting
Machine learning
Hybrid models

ABSTRACT

Energy consumption forecasting for buildings has immense value in energy efficiency and sustainability research. Accurate energy forecasting models have numerous implications in planning and energy optimization of buildings and campuses. For new buildings, where past recorded data is unavailable, computer simulation methods are used for energy analysis and forecasting future scenarios. However, for existing buildings with historically recorded time series energy data, statistical and machine learning techniques have proved to be more accurate and quick. This study presents a comprehensive review of the existing machine learning techniques for forecasting time series energy consumption. Although the emphasis is given to a single time series data analysis, the review is not just limited to it since energy data is often co-analyzed with other time series variables like outdoor weather and indoor environmental conditions. The nine most popular forecasting techniques that are based on the machine learning platform are analyzed. An in-depth review and analysis of the 'hybrid model', that combines two or more forecasting techniques is also presented. The various combinations of the hybrid model are found to be the most effective in time series energy forecasting for building.

1. Introduction

The International Energy Agency has identified energy efficiency in buildings as one of the five measures to secure long-term decarbonisation of the energy sector[1] [1]. Along with environmental benefits, building energy efficiency also presents vast economic benefits. Buildings with efficient energy systems and management strategies have much lower operating costs. Many countries have now accelerated the implementation of energy codes and regulations for various building types. These regulations outline basic requirements to achieve an energy efficient design for new buildings with a view to reduce the final energy consumption and related CO_2 emissions. In addition, many computer softwares have also been developed and widely implemented for energy efficient design of new buildings. Some of the most popular ones are EnergyPlus, DOE-2, eQUEST, IES, ECOTECT etc. A detailed study on the available computer-aided building energy analysis techniques and software tools is available in [2,3]. These regulations and computer-aided tools pertain to new buildings and are indeed very effective. However, once the building is functional, many factors govern the energy behavior of a building such as weather conditions, occupancy schedule, thermal properties of building materials, complex interactions of the energy systems like HVAC and lighting etc. Due to these complex interactions, accurate computation of energy consumption through computer simulation modeling is very difficult. For these reasons, data driven techniques for building energy analysis of existing buildings are very crucial. These techniques rely on past recorded data and attempt to model the energy consumption based on previous energy use patterns. Other factors influencing energy consumption can be used to improve the accuracy of such time series models. These techniques that make use of past data often fall under the of 'machine learning' and have been actively applied to building energy forecasting studies in the last two decades. The advantages and disadvantages of such data-driven techniques are presented in Table 1. This paper presents a comprehensive review of nine of the most popular machine learning techniques. The details of the techniques are described in later sections.

1.1. Significance to building performance optimization

To achieve an optimum level of energy performance in buildings, installation of efficient energy systems should be followed by appropriate operational and management strategies. This requires contin-

* Corresponding author.
E-mail address: bdgyj@nus.edu.sg (J. Yang).
[1] © OECD/IEA 2015 *World Energy Outlook Special Report*, IEA Publishing.
License: [http://www.iea.org/t & c/termsandconditions/].

http://dx.doi.org/10.1016/j.rser.2017.02.085
Received 4 July 2016; Received in revised form 25 January 2017; Accepted 27 February 2017
1364-0321/

Table 1
Comparison between data-driven and deterministic techniques for forecasting.

Technique	Advantages	Disadvantages
Data-driven	1. Very fast in computation with real-time data	1. Requires past recorded data
	1. Suitable for non-linear modeling	1. Non-transparent and confined
	1. Often more accurate than deterministic models	1. Difficult to generalize
Deterministic	1. Based on the science of building physics	1. Difficult to model real scenarios
	1. Transparent and no training data needed	1. Data unavailability of building properties
	1. Easy to generalize	1. Not very accurate

uous monitoring and management of the time series energy data along with other factors influencing energy performance of buildings. In relation to continuous monitoring and management of energy consumption in existing buildings, forecasting plays a significant role. It can provide a set of boundary conditions and targets for the building facility managers and owners within which the building's energy consumption should ideally fall (daily, weekly, monthly, and annual targets). As the time series forecasting model learns from previous energy consumption usage patterns, a gradual increase in the forecasted energy consumption values over a period of time may also notify the facility managers on the maintenance aspects of the building and energy systems. Besides the time series forecasting approach, other non-time series approaches can be adopted for building optimization purposes and they can also be combined with other computer simulation models like EnergyPlus to derive occupancy and other operational factors [4]. Yang et al. applied simulation based energy optimization for a test building in Spain. A web-based parallel genetic algorithm (GA) optimization framework making use of distributed computing resources was used to reduce the computation time [5]. Petri et al. presented a modular based optimization system that combines energy simulation and optimization using artificial neural network [6]. This helped in generating optimal set-points for a large scale building facility. The application showed significant energy reduction (kWh) in a real scenario. However, this required the building to be equipped with sensors and actuators for monitoring, control, and optimization. This may not be the case with most of the existing building infrastructure. Such challenges are discussed in detail by Aste et al. [7]. However, it is also noted that a relevant energy saving potential in buildings is related to commissioning, performance tracking and advanced control strategies. This being dependent on many factors including financial resources, policy support, green awareness, green material and technology etc. [8]. Zong et al. discussed on the challenges of implementing an economic model predictive control strategy (EMPC) for smart buildings [9]. It was observed that there are still challenges in application of model predictive control including the compromise between simplification and complication of building thermal dynamic modeling and balance between multi-energy systems. Realizing the challenges in integration of building performance data with other building related data, Hu et al. presented a novel method of linking traditionally disconnected data to construct data sources to enable in-depth and insightful building performance assessment [10].

Time series forecasting is integral to building performance optimization. Any optimization technique requires information either on future scenarios or in finding the best solutions against a test criterion. Machine learning techniques are useful in this regard and are often made use in solving these two problems. However, this review focuses on the time series forecasting aspects of building optimization rather than looking at the optimization problem in total (Fig. 1). An integration of these two shall be dealt with in a separate review.

1.2. Objectives of the review

Recent review studies on energy forecasting provide detailed accounts of the existing forecasting models and their classification. Zhao and Magoules reviewed and classified the existing methods for building energy consumption prediction into five categories [11]. Hippert et al. presented a review on short-term load forecasting [12]. Suganthi and Samuel presented a review on energy demand models for demand forecasting [13]. Fumo presented a review on building energy estimation and also studied the way the estimation models are classified [14]. Martinez-Alvarez et al. presented a survey on data mining techniques for time series forecasting of electricity [15]. The survey focused on the characteristics of the models and their configurations. Raza and Khosravi presented a review on short-term load forecasting techniques based on Artificial Intelligence (AI) techniques [16]. A recent study by Mat Daut et al. presented a review on building electrical energy consumption forecasting analysis using conventional and AI methods [17]. It was observed that the hybrid combination of SVM and swarn intelligence (SI) method has superior performance compared to other methods. The superior performance of the hybrid and ensemble models was also highlighted in recent review studies [18,19].

All these reviews provide vital information on energy forecasting models on different scales and emphasize on the superior performance of the hybrid models. A forecasting model can either be based on static data that usually fits a dependent variable to a set of independent variable, or it can make use of a single or parallel time series data. This study emphasizes on the forecasting techniques using time series data, which is also reflected in the title of this review. The significance to time series analysis is due to the increased awareness in real-time data collection and monitoring. Time series energy consumption can also be clubbed with time series data of indoor environmental conditions. With more sensors being deployed in buildings and more time series data being gathered, a suitable framework to analyze and to identify the forecasting capabilities is important. This review aims to understand the existing time series forecasting techniques and present their advantages and challenges. A detailed assessment of the hybrid model is also presented due to it being increasingly used in the literature. Since the hybrid model combinations are many, these are critically reviewed in a later section following the critical review of major, established techniques like ANN and ARIMA. This review paper shall also provide a basis of qualitative and quantitative comparison for all the 9 techniques mentioned here. It is to be noted that the hybrid model is considered as one technique among the 9 techniques presented. Within the hybrid model, there are a total of 29 combinations that are covered in this review. The objectives of this review paper are:

- To provide a collective and exhaustive review of the 9 major time series forecasting techniques with respect to building energy consumption.
- To perform a comparative analysis that includes both qualitative and quantitative aspects of these techniques.
- To elaborate on the various combinations of the hybrid model while assessing their performance and novelty.

1.3. Summary of papers reviewed

This section describes the structure of this paper and summarizes information on the papers that have been reviewed (Table 2). The structure of the reviews is as follows - first, an overview of each machine learning technique is provided, followed by a review of a set of studies exemplifying the application of this technique. For each paper reviewed, emphasis is given to the objective of the study, details of the time series data, and the accuracy of the model which is generally measured in Mean Absolute Percentage Error (MAPE). In the next half

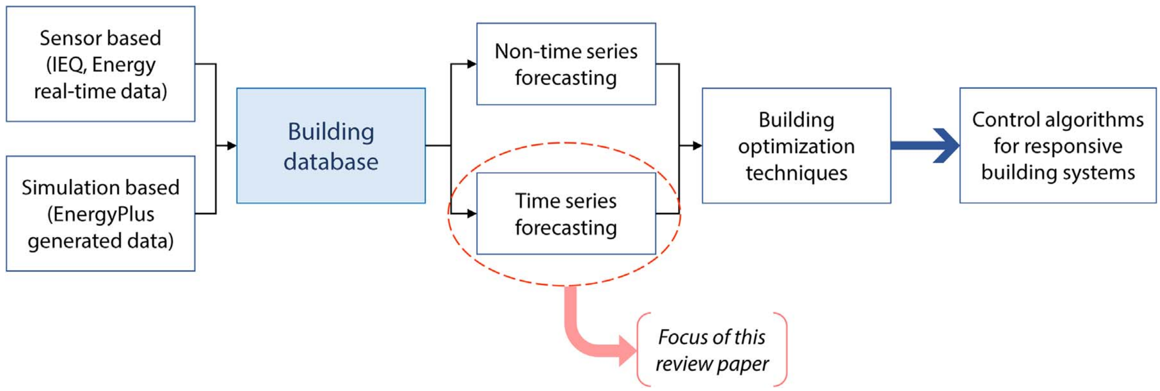

Fig. 1. Illustration showing the focus of this review paper within the domain of building optimization.

Table 2
Summary of papers reviewed.

Model	No. of papers reviewed	Publication year range	Year with maximum publications
ANN	17	1996–2015	2015
ARIMA	22	1992–2015	2011
SVM	19	2004–2015	2009
CBR	2	2013–2014	2013
Fuzzy	12	1998–2015	2015
Grey	13	2003–2015	2012
MA & ES	11	1971–2015	2013
NN	9	2004–2015	2014
Hybrid	61	1996–2015	2015

of this paper, a detailed discussion on the hybrid model is provided. The hybrid model is a combination of two or more machine learning techniques and has gained momentum in recent years. Most of the time series forecasting techniques currently employ the hybrid model. Therefore, a thorough investigation on the advantages and disadvantages of various combinations of models within the hybrid model is presented in later sections.

2. Time series techniques

A time series is an ordered sequence of values recorded over equal intervals of time. The analysis of time series can be divided into two parts. The first part is to obtain the structure and underlying pattern of the observed data. The second part concerns with fitting a model to make future predictions. Time series analysis is used for many applications including economic forecasting, process and quality control, census analysis etc. Recently, the application of time series analysis has been extensively applied to building energy consumption forecasting as increasingly buildings are being monitored on real-time basis. Such monitoring also provides past recorded data which can be used for crucial analysis and forecasting of the building's energy consumption. A usual approach in analyzing time series is to decompose the series into the following three components [20].

1. Trend – The general movement that the variable exhibits during the observation period without taking the seasonality and irregularities into account.
2. Seasonality – This is the periodic fluctuation of the variable subjected to analysis. It consists of effects that are stable along with time, magnitude and direction.
3. Residual – This is the remaining, mostly unexplainable part of the time series. These can be sometimes high enough to mask the trend and seasonality.

The fitting part of the time series model is a complex and challenging process involving detailed mathematical calculations. Usually, time series analysis can be divided into univariate and multivariate analysis. Univariate time series refers to a time series containing a single observation recorded sequentially over time, for example, hourly energy consumption. Multivariate time series is used when a group of time series variables are involved and their interactions are to be considered. This paper mostly deals with univariate time series analysis. Nine of the most popular time series forecasting techniques as observed in the case of building energy consumption forecasting are presented in the following sections.

2.1. Artificial neural network (ANN)

2.1.1. Overview of ANN

Artificial neural networks are modeling techniques that are somewhat similar in functioning to that of the human brain. The human brain contains a large collection of processing units called neurons that act in parallel for data processing and recollection. These neurons are connected with synaptic weights. It is found that these connections or 'weights' are capable of storing some kind of information that can be later retrieved. This concept was first introduced by McCulloch and Pitts in 1943 [21]. In the mathematical model of the neural network, the main purpose is to identify these 'weights' by training the model with past recorded data which is often arranged in sets of inputs and outputs. Hebb presented the adoption laws involved in the mathematical neural networks [22]. The calculations governing a single mathematical neuron were presented by Rosenblatt [23]. The training of the network is done with the idea of reducing the squared difference between the measured output and those predicted by the ANN model. The backpropagation algorithm which is based on reducing the outcome error by tuning in the right combination of 'weights' was put forward by Rumelhart et al. [24]. This algorithm revolutionized the use of ANNs for prediction as it reduced the computation time and increased the prediction accuracy. In a sense, for the simplest neural networks, the backpropagation algorithm works similar to the least square estimation method for simple linear regression. For building energy consumption forecasting, ANNs have been widely implemented for short, medium and long term energy forecasting. Often, the energy consumption time series is combined with other time series data such as weather and occupancy. These are summarized in the review section below.

2.1.2. Review of application studies on ANN

An early review on the effectiveness of ANN at forecasting and prediction was presented by Adya and Collopy [25]. Even as early as this, it was found that a majority of the studies supported the potential of ANN for forecasting and prediction. It was also noted that most of the studies lacked validation. This has improved in recent times where a high importance is given to validation. This is visible in the studies presented in this section.

Nasr et al. applied ANN to forecast electrical energy consumption [26]. Electrical energy data for each month from January 1995 to December 1997 was used to train the model and data from January 1998 to December 1999 was used to test and validate the model. The two models that were developed showed MAPE values of 5.03 and 4.43 respectively. Another early study on the application of ANN to forecast electrical energy consumption was performed by NIzami and Al-Garni [27]. In their study on the Eastern Province of Saudi Arabia, they related the electric energy consumption to weather data of air temperature, humidity, solar radiation and population. A two-layered feedforward ANN was used for the purpose. The results showed that the ANN, with an R^2 of 0.81 performs better than a regression model.

Gonzalez and Zamarreno [28] predicted short term electricity load by taking a part of the output back as an input in the feedback ANN model. The prediction error with respect to the measured output is used to train the network. The model produced impressive results for hourly load forecasting with the maximum Mean Absolute Percentage Error (MAPE) of 2.88. The feedback model used in this study was part of the Ph.D. dissertation of Schenker [29]. It was also suggested to quantify three aspects of the ANN – 1) the number of neurons in hidden layers, 2) the optimum size of the data set and 3) the training algorithm to be used. In this regard, it is to be noted that too many neurons may not be required in order to get a good result in the study on prediction and control models [30].

The application of ANN to electricity price forecasting was presented by Panapakidis and Dagoumas [31]. The study combined cluster analysis with different ANN topologies. The results showed that the Mean Absolute Error for day-ahead price forecasting was below 7.18%. It was also noted the use of raw data increased the error to about 20%.

There are also instances of combination of ANN with statistical techniques. For example, Karatasou et al. [32] combined the application of ANN and statistical analysis in predicting building energy consumption based on the methods described by Rivals and Personnaz [33]. The two ANN models that were proposed contained inputs as climatic variables and hour, day and week of the year as well as past values of energy consumption at *t-1*, *t-2* and *t-3*. The electricity consumption of industrial locations was explored by Azadeh et al. [34]. They used ANN to forecast long-term electricity consumption in energy intensive manufacturing industries of Iran using a feedforward neural network with back propagation error minimizing algorithm. The model took data from 1979 to 1999 for training and that from 2000 to 2003 for testing. The estimated results for three different model outputs were compared by a one-way analysis of variance (ANOVA). Another study on long-term energy consumption prediction using feedforward neural network was performed by Ekonomou [35]. The network had 4 neurons in the input layer, 1 neuron in the output layer and 20 and 17 neurons respectively in two hidden layers between input and the output layers. To train the model, past recorded data for 13 years (1992–2004) was used, while data from 2005 to 2008 was used for testing. The predicted results by the ANN model were very close to the test values.

Ann forecasting models based on recorded data are also compared with results from computer simulations. For example, Neto and Fiorelli [36] compared between an EnergyPlus simulation and an ANN forecasting model. Their study was based on data from an administrative building in the University of Sao Paulo. The feedforward neural network that they had used had origins in the works of Freeman and Skapura [37]. The results show that the EnergyPlus forecasts presented an error of ±13% for 80% of the tested database, whereas the ANN model showed an average error of about 10% with different networks for working days and weekends.

Yokoyama [38] predicted the cooling demand for a commercial building while proposing a global optimization method, called 'Modal Trimming Method' [39]. The global optimization method is used to assess the effect of numbers of neurons on the prediction accuracy. The study found that error for predicting cooling demand was 8.2%. It was also noted that increasing the number of neurons resulted in reduction of the prediction accuracy.

Mehdi et al. [40] combined multi-layer perceptron and radial basis function for hourly air temperature prediction. They developed four models with output as the 24 h time series of air temperature in a day with the output layer consisting of 24 neurons. The results showed the consistency of the MLP model to accurately forecast, whereas the RBF model was not accurate enough to predict hourly air temperature.

Pao [41] used ANN and ARMAX models to develop monthly electricity prediction model for Taiwan. They used monthly data from January 1990 to December 2002 for training and testing the model. The four inputs corresponded to the national income, population, gross domestic product and consumer price index. While comparing these models to linear models, it was found that ANN and ARMAX models were superior in the prediction accuracy. The ANN and ARMAX models displayed MAPE of 3.19% and 4.83% respectively for the testing dataset.

Hamzacebi [42] developed forecasting models using ANN for net electricity consumption of Turkey. The ANN model forecasted the sectoral electricity consumption of the four sectors which were the transportation, agriculture, residential and the industry sectors respectively. The data used to develop, validate and test the model was for a period between 1970 and 2004. The MAPE computed for the four sector's electricity consumption were 23.59%, 3.56%, 3.26% and 2.25% respectively.

Kelo and Dudul [43] developed a wavelet Elman neural network to forecast short-term electrical load prediction under the influence of ambient air temperature. The hourly total load demand for ten months of the year 2009 along with ambient air temperature data were used to develop the model. The load was decomposed into three sub-series which sow better regularities than the actual load profile. The Elman recurrent neural network is chosen for modeling and the model exhibited less MAPE values for the three seasons of summer, winter and rainy season respectively.

Chitsaz et al. [44] developed a prediction method using self-recurrent wavelet neural network (SRWNN) for short-term electricity load forecasting in micro-grids. The study was conducted for an institutional building in Vancouver for which hourly load for a period of one year has been analyzed. The SRWNN is a modified model of wavelet neural network (WNN). In WNN, a wavelet function is used to activate the hidden neurons of a feedforward neural network (FFNN). Both the models, WNN and SRWNN were developed and exhibited normalized MAE of 1.81% and 1.69% respectively. The computation time of the SRWNN model for the training phase is less than 35 s for one day prediction for the test cases of his paper, which is measured on a hardware set of Mac IntelCore i5 2.7 GHz with 12 GB RAM. While the computation time of WNN is less than 11 s.

Deb et al. [45,46] developed a feedforward ANN to forecast cooling loads for three institutional buildings in Singapore. Past two years daily energy consumption data was used for this study. To reduce the high variation in the energy consumption data, the data was divided into classes. Previous five days' energy classes were taken as the output to predict the next day. The results show that ANN is successful to predict next day energy consumption class with an R^2 value of 0.9794.

Tae Chae et al. utilized ANN to forecast sub-hourly electricity usage in a commercial building complex [47]. It was seen that the ANN with 15-min interval data forecasted accurately with more training data. The Mean Bias Error by taking 1, 2, 3 and 4 weeks for training are −10.4, −5.6, −2.3 and 0.03 respectively.

Chae et al. [48] developed forecasting models using ANN for sub-hourly electricity usage for a commercial building complex. The input variables that were selected for model development were the day type, time of day, HVAC set-point temperature schedules, outdoor air dry-bulb temperature and humidity. The results show that the daily peak prediction with past four weeks' training data performs the best. Three training methods that were used were Static, accumulative and sliding window with average APE of 3% and 4.5% for weekdays of August and September.

2.2. Autoregressive Integrated Moving Average (ARIMA)

2.2.1. Overview of ARIMA

ARIMA models are the basic and most general form of time series forecasting techniques. These are based on the idea of transforming the time series to be stationary by the differencing process. The time series can be assumed to be 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 input consists of lags of the dependent variable along with lags of the forecast error. It can be represented as in Eq. (1).

$$\begin{aligned} Output\ 'y' = {} & constant + weighted\ sum\ of\ one\ or\ more\ past\ values\ of\ \ 'y' \\ & + weighted\ sum\ of\ one\ or\ more\ past\ values\ of\ error\ \ 'e' \end{aligned} \tag{1}$$

The lags of the error term cannot be taken as the independent variables as they are not linear functions of the coefficients. Hence, the coefficients of ARIMA models with lagged errors must be computed by other techniques such as nonlinear optimization. The lags terms of the stationary time series are referred to as "autoregressive", whereas the lags of the forecasted error terms are referred as "moving average". A time series which requires to be differenced for the purpose of making it stationary is said to be an "integrated" version of a stationary series. These models are denoted as ARIMA (*p,d,q*), where '*p*' represents the order of the autoregressive part, '*d*' denotes the degree of first differencing involved and '*q*' denotes the order of the moving average part. The autoregressive part of the model with order '*p*' is represented as in Eq. (2).

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2+\ldots+} \phi_p Y_{t-p+} e_t \tag{2}$$

The moving average model of order '*q*' is represented as in Eq. (3).

$$Y_t = c + \Theta_1 e_{t-1} + \Theta_2 e_{t-2+\ldots+} \Theta_p e_{t-p+} e_t \tag{3}$$

Where '*Y*' is the output of a time series like electricity consumption data and 'e_t' is the error series. These models follow a common methodology which can be found in details in the work of Box and Jenkins [49].

2.2.2. Review of application studies on ARIMA

Newsham and Birt [50] developed an ARIMAX (Auto Regressive Integrated Moving Average with eXternal (or eXogenous) input) model to forecast the power demand for an office building. They used occupancy data as an external predictor to improve the model. The hourly power consumption data was collected for 79 days of which 5 complete days of network login data and 17 complete days of motion sensor data were missing. These values were imputed by using the mean of the non-missing values for that hour and week. The MAPE with occupancy and without occupancy data were 1.217 and 1.244 respectively, showing that when occupancy is considered, there is a slight improvement in forecasting accuracy.

Yao et al. [51] used a combined forecasting model based on an Analytical Hierarchy Process (AHP) to predict hourly cooling load. The AHP was used to deduce the weights of several models that were integrated to improve the forecasting accuracy. More details on AHP can be found in the work of Saaty [52,53]. In this study, three elements were used to determine the 'local priority'. These were the degree of fitting to the historical data, adaptability and reliability. The weights obtained for ARIMA, GM and ANN model were 0.564, 0.218 and 0.219 respectively.

Wang and Meng [54] presented a hybrid neural network and ARIMA model to forecast energy consumption for the entire Hebei province in China. They used annual energy consumption data from 1980 to 2008 for developing and testing the model. The ARIMA model was used to take care of the linear aspects, whereas, ANN was used to take care of the non-linear part. The results from the ANN were used to predict the error term of the ARIMA model. The MAPE for the hybrid model was an impressive 0.311%.

Kandananond [55] used ANN, ARIMA and multiple linear regression (MLR) to forecast electricity demand for Thailand. The data used was from 1986 to 2010. The results showed that the MAPE was 0.996% for ANN while those for ARIMA and MLR were 2.809% and 3.260% respectively. However, the paired test showed that there was no significant difference among these methods at α=0.05 and so, the ARIMA and MLR methods might be preferable to the ANN for the simplicity of their structure.

Chujai et al. [56] performed time series analysis for household electric consumption with ARIMA and ARMA models. The data used to develop the model was from December 2006 to November 2010. The results showed that the ARIMA model can represent the most suitable forecasting periods in monthly and quarterly and ARMA model can represent the most suitable forecasting periods in daily and weekly basis respectively. The most suitable forecasting periods found were for short term periods ranging from 28 days, 5 weeks and 6 months to 2 quarters for long-term.

Roken and Badri [57] used multivariate techniques such as ARIMA and dynamic regression econometrics techniques to forecast monthly electricity peak load for Dubai. They utilized the monthly data from January 1985 to March 2007 for 267 cases. The forecasting accuracy (coefficient of determination, R^2) between the actual data and the model results was 0.997. This was for the data between April 2006 and March 2007.

Zhuang et al. [58] studied on building load prediction based on time series method and neural networks. Initially, the seasonal model focusing on the analysis of periodicity of the sequence is developed. Then the residuals of the ARIMA model are modeled by ANN. The analysis was done over one week of data and the results showed that the combined model is superior to a simple time series model. However, no error analysis was provided.

Abdel-Aal and Al-Garni [59] used ARIMA models to forecast monthly electric consumption. They used data from past five years and forecasted new data for the sixth year. The derived model is a multiplicative combination of seasonal and non-seasonal levels. The model structure was ARIMA (1, 1, 0) $(1, 1, 0)_{12}$. The ARIMA models were found to be more accurate, requiring less data when compared to regression and abductive network machine-learning models. The average percentage error for the ARIMA model was 3.8%.

Abosedra et al. [60] estimated the electricity demand in Lebanon using Ordinary least Square (OLS), ARIMA and exponential smoothing. Monthly electricity data for one full year was used to develop an ARIMA (0, 1, 3) $(1, 0, 0)^{12}$ model. It was found that the ARIMA model is superior to the OLS and exponential models and presented a RMSE of 42.06 for monthly electricity consumption.

Almeshaiei and Soltan [61] developed a methodology for forecasting daily electric power load by utilizing the daily electricity data for Kuwait. The methodology to forecast was divided into five parts – (1) Primal visual and descriptive statistical analysis; (2) contour construction; (3) Load pattern decomposition; (4) Load pattern segmentation; (5) Future load forecasting. The model was developed using data from three years and the MAPE was 0.0384%.

Tserkezos [62] developed Box-Jenkins ARIMA based models to forecast residential electricity consumption in Greece using monthly and quarterly data. The data was obtained for a period of fifteen years with a total of 180 observations. The data was divided into two parts, 156 observations were used to develop the model and the remaining 24 observations were used to measure the performance of the model. The MAPE obtained for the monthly model was 3.78% and that obtained for the quarterly model was 7.69% respectively.

Erdogdu [63] performed electricity demand analysis using co-integration and ARIMA modeling for the electricity demand of Turkey. The data used in the estimation process was the quarterly time series net electricity consumption per capita for a period of twenty years with a total of 84 observations. Several tests were performed to determine if the partial adjustment model results are meaningful or not. The ARIMA model was used to project the electricity consumption based on two different scenarios and compared with the official projections which showed that for the next ten years' projections, Scenarios 1 and 2 inflate electricity demand by 60% and 345 respectively.

Twanabasu and Brendal [64] studied load forecasting in a smart grid oriented building. Three methods of forecasting were investigated, which were the ARIMA, ANN and SVM. The data used was the hourly energy consumption of Ostfold University College in Halden (HIOF), Norway and was for a period from 2008 to 2010. The holidays during that period were considered separately. The results showed that the ARIMA model displayed prediction accuracy MAPE of 5.67% when compared to 5.31% for the ANN model and 7.68% for SVM respectively. However, the ARIMA model was selected due to its transparency.

Katara et al. [65] used ARIMA model to forecast electricity demand in Tamale, Ghana. Yearly forecasts were made by utilizing data from 1990 to 2013. The data was divided into domestic, commercial and industrial usage. The model structure selected was ARIMA (1, 1, 3) based on several factors including a low MAPE of 10.42%.

Mohamed et al. [66] presented short term load forecasting using double seasonal ARIMA model. The data used for this study is half hourly load demand of Malaysia for a period of one year, measured in Megawatt (MW). The proposed model structure was ARIMA (0, 1, 1) $(0, 1, 1)^{48}$ $(0, 1, 1)^{336}$. The estimated MAPE was 0.99%.

Abledu [67] presented energy forecasting model for Ghana using average monthly maximum energy consumption data for a period of ten years. Out of this, one year data was selected to validate the model. The chosen model structure is SARIMA (1, 1, 1) (0, 1, 2). It was concluded that the SARIMA models were adequate for data sets of monthly energy consumption in general, and not just those with large sample size.

Rallapalli and Ghosh [68] used Multiplicative Seasonal Autoregressive Integrated Moving Average (MSARIMA) model to forecast monthly peak electricity demand for India. The data is collected for the period of April 2005 to march 2011 from five different regions in India. The MAPE for all the five regions ranged between 1.59% and 2.05%.

Wang et al. [69] proposed residual modification models with a view to improve the seasonal ARIMA model for electricity demand forecasting. Monthly data from past five years was used to develop the forecasting model from data obtained from Northwestern electricity grid in China. It was observed that the combination of the PSO optimal Fourier approach with the S-ARIMA model yielded the best MAPE of 2.19%.

Yasmeen and Sharif [70] developed ARIMA models to forecast electricity consumption for Pakistan. The monthly electricity consumption data of Pakistan from 1990 to 2011 was used. The results showed that ARIMA (3, 1, 2) model is the most appropriate in forecasting electricity consumption with a MAPE of 5.99%.

2.3. Support Vector Machines (SVM)

2.3.1. Overview of SVM

Support Vector Machines (SVMs) are developed on the concept of decision hyperplanes distributing data into two sets. The idea of determining this hyperplane is based on finding the largest margin between the two sets. Margin refers to the maximum width of the plane parallel to the hyperplane containing no interior data points. These were first introduced by Vapnik [71] in 1995. The details of the theory of SVM can be found in [72]. Along with classifying data, SVM can also be used for the case of developing a regression model. In this case, the goal is to determine a function which deviates from the measured outputs by a value no greater than an error term for each combination of the input values. For nonlinear regression problems, the data is transformed using a nonlinear kernel function which maps the inputs to a high-dimensional feature space. So the general performance of the SVM regression model depends on the adequate selection of the kernel parameters. This is often a challenge while developing SVM regression models and are often solved by using intelligent optimization techniques. Fan et al. [73] used differential evolution (DE) optimization techniques to find suitable kernel parameters for a SVM model to forecast energy consumption in a building.

2.3.2. Review of application studies on SVM

Chen et al. [74] applied SVM to load forecasting for a EUNITE competition database. The goal was to predict daily load demand for a month based on past two-year load demand data. It was seen that the treatment of the dataset as pure time series data gave better results rather than considering other factors like ambient temperature to predict the load demand. The MAPE without considering temperature was 1.95% winter season and 2.54% for the months of January and February.

Hong [75] developed a SVM regression model with immune algorithm (IA) to forecast a regional annual electric load in Taiwan. Data from 1981 to 2000 were utilized to train, validate and test the model. The suitable parameters to develop SVR load forecasting model are presented in another study by the same author [76]. The results show that the forecasting MAPE for the four regions are all below 2.45% for the SVRIA model.

Dong et al. [77] applied SVM to predict building energy consumption in a tropical region. Weather data with monthly mean outdoor dry-bulb temperature, relative humidity and global solar radiation are taken as the three inputs. These are supplemented by the mean monthly landlord utility bills to develop and test the model. The prediction results are found to have a CV less than 3% and percentage error less than 4%.

Li et al. [78] presented four modeling techniques for the prediction of hourly cooling load for an office building in Guangzhou, China. The six input parameters that are selected are based on three climatic variables - outdoor dry-bulb temperature, relative humidity and solar radiation intensity. One full month hourly data was used to develop the model. The mean relative error (MRE) for the SVR model tested on four different months were below 1.02%. A similar study by the same authors was also published in [79].

Xuemei et al. [80] proposed Least Square Support Vector Machine (LS-SVM) for cooling load forecasting for an office building in Guangzhou, China. The hourly climate data and building cooling load for five months is taken to develop the model. The results were compared to a back-propagation ANN and showed that LS-SVM performed superior with a Mean Absolute Relative Error of 1.65%.

Bozic et al. [81] proposed Least Square Support Vector Machine for short-term load forecasting. The hourly data for a week was utilized and hourly as well as daily load forecasting was performed. The results displayed MAPE values between 0.93% and 3.04% for daily forecasts. For the hourly forecasts, the APE was between −1.53% and 3.96%.

Fernandez et al. [82] presented improvements like variable data learning window and diverse learning data weighting combinations to existing short-term forecasting methods. The model was tested and compared with various entries for the donosti1, donosti2, ASHRAE and EUNITE competition. The results for next day load forecasting were better achieved with the new model, exhibiting a MAPE of 6.69%. In a similar attempt to improve the modeling methods by utilize the data from the above mentioned competition, Edwards et al. [83] proposed that Least Squares Support Vector Machines (LS-SVM) performs the best for residential buildings. The corresponding MAPE was 6.95 ±

0.21%.

Ogcu et al. [84] used ANN and SVM to forecast the electricity consumption of Turkey. Monthly electric energy consumption data for two years were utilized to develop the models. The MAPE for ANN and SVM for testing dataset were 3.9% and 3.3% respectively.

Zhao et al. [85] presented a combinational model based on SVM. Initially, a model group was developed using five forecasting models. The outputs of these five models formed the inputs for the SVM model. The data used to develop and test the model was from 1986 to 2000. The combined model exhibited better forecasting performance with a MAPE of 2.139%.

Ceperic et al. [86] presented a generic strategy for short-term load forecasting based on SVM. Feature selection algorithms for automatic model input selection and the use of particle swarm global optimization based technique for the optimization of SVR were explored. The model was trained and tested on two datasets. The modeling results exhibited an improvement from 20% to 23.4% for one dataset and from 2.5% to 34.2% for the other dataset respectively. The overall training time of the proposed model is at maximum 40 min, using Linux based work station with two Intel Xeon X5675 CPUs.

Jain et al. [87] developed building energy forecasting model using SVM for multi-family residential building in New York City. The study also examined the impact of temporal and spatial granularity on model accuracy. The results show that the most effective models are built with hourly consumption at the floor level.

Kaytez et al. [88] compared regression analysis, ANN and LS-SVM to forecast electricity energy consumption in Turkey. Gross electricity generation, installed capacity, total subscribership and population are used as the inputs using recorded data from 1970 to 2009. The testing result for LS-SVM was a MAPE of 1.004% when compared to 1.19% for ANN and 3.34% for multi-linear regression model respectively.

Liu et al. [89] studied time series prediction method for building energy consumption forecasting using SVM. Hourly energy consumption data for a month was studied and the first three weeks samples were taken as the training dataset. The model was tested on two buildings and exhibited MSE of 0.0186 and 0.091 for the two buildings respectively.

Fu et al. [90] used SVM to predict next day electricity load for public buildings. Different models were developed for each sub-system in the building. These were divided as air conditioning, lighting, power and other equipment. Three and a half months of hourly data was used to train the model and half month data was used to test the model. The results show that the SVM is superior to other machine learning models at system level energy forecasting as well as on building level (total) energy forecasting. The exhibited CV_RMSE for total energy forecasting by the SVM model was 15.2%.

2.4. Case-Based Reasoning (CBR)

2.4.1. Overview of CBR

Case-Based Reasoning is based on the recalling of information from a prior case to solve a new case. It is opposed to the idea of deriving rules from observations that can be applied to a generic set. Rather, this process relies on similar cases that has been tackled in the past and not so much on the methodology of solving them. The primary knowledge source in CBR is a memory of stored cases recording specific prior episodes [91]. The original formulation of this concept is derived from the study on the role of reminding in human reasoning [92]. More details on the theoretical fundamentals on CBR can be found in [93]. In its application in building energy consumption prediction, CBR takes into account cases that have similar types of input variables and tries to model based on such similar previous scenarios.

2.4.2. Review of application studies using CBR

Monfet et al. [94] presented a CBR based method for the energy demand of commercial buildings. To predict the total energy demand (dependent variable), the database of cases is searched for cases having similar independent variables such as total building electricity demand, total electrical cooling demand, total electrical heating demand, total fan demand, etc. Data for the previous three hours and available data for the next three hours are used to retrieve cases. First, the results of the monthly predictions for April to December 2012 where the weights, updated on a monthly basis, are presented. Secondly, the prediction made over the first three months of 2013 using the weights selected using a full year of data (data for 2012) are presented to forecast the energy demand every hour for the next three hours for the first three months of 2013, from January 1st to April 3rd. 2012 monthly results show that CV-RMSE and RMSE of the proposed method are within 41.98% and 20.4 kW. 2013 hourly results show that CV-RMSE and RMSE of proposed method are lower than 13.88% and 17.9 kW.

Monfet et al. [95] proposed Case-Based Reasoning (CBR) for hourly energy demand of commercial buildings. Two different approaches are considered to evaluate the performance of the CBR to predict the energy demand: The first evaluation set uses a case library including data from May 2008 to April 2009. This was used to predict the electricity demand from May 2009 to December 2009. For this set, no forecast weather information is available. New cases are created every hour based on hourly averaged monitored building data. These new cases are added in a continuous manner to the library of cases as they become available. Results show that the CV-RMSE vary between 13% and 19%, on an hourly basis. An additional set that includes data monitored in 2012, including weather forecast is used to evaluate the performance of CBR to predict whole building electric demand for the next 3 h and results show that the CV-RMSE are below 22%.

2.5. Fuzzy time series

2.5.1. Overview of Fuzzy time series

Fuzzy time series are time series observations with linguistic values rather than the conventional numerical or crisp values of observations. Therefore, the conventional rules of time series analysis cannot be applied to study these. Song and Chissom [96] introduced this concept and have laid down several formal definitions for fuzzy time series. Later studies focused on partitioning the universe of discourse, followed by constructing fuzzy relationships. The values are then forecasted and defuzzified to obtain the numerical output. It is to be noted that an adequate choice of the length of each fuzzy interval can highly improve the forecasting accuracy. A detailed description of the fuzzy time series analysis is beyond the scope of this paper. For building energy forecasting studies, the challenge lies in deriving the weights of the fuzzy logic relationships, membership functions and the rules. The next section provides a review of application studies on fuzzy time series technique.

2.5.2. Review of application studies on Fuzzy time series

Azadeh et al. [97,98] introduced fuzzy regression which is a combination of classical regression and fuzzy techniques. The monthly electricity consumption data from April 1992 to February 2004 was used to construct the model. The impact of data preprocessing and post processing on the fuzzy regression performance is studied. The proposed model shows superior results when compared to other machine learning models like genetic algorithm (GA) and ANN and exhibits a MAPE of 0.0082%.

Bolturk et al. [99] used fuzzy time series with Singh's method to forecast electricity consumption for a commercial building in Turkey. Singh's method [100] is a model of order three that utilizes the historical data of year n−2, n−1, n for framing rules to implement fuzzy logical relation, $A_i \rightarrow A_j$, where A_j, the current state, is fuzzified enrollments for year n+1. Two approached were considered in which the first approach entailed electricity consumption values for three different time periods in a day. The other approach considered the total electricity consumption. The RMSE for both these approaches were 467.567 and 490.310 respectively with a difference of 5% between them.

Efendi et al. [101] applied linguistic out-sample approach for fuzzy time series for daily electricity load demand forecasting for Malaysia. This study described a new rule in determining the weights of the fuzzy logical relationship (FLR) by using index number of close relationship in the fuzzy logical group. Daily electrical load data for 8 months was considered for the analysis and the MAPE for varying number of testing datasets was below 1.63%.

Ismail and Mansor [102] presented a fuzzy logic approach for forecasting half-hourly electricity load demand for Malaysia. Rules are developed for each day factors (working, weekend and holidays), temperatures and load using current load, current temperature, previous load and previous temperature. Four defuzzification methods are employed. The data used is half-hourly electricity load for a week. The results exhibit a MAPE between 1.645% and 5.81% for the four defuzzification methods chosen.

Lim et al. [103] proposed the chaos fuzzy controller to predict the short-term electric power demand of a power plant. The input data of the controller are weather-related ones, such as temperature, climate and the increase of temperature. The results show that the average error between predicted data and actual recorded data was 5.65%. I was suggested that more weather related data would improve the accuracy of the model.

Liu et al. [104] proposed short-term electric load forecasting method using slide window fuzzy time series to train the trend predictor in the training phase, and use these trend predictors to generate forecasting values in the forecasting phase. Hourly electric load for four days was chosen as the working examples. The maximum MAPE is found to be 7.74% for the testing dataset.

Pei [105] proposed an improved fuzzy time series approach for load forecasting. The approach utilized unequal-sized intervals partitioning based on K-means algorithm along with an improved fuzzification method. Hourly electric load data for a day was analyzed to build the forecasting model. The results show that the prediction accuracy is the highest by using 4 order fuzzy time series forecasting model with 8 clustering partition on domain.

Otto and Schunk [106] presented fuzzy-processing algorithm to forecast time series of electric load. The model is developed based on load values collected over a period of two years. The forecast load for three weeks has a standard deviation of 6264 MW and 4487 MW for 1996 and 1997 respectively.

Ozawa et al. [107] presented a new approach based on possibility theory and fuzzy auto-regression (AR) to model the time series data of electric power consumption. The maximum daily consumption of electrical energy (MW h) recorded for 52 weeks is taken for analysis. The parameters of fuzzy sets that construct the target fuzzy AR are determined by linear programming as proposed by Tanaka et al. [108]. It is seen that the proposed fuzzy AR model is successful to forecast possibility regions for winter and spring periods.

Pereira et al. [109] compared two types of modeling techniques to predict electric load time series The two models were Seasonal Autoregressive Integrated Moving Average with Exogenous Variables model (SARIMAX) and Fuzzy Interface System (FIS). Three exogenous variables are also considered, and they are the number of customers connected to the electricity distribution network, the temperature and the precipitation of rain. The minimum MAPE obtained by the SARIMAX model was 2.69% as compared to 1.99% for the FIS model.

Kurniawan [110] presented electrical load time series using interval type-2 fuzzy logic system. Electrical load data for about 2 years for Indonesia was used to develop and test the model. Various sensitivity analysis were performed for the variation in the number of input, learning rate value and variance parameter of Gaussian membership function. The results show that the model with two inputs interval type-2 fuzzy logic is the best with a RMSE value of 0.082691.

2.6. Grey prediction model

2.6.1. Overview of Grey prediction model

Grey system has gained popularity in solving uncertainty problems under discrete data and where information is not sufficiently available. This concept was introduced by Deng [111] in the early 1980s. The main purpose of the theory is to predict the behavior of systems which cannot be detected with stochastic or fuzzy methods with limited data. The basic Grey prediction model is the GM (1,1), which is a time series forecasting model and implies a first order single variable prediction model. GM (1,1) does not require any prior knowledge like the probability distribution of the input data and it can be used when the amount of input data is limited. This is the main benefit of this model that it can also be used with minimum data. This model entails a set of differential equations adapted for parameter variance as well as first-order differential equation. GM (1,1) is denoted as in Eq. (4).

$$\hat{x}_0^{(0)}(k) = \left[x^{(0)}(1) - \frac{u}{a}\right](1-e^{a})e^{-a(k-1)} \tag{4}$$

Here, $x^{(0)}$ is a non-negative sequence, 'a' and 'u' is estimated using the Ordinary Least Square (OLS) method. More details on the theory of Grey prediction models can be found in [112,113].

2.6.2. Review of application studies on Grey prediction model

Jiang et al. [114] applied grey forecasting technique to predict the operating energy performance of air cooled water chillers. The total data considered was from 8:30 to 13:00 with data recorded every 30 min interval. A grey and a residual grey model were developed with 5 and 4 entries respectively. The predicted values of coefficient of performance (ΔCOP) matched well with the measured ΔCOP. It was seen that the model can predict the ΔCOP values for several future hours with a maximum percentage error observed between actual and predicted values being 1.097%.

Lee and Tong [115] improved a grey forecasting model by incorporating genetic programming. The genetic programming sign estimation was used to minimize the error that grey forecasting technique usually yields. However, grey forecasting technique is highly rated for its applicability to small data sets or ones with limited information. The annual energy consumption data for China from 1990 to 2003 is considered for model-fitting and 2004–2007 is utilized as ex post testing. The MAPE for the combined genetic programming and grey modeling was 20.23 as compared to 26.21 with just the grey model.

Yao et al. [116] presented an improved grey based forecasting approach for electricity demand forecasting for Taiwan. Data for one week was used for this analysis. The prediction improved significantly by applying the transformed Grey model and average system slope concept. The adaptive value of 'a' in the Grey differential equation is obtained quickly with the average system slope technique. It showed that the absolute error was reduced to 4.88% by adopting this method. A year later, in a study by Yao and Chi [117] showed that a Taguchi-Grey based electricity demand model had a prediction error of 4.96%. Taguchi's method enables scientific analysis of the design of parameters, variance analysis and the determination of optimal parameters for a process.

Akay and Atak [118] applied grey prediction with rolling mechanism for electricity demand forecasting of Turkey. The annual data from 1970 to 2004 was used to develop the model. The total electricity consumption analysis generated an error of 3.69% for the entire period, whereas for just the testing part (year 1994–2004), the error was 3.43%. The same was 4.36% when only industrial electricity consumption was taken into account.

Pao et al. [119] forecasted CO_2 emissions, energy consumption and economic growth in China using an improved grey model. They employed the nonlinear grey Bernoulli model (NGBM) to predict these indicators and proposed a numerical iterative method to optimize the parameter of NGBM. The annual data for energy consumption was collected from 1980 to 2009 and the period between 2003 and 2008 was used to test the model. The MAPE for energy consumption was 6.26 and that for CO_2 was 4.72 respectively.

Zhou et al. [120] developed a grey box model for next day building thermal load prediction. The hourly building load for next day was determined based on the air temperature, relative humidity and solar radiation data which in turn were predicted using a regressive solar radiation module and a dynamic grey module respectively. The data used was on hourly basis for few days in March and September 2007. The results showed that the overall performance of building load prediction in MAPE is below 8.

Hsu and Chen [121] presented an improved grey prediction model for power demand forecasting. The improvement was done by residual modification using ANN sign estimation. Power demand data for Taiwan from 1985 to 2000 was utilized of which the data from 1985 to 1998 was used for model fitting and that from 1999 to 2000 was reserved for ex post testing. The MAPE for grey model and improved grey model was 3.88 and 1.29 respectively.

Bianco et al. [122] analyzed two forecast models for non-residential electricity consumption in Romania. The annual values of non-residential electricity consumptions were obtained for a period of 1975–1985. A trigonometric residual technique was employed to the Grey model with rolling mechanism and the resulting model (TGMRM) had a MAPE of 0.6 and 4.4 for the model building and testing stage respectively.

Hamzacebi and Es [123] developed forecasting model to forecast annual electricity consumption of Turkey using Optimized Grey Modeling. The total electricity consumption from 1945 to 2010 was taken for the study, of which, data from 1945 to 2005 was taken for modeling and that from 2006 to 2010 was taken for testing respectively. Two approaches were tested, the direct forecasting and the iterative forecasting. It was seen that direct OGM was better with a MAPE of 3.28% when compared to 5.36% for iterative forecasting.

Jin et al. [124] proposed a new grey model with grey correlation contest for short-term power load forecasting. The hourly data from January to June of 2009 was utilized to develop a hybrid optimization grey model (HOGM) based on segmented grey correlation and multi-strategy contest. The results showed that the maximum forecasting error based on HOGM is 4.91%.

Kang and Zhao [125] developed an improved grey model long-term load forecasting by connecting moving average method and Markov model with the grey model. The annual electrical load for Qingdao, China was obtained from 2000 to 2009 to develop the model. It was seen that the improved grey model had a prediction error of 1.69% when compared to the general grey model GM (1, 1) which had an error of 3.84%. In this study, however, no testing data was separated from the dataset.

Zhou et al. [126] presented a trigonometric grey prediction approach by combining the traditional grey model GM (1, 1) with the trigonometric residual modification technique. The China annual electricity consumption from 1981 to 1998 was used to develop the model and the data from 1999 to 2002 was used to test the model. The MAPE was found to be 2.12% for the model building stage and 2.37% for the testing stage.

Li et al. [127,128] presented research on short term load forecasting using improved grey dynamic model. The data used was hourly electrical load for four days to forecast the load for a future day. A cubical spline function is presented to calculate the derivative and background value in grey number interval. In addition, Taylor approximation method is applied to achieve high forecast accuracy. The forecasting accuracy for ordinary day and for special days was found to be 0.855% and 0.585% respectively.

2.7. Moving average and exponential smoothing (MA & ES)

2.7.1. Overview of MA & ES

Moving average and exponential smoothing are two different forecasting methods but are similar in the way that these models consider the time series locally stationary with a slowly moving average. However, the exponential smoothing method gives a higher weighting to recent values while the moving average method assigns equal weights to all values. These decomposition components are the basic underlying foundation of many of the time series methods. There exist various extensions to the basic moving average model such as ARMA, ARIMA etc. The principle behind moving averaging is that demand observations that are close to one another are also likely to be similar in value.

A simple equally weighted moving average is given as in Eq. (5).

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-m+1}}{m} \tag{5}$$

Here, the average is centered at period 't−(m+1)/2'. The forecast for the value of 'Y' at time 't+1' that is made at time 't' equals the simple average of the previous recent 'm' observations.

The concept of simple exponential smoothing is to attach larger weights to more recent observations. The forecasts are calculated using weighted averages which decrease exponentially as observations are taken further from the past. These can be mathematically represented as shown in Eqs. (6) and (7).

$$l_t = \hat{y}_{t+1|t} = \alpha y_t + (1-\alpha) l_{t-1} \tag{6}$$

$$\hat{y}_{T+h|t} = l_T \tag{7}$$

Where 'l_t' is the estimate of the level at time 't', 'y_t' is the observed time series at time 't', $\hat{y}_t$ is the forecasted value at time 't' and 'α' is the smoothing parameter of the level, $0 \le \alpha \le 1$. An 'α' close to 1 will a substantial adjustment value and make the forecasts more sensitive to swings in previously recorded values based on previous period's error. More details on the fundamentals of exponential smoothing can be found in [129].

2.7.2. Review of application of MA & ES

Christiaanse [130] applied general exponential smoothing for short-term hourly MWH load forecasting based on observed values of integrated hourly demand. Two years of hourly load data are assembled for this study. The data represented hourly MWH values for the AEP conforming load during 1966–1967, a period of 17,250 h. Input of the model is past values of observed load, measured in hourly MWH and output is hourly MWH load with lead times of one hour to one week. From the stimulation results, a model includes nine frequencies and a beta of 0.994 for the smoothing constant was selected, the standard error in MW of the selected is 224.

Taylor et al. [131] compared the accuracy of six univariate methods: double seasonal ARMA modeling, exponential smoothing for double seasonality, ANN, a regression method with principal component analysis and two simplistic benchmarking methods for short-term electricity demand forecasting for lead times up to a day ahead. A time series of hourly demand for the city of Rio de Janeiro and a series of half-hourly demand for England and Wales are used for comparison. The Rio data covers the period from Sunday 5 May 1996 to Saturday 30 November 1996. The England and Wales data is from the period between 27th March 2000 and 22nd October 2000. The results show that exponential smoothing achieved the lowest MAPE, below 3% for Rio dataset and below 1.5% for the England and Wales dataset. A highlight of this study is the success of the exponential smoothing method. In addition to its forecasting performance, it is important to note that, of the other four sophisticated methods considered, this method is the simplest and quickest to implement.

Taylor [132] presented an adaptation of Holt-Winters' exponential smoothing, ARIMA modeling, an intra-day cycle exponential smoothing model, methods that do not attempt to model the seasonality as well as an approach based on weather forecasts for very short-term load prediction which is between 10 and 30 min ahead. The data used in this study was obtained from the National Grid, (a transmission company in Great Britain). The data consists of 30 weeks (one minute interval) of observations of electricity demand in Great Britain, from 2nd April 2006 to 28th October 2006. For very short-term prediction, the best results are achieved using double seasonal Holt-Winters' exponential smoothing method and the restricted intra-day cycle exponential smoothing method with parameters for both methods optimized for a 30-min ahead prediction. The MAPE values of these two methods were below 0.55%. For predictions beyond 30 min ahead, combination of weather-based and Holt-Winters' exponential smoothing outperformed other methods with a MAPE below 1.2%.

Bindiu and Chindriu [133] presented a day-ahead load forecasting for a period of one week using Holt and Winters Exponential Smoothing method, a fittings manufacturer from Cluj-Napoca is chosen as case study and the history of the consumption data was taken as the input. The statistical measures indicate good results for the MAD to a very low value, approximately 7% of the maximum load for the forecasted period. MSE value indicates good results and low load forecast error. However, the MAPE of the chosen model is large.

Taylor [134] presented exponential smoothing formulations for single, double and triple seasonality based on the Holt-Winters method, intraday cycle exponential smoothing, ARMA model and neural network for short-term load forecasting. In this study, two load series are used; one is for Great Britain and the other is for France. Each consists of six full years of half-hourly observations for electricity demand from 2001 to 2006. The result show that the triple seasonal versions of all the methods were more accurate than the double seasonal methods, triple seasonal HWT exponential smoothing and triple seasonal ARMA outperformed other single methods with MAPE lower than 1.75%. However, the best result was achieved by combining ARMA and HWT exponential smoothing for triple seasonality with MAPE below 1.7%.

Kotillova [135] performed short-term electricity load forecasting using exponential smoothing, seasonal ARIMA and industry model. The electricity load data for every 30 min for the state New South Wales is used in the study. The monthly data for June and July 2010 is chosen for training and monthly data for August 2010 is used for testing. The MAPE of ARIMA and industry model are 1.468% and 1.421%. The proposed exponential smoothing model achieved the lowest MAPE which was lower than 0.5%.

Taylor [136] presented seven exponentially weighted methods: HWT Exponential Smoothing, Intraday cycle (IC) exponential smoothing, Total and Split Exponential Smoothing, DWR with Trigonometric Terms, DWR Spline, Spline-Based Exponential Smoothing and SVD-Based Exponential Smoothing to predict short-term load up to one day ahead. Moreover, ARMA, ANN and a weather-based approach are also described in this study. A British and French load series are used in the empirical analysis. Both consist of all half-hourly observations in 2007, 2008 and 2009. The best result is achieved by combining HWT, new SVD-based and weather-based model with MAPE lower than 1.15%.

Adilah et al. [137] presented five exponential smoothing methods: Holt-Winters for Intraday Seasonality, Holt-Winters for Intraweek Seasonality, HWT Exponential Smoothing (The Holt-Winters method for double multiplicative seasonality), and two Modified Holt-winters Exponential Smoothing by using square rooting the k index and cube-rooting the k index when forecasting the future load for lead times from a half-hour-ahead to a year-ahead. A half-hourly electricity load demand of Malaysia for one year, from September 01, 2005 to August 31, 2006 is used in this study for both k-step and one-step ahead forecasts. The result indicated that one-step ahead forecasts are not greatly influenced by the lead times, thus it is more suitable for electricity load demand forecasting. The Holt-Winters Taylor (HWT) exponential smoothing outperformed traditional Holt-Winters and modified Holt-Winters exponential smoothing methods in terms of MAPE.

Ariffin et al. [138] applied the moving averages (MA) and Exponential smoothing techniques (EST) for short- term load forecasting. The data used in his study is divided by two types of load namely Semester On (SOn) and Semester Off (SOf) in Universiti Teknologi PETRONAS (UTP), Malaysia. The load in UTP for the year 2011 will be forecasted based on the data of year 2010. It can be concluded from the results that for Semester ON, exponential smoothing technique exhibits better results when α=0.3 with MAPE around 12.9994%. In terms of Semester OFF, that exponential smoothing with α=0.3 exhibits better results. The MAPE is around 14.4330%.

Maçaira et al. [139] adopted exponential smoothing methods for forecasting the Brazilian residential energy consumption till 2050. In particular with Pegels' techniques, called Standard and Damped, hyper parameters optimizations are performed to adjust the official projections expected by Energy Research Company (ERC), through the "The Ten Year Energy Planning" (PDE) and "The National Energy Planning" (PNE). Annual data measured by TWH between 1995 and 2013 obtained from the ERC site is used in the study. It is found that the best performance is achieved by optimized model equations with Damped Pegels, the MAPE of PDE 2014–2023 and PNE 2050 are 2.68% and 3.11% respectively.

2.8. K – Nearest Neighbor prediction method (kNN)

2.8.1. Overview of kNN

KNN algorithms are techniques that are popular for pattern classification and are based on the similarity of the individuals for a given group. The members of a group are surrounded of individuals which have similar properties. This forms the basic learning rule for the KNN based classification. For new unclassified sample points, this rule assigns the classification to the nearest for a set of previously classified points. Unlike most statistical methods which builds a model from the information available in the data base, the KNN method considers the training set as the model itself. A KNN algorithm is characterized by issues such as number of neighbors, adopted distance, etc. More details on the kNN fundamental theory can be found in [140]. An important parameters defining the KNN classifier are the choice of a metric. The choice of the metric to measure the similarity between two time series depends mostly on the specific features of the series. The most common metric is the square of the Euclidean distance. The second important parameter is the number of neighbors. In practice, the optimal value of k is usually small for noise-free time series. Often, k is determined by minimizing the mean squared error of the training data.

2.8.2. Review of application studies on kNN

Lora et al. [141] proposed a time series prediction method based on the kNN technique. The electric energy data for Spain from January 2000 to May 2001 is taken to determine the optimal number of neighbors, whereas the data from June to November 2001 is taken to test and validate the model. The mean forecasting error is found to be 2.3% for the test data set. The results are compared to a conventional regression model whose parameters are obtained by solving a least squares problem. It is observed that the kNN classifier is more accurate.

Lei et al. [142] proposed the one-rank local linear model based on multivariate time series. The daily electrical load data from district in Chongqing is taken for analysis. The time series temperature data is also taken as an input along with daily electric data from January to March 2003. The results of the multivariate time series model with temperature also considered show that the maximum forecasting error is 6.16% and the average error is 0.97% respectively.

Brown et al. [143] proposed a new technique for real-time building energy modeling and prediction technique using kernel regression with kNN. Hourly power data from 4 buildings spanning for a period of 1.5 years starting from January 1st, 2007 was utilized for modeling. The results show that the RMS for all the four buildings are under 10.86.

Al-Qahtani and Crone [144] proposed a multivariate k-NN regression method for forecasting electricity demand in the U.K. The hourly electrical load data for 2004 is used to train the model and to optimize the model parameters. The model is then used to predict all days for the year 2005. The error is estimated using an average on the 24-h ahead trace forecasts. The results show that the proposed multivariate K-NN model exhibits a MAPE of 1.8133%.

Lachut et al. [145] developed prediction models to forecast appliance-level power draw and home-level energy consumption for residential buildings in Arkansas, Maryland and California. For the analysis, three machine learning algorithms including kNN and one time series based (ARMA) model was used. It is observed that kNN performed better for predictions on a weekly basis. In addition, this study also demonstrates that a simple statistic-based algorithm which uses the average power draw in previous time cycle as the predicted average draw in the next time cycle performs well and can significantly reduce computational complexities.

Roger et al. [146] used a Multiple-Input Multiple-Output for long-term power forecasting in Cameroon. The model was based on KNN learning and training. The dataset consists of power consumption values for several countries from 1972 to 2009. The results with respect to the MAPE values show that the model is best in forecasting for next 10 and 15 years in advance. The model is then used to forecast the electricity consumption in Cameroon till 2035 using forecasted population growth.

Colombo et al. [147] presented a Maximum Length Weighted Nearest Neighbor (MLWNN) approach for next electricity load forecasting. Two years' hourly data from three countries is used for the analysis. The results show that the proposed model outperformed Weighted Nearest Neighbor (WNN) on two of the datasets with a similar performance on the third one.

A detailed comparison on the above-mentioned 8 time series forecasting techniques is presented under the section on 'discussion'. As a brief comparative analysis with respect to the time consumed, 6 studies are identified that have comparable input data configurations and utilizes different forecasting technique. It is hard to find a single study that applies all these forecasting techniques on a common dataset. Therefore, the 6 studies that are selected for comparison have input data configurations as close as possible. The results of the comparison are presented in Table 3. Since the dataset is not common, any derivation on the capability of these should be avoided. For example, the MAPE values for ANN are lower than that for ARIMA, however, the length of training data for ANN is also less than the data taken to train the ARIMA model. In most cases, it is seen that the accuracy of machine learning techniques for building energy forecasting are much comparable to each other.

2.9. Hybrid models

2.9.1. Overview of hybrid models

Hybrid models are combinations of two or more machine learning techniques. These models are more robust as they often compliment the advantages of the individual techniques involved and improve the forecasting accuracy. A detailed discussion on hybrid methods and the advantages and disadvantages of the various combinations involved is discussed in the following sections.

2.9.2. Review of application studies on hybrid models

The review papers in this section pertain to different combinations of the machine learning techniques under hybrid models and are dealt separately.

2.9.2.1. ANN+Evolutionary algorithms. Azadeh et al. [148] presented an integrated GA and ANN approach to estimate and predict electricity demand using stochastic procedures. The data used is the electricity consumption of the Iranian agriculture sector from 1981 to 2005, of which 21 years data is used to estimate the model parameters and 4 years data is used to test the model. The independent variables of price, value added, number of customers and electricity consumption in the precious periods were obtained by ANN and used in GA to predict future electricity consumption. Three ANN models are developed and the results show that all the models have a MAPE of less than 0.132%.

El-naggar and Al-rumaih [149] presented a performance comparison of three estimation techniques for peak load forecasting in power systems. The three estimation techniques used are the GA, LS and least absolute value filtering (LAVF). The data used is the annual peak electricity load from the unified Egyptian network for a period of 1977–1993. The first twelve years in the dataset are used to evaluate the estimation process and the remaining is use to test the model. The results show that GA performed the best among the three estimation methods with maximum error of 2.75%.

Bashir and El-Hawary [150] employed particle swarm optimization (PSO) algorithm to improve an ANN model for short-term load forecasting. The hourly data and weather information for an area in the State of New York for the year 2001 and 2003 is taken for analysis. This paper introduces wavelet transformation technique to characterize electricity load. The results show that the MAPE for the testing dataset is 3.9695%.

Lee and Ko [151] developed radial basis neural network (RBFNN) with nonlinear time-varying evolution particle swarm optimization (NTVE-PSO) algorithm to predict hourly electricity load. The data is obtained from Taiwan Power Company for 2007. Three schemes were developed for the prediction model based on different seasons. The results show that the NTVE-PSO-RBFNN has the best forecasting accuracy.

Donate et al. [152] proposed two new approaches to improve automatic design of an ANN. In addition to the existing GA approach, this study proposes differential evolution algorithm (DE) and estimation of distribution algorithms (EDA) as means for automatic design of an ANN. Five different datasets were used to develop and test the estimation approaches. The results show that for over 200 generations, DE and EDA perform better than GA.

Xue et al. [153] proposes a real-valued genetic algorithm (RGA)-based neural network with support vector machine (NN-SVM) model

Table 3
Quantitative comparison for some of the best results of the 6 major time series forecasting techniques.

Model	Data type	Length of training data	Accuracy MAPE (%)	Time consumed	Ref.
ANN	Hourly electricity load	1 year	1.69–1.81	35 s	[44]
ARIMA	Monthly peak electricity demand	6 years	1.05–2.59	N.A.	[68]
SVM	Hourly cooling load and climate data	1 month	1.001–1.016	< 1 min	[79]
Fuzzy	Daily electricity load	6 months	1.23–1.63	N.A.	[101]
Grey	Half-hourly cooling load	4 months	0.416–1.097	N.A.	[114]
MA & ES	10–30-min electricity demand	7 months	1.2	N.A.	[132]
Hybrid	*Covered in detail in the next section*				

to predict the power load in both short-term and mid-term forecasting by using a radial-basis-function neural network (RBFNN), SVM and RGA. Historical load data from year 1996 to 2009 given half-hourly by RTE institution in France is used in the experiment. The data from year 1996 to 2008 is used for training, and year 2009 is for testing. Several experiments including one-week daily curve forecast, one-month peak load consumption forecast were performed and results of RGA-SVM, Kalman Filter and RBFNN were compared. The result showed that solely using RGA-SVM is not suitable for hourly forecasting for daily and weekly but is reliable for predicting next month consumption.

Kouhi et al. [154] presented a new forecasting method based on ANN and chaotic intelligent feature selection. The feature selection method selects the best set of candidate input which is used as inputs for the ANN. Hourly electricity for 40 days is used to train, validate and test the model. The results show the proposed method has an error of 1.84 for July 2006 data of New England market.

Defilippo et al. [155] studied the application of GA in selecting proper ANN architecture and training parameters for short-term load forecasting. The hourly electrical load (MWh) from a power company in Rio de Janeiro for a period of two years is used for the analysis. A total of 6 GA models were developed based on 2 cost functions. The cost function minimizes the autocovariances present in the error series. The median MAPE in all 6 GA runs is found to be 2.24% which is 0.2% below the median of the ANN model without GA.

Li et al. [156] presented an optimized ANN model using Particle Swarm Optimization algorithm (iPSO) for hourly prediction of building electricity consumption. In addition, principal component analysis (PCA) is used to select the significant modeling inputs. Two datasets were used to perform the analysis. The first dataset used is from the Great Building Energy Predictor Shootout I, organized by ASHRAE and the second dataset is from a library building located in Hangzhou, China. The results show that iPSO-ANN is the best for both the datasets and exhibited average MAPE values of 0.0166% and 0.0596% respectively.

2.9.2.2. ARIMA+ANN. Wang et al. [157] proposed an approach for short-term load forecasting by applying wavelet de-noising in a combined model of ARIMA and ANN. The wavelet transformation first categorizes an approximate part associated with low frequency and a detailed part associated with high frequencies. The hourly electrical load data of New South Wales (Australia) for 4 weeks was used to train and validate the model. The results show that the proposed wavelet de-noising based combined model (WDCM) exhibits a MAPE of 0.016% which is lower than the individual SARIMA, BPNN and the combined model without wavelet de-noising.

Zhuang et al. [158] presented a combined prediction method by combining ARIMA and ANN. The dataset is obtained by building simulation software, DeST. The load data from July 1 to July 31 is used to develop the model and the data of first week of August is used to test the model. The result shows that the combined model is superior to simple time series model and solves the problem of nonlinearity fitting.

2.9.2.3. ARIMA+Evolutionary algorithms. Yang et al. [159] proposed evolutionary programming approach to identify the ARMAX model for short-term load forecasting. Hourly data for four seasons along with temperature data for three Taiwan cities are taken for analysis. The results show that for all the four seasons, the proposed model had a maximum error of 2.27% and was superior when compared to traditional gradient search based approach.

Huang et al. [160] proposed a particle swarm optimization (PSO) approach to identify the ARMAX model for short-term load forecasting. Hourly data for four seasons along with temperature data for three Taiwan cities are taken for analysis. The results show that for all seasons, the proposed PSO based model has error lower than 2.55%.

Wang et al. [161] proposed a ARMAX model based on evolutionary algorithm and particle swarm optimization for short-term load forecasting. The data used for this study is real-time hourly load in summer for a period of 2 months in 2005. The weekends and weekdays are analyzed separately. For the 8 testing days, the proposed model has MAPE lower than 3.418%.

2.9.2.4. ARIMA+SVD+CONVEX HULL methods. Lu et al. [162] proposed a new method based on physical-statistical approach for building energy consumption forecasting. The physical model provides theoretical inputs underlying the energy flow mechanisms. Following this, stochastic parameters are introduced and a time series model is constructed which is then generalized based on convex hull technique. The data used for analysis is from 4 sports hall for a period of 5 years between 2009 and 2013. The results show that the RMSE for electricity consumption for all 4 sports hall was less than 10.15.

2.9.2.5. ARIMA+SVM. Nie et al. [163] proposed a hybrid method based on ARIMA and SVM to forecast short-term electricity load. The hourly load data is obtained from an electric power company in Heilongjiang, China for a period between March 1st and May 31st, 1999. The MAPE for the proposed hybrid method is 3.85% when compared to MAPE of 4.5% and 4% from ARIMA and SVM models respectively.

2.9.2.6. Combined Kernel-based SVRs. Che and Wang [164] proposed a kernel-based support vector regression combination model by using an individual model selection algorithm. Half-hourly data from New South Wales, Australia and California, USA are obtained for 1 week and 1 month respectively. For the NSW dataset, the MAPE is 2.13% and that for California is 2.37% respectively.

2.9.2.7. DEMD+SVR+AR. Fan et al. [165] presented a SVR model hybridized with differential empirical mode decomposition (DEMD) method and auto regression (AR) for electric load forecasting. The DEMD method is used to decompose the electric load into parts associated with high frequencies and parts with low frequencies. The data is collected for two cases: For Case 1, electric load is obtained from 2nd to 7th May 2007 for New South Wales, Australia. For Case 2, electric load from 1st to 12th January 2015 from New York Independent System Operator, USA is obtained for analysis. The results show that for Case 1, the MAPE is 9.71% for small sample size and 5.1% for large sample size. The result for Case 2 is 8.19% for small sample size and 5.37% for large sample size.

2.9.2.8. Dynamic model+Fuzzy. Lee and Hong [166] proposed a hybrid model to forecast electric power for several months ahead based on dynamic and fuzzy time series approach. Data from four sectors for Seoul between 2000 and 2011 is taken for analysis. The four sectors were – household, public, service and industrial. The results show that the MAPE for all the sectors is less than 5.1%.

2.9.2.9. Ensemble/Hybrid ANN. Pao [167] proposes two hybrid nonlinear models that combine a linear model with ANN. Several linear models are discussed and the seasonal general autoregressive conditional heteroscedasticity (SEGARCH) and Winters with volatility (WARCH) models are combined with ANN to predict the electricity and petroleum energy consumption of Taiwan. The data used extends from January 1993 to December 2005 for estimating the model parameters and the subsequent period, from January 2006 to December 2007 is used for testing. The results show that the MAPE for electricity consumption is 2.56% while that of petroleum energy consumption is 3.51%.

Li et al. [168] proposed a hybrid quantized Elman neural network (HQENN) with three quantized inputs for short-term load forecasting. This study also combines GA to obtain the optimal structure of the HQENN model. Hourly load for September 2008 for Chongqing was divided into three seasons and taken for analysis. The results show that for all the three seasons, the MAPE was less than 1.66%.

2.9.2.10. Fuzzy ANN+Bilevel optimization. Mao et al. [169] proposed a combination of self-organizing fuzzy neural network (SOFNN) learning method with a bilevel optimization method for short and mid-term load forecasting. Two datasets were examined. The first dataset is the 15-min load demand from January to July 2006 from Hebei Province, China and the second dataset is half-hourly load from 1997 to 1998 from EUNITE competition. For the first dataset, the results show that the MAPE is less than 3.02%. The best result obtained by the bilevel optimization model is a MAPE of 1.4006% for the EUNITE dataset.

2.9.2.11. Fuzzy neuro. Chang et al. [170] developed a weighted evolving fuzzy neural network (WEFuNN) for monthly electricity demand forecasting in Taiwan. This study adopts a weighted factor to calculate the importance of each factor among the different rules. Seven factors as specified by the Taiwan Power Company are taken as inputs to the WEFuNN model. The monthly data is collected for a period between January 1997 and December 2006. The results show that the proposed model has a MAPE of 6.43%, which is better than other approaches.

Chen and Wang [171] proposed a collaborative principal component analysis and fuzzy feedforward neural network (PCA-FFNN) for long-term load forecasting. The annual energy consumption from 1945 to 2008 are analyzed. The results show that proposed model has a MAPE of 0.5% for the testing dataset which is 12–91% better than the other approaches.

Chaturvedi et al. [172] proposed an integration of wavelet transform, adaptive GA and fuzzy system with Generalized Neural Network (GNN) for short-term weekday electrical load forecasting. One week's load data from a substation in Agra, India is used for analysis. The RMSE for the proposed method is 0.0486 kW.

2.9.2.12. Fuzzy+ARIMA/ARMAX. Yang and Huang [173] proposed a self-organizing model of fuzzy autoregressive moving average with exogenous input variables (FARMAX) for one day ahead hourly load forecasting. The hourly load data for 4 seasons in the year 1992 is collected from Taipower, Taiwan. The weather variables of temperature, humidity, wind speed and precipitation are also considered for the analysis. The results show that the MAPE for all the 4 seasons was less than 2.58% for both weekdays and weekends.

Azadeh et al. [174] presented an integrated fuzzy system, data mining and time series framework to predict the electricity demand. Monthly electricity consumption of Iran from 1995 to 2005 is considered for the study. The proposed model has a MAPE of 0.02% for the next twelve months electricity consumption.

2.9.2.13. Fuzzy+Evolutionary algorithms. Enayatifar et al. [175] proposed a hybrid algorithm based on a refined high-order weighted fuzzy algorithm and an imperialist competitive algorithm (RHWFTS-ICA) for short-term lead forecasting. Half-hourly load data of U.K. and France collected during 2003 and 2004 are used and divided into eight groups for the analysis. The results show that the MAPE for the proposed model is 0.79%.

Sadaei et al. [176] proposed an enhanced hybrid algorithm with sophisticated exponentially weighted fuzzy algorithm for short-term load forecasting. Half hourly load data from France and Britain for the year 2005 is used and divided into eight groups for the analysis. The results show that the MAPE values for the proposed model is 0.94%.

2.9.2.14. Fuzzy+SOM+SVM. Che et al. [177] presents an adaptive fuzzy combination model based on the self-organizing map (SOM), the SVR and fuzzy inference method. In the study, the proposed new algorithm is trained by electric load in New South Wales (NSW) from May 2, 2007 to May 7, 2007, and tested using electric load of May 8, 2007. The electric load data in NSW were collected on a half-hourly basis (48 data points per day) for 7 days. Two experiments are performed, in the first experiment, the electric load values (starting from 0:00 A.M. to 23:30 P.M.) of May 8, 2007 have been predicted using the proposed forecasting model and three comparison models. In the second experiment, one-week prediction from May 18th, 2007 to May 24th, 2007 was performed by these models. The MAPE of proposed AFCM is 9.9524% for the first experiment and 11.1019 for the second experiment which are smaller than MAPE of standard SVR, PSO-SVR and PSO-BP.

2.9.2.15. Fuzzy+SVM+Evolutionary algorithms. Son and Kim [178] proposed a model for one month ahead forecast of electricity demand in residential sector based on SVR and fuzzy-rough feature selection with particle swarm optimization. The monthly electricity load from January 1991 to December 2012 is used for the analysis. The first 240 months are used to train the model and the remaining 24 are used to test the model. The proposed model has a MAPE of 2.13% and is significantly less when compared to other models.

2.9.2.16. Grey+evolutionary algorithms. Li and Zhao [179] proposes a hybrid model based on Grey and improved decimal-code genetic algorithm (GA) with a one-point linearity arithmetical crossover, which can greatly improve the processing speed of crossover and mutation for short-term load forecasting. A daily load forecasting example is used to test the IGA model. Compared with GM(1,1) model, the proposed GM(1,1)-IGA model has better accuracy, the error of proposed GM(1,1)-IGA is smaller than 2%.

2.9.2.17. Grey+Fuzzy+MARKOV. Asrari et al. [180] presented a hybrid method called Grey-Fuzzy-Markov Chain Method (GFMCM) which comprised of three stages. In the first stage, daily load is forecasted by Grey model with its training deviations classified. In the second stage fuzzy-set theory was used to forecast and finally it was fed into the Markov chain model to predict future relative errors that might be supplied by the Grey model for Short-term load forecasting (STLF). Real electric load data collected from three electricity markets including Ontario and PJM electricity markets (EMs) for 2004 and Iran EM for 2009–2010. The predictions have been calculated for four separate weeks, each of which corresponding to one of the four seasons of the year. For these markets, the hourly measured load data of 20 days prior to the forecasting day have been used to build the forecasting model. Average WMAPE of the proposed method for Ontario, PJM and Iran markets are 2.915, 3.586 and 2.390, which are smaller than values of other compared MPLNN, wavelet-ARIMA, ARIMA and GM models.

2.9.2.18. KNN+neuro fuzzy. Wei et al. [181] presents a k-NN based neuro-fuzzy for time series prediction. The Poland Electricity Load data set is used in the experiment, this time series represent the daily load of Poland during around 1500 days in the 1990s. 1000 first values for training, and the rest, over 500 values, for testing. In the study, k is set to be 150 and 3 previous data are used to do prediction for a given time. The result shows that the proposed kNF method has the lowest MAE of 0.0364% and setting 'k' to be around 150 provides the best result.

2.9.2.19. KNN+SVM. Sorjamaa et al. [182] presented a least squares Support Vector Machines (LS-SVM) method combined with input selection criteria; k-nearest neighbor approximation method (*k*-NN), mutual information (MI) and nonparametric noise estimation (NNE) for long-term electricity load forecasting. Poland electricity load dataset is used in the study, and it represents two periods of the daily electricity load of Poland during around 1500 days in the 1990s. The first 1000 values are used for training, and the remaining data for testing. From the result, MSE of LS-SVM with kNN selected inputs is smaller than 0.006 and the kNN selection criterion is the fastest, because the selection of hyper parameters is not needed. This makes kNN around 10 times faster than MI and 20 times faster than NNE.

Chen and Lee [183] presents a weighted Least Squares Support Vector Machine (LS-SVM) combined with k-nearest neighbors and mutual information approach for time series forecasting. The Poland Electricity dataset records the electricity load of Poland, covering 1500 days in 1990s is used in the study for one-step forecasting. The first 1000 values are used for training, and the remaining 500 values are used for testing. From the result, the proposed method achieved the smallest MAE of 0.0245%. For multi-step forecasting, the EUNITE dataset which contains the electricity load demand every half an hour from January 1997 through December 1998 is used. MAPE of the proposed method is 1.71%.

2.9.2.20. Neuro fuzzy. Azadeh et al. [184] presented a hybrid adaptive neural network based fuzzy inference system (ANFIS) for electricity consumption forecasting. This was the first study of its kind with a hybrid simulation-adaptive network fuzzy inference system (ANFIS) for improvement of electricity consumption estimation. The unique features of the proposed algorithm are two folds. Firstly, ANFIS is ideal for complex and uncertain data because it is composed of both ANN and fuzzy systems. Secondly, Monte Carlo simulation is used to generate input variables whereas the conventional methods use deterministic data. The proposed algorithm is applied to 130 set of data which are the monthly consumption in Iran from April 1994 to February 2005. MAPE of monthly electrical energy consumption prediction (2005–2006) is 0.0155% for the proposed method.

Akdemir and Cetinkaya [185] studies an adaptive neural fuzzy inference system for long term load forecasting. In this study, ANFIS is used to forecast long term energy demands up to 2025. The data includes population, reached maximum power demand in an hour and total energy level given in giga-watt hours (GW h). The estimations for energy demands for future years were implemented using the projections of official distributing energy service called Turkey Electric Transmission Company (TEIAS). The results displayed a total average MAPE of ANFIS as 0.82439%.

Hooshmand et al. [186] proposed a new two-step algorithm for short-term load forecasting (STLF). In the first step, a wavelet transform (WT) and an ANN were used for the primary forecasting of the load over the next 24 h. In the second step, a WT, the similar-hour method and ANFIS were used to improve the results of primary load forecasting. Iran's electricity load and electricity load from New South Wales State of Australian were used in the study. The results showed a MAPE for the proposed method as 1.603%, which has a significant reduction in comparison with the other 4 methods.

2.9.2.21. Random forest. Grzegorz Dudek [187] proposed a random forest model for short-term electricity load forecasting. This is an ensemble learning method that generates many regression trees (CART) aggregating their results. The RF forecasting model is characterized by simplicity. The number of parameters to be estimated is small, which implies a simple procedure of the model optimization. Hourly electrical load data of the Polish power system for the period between 2002 and 2004 is used in this study to forecast the load curve for the next day. Alternative models such as CART, ARIMA, exponential smoothing and neural networks are compared with RF. From the result, the proposed RF model provided as accurate forecasts as ANN and outperformed the CART, ARIMA and ES models. MAPE of the proposed method is lower than 1.42%.

2.9.2.22. SEAM+the regression model. Wu et al. [188] proposes a hybrid models by combining seasonal exponential adjustment method (SEAM) with the regression methods where SEAM and the regression models are employed to seasonal and trend items forecasting respectively for 1-week-ahead daily load forecasting. Data used in this study are sampled in Victoria State grid in Australia at intervals of thirty minutes. This led to a total number of load demand in one day as 48. The load demand on May 1, May 8 and May 15 are extracted to forecast the one on May 22. Eleven regression models are developed and compared. From the results, MAPE of the proposed method is lower than 7.77% and all of the combined models work better than the single ones for values of MAPE.

2.9.2.23. SOM+multi SVMs. Cao [189] presents the use of Support Vector Machines (SVMs) experts for time series forecasting. The generalized SVMs experts have two-stage neural network architecture. In the first stage, self-organizing feature map (SOM) is used to partition the whole input space into several disjointed regions. Building data sets taken from the contest of "The Great Energy Predictor Shootout-the First Data Analysis and Prediction" are evaluated in the experiment. Data set 1 contained time record of whole building electricity, hourly chilled water and hot water usage for a four-month period in an institutional building. It is sequentially divided into two parts of 2326 data patterns that are used for training SVMs, and the remaining 600 data patterns that are used as the validation set. From the result, coefficient of variation (CV) of proposed method is 16.35 for hourly whole building electricity forecasting. Comparing the results of the SVM expert's models with those of the best single SVM model, it can be observed that the SVMs experts achieve a smaller CV value than the best single SVM model.

2.9.2.24. SVM+evolutionary algorithms. Pai and Hong [190] proposed a recurrent Support Vector Machines with genetic algorithms (RSVMG) to forecast electricity load. Taiwan regional electricity load values from 1981 to 2000 serve as experimental data to show the forecasting performances of RSVMG models and compared with ANN and regression models. The data are divided into three data sets: the training data set from ranging from 1981 to 1992, the validation data set ranging from 1993 to 1996 and the testing data set ranging from 1997 to 2000. Results show that the proposed RSVMG model has smaller MAPE values than SVMG, ANN and regression models. The MAPE of RSVMG model is lower than 1.8955% for Northern, Central, Southern and Eastern regions.

Hong [191] presented a SVR model with a hybrid evolutionary algorithm (chaotic genetic algorithm, CGA) to forecast electric loads. This study uses Taiwan regional electric load data to compare the forecasting performances of SVRCGA models with ANN and regression models. Totally, there are 20 data points for electric load ranging from 1981 to 2000 for a region in Taiwan. The data are divided into three data sets: the training data set ranging from 1981 to 1992, validation data set ranging from 1993 to 1996, and the testing data set ranging from 1997 to 2000. The MAPE of the proposed SVRCGA model is lower than 2.57% for Northern, Central, Southern and Eastern regions.

Hong [192] presented a new hybrid method based on SVR and chaotic particle swarm optimization (CPSO) for electric load forecasting. For regional load forecasting, Taiwan regional electric load values from 1981 to 2000 serve as experimental data which is divided into the same three subsets: the training data set ranging from 1981 to 1992, the validation data set ranging from 1993 to 1996 and the testing data set ranging from 1997 to 2000. For annual load forecasting, Taiwanese annual electric load from 1945 to 2003 serve as experimental data which is divided into three data sets: the training data set ranging from 1945 to 1984, the validation data set ranging from 1985 to 1994 and the testing data set ranging from 1995 to 2003. MAPE of SVRCPSO is 2.1860% for regional load forecasting and 1.6134% for annual load forecasting. Results prove that SVRCPSO model yields improved forecast results and outperforms the other compared forecasting models.

Li et al. [193] proposed a new hourly building cooling load prediction method which combines SVM and Stimulated Annealing Particle Swarm Optimization (SAPSO) which has the advantages of both the PSO algorithm and the SA algorithm. The historical data for Zhongkai University scientific building is used in the study. The cooling load data from 0:00 on 4/1/2006 to 12:00 10/31/2006 are selected as training sample and the cooling load data from 0:00 on 4/1/2006 to 12:00 8/31/2006 is selected as testing sample. Result shows that SAPSO based SVR has a higher accuracy than the traditional SVR, root mean squared error of the proposed method is 3.217%.

Niu et al. [194] proposed a new feature selection mechanism based on ant colony optimization combined with SVM for short-term hourly load forecasting. Data are chosen from the database of Inner Mongolia region. In the study, the power load data from 0:00 on 5/1/2004 to 12:00 on 3/31/2006 are selected as training sample to establish the single-variable time series. And the power load data from 13:00 on 3/31/2006 to 24:00 on 5/28/2006 as testing sample. RMSRE of proposed ACO–SVM is 1.50%. Comparing with the ANN and SVM method, the proposed ACO-SVM model has higher forecast precision.

Wang et al. [195] proposed a Hybrid Particle Swarm Optimization (HPSO) with genetic algorithm (GA) mutation to optimize the SVM forecasting model for hourly power load forecasting. Power loading data in Baoding region is used to prove the effectiveness of the model, the power load data from 0:00 at 06/10/2006 to 12:00 at 6/9/2007 are as training sample and used to establish the single-variable time series. And the power load data from 13:00 at 10/06/2007 to 24:00 at 25/06/2007 as testing sample, and the power load data from 13:00 at 27/06/2007 to 24:00 at 27/06/2007 as forecasting points. The results show that the model is effective in the forecasting of short-term power load than the other models. The root-mean-square relative error (RMSRE) of the new model was found to be 1.82%. On the other hand, the single SVM model and BP network's accuracy was found to be 2.43% and 4.10% respectively.

Hong [196] presents an electric load forecasting model combining the seasonal recurrent support vector regression model with chaotic artificial bee colony algorithm (namely SRSVRCABC). The proposed SRSVRCABC employs the chaotic behavior of honey bees which is has a better performance in function optimization to overcome premature local optimum. The study uses historical monthly electric load data of Northeast China to evaluate the forecasting performance of the proposed SRSVRCABC, the employed data are divided into three data sets, the training data set ranging from December 2004 to July 2007, the validation data set ranging from August 2007 to September 2008, and the testing data set ranging from October 2008 to April 2009. MAPE of the proposed SRSVRCABC model is 2.387% which has obtained significant smaller MAPE values than other alternative models (ARIMA (1, 1, 1), TF-ε-SVR-SA, and SSVRCABC models).

Zhang et al. [197] presented a novel hybrid algorithm to improve the accuracy of monthly load forecasting. It is known as the chaotic genetic algorithm simulated annealing algorithm (CGASA) with an SVR model. This study employs 53 months data historical electric load data (from December 2004 to April 2009) of Northeast China region to compare the forecasting performances for the proposed SSVRCGASA model with other models (ARIMA and TF-ε-SVR-SA models). The data is divided into three sub-sets in the following way. The training set containing the preceding 32 months, the validation set with the middle 14 months and the testing set containing the last 7 months. From the result, MAPE of the proposed SSVRCGASA model is 1.901% which is smaller than other models.

Ju and Hong [198] presented a SVR-based electricity forecasting model which applied chaotic gravitational search algorithm (CGSA) to improve the performance of forecasting. This study employs 53 months data historical electric load data (from December 2004 to April 2009) of Northeast China to compare the forecasting performances among the proposed SSVRCGSA model and other models (ARIMA and TF-ε-SVR-SA models). The employed data are divided into three data sets with the training data set containing 32 month from December 2004 to July 2007, the validation data set with 14 months from August 2007 to September 2008 and the testing data set containing 7 months from October 2008 to April 2009. The MAPE of the proposed method for monthly load forecasting is 2.587% showing that the proposed SSVRCGSA model is superior to other alternatives.

Hong et al. [199] presented a SVR-based electric load forecasting model which applied chaotic genetic algorithm (CGA) to improve the forecasting performance. This study employs 53 months of historical electric load data (December 2004–April 2009) of Northeast China region to compare the forecasting performances among the proposed SSVRCGA model and other models (ARIMA and TF-ε-SVR-SA models). The employed data are divided into three data sets with the training data set containing 32 months from December 2004 to July 2007, the validation data set with 14 months from August 2007 to September 2008 and the testing data set of 7 months from October 2008 to April 2009. The MAPE of the proposed SSVRCGA model is 2.695% which is smaller than other alternative models.

Kavousi-Fard et al. [200] proposed a hybrid prediction algorithm comprised of Support Vector Regression (SVR) and Modified Firefly Algorithm (MFA) to forecast the short term electrical load. The proposed approach is applied to the electrical load published by Fars Electrical Power Company, Iran. The diurnal data was collected for five different cities with different electric load patterns from 21st March 2007 to 20th February 2010. The testing data set comprised of electrical load measured between 21st January 2010 and 20th February 2010 for each city. The rest of the data was used for to train the model. The MAPE of the proposed SVR-MFA method was found to be 1.6909%. When compared with the MAPE obtained from the ARMA model, ANN, SVR-GA, SVR-HBMO, SVR-PSO and SVR-FA, the experimental results showed that the proposed algorithm outperforms other techniques.

Chen et al. [201] proposed a new combined forecasting method (ESPLSSVM) based on empirical mode decomposition, seasonal adjustment, particle swarm optimization (PSO) and least squares support vector machine (LSSVM) model for electric load forecasting. The data used in this paper is the electric load data of South Australia (SA), the data period covered from 2nd May 2011 to 3rd July 2011 (9 weeks in all). Among these data, the electric load data from 2nd May 2011 to 26th June 2011 were used to train the model, and data from 27th June 2011 to 3rd July 2011 were used to test the models. The MAPE for the

ESPLSSVM model is smaller than 3.05% for daily load forecasting and 1.79% for whole week load forecasting. Results show ESPLSSVM performed better than basic LSSVM (LSSVM), empirical mode decomposition-based signal filtering method processed by LSSVM (ELSSVM) and seasonal adjustment processed by LSSVM (SLSSVM) load forecasting approaches.

Baziar and Kavousi-fard [202] proposed a new hybrid method based on support vector regression (SVR) combined with Krill Herd (KH) Algorithm for short term load forecasting. The load data for a 400 kV substation of Fars Regional Company in 2009 were used in this study. The input data set is the diurnal peak load of eight weeks earlier the forecasting day. The test data is the last month of the data sample. From the results, MAPE of the proposed method was found to be 1.27046%, which is lower compared with ARMA, ANN, SVR and SVR-PSO model.

2.9.2.25. Wavelet+ANN. Ghofrani et al. [203] presented a hybrid WT-BNN (Wavelet Transform-Neural Network) load forecasting framework. A new input selection method which combines correlation analysis and ℓ^2-norm is developed to determine the input sub-series of the BNNs which are most correlated to the output series. New England hourly load data are utilized to evaluate the performance of the proposed method; the data include the hourly historical load for a 5-year period between 2008 and 2013. 80% of the load data and the remaining 20% is used for the test. From the result, MAPE for the proposed method is 0.438 which outperforms the ANN, SIWNN and SSA-SVR methods by 79.5%, 75.7% and 69.6%, respectively.

Kouhi and Keynia [204] proposed a new hybrid short-term load forecasting method which consists of wavelet transform, an intelligent two-stage feature selection, and cascaded neural network. The proposed STLF has been tested for load forecast of day ahead electricity market of PJM, New York and New South Wales. First experiment of proposed STLF was performed on PJM electricity market. For PJM electricity market, two kind of forecasting has been presented: 1-h ahead (hourly) and 1-day ahead (daily) forecasting. Average MAPE of the proposed method for hourly and daily forecasting is 1.3014%, 1.2348%. For New York electricity market, MAPE of the proposed method is lower than 3.5604% for hourly forecast and 1.3093% for daily forecast. For New South Wales load data, average MAPE of the proposed method for daily load forecast 1.44%.

2.9.2.26. Wavelet+ANN+evolutionary algorithms. Moazzami et al. [205] presented a new hybrid method based on Wavelet decomposition, neural network and GA for day-ahead peak load prediction in Iran National Grid (ING). In this study ING peak load data in time horizon of February 4, 2006–July 22, 2011 has been used. Also the weather data in the same period for three large cities of Iran, Tehran, Tabriz and Ahvaz with mild, cold and warm climates respectively were used in this study. The proposed method with GFF ANN model has generated the lowest MAPE of 1.2%. For the EUNITE test case, MAPE of the proposed method is 1.076%.

2.9.2.27. Wavelet+ANN+Grey. Meng et al. [206] presented a new method which combines discrete wavelet transform (DWT), neural network and Grey model for forecasting monthly electric energy consumption. The data used in the study was obtained from the website of the Chinese Economic and Financial Database of the China Center for Economic Research (CCER). The actual values of the monthly electric energy consumption (108 kW h) in China from January 1990 to December 2006 were selected to validate the proposed methods. The result shows that the MAPE of proposed method is 2.67% which outperforms other methods.

2.9.2.28. Wavelet+exponential smoothing+weighted nearest neighbor. Sudheer and Suseelatha [207] proposed a hybrid method based on wavelet transform, Triple Exponential Smoothing (TES) model and weighted nearest neighbor (WNN) model for forecasting the 24 h ahead electricity load. The hourly load data of California and Spain energy markets are used in the study. The electrical load for every hour of the year 2000 is considered for California market and of the year 2002 for the Spain market. The daily forecast errors (MAPE) for the analyzed weeks in Spain energy market is less than 2.7681%. The daily forecast errors (MAPE) for the analyzed weeks in California energy market is less than 1.8400%. Weekly forecast errors (MAPE) for the analyzed weeks of California energy market is 1.0967%, weekly forecast errors (MAPE) for the analyzed weeks of Spain energy market is 1.2497%.

2.9.2.29. Wavelet+Grey+evolutionary algorithms. Bahrami [208] proposed a new model, WGMIPSO based on combination of the WT (wavelet transform) and GM (grey model) for STLF (short term electric load forecasting) and is improved by PSO (particle swarm optimization) algorithm. Iran and New York network load of load data is used in the study. The MAPE of daily and monthly load forecast for New York network data are 0.6826% and 1.8205%. The WME (weekly mean error) and DME (daily mean error) for Iran data are 1.02% and 1.07%. From the result, the proposed method has better performance in all of the understudied cases in comparison with other understudied methods in the paper.

3. Discussion

The above-mentioned time series prediction techniques have been successfully applied to building energy consumption forecasting. Each technique possesses certain advantageous characteristics which should be suitably applied to the case. This section throws light on the advantages and disadvantages of each for the techniques discussed previously.

ANN has more advantages than statistical models, as it can map the input and output relationship without making complex dependency among the inputs. It is seen that ANN provides much better performance as compared to previous implemented techniques for non-linear mapping. ANN can perform nonlinear modeling without any prior knowledge about the relationships between the input and output variables. Therefore, these are more general and flexible modeling technique for forecasting. However, ANN is dependent on initialization of weight values, exhibits local minima and slow convergence problem. In addition, it is always a challenge to obtain a balance between overfitting and generalization for ANNs. On the other hand, ARIMA is a universal approximator with enough elements regressed and averaged such that the approximation can be made to fit any time series. Yet, ARIMA identification is complex and time consuming and many ARIMA models have no structural interpretation. Identification and estimation can be badly distorted by outlier effects. For building energy forecasting problems, SVM can deal with issues such as small sample, nonlinear, high dimension and local minimum points. Usually, the long-term data has the characteristics of small samples and this makes SVM suitable for forecasting long-term data. In addition, the SVM regression method has good capability of fitting and generalization just like the ANN. A special strength is the use of a kernel function to introduce nonlinearity and to deal with arbitrarily structured data. However, just like the other two above-mentioned techniques, SVM lacks transparency of results and cannot be interpreted easily. Usually the kernel function depends on certain parameters, which have to be optimized to achieve good results. This optimization is usually done using evolutionary algorithms like GA, PSO, DE etc. However, finding good parameters can become a computational challenge as the number of parameters and size of the dataset increases.

As for the CBR technique, the main advantage is that it can be applied to almost any domain. The CBR system does not try to find rules between the parameters of the problem, rather it just tries to find similar problems in the data and to use solutions of these problems as a solution to the case under study. The second advantage is that the approach of CBR to learning and problem solving is very like the human cognitive processes. CBR recognizes the connection between similar events in the past and future. However, CBR is rarely used for time series forecasting. The reason for this is that the use of CBR for time series processing introduces new aspects for which the sequences can be very long and with different lengths. Moreover, the huge amount of data and the presence of noise in these data associated to real time constraints makes the CBR system unnecessary. In contrast, fuzzy prediction techniques are very good at solving uncertainties in load forecast. However, the temporal patterns are defined by rigid regions which are hard to adjust when there is noise in the dataset. It often acquires high computational complexity and lacks stability. For fuzzy prediction technique, the time series should be converted to stationary and periodic series in order to derive patterns in the time series.

Grey prediction theory deals with systems with lack of available information and can find solutions to problems with small samples. The main purpose of the theory is to predict the behavior of systems which cannot be detected with stochastic or fuzzy methods with limited data. They entail easy calculations and are well applied to short-term load predictions. However, this model exhibits the problem of recognizing the random or noise component. Moving average and exponential smoothing provide simplicity in calculations, and works well even with a low number of observations. It considers the time series to be stationary locally with a fairly stable value of the mean. The nearest neighbor technique does not explicitly require training and is simple and intuitive with an ease of implementation. However, it is often challenging to compute the precise number of 'nearest neighbor' value which is a key parameter in this method. A large number will lead to a smoother fit, and a lower variance, and vice versa for a small number. A summary of the advantages and disadvantages of the techniques discussed here is presented in Table 4. The characteristics of the hybrid prediction technique is dealt in detail in the following section.

3.1. Special note on hybrid models

Hybrid models are apt to deal with real-world problems which are often complex in nature. A single machine learning model may not be able to capture the complexities in building energy and operational data. In such cases, utilization of the hybrid model can be of benefit. Hybrid methods are robust analysis tools for a large class of complex problems not amenable to traditional classical methods. By combining different methods, complex autocorrelation structures in the data can be modeled more accurately. In addition, the problem of model selection can be eased with a little extra effort. For example, one of the most important and widely used time series models is the ANNARIMA model. As a hybrid of ARIMA and ANN, this model takes advantage of the unique strength of ANN and ARIMA in nonlinear and linear modeling respectively. The benefits of such methods appear to be substantial when dealing with non-stationary series. The non-stationary nonlinear component can be modeled using the ANN model and the stationary, residual linear component can be modeled by ARIMA model. In a similar way, other methods can be combined to improve the forecasting accuracy. This paper has reviewed 28 different combinations of hybrid models for time series forecasting and a summary of the novelty of each model is presented in Table 5.

Hybrid methods have also been used effectively in non-time series building optimization. Multi objective optimization involving building simulation, GA (NSGA-II) and ANN was performed by Magnier and Haghighat [209]. The main objective of such optimization techniques is to reduce the computational time. The authors used a Response Surface Approximation Model (RSA) to learn the behavior of the base building model and then use this for the GA for evaluating the individuals. The RSA method used was the multilayer feed-forward ANN. Although the computation time is reduced for the GA optimization, a considerable time is required to create base cases for running the building simulations to train the ANN model. A similar combination of GA and ANN for building optimization was performed by Petri et al. [6]. Here also, ANN was used to predict the simulation results which then were used for optimization. However, the time taken for running the simulations to train the ANN is not mentioned. The ANN model presents the advantage of being fast and adaptive for use in conjunction with a given Building Management System (BMS). The ANN based optimization module can function independently for solving an optimization process

Table 4
Summary of qualitative comparison for the 9 major time series forecasting techniques.

Model	Advantages	Disadvantages
ANN	1. Ability to precisely map input and output relationships 2. Performance well for non-linear time series 3. More general and flexible	1. Depends on initialization of weight values 2. Problem of the local minima 3. Overfitting and difficult to generalize
ARIMA	1. Uses lag and shift of historical data 2. Regression model with a moving average (improves accuracy) 3. Provides confidence intervals on predictions with reliability	1. Model identification is difficult 2. Not suitable for long-term prediction 3. Does not fully capture the non-linear patterns of the series
SVM	1. Good for fitting and generalization 2. Performs well for long-term time series 3. Use of a kernel function introduces nonlinearity and deals with arbitrarily structured data	1. Lack of transparency of results 2. Finding optimum parameters can be a computational burden as number of parameters and size of dataset increases
CBR	1. Similar to human cognitive processes 2. Doesn't need to find rules between parameters of the problem	1. Needs introduction of new aspects, e.g. case representation for time series processing 2. Needs huge data
Fuzzy	1. Close to human experience via membership functions and rules. 2. Good for solving uncertainties in load forecasting	1. Temporal patterns are defined by rigid regions, hard to adjust with noise 2. High computational complexity and lacks stability
Grey	1. Capable of predicting with limited data and incomplete information 2. Easy to compute and calculate	1. Inadequate in recognizing random component 2. Problem with conventional approach of model validation
MA & ES	1. Simplicity in calculations 2. The use of low number of observations 3. Transparency in approach	1. Poor results compared to sophisticated techniques 2. Not suitable for long-term and non-linear prediction
NN	1. Simple process with no explicit training step required 2. Intuitive and ease of implementation	1. Function is often approximated only locally 2. Challenging to compute exact number of nearest neighbors
Hybrid	1. Complimentary combination of different machine learning methods 2. Robust for complex problems and often improves performance	1. High model complexity 2. Computational intensive 3. Often difficult to identify which methods to combine

Table 5
Summary of different combination of methods pertaining to hybrid models.

Combination of methods	Novelty
ANN+Evolutionary Algorithms	Neural network is very likely to get stuck in the local optima. The evolutionary algorithms have the ability for global optima searching with fast convergence.
ARIMA+ANN	The periodicity and linearity of the load is described by the ARIMA model, and the residual error is described by ANN that fits the nonlinear characteristic.
ARIMA+SVMs	Neither SVM nor ANN can forecast the linear part of the load accurately. The proposed hybrid method firstly uses ARIMA to forecast the load, and then uses SVMs to correct the deviation.
ARIMA+Evolutionary Algorithms	Evolutionary algorithms like genetic programming (GP) in linear or nonlinear modeling can obtain the residual. The periodicity and linearity is described by the ARIMA model.
ARIMA+SVD+Convex Hull Methods	The physical model provides mechanism of energy flows. Convex hull technique is derived to parameterize the individual level model parameters and improve accuracy.
Combined Kernel – based SVRs	The ideal kernel is hard to determine for a regression task. The combination selection algorithm selects the optimal subset of individual support vector regression (SVR) kernel models using the goodness measurement.
DEMD+SVR+AR	Hybridization of support vector regression (SVR) with auto regression (AR) and different empirical mode decomposition (DEMD) method. DEMD is used to decompose the data into several detail parts with high frequencies and an approximate part associated with low frequencies.
Dynamic model+Fuzzy	This hybrid model combines dynamic and fuzzy time series approaches and allows for distinct responses from different individual sectors.
Ensemble/Hybrid ANN	Ensemble ANN is capable of nonlinear mapping, learning complex dynamics of the time series and extracting more information than a single ANN. Generally, ensemble ANN improves forecasting accuracy.
Fuzzy ANN+Bilevel Optimization	Combination of a self-organizing fuzzy neural network (SOFNN) learning method can automatically determine both the model structure and parameters with a bilevel optimization method. This automatically selects the best pre-training parameters to identify the best fuzzy neural networks.
Fuzzy neuro	Parameters of ANN (e.g, connection weights) are modified by fuzzy rules which are developed from experience or learned from data to improve the model performance. The model is more robust in solving uncertainty.
Fuzzy+ARIMA/ARMAX	Overcomes the slow convergence problem of hybrid fuzzy models. Fuzzy logic approach utilizes the expertise and experience to select a linear model to approximate the inherently non-linear dynamic loads. Autocorrelation function (ACF) is used to select input variables for fuzzy model.
Fuzzy+Evolutionary Algorithms	Evolutionary algorithms are employed to optimize the parameters of fuzzy method.
Fuzzy+SVM+Evolutionary Algorithms	Based on SVR formulation, fuzzy feature selection method is adopted to improve the forecasting capability of trained SVR model by searching the optimal variable subset. Evolutionary algorithms utilize fuzzy dependency degree as an evaluation measure.
Grey+Evolutionary algorithms	Traditional Grey forecasting model is not accurate and the value of its parameter ‘ a ’ is constant. Evolutionary Algorithms search the optimal ‘ a ’ value, thereby enhancing the accuracy of forecasting.
Grey+Markov	The forecasting precision of Grey forecasting model for data sequence with large random fluctuations is low. Markov-chain forecasting model can be used to forecast a system with randomly varying time series and can take care of the large fluctuations.
KNN+Neuro fuzzy	Instead of using all historical data to train a global model, ‘k’ most similar data is chosen as the training set to create a local model. The training set is more relevant to the prediction each time than the global model.
KNN+SVM	The proposed model utilizes k-nearest neighbors based selection criteria to determine optimal set of inputs combined with SVM which avoids local minima problems and achieve high accuracy.
Neuro fuzzy	Combination of ANN and fuzzy logic enhances the ability to automatically learn and adapt. The other two characteristics of this hybrid method are quick convergence and high accuracy.
Random forest	This is an ensemble learning method that generates many regression trees (CART) and aggregates their results. The RF forecasting model is characterized by simplicity, while the number of parameters is small.
SEAM+Regression	Treats the trend and seasonal items as two separated forecast processes. By the use of seasonal exponential adjustment method (SEAM), seasonal item in the original load is eliminated, and then regression model is used to forecast the trend item.
SOM+Multi SVMs	Combination of SVMs with SOM using a two-stage architecture. Different input regions are separately learned by the most appropriate SVMs experts. The time complexity of training SVMs is reduced.
SVM+Evolutionary algorithms	SVR models utilize nonlinear mapping feature to deal with nonlinear regressions but lack in a methodical algorithm for obtaining optimal model parameters. Evolutionary algorithms can obtain SVR parameters accurately and effectively.
Wavelet+ANN	The cascaded ANN (CANN) structure can extract input/output mapping function of the nonlinear data more efficiently. Combined with wavelet transform (WT) the time series can be decomposed into several sub-series with more detailed periodic information, which are easier to predict.
Wavelet+ANN+Evolutionary algorithms	In addition to the advantages of wavelet + ANN, evolutionary algorithms adjust parameters like input selection, step size, momentum values and the number of neurons in hidden layers of ANN.
Wavelet+ANN+Grey	Adopts DWT to extract features of data into several simple series. Grey model is then selected to mitigate the stochastic effect on the primary trend. GM, DWT and ANN together overcome difficulties of excessive information.
Wavelet+Grey+Evolutionary algorithms	Only few sample data is required for Grey models but parameters of the Grey model are most likely to be constant. Evolutionary algorithms determines the optimal parameters. Wavelet transform filters the data and reduces irrelevant information.
Wavelet+Exponential smoothing+Weighted Nearest Neighbor	Wavelet transform acts as a preprocessor to decompose the original load series into deterministic and fluctuation parts. The seasonal and trend factors in the deterministic component are modeled using exponential smoothing method and the faster dynamics of data in the fluctuation component are modeled using the WNN technique.

or can be embedded within the GA based optimization module. ANN can also be calibrated for real cases and used as a cost function in optimization programs to achieve energy saving targets [210].

4. Conclusions

This study presents a comprehensive review on 9 time series forecasting techniques for building energy consumption. These techniques have been widely used in other fields like economic forecasting, quality control, stock market, weather forecasting etc. Recently, these techniques have been experimented and applied to building energy consumption data. Such forecasting models are of immense importance for real-time energy monitoring and efficient building operation and optimization. These models can also be used to detect faults over time in building systems, provide future consumption scenarios and also form a basis for analyzing the energy consumption in relation to other building variables like occupancy scheduling.

Several successful machine learning models have been developed using past recorded energy data for short, medium and long-term forecasting. It is observed that each of the described techniques possesses a set of advantages and disadvantages. These have been analyzed and presented in detail with respect to the analysis of building energy data. Special emphasis is given to the 'Hybrid' model, which is a combination of two or more machine learning techniques in a way that each model compliments the strength of the other. For example, a hybrid model that considers ARIMA and evolutionary algorithms (EA) can take advantage of the ARIMA model to determine the periodicity and linearity, whereas the EA can efficiently determine the residuals. Various combinations of the hybrid model and their novelty are identified in the literature and systematically presented in this paper. It is seen that the combination of time series forecasting techniques like ANN, ARIMA combine very well with optimization techniques like GA, PSO etc. Such combinations have been widely explored in research dedicated to building optimization. The growing trend in research in building energy efficiency is expected to continue in the light of global sustainability drive. This makes real-time energy data monitoring and forecasting relevant and vital in this field. This paper provides a comprehensive summary of the existing forecasting techniques along with combinations of the hybrid model and paves way for future research in the field of building energy consumption.

4.1. Cost implications and future scope

The domain of building optimization is based on an extensive data collection, monitoring, forecasting, optimization, and controls network. All such infrastructure invariably adds to the total operating cost of the building. The challenge here is to study the added benefits in terms of cost of investment and cost regained due to energy savings arising from building optimization. There are few studies that highlight the financial part for building performance control and optimization. This limitation in cost data sharing is either because of the market forces involved or due to the confidentiality of the nature of data involved. Labeodan et al. discussed the applications of a low-cost Wireless Sensors and Actuators Network for occupancy modeling and lighting control in an office building [211]. The total cost of system was approximately € 2575 for 12 workstations that included wireless motion sensors and chair sensors. The results show an average reduction of 24% in lighting energy usage for a two-week period with the given cost of implementation to be about € 215 per workstation. The authors noted that the higher initial cost and lack of awareness are contributory factors for the reduced pace of deployment of sensors. However, the energy saving obtained, ease of deployment and improved environmental sensing demonstrate this as a viable solution for achieving improved building performance. Kumar et al. also realized that the cost of IAQ monitoring sensors does not meet the requirements for large scale deployment for control and automation [212]. Lilis et al. mention that the interest in modern Internet of Things (IoT) based solutions for building optimization is inhibited due to lack of estimation of the cost benefits [213]. Chen et al. noted that the cost related to energy consumption of a random set of buildings in China with Building Automation (BA) systems is almost twice of that of buildings without the BA system [214]. This is due to the sensor faults and control strategy flaws that result in a significant increase in final energy consumption. The authors note that the capital investment of intelligent systems for a public building in China is about 100–300 RMB/m^2 (US$ 14.4–43.2/m^2). The intelligent systems refer to the functional systems of air-conditioning, building automation, security, fire control, etc. The authors also calculated a reduction of 6% in energy savings with the intelligent systems installed for a case study building in Beijing.

A few review studies have covered the wireless sensor network enabled building energy management system for smart home/building applications [215,216]. Since the model predictive control technology is still in a developmental stage and requires heavy optimization based on the building type and function, the implementation currently focusses more on test-bedding and validation. At the same time, the technology is not yet available at the required cost of deployment at large scales. These challenges are leading to the slow penetration of building automation and optimization at large. The scope of this review paper, however, is limited to the study of time series forecasting techniques for building energy consumption which form an integral part of the building optimization and control process. A detailed review on the related costs of sensor deployment and monitoring challenges shall be covered in future studies.

References

[1] IEA. Energy and climate change; 2015. ⟨http://doi.org/10.1038/479267b⟩.
[2] Al-Homoud MS. Computer-aided building energy analysis techniques. Build Environ 2001;36:421–33.
[3] Crawley DB, Hand JW, Kummert M, Griffith BT. Contrasting the capabilities of building energy performance simulation programs. Build Environ 2008;43:661–73.
[4] Yang J, Santamouris M, Lee SE, Deb C. Energy performance model development and occupancy number identification of institutional buildings. Energy Build 2015. http://dx.doi.org/10.1016/j.enbuild.2015.12.018.
[5] 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. http://dx.doi.org/10.1016/j.enbuild.2014.02.053.
[6] Petri I, Li H, Rezgui Y, Chunfeng Y, Yuce B, Jayan B. A modular optimisation model for reducing energy consumption in large scale building facilities. Renew Sustain Energy Rev 2014;38:990–1002. http://dx.doi.org/10.1016/j.rser.2014.07.044.
[7] Aste N, Manfren M, Marenzi G. Building automation and control systems and performance optimization: a framework for analysis. Renew Sustain Energy Rev 2016. http://dx.doi.org/10.1016/j.rser.2016.10.072.
[8] Jagarajan R, Abdullah Mohd , Asmoni MN, Mohammed AH, Jaafar MN, Lee Yim Mei J, Baba M. Green retrofitting – a review of current status, implementations and challenges. Renew Sustain Energy Rev 2017;67:1360–8. http://dx.doi.org/10.1016/j.rser.2016.09.091.
[9] Zong Y, Böning GM, Santos RM, You S, Hu J, Han X. Challenges of implementing economic model predictive control strategy for buildings interacting with smart energy systems. Appl Therm Eng 2016. http://dx.doi.org/10.1016/j.applthermaleng.2016.11.141.
[10] Hu S, Corry E, Curry E, Turner WJN, O'Donnell J. Building performance optimisation: a hybrid architecture for the integration of contextual information and time-series data. Autom Constr 2016;70:51–61. http://dx.doi.org/10.1016/j.autcon.2016.05.018.
[11] Zhao H, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16:3586–92. http://dx.doi.org/10.1016/j.rser.2012.02.049.
[12] Hippert HS, Pedreira CE, Souza RC. Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans Power Syst 2001;16:44–55. http://dx.doi.org/10.1109/59.910780.
[13] Suganthi L, Samuel A a. Energy models for demand forecasting – a review. Renew Sustain Energy Rev 2012;16:1223–40. http://dx.doi.org/10.1016/j.rser.2011.08.014.
[14] Fumo N. A review on the basics of building energy estimation. Renew Sustain Energy Rev 2014;31:53–60. http://dx.doi.org/10.1016/j.rser.2013.11.040.
[15] 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;8:13162–93. http://dx.doi.org/10.3390/en81112361.
[16] 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. http://dx.doi.org/10.1016/j.rser.2015.04.065.
[17] 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 2016. http://dx.doi.org/10.1016/j.rser.2016.12.015.
[18] Ahmad aS, Hassan MY, Abdullah MP, Rahman H a, 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. http://dx.doi.org/10.1016/j.rser.2014.01.069.
[19] 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 2016. http://dx.doi.org/10.1016/j.rser.2016.10.079.
[20] Brockwell P, Davis R. Introduction to Time Series and Forecasting. Heidelberg, Germany: Springer; 2002.
[21] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;5:115–33. http://dx.doi.org/10.1007/BF02478259.
[22] Morris RGDO. Hebb: The Organization of Behavior, 50. New York: Wiley; 1949. p.

437.
[23] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev 1958;65:386–408.
[24] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6. http://dx.doi.org/10.1038/323533a0.
[25] Adya M, Collopy F How effective are neural networks at forecasting and prediction? A review and evaluation; n.d.
[26] Nasr GE, Badr EA, Younes MR. Neural networks in forecasting electrical energy consumption. FLAIRS-01 .In: Proc., American association for artificial intelligence; 2001.
[27] Nizami J, Ai-Garni AZ. Forecasting electric energy consumption using neural networks. Energy Policy 1995;23:1097–104.
[28] González PA, Zamarreño JM. Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. Energy Build 2005;37:595–601. http://dx.doi.org/10.1016/j.enbuild.2004.09.006.
[29] G. Schenker, Prediction and control using feedback neural networks and partial models. Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, 1996.
[30] Eric S. Predicion and control using feedback neural networks and partial models. Zurich; 1996.
[31] Panapakidis IP, Dagoumas AS. Day-ahead electricity price forecasting via the application of artificial neural network based models; 2016. ⟨http://doi.org/10.1016/j.apenergy.2016.03.089⟩.
[32] Karatasou S, Santamouris M, Geros V. Modeling and predicting building's energy use with artificial neural networks: methods and results. Energy Build 2006;38:949–58. http://dx.doi.org/10.1016/j.enbuild.2005.11.005.
[33] Rivals I, Personnaz L. Neural-network construction and selection in nonlinear modeling. IEEE Trans Neural Netw 2003;14:804–19. http://dx.doi.org/10.1109/TNN.2003.811356.
[34] Azadeh A, Ghaderi SF, Sohrabkhani S. Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. Energy Convers Manag 2008;49:2272–8. http://dx.doi.org/10.1016/j.enconman.2008.01.035.
[35] Ekonomou L. Greek long-term energy consumption prediction using artificial neural networks. Energy 2010;35:512–7. http://dx.doi.org/10.1016/j.energy.2009.10.018.
[36] Neto AH, Fiorelli FAS. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy Build 2008;40:2169–76. http://dx.doi.org/10.1016/j.enbuild.2008.06.013.
[37] Freeman JA, Skapura DM. Neural networks: algorithms, applications, and programming techniques; 1991.
[38] Yokoyama R, Wakui T, Satake R. Prediction of energy demands using neural network with model identification by global optimization. Energy Convers Manag 2009;50:319–27. http://dx.doi.org/10.1016/j.enconman.2008.09.017.
[39] Yokoyama R, Ito K. Capability of global search and improvement in modal trimming method for global optimization. JSME Int J Ser C 2005;48:730–7. http://dx.doi.org/10.1299/jsmec.48.730.
[40] Rezaeian-Zadeh M, Zand-Parsa S, Abghari H, Zolghadr M, Singh VP. Hourly air temperature driven using multi-layer perceptron and radial basis function networks in arid and semi-arid regions. Theor Appl Clim 2012;109:519–28. http://dx.doi.org/10.1007/s00704-012-0595-0.
[41] Pao H. Comparing linear and nonlinear forecasts for Taiwan's electricity consumption. Energy 2006;31:2129–41. http://dx.doi.org/10.1016/j.energy.2005.08.010.
[42] Hamzaçebi C. Forecasting of Turkey's net electricity energy consumption on sectoral bases. Energy Policy 2007;35:2009–16. http://dx.doi.org/10.1016/j.enpol.2006.03.014.
[43] Kelo S, Dudul S. A wavelet Elman neural network for short-term electrical load prediction under the influence of temperature. Int J Electr Power Energy Syst 2012;43:1063–71. http://dx.doi.org/10.1016/j.ijepes.2012.06.009.
[44] Chitsaz H, Shaker H, Zareipour H, Wood D, Amjady N. Short-term electricity load forecasting of buildings in microgrids. Energy Build 2015;99:50–60. http://dx.doi.org/10.1016/j.enbuild.2015.04.011.
[45] Deb C, Eang LS, Yang J, Santamouris M. Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. Energy Build 2015. http://dx.doi.org/10.1016/j.enbuild.2015.12.050.
[46] Deb C, Eang LS, Yang J, Santamouris M. Forecasting energy consumption of institutional buildings in Singapore. Procedia Eng 2015;121:1734–40. http://dx.doi.org/10.1016/j.proeng.2015.09.144.
[47] Chae YT, Horesh R, Hwang Y, Lee YM. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. Energy Build 2016;111:184–94. http://dx.doi.org/10.1016/j.enbuild.2015.11.045.
[48] Chae YT, Horesh R, Hwang Y, Lee YM. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. Energy Build 2016;111:184–94. http://dx.doi.org/10.1016/j.enbuild.2015.11.045.
[49] Box GEP, Jenkins GM. Time series analysis: forecasting and control. Hoboken, NJ, USA: John Wiley and Sons; 2008.
[50] Newsham GR, Birt BJ. Building-level occupancy data to improve ARIMA-based electricity use forecasts. BuildSys 2010;2010. http://dx.doi.org/10.1145/1878431.1878435.
[51] Yao Y, Lian Z, Liu S, Hou Z. Hourly cooling load prediction by a combined forecasting model based on Analytic Hierarchy Process. Int J Therm Sci 2004;43:1107–18. http://dx.doi.org/10.1016/j.ijthermalsci.2004.02.009.
[52] Saaty RW. The analytic hierarchy process–what it is and how it is used. Math Model 1987;9:161–76. http://dx.doi.org/10.1016/0270-0255(87)90473-8.
[53] Saaty TL. The analytical hierarchy process: planning, priority setting and resource allocation. New York: McGraw-Hill; 1980.
[54] Wang X, Meng M. A hybrid neural network and ARIMA model for energy consumption forcasting. J Comput 2012;7:1184–90. http://dx.doi.org/10.4304/jcp.7.5.1184-1190.
[55] Kandananond K. Forecasting electricity demand in Thailand with an artificial neural network approach. Energies 2011;4:1246–57. http://dx.doi.org/10.3390/en4081246.
[56] Chujai P, Kerdprasop N, Kerdprasop K. Time series analysis of household electric consumption with ARIMA and ARMA models. IMECS 2013;2013, [I].
[57] Roken R, Badri M. Time series models for forecasting monthly electricity peak-load for Dubai. Chancellor's Undergrad Res Award; 2006, p.1–14.
[58] Zhuang J, Chen Y, Shi X, Wei D. Building cooling load prediction based on time series method and neural networks. Vol. 8; 2015, p. 105–14.
[59] Abdel-Aal RE, Al-Garni aZ. Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis. Energy 1997;22:1059–69. http://dx.doi.org/10.1016/S0360-5442(97)00032-7.
[60] Abosedra S, Campus B. Demand for electricity in Lebanon. Int Bus Econ Res J 2009;8:11–8.
[61] Almeshaiei E, Soltan H. A methodology for electric power load forecasting. Alex Eng J 2011;50:137–44. http://dx.doi.org/10.1016/j.aej.2011.01.015.
[62] Dikaios TE. Forecasting residential electricity consumption in Greece using monthly and quarterly data. Energy Econ 1992;14:226–32. http://dx.doi.org/10.1016/0140-9883(92)90016-7.
[63] Erdogdu E. Electricity demand analysis using cointegration and ARIMA modelling: a case study of Turkey. Energy Policy 2007;35:1129–46. http://dx.doi.org/10.1016/j.enpol.2006.02.013.
[64] Twanabasu SR, Bremdal BA. Load forecasting in a smart grid oriented building. In: 22nd int. conf. exhib. electr. distrib. (CIRED 2013), Institution of Engineering and Technology; 2013, p. 0907–0907.
[65] Katara S, Faisal A, Engmann GM. A time series analysis of electricity demand in Tamale. Ghana 2014;4:269–75. http://dx.doi.org/10.5923/j.statistics.20140406.03.
[66] Mohamed N, Ahmad MH, Ismail Z. Improving short term load forecasting using double seasonal arima model. Appl Sci 2011;15:223–31.
[67] Godfred Kwame A. Modeling and forecasting energy consumption in Ghana. J Energy Technol Policy 2013;3:1–10.
[68] Rallapalli SR, Ghosh S. Forecasting monthly peak demand of electricity in India—a critique. Energy Policy 2012;45:516–20. http://dx.doi.org/10.1016/j.enpol.2012.02.064.
[69] Wang Y, Wang J, Zhao G, Dong Y. Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: a case study of China. Energy Policy 2012;48:284–94. http://dx.doi.org/10.1016/j.enpol.2012.05.026.
[70] Yasmeen F, Sharif M. Forecasting electricity consumption for Pakistan. Int J Emerg Technol Adv Eng 2014;4:496–503.
[71] Vapnik V. The nature of statistical learning theory. New York, USA: Springer; 1995.
[72] Vapnik V, Golowich SE, Smola A. A support vector method for function approximation, regression estimation, and signal processing. Adv Neural Inf Process Syst 1996;9:281–7.
[73] Zhang F, Deb C, Lee SE, Yang J, Shah KW. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. Energy Build 2016;126:94–103. http://dx.doi.org/10.1016/j.enbuild.2016.05.028.
[74] Chen B-J, Chang M-W, Lin C-J. Load forecasting using support vector machines: a study on EUNITE competition 2001. Power Syst IEEE Trans 2004;19:1821–30. http://dx.doi.org/10.1109/TPWRS.2004.835679.
[75] Hong W-C. Electric load forecasting by support vector model. Appl Math Model 2009;33:2444–54. http://dx.doi.org/10.1016/j.apm.2008.07.010.
[76] Pai P-F, Hong W-C. Support vector machines with simulated annealing algorithms in electricity load forecasting. Energy Convers Manag 2005;46:2669–88.
[77] Dong B, Cao C, Lee SE. Applying support vector machines to predict building energy consumption in tropical region. Energy Build 2005;37:545–53. http://dx.doi.org/10.1016/j.enbuild.2004.09.009.
[78] Li Q, Meng Q, Cai J, Yoshino H, Mochida A. Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks. Energy Convers Manag 2009;50:90–6. http://dx.doi.org/10.1016/j.enconman.2008.08.033.
[79] Li Q, Meng Q, Cai J, Yoshino H, Mochida A. Applying support vector machine to predict hourly cooling load in the building. Appl Energy 2009;86:2249–56. http://dx.doi.org/10.1016/j.apenergy.2008.11.035.
[80] Xuemei LXL, Jin-hu LJL, Lixing DLD, Gang XGX, Jibin LJL. Building cooling load forecasting model based on LS-SVM. Asia-Pac Conf Inf Process 2009;1:55–8. http://dx.doi.org/10.1109/APCIP.2009.22.
[81] Božic M, Stojanovi M, Stajic Z. Short-term electric load forecasting using least square support vector machines. Autom Control Robot 2010;9:141–50.
[82] Fernandez I, Borges CE, Penya YK. Efficient building load forecasting. ETFA2011, IEEE; 2011, p. 1–8.
[83] Edwards RE, New J, Parker LE. Predicting future hourly residential electrical consumption: a machine learning case study. Energy Build 2012;49:591–603. http://dx.doi.org/10.1016/j.enbuild.2012.03.010.
[84] Oğcu G, Demirel OF, Zaim S. Forecasting electricity consumption with neural networks and support vector regression. Procedia – Soc Behav Sci 2012;58:1576–85. http://dx.doi.org/10.1016/j.sbspro.2012.09.1144.
[85] Zhao W, Wang F, Niu D. The application of support vector machine in load forecasting. J Comput 2012;7:1615–22. http://dx.doi.org/10.4304/jcp.7.7.1615-1622.
[86] Ceperic E, Ceperic V, Member S, Baric A. A strategy for short-term load

forecasting by support vector regression machines. IEEE Trans Power Syst 2013:1–9. http://dx.doi.org/10.1109/TPWRS.2013.2269803.
[87] Jain RK, Smith KM, Culligan PJ, Taylor JE. Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl Energy 2014;123:168–78. http://dx.doi.org/10.1016/j.apenergy.2014.02.057.
[88] Kaytez F, Taplamacioglu MC, Cam E, Hardalac F. Forecasting electricity consumption: a comparison of regression analysis, neural networks and least squares support vector machines. Int J Electr Power Energy Syst 2015;67:431–8. http://dx.doi.org/10.1016/j.ijepes.2014.12.036.
[89] Liu D, Chen Q, Mori K. Time series forecasting method of building energy consumption using support vector regression. In: International conference inf. autom., Lijiang, China; 2015, p. 1628–32.
[90] Fu Y, Li Z, Zhang H, Xu P. Using support vector machine to predict next day electricity load of public buildings with sub-metering devices. Procedia Eng 2015;121:1016–22. http://dx.doi.org/10.1016/j.proeng.2015.09.097.
[91] Leake D. Case-based reasoning: experiences, lessons, and future directions. Cambridge MA: MIT press; 1996.
[92] Schank R. Dynamic memory: a theory of learning in computers and people. Cambridge, England: Cambridge University Press; 1982.
[93] Aamodt A, Plaza E. Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun 1994;7:39–59.
[94] Monfet D, Arkhipova E, Choiniere D Evaluation of a case-based reasoning energy prediction tool for commercial buildings. In: Proceedings of the 13th international conference enhanc. build. oper., Montreal, Quebec, Canada; 2013, p. 1–7.
[95] 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. http://dx.doi.org/10.1016/j.enbuild.2014.06.017.
[96] Song Q, Chissom BS. Fuzzy time series and its models. Fuzzy Sets Syst 1993;54:269–77.
[97] 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;35:2351–66. http://dx.doi.org/10.1016/j.energy.2009.12.023.
[98] Azadeh A, Seraj O, Saberi MA. Total fuzzy regression algorithm for energy consumption estimation. 2008 In: Proceedings of the 6th IEEE international conference ind informatics; 2008, p. 1562–8. ⟨http://doi.org/10.1109/INDIN.2008.4618353⟩.
[99] Bolturk E, Oztaysi B, Sari IU. Electricity consumption forecasting using fuzzy time series. In: 2012 IEEE 13th international symp. comput. intell. informatics, IEEE; 2012, p. 245–249.
[100] Singh SR. A simple method of forecasting based on fuzzy time series. Appl Math Comput 2007;186:330–9.
[101] Efendi R, Ismail Z, Mat M. A new linguistic out-sample approach of fuzzy time series for daily forecasting of Malaysian electricity load demand. Appl Soft Comput J 2015;28:422–30. http://dx.doi.org/10.1016/j.asoc.2014.11.043.
[102] Ismail Z, Mansor R. Fuzzy logic approach for forecasting half-hourly Malaysia electricity load demand. In: Int inst forecast 2011 proc; 2011, p. 1–17.
[103] Lim Y-D, Kim S-C, Park J-O, Choo Y-G. A predictive demand of the maximum electric power using chaos-fuzzy. In: Proceedings of the 10th IEEE international conference fuzzy syst. (Cat. no.01CH37297). Vol. 2, IEEE; 2001, p. 1511–1514.
[104] Liu X, Bai E, Fang J, Luo L. Time-variant slide fuzzy time-series method for short-term load forecasting. In: Proceedings 2010 IEEE international conference intell comput intell syst ICIS; 2010, p. 1:65–8.
[105] Pei AO. Load forecasting based on fuzzy time series. In: Proceedings of the 3rd international conference mater. mech. manuf. eng. (IC3ME 2015), Guangzhou, China; 2015, p. 715–9.
[106] Otto P, Schunk T. Fuzzy based time series forecasting of electric load. In: Eur. control conference, Karlsruhe, Germany; 1999, p. 4799–804.
[107] Ozawa K, Niimura T, Nakashima T. Fuzzy time-series model of electric power consumption. In: IEEE can. conference electr. comput. eng., Alberta, Canada; 1999, p. 1195–8. ⟨http://doi.org/10.1109/CCECE.1999.808235⟩.
[108] Tanaka H, Uejima S, Asai K. Linear regression analysis with fuzzy model. IEEE Trans Syst Man Cyber 1982;12:903–7.
[109] Pereira CM, Almeida NN De, Velloso MLF. Fuzzy modeling to forecast an electric load time series. Procedia Comput Sci 2015;55:395–404. http://dx.doi.org/10.1016/j.procs.2015.07.089.
[110] Thiang YK. Electrical load time series data forecasting using interval type-2 fuzzy logic system. In: Proceedings of the 3rd international conference comput. sci. inf. technol., Vol. 5, IEEE; 2010, p. 527–531.
[111] Deng J-L. Control problems of grey systems. Syst Control Lett 1982;1:288–94. http://dx.doi.org/10.1016/S0167-6911(82)80025-X.
[112] Chiang JS, Wu PL, Chiang SD, Chang TJ, Chang ST, Wen KL. Introduction to grey system theory. Taiwan: GAO-Li Publication; 1998.
[113] Deng J-L. Introduction to grey system theory. J Grey Syst 1989;1:1–24.
[114] Jiang Y, Yao Y, Deng S, Ma Z. Applying grey forecasting to predicting the operating energy performance of air cooled water chillers. Int J Refrig 2004;27:385–92. http://dx.doi.org/10.1016/j.ijrefrig.2003.12.001.
[115] Lee Y-S, Tong L-I. Forecasting energy consumption using a grey model improved by incorporating genetic programming. Energy Convers Manag 2011;52:147–52. http://dx.doi.org/10.1016/j.enconman.2010.06.053.
[116] Yao AWL, Chi SC, Chen JH. An improved grey-based approach for electricity demand forecasting. Electr Power Syst Res 2003;67:217–24. http://dx.doi.org/10.1016/S0378-7796(03)00112-3.
[117] Yao AWL, Chi SC. Analysis and design of a Taguchi–Grey based electricity demand predictor for energy management systems. Energy Convers Manag 2004;45:1205–17. http://dx.doi.org/10.1016/j.enconman.2003.08.008.
[118] Akay D, Atak M. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. Energy 2007;32:1670–5. http://dx.doi.org/10.1016/j.energy.2006.11.014.
[119] Pao H-T, Fu H-C, Tseng C-L. Forecasting of CO_2 emissions, energy consumption and economic growth in China using an improved grey model. Energy 2012;40:400–9. http://dx.doi.org/10.1016/j.energy.2012.01.037.
[120] Zhou Q, Wang S, Xu X, Xiao F. A grey-box model of next-day building thermal load prediction for energy-efficient control. Int J Energy Res 2008;32:1418–31. http://dx.doi.org/10.1002/er.1458.
[121] Hsu CC, Chen CY. Applications of improved grey prediction model for power demand forecasting. Energy Convers Manag 2003;44:2241–9. http://dx.doi.org/10.1016/S0196-8904(02)00248-0.
[122] Bianco V, Manca O, Nardini S, Minea A a. Analysis and forecasting of nonresidential electricity consumption in Romania. Appl Energy 2010;87:3584–90. http://dx.doi.org/10.1016/j.apenergy.2010.05.018.
[123] Hamzacebi C, Es HA. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 2014;70:165–71. http://dx.doi.org/10.1016/j.energy.2014.03.105.
[124] Jin M, Zhou X, Zhang ZM, Tentzeris MM. Short-term power load forecasting using grey correlation contest modeling. Expert Syst Appl 2012;39:773–9. http://dx.doi.org/10.1016/j.eswa.2011.07.072.
[125] Kang J, Zhao H. Application of improved grey model in long-term load forecasting of power engineering. Syst Eng Procedia 2012;3:85–91. http://dx.doi.org/10.1016/j.sepro.2011.11.012.
[126] Zhou P, Ang B, Poh K. A trigonometric grey prediction approach to forecasting electricity demand. Energy 2006;31:2839–47. http://dx.doi.org/10.1016/j.energy.2005.12.002.
[127] Li G-D, Wang C-H, Masuda S, Nagai M. A research on short term load forecasting problem applying improved grey dynamic model. Int J Electr Power Energy Syst 2011;33:809–16. http://dx.doi.org/10.1016/j.ijepes.2010.11.005.
[128] Li G, Yamaguchi D, Nagai M. Application of improved grey prediction model to short term load forecasting. In: International conference electr. eng.; 2006, p. 1–6.
[129] Winters PR. Forecasting sales by exponentially weighted moving averages. Manag Sci 1960;6:324–42.
[130] Christiaanse WR. Short-term load forecasting using general exponential smoothing. IEEE Trans Power Appar Syst 1971:900–11. http://dx.doi.org/10.1109/TPAS.1971.293123.
[131] Taylor JW, de Menezes LM, McSharry PE. A comparison of univariate methods for forecasting electricity demand up to a day ahead. Int J Forecast 2006;22:1–16. http://dx.doi.org/10.1016/j.ijforecast.2005.06.006.
[132] Taylor JW. An evaluation of methods for very short-term load forecasting using minute-by-minute British data. Int J Forecast 2008;24:645–58. http://dx.doi.org/10.1016/j.ijforecast.2008.07.007.
[133] Bindiu PDSR, Chindriú EM, Pop PDSG V. Day-ahead load forecasting using exponential smoothing, 6; 2009.
[134] Taylor JW. Triple seasonal methods for short-term electricity demand forecasting. Eur J Oper Res 2010;204:139–52. http://dx.doi.org/10.1016/j.ejor.2009.10.003.
[135] Kotillová a. Very short-term load forecasting using exponential smoothing and Arima models. J Inf, Control Manag 2011;9:85–92.
[136] Taylor JW. Short-term load forecasting with exponentially weighted methods. IEEE Trans Power Syst 2012;27:458–64.
[137] Adilah N, Jalil A, Ahmad MH, Mohamed N. Electricity load demand forecasting using exponential smoothing methods. World Appl Sci J 2013;22:1540–3. http://dx.doi.org/10.5829/idosi.wasj.2013.22.11.2891.
[138] Ariffin S, Karim A, Alwi SA. Electricity load forecasting in UTP using moving averages and exponential smoothing techniques. Appl Math Sci 2013;7:4003–14. http://dx.doi.org/10.12988/ams.2013.33149.
[139] Maçaira PM, Souza RC, Oliveira FLC. Modelling and forecasting the residential electricity consumption in Brazil with pegels exponential smoothing techniques. Procedia Comput Sci 2015;55:328–35. http://dx.doi.org/10.1016/j.procs.2015.07.057.
[140] Dasarathy BV. Nearest neighbour (NN) Norms: NN pattern classification techniques. Los Alamitos: IEEE Computer Society Press; 1991.
[141] Lora AT, Santos JMR, Riquelme JC, Exposito AG, Ramos JLM. Time-series prediction: application to the short-term electric energy demand. In: 10th Conf. Spanish Assoc. Artif. Intell. (CAEPIA), 5th Conf. Technol. Transf., Albacete, Spain; 2004, p. 577–86.
[142] Lei S, Sun C, Zhou Q, Zhang X. The research of local linear model of short term electrical load on multivariate time series. In: 2005 IEEE russ power tech; 2005, p. 1–5. ⟨http://doi.org/10.1109/PTC.2005.4524543⟩.
[143] Brown M, Barrington-Leigh C, Brown Z. Kernel regression for real-time building energy analysis. J Build 2012;5:263–76. http://dx.doi.org/10.1080/19401493.2011.577539.
[144] Al-Qahtani FH, Crone SF. Multivariate k-nearest neighbour regression for time series data — a novel algorithm for forecasting UK electricity demand. In: 2013 International Jt. conference neural networks, Dallas, USA; 2013, p. 1–8. ⟨http://doi.org/10.1109/IJCNN.2013.6706742⟩.
[145] Lachut D, Banerjee N, Rollins S. Predictability of energy use in homes. In: Proceedings international green comput. conference (IGCC 14), Dallas, USA; 2014, p. 1–10. ⟨http://doi.org/10.1109/IGCC.2014.7039146⟩.
[146] Roger T, Oumarou H, Gérard M, Moses T. Improved electric power demand forecasting by adapting the weighted average to the MISMO strategy. Int J Electr Eng Inform 2014;6:237–49.
[147] Colombo T, Koprinska I, Panella M. Maximum length weighted nearest neighbor

approach for electricity load forecasting. In: International jt. conference neural networks, Killarney, Ireland; 2015.
[148] Azadeh A, Ghaderi SF, Tarverdian S, Saberi M. Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. Appl Math Comput 2007;186:1731–41. http://dx.doi.org/10.1016/j.amc.2006.08.093.
[149] El-naggar KM, Al-rumaih K a. Algorithm, optimal filter estimator and least error squares technique: comparative study. Int'l J Electr Robot Electron Commun Eng 2007;1:941–5.
[150] Bashir Z a, El-Hawary ME. Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Trans Power Syst 2009;24:20–7. http://dx.doi.org/10.1109/TPWRS.2008.2008606.
[151] Lee C-M, Ko C-N. Time series prediction using RBF neural networks with a nonlinear time-varying evolution PSO algorithm. Neurocomputing 2009;73:449–60. http://dx.doi.org/10.1016/j.neucom.2009.07.005.
[152] Donate JP, Li X, Sánchez GG, Miguel AS. Time series forecasting by evolving artificial neural networks with genetic algorithms, differential evolution and estimation of distribution algorithm. Neural Comput Appl 2011:11–20. http://dx.doi.org/10.1007/s00521-011-0741-0.
[153] Xue J, Xu Z, Watada J. Building an integrated hybrid model for short-term and mid-term load foreacsting with genetic optimization. Int J Innov Comput Inf Control 2012;8:7381–91.
[154] Kouhi S, Keynia F, Najafi Ravadanegh S. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection. Int J Electr Power Energy Syst 2014;62:862–7. http://dx.doi.org/10.1016/j.ijepes.2014.05.036.
[155] Defilippo SB, Neto GG, Hippert HS. Short-term load forecasting by artificial neural networks specified by genetic algorithms – a simulation study over a Brazilian dataset. SIO 2015, 13° Simp. Argentino Investig. Oper., Rosario, Argentina; 2015, p. 57–66.
[156] 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. http://dx.doi.org/10.1016/j.enbuild.2015.09.002.
[157] Wang J, Wang J, Li Y, Zhu S, Zhao J. Techniques of applying wavelet transform into combined model for short-term load forecasting. Electr Power Energy Syst 2014;62:816–24. http://dx.doi.org/10.1016/j.epsr.2005.07.003.
[158] Zhuang J, Chen Y, Shi X, Wei D. Building cooling load prediction based on time series method and neural networks. Int J Grid Distrib Comput 2015;8:105–14.
[159] Hong-Tzer Y, Chao-Ming H, Ching-Lien H. Identification of ARMAX model for short term load forecasting: an evolutionary programming approach. IEEE Trans Power Syst 1996;11:403–8. http://dx.doi.org/10.1109/59.486125.
[160] Chao-Ming H, Chi-Jen H, Ming-Li W. A particle swarm optimization to identifying the ARMAX model for short-term load forecasting. Power Syst IEEE Trans 2005;20:1126–33. http://dx.doi.org/10.1109/TPWRS.2005.846106.
[161] Wang B, Tai N, Zhai H, Ye J, Zhu J, Qi L. A new ARMAX model based on evolutionary algorithm and particle swarm optimization for short-term load forecasting. Electr Power Syst Res 2008;78:1679–85. http://dx.doi.org/10.1016/j.epsr.2008.02.009.
[162] 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. http://dx.doi.org/10.1016/j.apenergy.2014.12.019.
[163] Nie H, Liu G, Liu X, Wang Y. Hybrid of ARIMA and SVMs for short-term load forecasting. Energy Procedia 2011;16:1455–60. http://dx.doi.org/10.1016/j.egypro.2012.01.229.
[164] Che J, Wang J. Short-term load forecasting using a kernel-based support vector regression combination model. Appl Energy 2014;132:602–9. http://dx.doi.org/10.1016/j.apenergy.2014.07.064.
[165] Fan G-F, Peng L-L, Hong W-C, Sun F. Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. Neurocomputing 2015;173:958–70. http://dx.doi.org/10.1016/j.neucom.2015.08.051.
[166] Lee W-J, Hong J. A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. Int J Electr Power Energy Syst 2015;64:1057–62. http://dx.doi.org/10.1016/j.ijepes.2014.08.006.
[167] Pao HT. Forecasting energy consumption in Taiwan using hybrid nonlinear models. Energy 2009;34:1438–46. http://dx.doi.org/10.1016/j.energy.2009.04.026.
[168] Li P, Li Y, Xiong Q, Chai Y, Zhang Y. Application of a hybrid quantized Elman neural network in short-term load forecasting. Int J Electr Power Energy Syst 2014;55:749–59. http://dx.doi.org/10.1016/j.ijepes.2013.10.020.
[169] Mao H, Zeng X-J, Leng G, Zhai Y, Keane J a. Short and mid-term load forecasting using a bilevel optimization model. IEEE Trans Power Syst 2009;24:1080–90.
[170] Chang P-C, Fan C-Y, Lin J-J. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. Int J Electr Power Energy Syst 2011;33:17–27. http://dx.doi.org/10.1016/j.ijepes.2010.08.008.
[171] Chen T, Wang Y. Long-term load forecasting by a collaborative fuzzy-neural approach. Int J Electr Power Energy Syst 2012;43:454–64. http://dx.doi.org/10.1016/j.ijepes.2012.05.072.
[172] Chaturvedi DK, Sinha aP, Malik OP. Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network. Electr Power Energy Syst 2015;67:230–7. http://dx.doi.org/10.1016/j.ijepes.2014.11.027.
[173] Hong-Tzer Y, Huang C-M. A new short-term load forecasting approach using self-organizing fuzzy armax models. IEEE Trans Power Syst 1998;13:217–25. http://dx.doi.org/10.1109/59.651639.
[174] Azadeh A, Saberi M, Ghaderi SF, Gitiforouz A, Ebrahimipour V. Improved estimation of electricity demand function by integration of fuzzy system and data mining approach. Energy Convers Manag 2008;49:2165–77. http://dx.doi.org/10.1016/j.enconman.2008.02.021.
[175] Enayatifar R, Sadaei HJ, Abdullah AH, Gani A. Imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS–ICA) for short term load forecasting. Energy Convers Manag 2013;76:1104–16. http://dx.doi.org/10.1016/j.enconman.2013.08.039.
[176] Sadaei HJ, Enayatifar R, Abdullah AH, Gani A. Short-term load forecasting using a hybrid model with a refined exponentially weighted fuzzy time series and an improved harmony search. Int J Electr Power Energy Syst 2014;62:118–29. http://dx.doi.org/10.1016/j.ijepes.2014.04.026.
[177] Che J, Wang J, Wang G. An adaptive fuzzy combination model based on self-organizing map and support vector regression for electric load forecasting. Energy 2012;37:657–64. http://dx.doi.org/10.1016/j.energy.2011.10.034.
[178] Son H, Kim C. Forecasting short-term electricity demand in residential sector based on support vector regression and fuzzy-rough feature selection with particle swarm optimization. Procedia Eng 2015;118:1162–8. http://dx.doi.org/10.1016/j.proeng.2015.08.459.
[179] Li F, Zhao X. The application of genetic algorithm in power short-term load forecasting; 2012, p. 50. ⟨http://doi.org/10.7763/IPCSIT.2012.V50.28⟩.
[180] Asrari A, Javan DS, Javidi MH, Monfared M. Application of gray-fuzzy-markov chain method for day-ahead electric load forecasting; 2012, p. 228–37.
[181] Wei C-C, Chen T-T, Lee S-J. k-NN based neuro-fuzzy system for time series prediction. In: Proceedings of the 14th ACIS international conference softw eng artif intell netw parallel/distributed comput; 2013, p. 569–74. ⟨http://doi.org/10.1109/SNPD.2013.68⟩.
[182] Sorjamaa A, Hao J, Reyhani N, Ji Y, Lendasse A. Methodology for long-term prediction of time series. Neurocomputing 2007;70:2861–9. http://dx.doi.org/10.1016/j.neucom.2006.06.015.
[183] Chen T-T, Lee S-J. A weighted LS-SVM based learning system for time series forecasting. Inf Sci 2015;299:99–116. http://dx.doi.org/10.1016/j.ins.2014.12.031.
[184] Azadeh A, Saberi M, Gitiforouz A, Saberi Z. A hybrid simulation-adaptive network based fuzzy inference system for improvement of electricity consumption estimation. Expert Syst Appl 2009;36:11108–17. http://dx.doi.org/10.1016/j.eswa.2009.02.081.
[185] Akdemir B, Çetinkaya N. Long-term load forecasting based on adaptive neural fuzzy inference system using real energy data. Energy Procedia 2012;14:794–9. http://dx.doi.org/10.1016/j.egypro.2011.12.1013.
[186] Hooshmand R-A, Amooshahi H, Parastegari M. A hybrid intelligent algorithm based short-term load forecasting approach. Int J Electr Power Energy Syst 2013;45:313–24. http://dx.doi.org/10.1016/j.ijepes.2012.09.002.
[187] Dudek G. Short-term load forecasting using random forests. Switzerland: Springer International Publishing; 2015.
[188] Wu J, Wang J, Lu H, Dong Y, Lu X. Short term load forecasting technique based on the seasonal exponential adjustment method and the regression model. Energy Convers Manag 2013;70:1–9. http://dx.doi.org/10.1016/j.enconman.2013.02.010.
[189] Cao L. Support vector machines experts for time series forecasting. Neurocomputing 2003;51:321–39. http://dx.doi.org/10.1016/S0925-2312(02)00577-5.
[190] Pai P-F, Hong W-C. Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. Electr Power Syst Res 2005;74:417–25. http://dx.doi.org/10.1016/j.epsr.2005.01.006.
[191] Hong W-C. Hybrid evolutionary algorithms in a SVR-based electric load forecasting model. Int J Electr Power Energy Syst 2009;31:409–17. http://dx.doi.org/10.1016/j.ijepes.2009.03.020.
[192] Hong W-C. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. Energy Convers Manag 2009;50:105–17. http://dx.doi.org/10.1016/j.enconman.2008.08.031.
[193] Li X, Ding L, Li L. A novel building cooling load prediction based on SVR and SAPSO. In: 3CA 2010–2010 international symp comput commun control autom. Vol. 1; 2010, p. 528–532. ⟨http://doi.org/10.1109/3CA.2010.5533863⟩.
[194] Niu D, Wang Y, Wu DD. Power load forecasting using support vector machine and ant colony optimization. Expert Syst Appl 2010;37:2531–9. http://dx.doi.org/10.1016/j.eswa.2009.08.019.
[195] Wang Y, Niu D, Ma X. Optimizing of SVM with hybrid PSO and genetic algorithm in power load forecasting. J Netw 2010;5:1192–200. http://dx.doi.org/10.4304/jnw.5.10.1192-1200.
[196] Hong W-C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 2011;36:5568–78. http://dx.doi.org/10.1016/j.energy.2011.07.015.
[197] Zhang WY, Hong W-C, Dong Y, Tsai G, Sung J-T, Fan G. Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting. Energy 2012;45:850–8. http://dx.doi.org/10.1016/j.energy.2012.07.006.
[198] Ju F-Y, Hong W-C. Application of seasonal SVR with chaotic gravitational search algorithm in electricity forecasting. Appl Math Model 2013;37:9643–51. http://dx.doi.org/10.1016/j.apm.2013.05.016.
[199] Hong W-C, Dong Y, Zhang WY, Chen L-Y, Panigrahi B K. Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm. Int J Electr Power Energy Syst 2013;44:604–14. http://dx.doi.org/10.1016/j.ijepes.2012.08.010.
[200] Kavousi-Fard A, Samet H, Marzbani F. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. Expert Syst Appl 2014;41:6047–56. http://dx.doi.org/10.1016/j.eswa.2014.03.053.
[201] Chen Y, Yang Y, Liu C, Li C, Li L. A hybrid application algorithm based on the support vector machine and artificial intelligence: an example of electric load forecasting. Appl Math Model 2014;39:2617–32. http://dx.doi.org/10.1016/

j.apm.2014.10.065.
[202] Baziar A, Kavousi-fard A. Short term load forecasting using a hybrid model based on support vector regression. Vol. 4; 2015, p.189–95.
[203] Ghofrani M, Ghayekhloo M, Arabali a, Ghayekhloo a. A hybrid short-term load forecasting with a new input selection framework. Energy 2015;81:777–86. http://dx.doi.org/10.1016/j.energy.2015.01.028.
[204] Kouhi S, Keynia F. A new cascade NN based method to short-term load forecast in deregulated electricity market. Energy Convers Manag 2013;71:76–83. http://dx.doi.org/10.1016/j.enconman.2013.03.014.
[205] Moazzami M, Khodabakhshian a, Hooshmand R. A new hybrid day-ahead peak load forecasting method for Iran's National grid. Appl Energy 2013;101:489–501. http://dx.doi.org/10.1016/j.apenergy.2012.06.009.
[206] Meng M, Niu D, Sun W. Forecasting monthly electric energy consumption using feature extraction. Energies 2011;4:1495–507. http://dx.doi.org/10.3390/en4101495.
[207] Sudheer G, Suseelatha A. Short term load forecasting using wavelet transform combined with Holt–Winters and weighted nearest neighbor models. Int J Electr Power Energy Syst 2015;64:340–6. http://dx.doi.org/10.1016/j.ijepes.2014.07.043.
[208] Bahrami S, Hooshmand R-A, Parastegari M. Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm. Energy 2014;72:434–42. http://dx.doi.org/10.1016/j.energy.2014.05.065.
[209] Magnier L, Haghighat F. Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. Build Environ 2010;45:739–46. http://dx.doi.org/10.1016/j.buildenv.2009.08.016.
[210] Kalogirou SA. Applications of artificial neural-networks for energy systems. Appl Energy 2000;67:17–35. http://dx.doi.org/10.1016/S0306-2619(00)00005-2.
[211] Labeodan T, De Bakker C, Rosemann A, Zeiler W. On the application of wireless sensors and actuators network in existing buildings for occupancy detection and occupancy-driven lighting control. Energy Build 2016;127:75–83. http://dx.doi.org/10.1016/j.enbuild.2016.05.077.
[212] Kumar P, Skouloudis AN, Bell M, Viana M, Carotta MC, Biskos G, et al. Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban buildings. Sci Total Environ 2016;560:150–9. http://dx.doi.org/10.1016/j.scitotenv.2016.04.032.
[213] Lilis G, Conus G, Asadi N, Kayal M. Towards the next generation of intelligent building: an assessment study of current automation and future IoT based systems with a proposal for transitional design. Sustain Cities Soc 2017;28:473–81. http://dx.doi.org/10.1016/j.scs.2016.08.019.
[214] Chen Z, Wang F, Feng Q. Cost-benefit evaluation for building intelligent systems with special consideration on intangible benefits and energy consumption. Energy Build 2016;128:484–90. http://dx.doi.org/10.1016/j.enbuild.2016.07.021.
[215] Kazmi AH, O'grady MJ, Delaney DT, Ruzzelli AG, O'hare GMP. A review of wireless-sensor-network-enabled building energy management systems. ACM Trans Sens Netw 2014;10:1–43. http://dx.doi.org/10.1145/2532644.
[216] Kuzlu M, Pipattanasomporn M, Rahman S. Review of communication technologies for smart homes/building applications. In: 2015 IEEE innov. smart grid technol. – Asia (ISGT ASIA), IEEE; 2015, p. 1–6. ⟨http://doi.org/10.1109/ISGT-Asia.2015.7437036⟩.