Energy & Buildings 165 (2018) 301–320

Contents lists available at ScienceDirect

Energy & Buildings

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

A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review

Tanveer Ahmad, Huanxin Chen*, Yabin Guo, Jiangyu Wang

School of Energy & Power Engineering, Huazhong University of Science and Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 27 August 2017
Revised 25 December 2017
Accepted 14 January 2018
Available online 10 February 2018

Keywords:
Data driven based approaches
Energy mapping models
Forecasting models
Benchmarking
Statistical and machine learning patterns

ABSTRACT

Energy consumption models play an integral part in energy management and conservation, as it pertains to buildings. It can assist in evaluating building energy efficiency, in carrying out building commissioning, and in identifying and diagnosing building system faults. This review takes an in-depth look at energy-demand prediction models for buildings in that it delves into recent developments in building energy approaches used to predict energy usage. By enlisting current approaches to the modelling of buildings, methods for building energy simulations can be categorized into four level classes as follows: (i) data-driven approaches; (ii) physics-based approaches; (iii) large scale building energy forecasting approaches; and (iv) hybrid approaches. The focus of this review is to target the data-driven approach and large-scale building energy predicting-based approaches. Here the data driven approaches can be categorized by (1) artificial neural network based approaches; (2) clustering based approaches; (3) statistical and machine learning-based approaches; and (4) support vector machine based approaches. From there, the type of data-driven based approach is further grouped by (a) benchmarking models; (b) energy-mapping models; (c) energy forecasting models; and (d) energy profiling models. Large-scale building-energy prediction techniques is then categorized as follows: (1) white-box based approaches; (2) black-box based approaches, and (3) grey-box based approaches. The current study explores first-rate data-driven based approaches about building energy analysis for industrial, commercial, domestic, etc., within a rural and urban setting. This review paper is based on the necessity of identifying points of departure and research opportunities for urban and rural-level analyses of building level energy performance. A variety of issues are explored which include: energy performance metrics; end-use of different building types; multiple levels of granularity; and urban and rural scales. Each technique encompasses a variety of input information as well as varying calculations or simulation models along with furnishing contrasting outcomes that suggest a variety of usages. A thorough review of each technique is presented in this study. This review highlights strengths, shortcomings, and purpose of the methods of numerous data-mining based approaches. A comprehensive review of energy forecasting models that are specified in the literature part is also provided.

1. Introduction

Building energy use prediction models play an integral part in energy management and conservation. These models can assist in examining the energy efficiency of buildings; in the construction of commission activities; detecting building system faults; and in identifying those faults. According to how specific it is, predicted energy could be categorized into the following five categories: (1) whole building energy/electricity; (2) heating and cooling energy; (3) heating energy; (4) cooling energy; (4) and (5) all others. Fig. 1, differentiates the percentages of the energy as mentioned earlier types [1].

As depicted above, over 50% of the studies concentrate on the forecasting of whole building level energy usage, which captures the total performance of the building. The total amount of all studies for the heating and cooling category is 35%. This is because commercial or educational/research buildings are considered the most often as their heating substance as mentioned earlier and cooling energy uses comprise a massive part of the building's energy consumption. Of note is that, relative to climate zones and realized the needs of the studies, some of the studies selected heating or cooling [2] energy as outputs.

* Corresponding author: School of Energy & Power Engineering, Huazhong University of Science and Technology, Wuhan, China.

E-mail addresses: ahmad_tanveer@hust.edu.cn (T. Ahmad), chenhuanxin@tsinghua.org.cn (H. Chen).

https://doi.org/10.1016/j.enbuild.2018.01.017

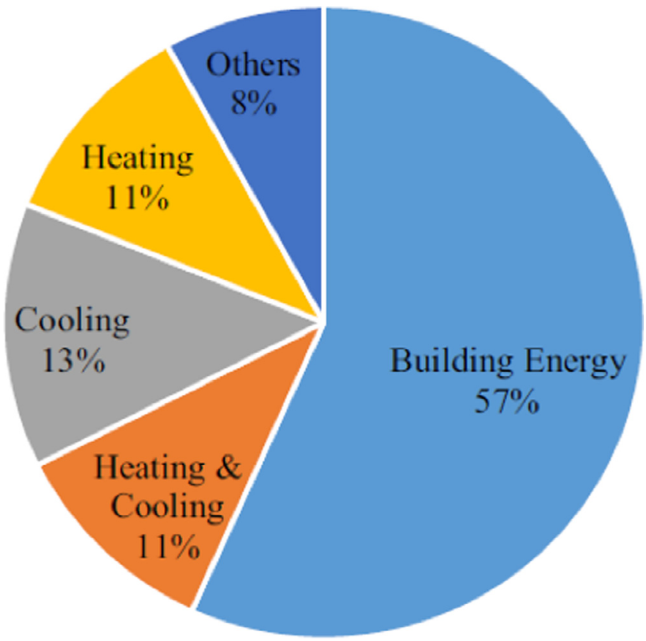

Fig. 1. The composition of energy type [1].

Building energy prediction models have loosely categorized into the following areas: (1) engineering; (2) Artificial Intelligence (AI) -based; (3) hybrid approaches, and (4) data-driven approaches [1]. To attain a maximum level of energy performance, installation of efficient energy-efficient systems could be put in place, along with followed by proper operation and management models [4]. Commercial buildings that contain modern metering and monitoring capabilities and systems coupled with efficient building management systems are the best methods in which to implement electricity load reduction activities. Moreover, potential economic benefits resulting in energy-demand reductions can be more pertinent for prosumers–customers who not only produce energy but who consume it [5].

Due to the ease of development and interpretation, and when differentiating forecasting methods, the regression model is the most commonly-used method enlisted for building load forecasting [6–9,59]. Current literature offers a formidable basis in which to classify the work as it pertains to the types of models, forecast horizon, and scale (single building to regional or national level) [1,10].

Since the 1990s, researchers have created a variety of simulation tools to estimate building energy use. These devices are identified as the following methods: engineering, AI-based, and hybrid [11]. The engineering method predicts the energy consumption by utilizing thermos-dynamic equations to account for the system's physical behavior as well as their interactions with the environment. This helps to predict energy use, i.e., energy consumption of individual building components, or that of the entire building [12] and it is defined as the 'white-box' method since the inner logic is known and evident.

Different from the engineering approach, the AI-based method is known as the 'black-box' method. This is because it estimates energy consumption without any knowledge of the building's internal relationship and its segments. The hybrid method, which is referred to as the 'grey box,' combines the white-box and the black-box methods to drive out the limitations associated with each method.

The white-box and grey-box methods each require detailed building information to simulate the inner relations utilized to predict energy use. Hence, model development requires tedious expertise, which is quite time-consuming for the existing buildings as it pertains to their energy consumption studies, utilizing the white-box and the grey-box methods are impractical if not impossible due to the lack of efficiency in doing so. Major challenges and difficulties can arise when attempting to glean building envelope specifications as well as mechanical systems. This results in an inability to use these methods comprehensively for existing building stock. However, a thorough review of energy consumption predictions in buildings that include the black box, white-box, and grey-box methods is available in [11].

Applied-learning algorithms within these methods may be pertinent when determining energy consumption. However, a limited number of studies used multiple prediction algorithms. Of those studies that utilized applied-learning algorithms, the robustness and capabilities were compared to ascertain their purposefulness in energy consumption forecasting [13,14]. Our review resulted in the following percentages of learning algorithm applications for energy use prediction: regression (26%); Acritical Neural Network (ANN) (41%); Support Vector Machine (SVR) (12%); and all others (21%). Of the four categories, we found ANN to be the most-used algorithm ANNs comprise, but are not very limited to the following elements: Multilayer Perceptron (MLP) [3]; Feed-Forward Neural Network (FFNN) [17]; Back-Propagation Neural Network (BPNN) [15,16]; and Radial Basis Function Network (RBFN) [18]. These ANNs were used in the most recent studies. They were widely favored due to user-friendly implementation and unequivocal prediction performance.

Using Multiple Linear Regression (MLR) in long-term energy consumption predictions resulted in advantages such that it was easy to use, as were the computation practices. Merely five studies enlisted SVR to forecast building energy consumption. With that said, SVR demonstrated its exceptional prediction accuracy in the construction of energy-use predictors when measured against other learning algorithms [16,19]. As well, the Autoregressive Moving Average (ARMAX) algorithm [13], Chi-Squared Automatic Interaction Detector (CHAID) algorithm [16], and Case-Based Reason (CBR) [20,21] algorithm was also enlisted for building energy use prediction. The researchers selected minute-by-minute [22], 15 min [23,24], weekly [25], and monthly [2] time scales to forecast building energy consumption.

In this review, several benefits of performing the large scale energy forecasting via simulation are found. These include the identification of: (i) energy outliers [27]; (ii) resources of energy (e.g., heat or waste power) in city in different buildings or districts located in the same area or district [26]; (iii) candidates for retrofit intervention [29]; (iv) local balancing and demand-side management operations [28]; (v) peak power demand [32]; (vi) large benchmarking analyses involving whole communities [30,31]; and (vi) improved urban planning within a designated area. Available data and the granularity level of the data must; maximize when analyzing urban-sale energy consumption ranking. Due to smart metering and increased awareness and comprehension of utilization data, the amount of data collection possible from single storey buildings has expanded during the previous few years. An additional reason for this upsurge is because of an increased usage.

It is important to note that, even if energy consumption data is available for analytical purposes, protection and privacy policies may exclude them as useful information sources. Therefore, it is essential to apply anonymization and aggregation methods. However, these methods can compromise data quality. [33]. Furthermore, building energy usage and evaluation of large scale can expend time allotments, especially when used with the single building based simulation approaches. This is largely due to the time-consuming process of data-gathering, the execution of monitoring techniques and simulation, and predictive factors involved in uncertainties [15]. Therefore, it is crucial to develop new data-collection tools that will ensure more timely results in collection processes, all the while preserving high-level granularity so as not to affect any final results adversely. These tools must fully support decision-making processes. In this review paper, we categorized the learning algorithms into four classifications as follows: (1) regression; (2) ANN; (3) SVR; and (4) all others.

Jani et al [230], has been conducting a study on the applicability of ANN for performance prediction of solid desiccant dehumidifier cooling (SDDC) methods. Various kinds of ANN are employed to represent the SDDC systems. Based on the effectiveness and use of test data, an artificial neural network method was proposed which is comprised of various models. The artificial neural network forecasting for these variables normally accepted with the highest correlation coefficient with the experimental values. The artificial neural networks can be employed with greater accuracy in calculating the efficiency and performance of SDDC method. This review was helpful for obtaining the different possibilities to promote and assist the research of artificial neural networks and their expediency which is flattering common in the upcoming period.

Kulkarni et al. [231] narrate the outcomes of research, analyzing the impact of numerous glazing methods on transoms and the decrease the cooling load demand in the building environment. Software (DesignBuilder) has been practiced for the simulation purpose of estimating the demand of the cooling load. Moreover, the research also shows the equivalent carbon credit and the potential of reducing the emission of CO2. In this study, retrofitting approaches actively affect the energy saving level, however, the period of payback is usually quite abundant of order eight-years.

Some studies have been conducted on, ANN algorithms for a vapor compression solid desiccant hybrid air-conditioning model is proposed to forecast the capacity of cooling load demand, the coefficient of performance and power input of the method. This research [232], is also conducts the practical analysis to set up for assembling the needed experimental data test. The results from the investigation narrate that the artificial neural networks algorithms can be implemented vigorously and can accommodate high reliability and accuracy for prognosticating the efficiency and performance of the SDDC systems [232–233].

Jani et al. [234] researched the impact of ambient temperature and outdoor humidity ratio of supply air temperature and COP have been estimated. The difference in COP has been acquired at different regeneration inlet and outlet temperatures. The results from the simulations explain the appropriateness of such models for cooling load requirement in the building environment in hot as well as humid weather circumstances.

Vildan [235] introduced a model based on simulation is rendered for the ascertainment and determination of the shading impact on buildings. The 1st part of the estimation scheme, solar radiation data for estimation were acquired during a technique by Kilic and Ozturk employed for ascertaining the relation of correlation, considering the data from Turkey. In the 2nd part, shaded spaces on vertical of the buildings, designed by adjacent buildings, practice the components of the shadow of a vertical pole on a three-dimensional coordinate method.

This review is comprised on the importance of identifying points of departure and research opportunities for urban and rural-level analyses of building level energy performance. A variety of issues are investigated which include: energy efficiency and performance metrics; end-use of different building types; multiple levels of granularity; and urban and rural scales. Each technique encompasses a variety of input information to train the validate the models as well as different calculations or simulation models along with furnishing contrasting outcomes that suggest a variety of usages. The capability to forecast the energy usage of building in a rural and urban background, practicing a different type of performance metrics for various building granularities and categories, across differing environmental and geographic locations, is significant for future energy perceptive planning.

The main benefits of implementing data driven and large scale (LS) energy forecasting into building simulation are numerous, for illustrate, the recognize of (i) better urban and rural planning in a specified area; (ii) local balancing and demand-side management; (iii) energy outliers; (iv) analysis of benching marking engaging entire communities; (v) retrofit invasion of candidates; (vi) peak demand of electricity usage and; (vi) energy sources (e.g., heat or waste power) in city level or various buildings of the similar city, country or district level. To examine the use of energy at different ranges, the granularity level and possible erudition of the energy consumption data necessity be assessed. The number of energy consumption and environmental data which is conceivable to assemble from the building's environment has progressed since several years because to the stabbing of 'smart metering' infrastructure, better understanding and greater access electric company or utility company's net energy consumption data and the enforcement of superintendence or management systems at the building level.

Though, if energy consumption data is apparently accessible for forecasting analysis, protection and privacy procedures may eliminate them as different authorizations (sources) of detailed information. Anonymization or aggregation methods are consequently needed, ever consists the essence of the various datasets available in the form of energy consumption data as well as environmental data. The estimation of data driven and large scale building energy usage might be notably consuming of time if implemented with one approach for building the simulation, because to data calculation methods, monitoring and simulation procedures and evaluation of contingencies or uncertainties. In this circumstance, the exploration of different techniques to use real building data and accumulate regarding time-saving and efficient behaviour while preserving a tremendous amount of quality that doesn't negotiate the consequence is a meaningful and notable augmentation for accessories (accessories) that would assist decisiveness comfort. Based on existing strategies for the modelling in buildings environment, it is conceivable to classify simulation of building techniques toward major three classifications as ensues or follows: (i) hybrid based approaches (ii) physics-based approaches and (iii) data-driven and large scale based approaches.

This review paper is witnessed with a consideration of data-driven and large scaled based energy prediction approaches employed to energy analysis of sector at a rural and urban level. The review is motivated by the requirement to recognise the points of departure and research opportunities for rural and urban level building analysis as well as energy efficiency and performance using large scale and data driven methods. Cognisance has considered of different of concerns including varying end-use and building types, multiple levels of granularity, different energy performance metrics and urban scales. Overview of the review is presented in Table 1.

The rest section of the review paper is designated as follows: Section-2 identifies the different data-driven base approaches which consist of clustering based, data-driven models, artificial neural networks, support vector machine, statistical and machine learning models. Section-3 outlines the large scale building's energy prediction techniques including grey-box based approaches, white-box based methods, and black-box based approaches. Sections-4 concludes this review paper.

2. Data-driven based approaches

Several undertakings have been enlisted to build large scale power usage algorithms at city-scale levels of functionality and

Table 1
Large scale and data-driven based energy prediction models.

Prediction models	Model category	Category distribution	Publications	References	Year range
Data-driven based approaches	Clustering-based data-driven models	Energy forecasting models	10	[44–53]	[1985, 2003, 2005, 2008, 2009, 2012, 2014]
		Energy mapping models	07	[54–60]	[1999, 2000, 2002, 2003, 2006, 2007]
		Energy profiling models	11	[61–72]	[1995, 1997, 2006, 2007, 2010, 2011]
	Statistical and regression based data-driven models	Forecasting prediction models	21	[73–83, 228–229]	[2001, 2004, 2005, 2009, 2010, 2011]
		Benchmarking models	09	[84–93]	[1999, 2002, 2005, 2008, 2009, 2010, 2013, 2015]
		Energy mapping models	13	[94–107]	[1980, 2001, 2008, 2009, 2011, 2012, 2013, 2014]
		Energy profiling models	21	[108–129]	[2000, 2001, 2004, 2005, 2007, 2008, 2009, 2011, 2013, 2014]
	ANN based data-driven models	Forecasting prediction models	23	[130–147, 230–235]	[2001, 2002, 2005, 2006, 2007, 2008, 2009, 2010, 2012, 2013, 2014, 2017]
		Benchmarking prediction models	07	[148–154]	[1988, 1996, 2000, 2002, 2003, 2005, 2006]
	SVM based data-driven models	Forecasting prediction models	22	[155–177]	[2004, 2006, 2007, 2008, 2009, 2010, 2012, 2014]
		Benching marking models	13	[178–191]	[2000, 2008, 2009, 2010, 2012, 2014]
		Energy mapping models	12	[192–203]	[1995, 2005, 2006, 2009, 2011, 2012, 2014]
Large scale based energy prediction approaches	Building sector energy forecasting using large scale based approaches	Energy forecasting white box data mining based approaches	04	[204–207]	2011, 2013, 2014
		Grey box data mining based approaches	05	[208–212]	2005, 2006, 2007,2014
		Black box data mining based approaches	14	[213–227]	1992, 2005, 2009, 2011, 2013, 2014, 2015, 2017

granularity. Data-driven based approaches (DDBA) furnish an equal equivalence among reducing the development period of the algorithm while sustaining an ample amount of precision [34]. The approaches outlined in this study were widespread at the building level. It has been discovered that the consolidation of different DDBA is incorporated at large-scale for forecasting of power consumption and to compensate for data deficiencies. This process demonstrated that some kind of well-known DDBA such as SVM and ANN continues to be overlooked at large scale. These data-driven models help furnish accurate information regarding the building stock (building profiling). The research presented a comprehensive discussion on the literature review of the current utilization and applications of DDBA at a rural and urban scale level, highlighting the importance of the future importance of research in this context.

2.1. Clustering-based data-driven models

Clustering-based data-driven models are an unsupervised energy consumption data analysis approach including the purpose of discovering unlabelled datasets and hidden information [35,49]. Clustering based approaches have been employed as beginning periods to develop models such as benchmarking that characterize common elements of illustrative buildings for baseline comparisons [37,38]. Clustering based algorithms separate sub-groups where every component in a assortment correlates to the different group within the similar cluster, yet distinct from the components in the different types of clusters [36]. The clustering based models are widely documented, but it's primarily practiced achieving categorization activities on a designated classification of buildings sector [40,41], for which limited studies exist at urban level [42,43]. Widespread use of clustering the building sector is evident in the literature for performing an initial step to identify representative buildings (centroids) and to create archetypes [39]. The most widely used clustering algorithms include model-based clustering (KMMBC); K-Means; K-Medoids (KM); and Hierarchical Agglomerative Clustering (HAC).

2.1.1. Energy forecasting models

Energy consumption forecasting is critical for energy policy and national economics. However, it is a complicated and uncertain problem brought on by the outer-environment as well as several uncertainty factors. Potential prediction models for the countries like as India, Turkey, UK, Iran, New Zealand, Finland, and China [44–47] were also discussed. Short-term predictors of future energy consumption (next 24 hours) are necessary for the support of decision for part or unit commitment, equating demand and supply with various network parameters, and selling/buying power in the day-ahead place exchanges. Massive number of material exists in form of literature on load prediction, including semi-parametric approaches (e.g., additive models), neural networks, seasonal time-series models, exponential smoothing (see, e.g., [48–51]) and exogenous factors (e.g., SARIMAX, PARX). The fundamental principle of set pair analysis (SPA) based clustering model is demonstrated as in reference [52]. This is followed by analyzing the uncertainty of identity Discrepancy Contrary Analysis (IDCA) along with forecasting energy usage within a specified year based on the calculated impact circumstances of each class. The flowchart is shown in Fig. 2, with procedure specifications [52].

Perform 'cluster-analysis' with a 'Fisher's optimal partition model' for the rising Dynamic Relative Indicator (DRI) of chronicled power usage. Further, the criteria to perform set pair analysis (SPA) and the correlating methods are specified to shift the mean amounts of Dynamic Relative Indicator (DRIs) inside intermediary amounts to determine and estimate the different degree, identical degree and contrary degree among reference sets and influence factors. Then, interspersed including substances (weights) of influential circumstances, attain combined connection numbers

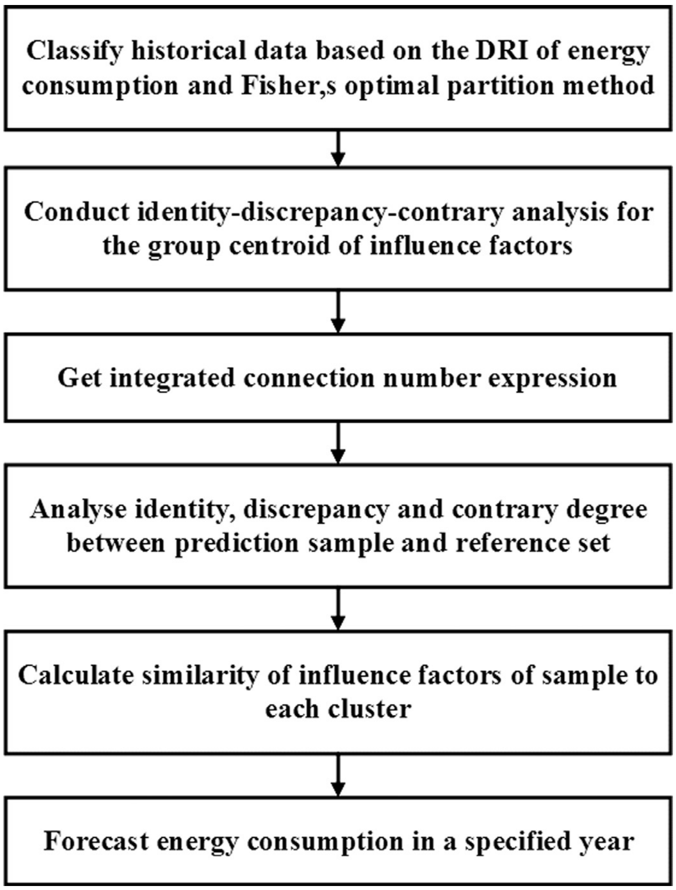

Fig. 2. Flow chart of SPA-based cluster forecast model.

that present the magnetism determinants of energy consumption of energy to be predicted in the selected year with the consolidation amounts gathered from calculating their similitudes with each cluster and Identity Discrepancy-Contrary (ID-C) analysis. Finally, build a prediction model and estimate the consumption of energy in a designated year.

Let $k(k=1, 2,\dots, k)$ depicts the total clusters. The factors of influence for each type is narrated as:

$$\bar{x}_n^{(k)} = \frac{1}{M}\sum_{m=1}^{M} x_{mn}^{(k)} \tag{1}$$

$\bar{x}_n^{(k)}$ is the average amount of *nth* for type k. $x_{mn}^{(k)}$ is the corresponding dynamic relative indicator amount of *nth* with influence factor on data sample $m(m=1, 2,\dots, M)$. Then, according to the PSA theory [53], the combined amount or number μ_k, employed to identify different growth elements can be presented as:

$$\mu_k = a_k + b_k i + c_k j = \sum_{n=1}^{N} a_{nk} w_n + i\sum_{n=1}^{N} b_{nk} w_n + j\sum_{n=1}^{N} c_{nk} w_n \tag{2}$$

$$a_{nk} = \lambda \bar{x}_n^{(k)} \tag{3}$$

$$c_{nk} = \frac{\eta'}{\bar{x}_n^{(k)}} \tag{4}$$

$$b_{nk} = 1 - a_{nk} - c_{nk} \tag{5}$$

where a_k, b_k, and c_k different-degree, identical-degree and contrary-degree, respectively; w_n is a weight of indicator b_{nk}, a_{nk}, and c_{nk} is discrepancy degree, identical degree of nth similar to the set of reference on type k, sequentially; j is the contrary coefficient and i is the discrepancy coefficient. λ and $\acute{\eta}$ shows the coefficients;

if $\bar{x}_n^{(k)} \in [x^L,\ x^U]$ λ and $\acute{\eta}$ will satisfy

$$0 < \lambda\eta \le 0.25 \tag{6}$$

$$(x^U)^2\lambda + \eta - x^U \ge 0 \tag{7}$$

$$(x^L)^2\lambda + \eta - x^L \ge 0 \tag{8}$$

where x^L and x^U are upper and lower amounts of DRI for previous (historical) record indicators. The various amount enclosures of λ and $\acute{\eta}$ are shown in Fig. 3 [52].

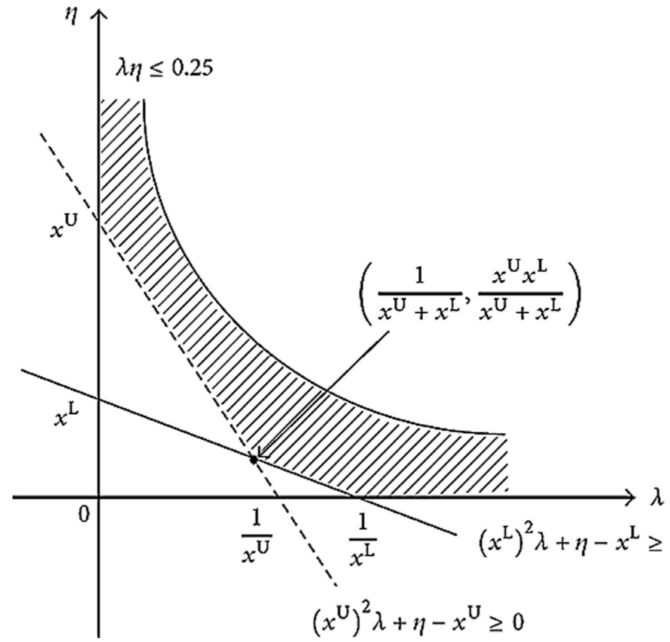

Fig. 3. Coefficient value scope of λ and $\acute{\eta}$ [52].

Two linear functions intersection point (9) is often used to notify λ and $\acute{\eta}$,

$$\lambda = \frac{1}{(x^U + x^L)},\quad \eta = \frac{x^U x^L}{(x^U + x^L)} \tag{9}$$

The degree of connection among the calculated indicator value of a prediction year and a set of reference $(x_n^k = 1)$ can also be established.

2.1.2. Energy mapping models

To contain the conception of the sustainable environment in building level, the expanded usage of particular forecasting tools is required, both at a several urban scales and at the building level. A plethora of energy-prediction and environmental enforcement tools exist at the building level. However, very few tools are capable of forecasting energy at an urban scale. In the United Kingdom, the construction sector is changed at approximately 1% per year [54], which penitentiary is relatively more common for a massive developed post-industrial land. If the construction conditions are to have increased sustainability, and if ‘carbon dioxide’ decreases are to be accomplished, then crucial measures that can enhance the present building conditions necessity be analyzed [55]. An electricity and environmental prognostication model have been constructed in partnership with regional executives in Southland Wales, UK, as a consolidated policy for increased sustainable environmental and to forecast and estimate for decreases in ‘carbon dioxide’ and different emanations [56]. The processor design provides strategies for incorporating residential power control and

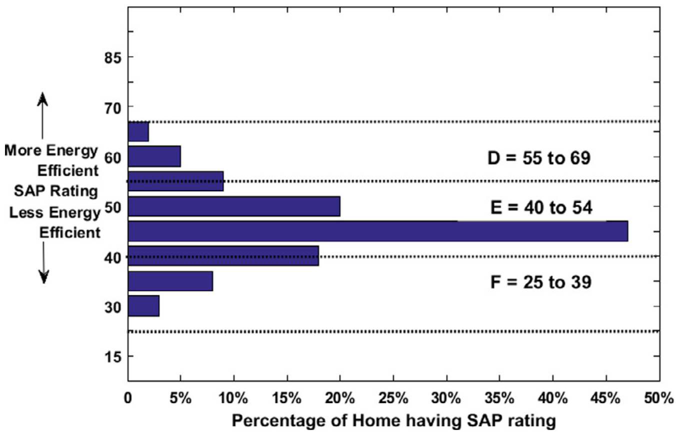

Fig. 4. Division of the SAP ranks with each power performance steps established with all regional authority owned assets [6].

Table 2
Welsh house property measures target SAP evaluation through story space.

Floor area (m^2)	SAP rating
Up to 35.0	58.0
36.0–40.0	59.0
41.0–45.0	60.0
46.0–50.0	61.0
51.0–55.0	62.0
56.0–60.0	63.0
61.0–70.0	64.0
71.0–80.0	65.0
81.0–90.0	66.0
91.0–100.0	67.0
101.0–110.0	68.0
111.0–120.0	39.0
Over 120.0	70.0

environmental preparation, empowering arrangement producers to contemplate increased power proficiency. Originally produced by 'Neath Port Talbot District Borough Council (NPTCBC)', Power and Environment Forecast (EEP) has been executed by different political governments in the United Kingdom as well as Australia [57].

The power performance of the regional administration headquarters assets NPTCBC has been documented to demonstrate the dimension of Standard Assessment Procedure (SAP) evaluations Fig. 4. Aforementioned reveals that the residences are frequently in band *E*, SAP rating of 40.0–54 as demonstrated by the measure interjected in the resort knowledge pack in reference [6,58,59].

The 'Welsh Assembly Government' has launched the 'Welsh Housing Quality Standard (WGQS)' [60], as a general-purpose criterion for friendly accommodations in Wales. The energy and environmental prediction model provide information to authenticate the power performance standard for regional administration and contracted quarters. WGQS recommends that a lodging necessity be appraised adopting the SAP evaluation practice, and it must produce a number within 70 and 58, depending on the ground field of the residence Table 2 [6].

2.1.3. Energy profiling models

Residences spend 40% of the overall fundamental power and 30% of the seasonal power, thus massively impacting arboretum gas emanations. As a result, domestic electricity performance and concepts such as the approximately zero power for the buildings are consistently obtaining attention and notoriety as measures to decrease carbon eruptions as well as the dominant reliance on specimen combustibles. The proposed methodologies of the energy profile are implemented to the charge trajectories of multiple homes, resulting in the establishment of the best clustering scheme. The methods can be implemented for any home. Profiling and grasping the electricity consumption trends of their facilities will encourage the implementation of energy efficiency measures, electricity waste elimination, and Demand Side Management (DSM) solutions [61–63]. Typical profiling algorithms applied in quantity of load profiling are the 'K-means', the Minimum Variance Criterion (MVM), the self-organizing map and the Fuzzy C-means. These types of algorithms are used primarily due to their effectiveness [64–66]. Other approaches with lesser use in the load curve clustering literature include the Hopfield neural network [67], the ISO DATA algorithm [68], and the Support Vector Clustering (SVC) [69].

Aside from the algorithms as mentioned earlier, hybrid combination utilizations are also taken into account. However, standard classification schemes before-mentioned as the SVM [70] or the Bayes classifier [71] are applied as a profiling tool, but they are not enlisted as a load profiling solution. This is because the categorization of the electricity consumption patterns of buildings is usually developed as an unsupervised machine learning task, i.e., the frequency of clusters and their respective composition are a priori unknown variable. For such challenges, ample discernment could be applied to the selection and training of the appropriate pattern recognition algorithms. Also, several methods have been proposed for load profile conversions to the frequency-domain as a compression measure of the measurement of the clustering information data anchored [72].

The initial grouping of the recorded data into subsets may enhance the efficiency of the clustering procedure and decrease the input data set size. Therefore, the frequency-domain techniques and utilized for the analysis of the load data. In Time-Domain (T-D) modelling, the data set containing the metered data in kilowatts (kW) is $P=\{p_j,\ j=1,\ \ldots,\ N\}$, where $N=365$ and respectively everyday measurement sweep is formulated as a vector p_j:

$$p_j = [p_{j1} \ldots, p_{j96}] \tag{10}$$

For each building, two individual data sets are envisaged that correspond to the years 2010 and 2011. The load curves are categorized according to their shape similarity since, at the main clustering stage, the load magnitude is irrelevant. Each member pattern of the set P is normalized in [0,1] with the largest value of the set. The resulting inclination with the data normalized patterns is expressed as $X=\{x_j,\ j=1,\ \ldots,\ N\}$. The second technique applies the recognized transformation into the Frequency-Domain (FD) [72] by using the Fast Fourier Transform (FFT). Follow up the Nyquist's hypothesis, the highest essential rotation positions $h_{max}=T_o/(2\ldots T_{sam})=48$ to $T_o=1440$ min as the base period (24 h). Next, all 'amplitudes' are assigned in restraining order, as references the indicator [72]:

$$\varsigma_h = \frac{1}{N}\sum_{J=1}^{N}|A_{hj} - \bar{A}_h| \tag{11}$$

where $h=1,\ldots,\ h_{max},\ A_{hj}$ and $\bar{A}_h$ is the amplitude of the jth load curve and the average value, respectively.

Vector $\emptyset_n$ includes the amplitudes and two phases-related features of the top-ranked n harmonic orders. Two different sets of harmonic-based coefficients are taken into consideration assuming a reduction of 50% and 25% of the original Time-Domain (TD) samples. Thus, the resulting data sets are $\emptyset_{16}$ and $\emptyset_{18}$. The phase-related features ψ_h' and ψ_h'' of the hth harmonic order of phase angle a_h are identified in (12) and (13) [72]:

$$\psi_h' = \frac{1-\cos(a_h)}{2}\frac{A_h}{\sqrt{\sum_{s\epsilon\emptyset j} A_s^2}} \tag{12}$$

Table 3
Summary of different clustering-based data-driven models.

Model category	Clustering-based data-driven models
Energy forecasting models	
Prediction model	K-means [48], Life cycle assessment [47], Fuzzy C-means, Bayes classifier [49], Periodic Autorepression (PAR) model [48], SOM (proximity)/FCM [49], K-means++ [49], Box-Jenkins [48], MVM [49], SOM (square)/K-means++ [49], Econometric model [52], FCM [49], SOM [49], SOM (proximity)/K-means++ [49], Integrated models [52], SOM (square)/MVM [49], SOM (proximity)/MVM [49], Weather-based models [50], SOM (square)/FCM [49], Clustering based on set pair analysis (SPA) model [52], Hybrid model [45].
Software	Matlab, Eneryplus, Trnsys, BLAST, ESP-r
Application and usage	Energy prediction in building environment, Daily load curves, Reveal significant opportunities of energy efficient improvements and efficient buildings operation during the whole day, statistically significant factors affecting energy use
Advantages	Enabling the efficient operation and management of the distribution grid in real-time by identifying and predicting the behavior of large building customers, Maximizes difference between clusters to allow for a meaningful comparison of the differences between building clusters
Energy mapping models	
Prediction model	Degree day method [59], Residential load factor method [59], Set pair analysis [54], Adaptive particle swarm optimization [56], Regression method [59], Temperature-based fourier series model [59], Principle Component Analysis [59], Auto-Regressive model [59], Auto-Regressive Integrated Moving Average [59], Conditional Demand Analysis [59], Back propagation neural network [59], Genetic algorithm [59], Multi-Layer Perceptrons [59], Fuzzy C-mean clustering [59]
Software	DOE-2, EnergyPlus, BLAST, ESP-r, MATLAB, Solvo@,
Application and usage	Data pre-processing, Building electricity usage prediction, Forecast power usage is achievable, efficient and suitable for effective applications
Advantages	Accurate and true predictions of electricity usage, Depression of obscure fuel oil tuberculosis at the different industrial and mill areas, Reduce problems of uncertainty from three different perspectives of identification, Contrary and discrepancy depicts and features comprehensively requisite properties of things
Energy profiling models	
Prediction model	Discrete Fourier Transform [72], Regression-based electricity load model [62], Curvilinear component analysis [66], Mathematical modeling [64], Classical model [62], Change-point models [62], Linear regression models [62], Classical k-means model [64], Fuzzy logic [66], WARD model [64], K-means [64], Hierarchical clustering [64], Hierarchical models [64], Kohonen self-organizing map [66], Hopfield–K-Means clustering algorithm [67], ISODATA algorithm [68], Fuzzy k-means [69], Quadratic programming model [69], Bottom-up approach [72]
Software	MATLAB, BLAST, ESP-r
Application and usage	Customer classification, Daily load curves forecasting
Advantages	Accurate energy forecasting for the industrial and commercial users, Decrease the uncertainty problem from three aspects of identity, High volatility, Presence of outliers

$$\psi_h'' = \frac{1-\sin(a_h)}{2}\frac{A_h}{\sqrt{\sum_{s\epsilon\emptyset j} A_s^2}} \tag{13}$$

For each Supervisory Control and Data Acquisition (SCADA) unit, a set of features f_n is defined in (14). Apart from the h harmonic components also the zero-order harmonic A_0 with phase angle equal to zero is included in each set.

$$f_n = \left\{ \left(A_p, \psi_p', \psi_p''\right), (p \neq 0) \cap \emptyset_n \right\} \cup \{(A_p)(p=0) \cap (p=\emptyset_n)\} \tag{14}$$

Table 3 shows the summer of model type, advantages, usage in different applications of energy forecasting, mapping and energy profiling methods.

2.2. Statistical and regression-based data driven models

In statistical modelling, regression investigation is a mathematical method used to forecast the interrelatedness between parameters. It encompasses many procedures for modelling and analyzing numerous parameters, concentrating on the association between one or more independent variables and a dependent variable. Further to the point, regression dissection demonstrates whence the expected amount of the minor variable modifications during any one of the free parameters is diverse, while the additional individual parameters are existed solidified. Several regression dissections calculate the dependent expectation of the subordinate parameter addressed the autonomous parameters.

2.2.1. Forecasting prediction models

Significant activity in much smart grid utilization, like as demand response to contingency superintendence is the short-term power capacity prediction at various orders, of a particular consumer to an entire array of consumers. Forecasting prediction models construct a quantitative-evaluation of several 'machine learning methods' for short-term energy demand prediction at the aggregate and the individual level. In this setting, energy consumption prediction for different time horizons (e.g., 1 h ahead, 1 day ahead, 1 month ahead), and space scales (e.g., distribution transformer, individual house-level meter) is increasingly paramount for many applications of the future power grid (or Smart Grid), which includes demand response, frequency and voltage regulation, or autonomous emergency management [73]. Machine-learning research has augmented several methods that apply to the energy consumption prediction. This includes linear regression and ARMA model [74] to neural networks [75] or SVM regressions [76]. Notably, these techniques have normally been implemented at huge space scales, such as predicting the electrical load of a market segment that provides energy for thousands of customers [77,78], or even an entire country [79].

In this section, define methodology adopted to decide the relevant features of the energy forecasting models. To decide the relevant features, contemplate two-time horizons: one hour onward and twenty-four hours ahead. In each case, the predicted value is the hourly consumption of unless a particular customer or the whole set of users. To quantitatively assess the magnitude of this relation, formulated the autocorrelation of the time series of consumption, which sheds light on how much the consumption at hour h is correlated with the consumption at an hour $h' < h$. The auto-correlation of a time series $S = s_1,\ s_2; \ldots,\ s_N$ for a certain lag τ is calculated as:

$$R(S,\tau) = \frac{E[(s_i - \mu(S)).(s_{i-\tau} - \mu(S))]}{\sigma(S)^2} \tag{15}$$

where E is the suspected amount speculator, S, and $\sigma(S)$ is the standard deviation and $\mu(S)$ is the arithmetic mean of time scale. Fig. 5, (a) shows the auto-correlation of the energy consumption time series for an individual customer. It is recognizing the

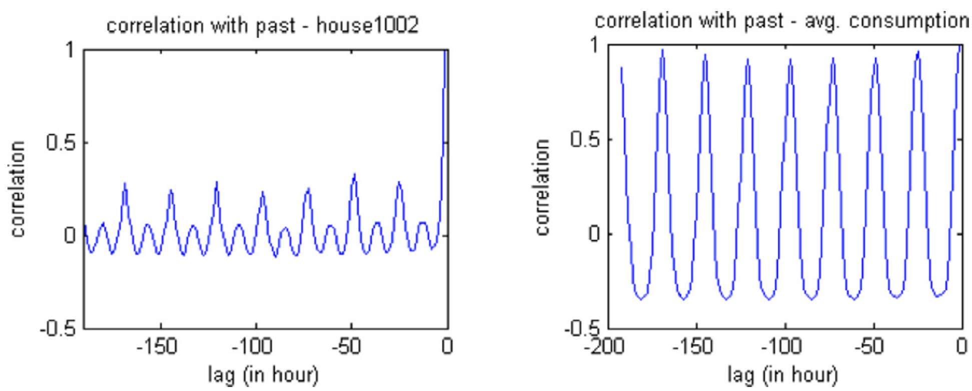

Fig. 5. Auto-correlation of consumption (a) Individual household (id 1002) (b) Aggregation (782 households) [80].

value of the use in that it is highly correlated with the previous consumption up to 3-hours before. For past hours, a local buying power maximum exists in correspondence of the hour $h' \epsilon \{h-23.9,\ h-47.9, h-71.9\ldots, h-167.9\}$. To demonstrate that the energy consumption of the previous 7-days at the same hour h possesses some predictive factors. At the aggregate level, the auto-correlation between the energy consumption at hour h and the past use at hour $h' \epsilon \{h-23.9,\ h-47.9, h-71.9\ldots, h-167.9\}$ is much higher than in the case of individual households, such as presented in Fig. 5(b) [80]. Hence, this implies that forecasting energy consumption when including historical data is anticipated to increase accuracy more so as the whole level than at the particular level.

Recent literature established SVM based regression methods SVR is capable models for prognosticating expected energy usage [77,79]. Other well-established methods have been identified as linear regression and Multi-Layer Perceptron (M-LP). Wu et al. [81] made a model to enhance the forecast of climate and financial time series. On Demand Response (DR), utility companies worked with large industrial customers to reduce their energy consumption during designated peak hours of energy usage. To ascertain customers' energy-use reduction, utility companies needed to estimate what consumers would become utilized in the absence of a DR situation (that is, the DR baseline). DR baselines vary from simple averaging to regression [82,83].

2.2.2. Benchmarking prediction models

In the contract-off that occurs when tendering in the next day power demand and the exception time balancing demand, ways require accurate predictions for stabilizing market costs to arrive at knowledgeable determinations. A variety of previous patterns that outlined the balancing of forecasting market needs, including a fewer augmentation, is benchmarked and the models individually benchmarked for one-hour ahead and one day-ahead prediction while comparing point and interim projections. The benchmarked patterns yielded informing one day ahead point predictions, supporting that accessible erudition efficiently the plugging of the one-day ahead demand is adequately indicated in the one-day ahead price of the market preferably than the balancing prices of market trend.

Misiorek and Weron [84] propose a major critique of one-day projection methods. In a benchmark limit on the time sequence approaches corresponding respective another. Different studies that benchmark 1-day prediction techniques are explored in the literature in reference [85], which examine three distinct time scale patterns, 'artificial neural models' and 'wavelet based approaches' for the inter-connection of PJM (Pennsylvania-New Jersey-Maryland) one day ahead price, and that the experiments conducted by Nogales et al. [86], which associate two separate time order patterns

Table 4
Review of number patterns for benchmarking.

Name	State algorithm	Prediction method volume
HIST.	Markove	Form of the random historical values
RND.	Markove	Forms of the random from distribution
CROST.	Arrival	ARI-algorithm for unevenly space times series
SARMA.	None	Seasonal ARNA model

for the 'Californian' and the 'Spanish' 1-day leading supermarket prices. To our knowledge, the survey results do not exist as they pertain to balancing market prices. Nonetheless, problem considerations do endure, including Skytte in this study [87], Pettersen and Fleten [88], Söder and Olsson [89], Jaehnertd et al. [90], Soder and Brolin [91] and Boomsmaa et al. [92], for which they are all of the 'Nordic' business market. The participation in aforementioned substance is a well-organized analysis and benchmarking methods based on time series for calculating demand price predictors. Importance is placed on the one day-ahead boundary, however, both and one-day and one-hour ahead predictions are benchmarked. A summary of the rules for the benchmarked amount can be determined in Table 4 [93].

To estimate the interim projections from unconditional different coverage: Sublet y_t be a recognized meaning in the out of specimen limit and give $L_{t\,(p)}$ and $U_{t\,(p)}$ be the uppermost and drop boundaries of the probability projections to cover the probability, p, sequentially. Later recognize a symbol variable that can be described as follows:

$$I_t = \{1,\ if\ y_t\ \epsilon \left[L_{t(p)}U_{t(p)}\right] \tag{16}$$

$$I_t = \{0,\ if\ y_t\ \neq \left[L_{t(p)}U_{t(p)}\right] \tag{17}$$

The overall conclusion is considered comparable for both patterns in that including the one hour ahead forecasts, maximum hours are prognosticated equitable accurately including numbers that reach from 0.70–0.90, whereas a assortment of the times shows complicated to prognosticate, thereby receiving a recondite amount [93]. The average means absolute error (MAE) of the 4 distinct designs can be detected in Table 5 [93]. The ranking of the various pattern is indisputable: toward the short-term prediction (next one-hour), the SARMA pattern exceeds whole other type in every week. For 1-day ahead prediction, the CROST pattern amidst intermittently apportioned event series (lined) is superior for all ahead weeks, excluding first week and week-8 wherever the SARMA pattern demonstrated immeasurable results. The responsibility of prognosticating the one day ahead comparing supermarket pamphlet is formidable; the SARMA pattern led to the weakest achievement during week-10 and week-12, whereas the CROST pattern is substandard in week-1.

Table 5
MAE of estimating market volume prediction for various patterns.

w	$\lvert\bar{v}_t\rvert$	1-Hour ahead forecast				Day ahead forecast			
		CROST	SARMA	HIST	RAND	CROST	SARMA	HIST	RAND
1	109.800	67.53	37.99	76.73	77.68	110.18	103.63	107.25	108.44
2	59.120	46.73	33.56	53.96	59.74	30.60	54.64	86.50	88.97
3	86.650	68.89	44.57	60.36	65.54	45.17	72.20	98.39	100.28
4	35.050	33.79	34.13	54.83	66.18	20.61	61.09	60.97	63.47
5	90.430	96.0	41.69	56.23	68.02	47.40	71.47	80.81	79.61
6	43.030	33.21	25.02	43.23	61.29	25.54	53.64	70.54	73.88
7	21.450	19.38	15.22	35.97	49.28	14.69	27.99	50.19	51.82
8	43.020	33.62	18.97	44.04	53.49	28.89	19.04	56.10	56.60
9	121.850	78.31	35.90	80.12	83.81	65.02	78.69	107.01	106.29
10	94.840	58.70	35.48	77.77	89.05	50.86	145.74	87.48	87.94
11	98.370	65.66	43.86	71.35	73.21	55.78	57.52	113.50	117.01
12	39.880	39.85	30.60	36.59	45.85	41.55	104.29	64.10	66.52

MWh-Weekly Average

2.2.3. Energy mapping models

Energy mapping introduces a bottom-up mathematical method consists of a Terrestrial Information System (TIS) to predict the energy of household properties beyond a whole country, province or city. The adoption of a multiple-linear regression model permits the down-scaling of measured natural gas and electricity usage from the aggregated post-code level to single-dwelling households, which is based on several descriptors, such as dwelling kind, duration of development, story facade, and some inhabitants. The power utilization is apportioned to different end-users and corrected for weather, and then the energy savings potential is estimated by including the implementation of common refurbishment measures.

According to the review of Swan and Ugursal [94], 2 characters of instituting assets power patterns occur at large scale: top-down approaches work to place a correlation among summed power consumption and related data before-mentioned as housing description statistics or economic data [95–97]; bottom-up approaches estimate energy consumptions of single building, or groups of buildings for different identity through a hierarchy of disaggregated input data. The results are then deduced for the entire building stock by way of proxy indicators. Engineering-based approaches utilize quantitative data that highlights features of the home property in which to calculate power utilization of a set of buildings that represent the stock with a numerical model [98–100]. Several statistical methods in literature analyze the connection between individual energy consumption of buildings and a range of variables related to building characteristics, such as household composition and occupants' behaviors as they pertain to a sample of buildings [101,102]. Although much the Geographic Information (GIS)-based engineering models have been generated at city scale [103], only a few GIS-based statistical models exist. Howard et al. [104] created a GIS-based analytical method to identify the arrangement of instituting power utilization intensities in the New York City.

The following equations are most normally enlisted for natural gas and electricity energy mapping models:

$$y_{gas} = \beta_{0,gas} + x_{floor} \cdot \beta_{floor,gas} + \sum_{i=1}^{25} \left(x_{type,i,gas} \cdot \beta_{type,i,gas} \right) + \varepsilon_{gas} \tag{18}$$

$$y_{ele} = \beta_{o,ele} + x_{floor} \cdot \beta_{floor,\ elec} + x_{people} \cdot \beta_{people,\ elec} + \sum_{i=1}^{5} \left(x_{type,i,elec} \cdot \beta_{tyoe,i,elec} \right) + \varepsilon_{elec} \tag{19}$$

The dependent variable y is designated by annual values of average-measured energy consumption per dwelling (natural gas or electricity) obtainable by post-code area. The independent variables are relative to the characteristics of dwellings and the components of households: floor surface x_{floor}, the number of occupants x_{people} and type of dwelling x_{type}. Based on the stepwise regression procedure, the variable floor surface x_{floor} was found to be significant for both natural gas and electricity models, while the variable number of occupants x_{people} were only pertinent for electricity [105–107].

2.2.4. Energy profiling models

A variety of corporeal factors establishes power requirement load characterization. Procurement of the ideal combination of stratagems and latest renewable energy system installation dictates a simplistic approach in which to practice at the immediate plan scaffold. Residential area, steaming load outline for several classifications of houses become formed by adopting the dynamic-thermal pattern, which has been created employing the thermal-resistant-network scheme. The everyday disruption power requirement load characterization of apparatus, residential space heating and hot water can be ascertained using this approach which can provide a daily load outline from a peculiar residence to residential inhabitants. In the United States profitable house division, Heating, Ventilation and Air Conditioning (HVAC) practices comprise 40.0% of power usage [108]. Nearly 34.0% of the building places provide interior heating, ventilation and air conditioning control [109,110] that produce cooling comfort applying chillers for temperature dismissal and disband it to air handling assemblages. Head is generated by, using boilers that move it to fan coil radiators, assemblages, or baseboard radiators [111–113]. The 'ANSI/ASHRAE 90.1-2007′ suggests heating, ventilation and air conditioning systems programs adhere to the inventories determined for representative structure classes [114] that could be manipulated relative to possession appropriations obtained while a building layout phase according to engineers' previous experiences as well as large scale inhabitant survey results [115]. According to certain directions, heating, ventilation and refrigeration practice should move in a continuous sequence of occupied spaces and could be cycled off and on to accommodate cooling and heating load demand while deserted hours [116].

Nevertheless, researchers have demonstrated that occupancy space or sequential is indeterminate in the universe; therefore, inhabitant departure and arrival circumstances are problematic to pre-destine [117–119]. To compensate for the inefficiencies that are aligned with utilizing predetermined heating, ventilation and air conditioning system records, many studies have practiced real-time possession or occupancy to refine HVAC processes [120,121]. The

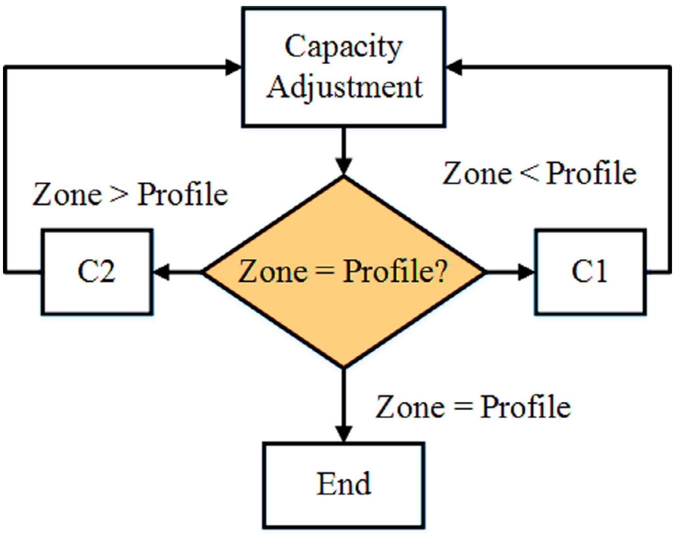

Fig. 6. *C* rules for profile clustering [129].

intention is that power consumption might be conserved by not encompassing HVAC system in the unoccupied area. Considerable power conservations become published in earlier studies halting motionless set points in the abandoned area [122–126] in which area heats are permitted to floating diverse varieties, compromising on whether the area is congested or opposing [127,128].

In reference [129], the 'Minkowski' length is picked to determine the connection among two characterizations where it is practiced as an imprecise purpose to contain clustering horizon.

$$\textit{Minkowski distance } d_{io} = \left(\sum_{p=1}^{n} \left| x_{ip} - x_{op} \right|^{r} \right)^{\frac{1}{r}} \tag{20}$$

d_{io} is the length builds on the ith centroid and item o; n shows the dimension of the vector, equal to 479.9 in this research; x_{ip} represents the amount of probability at time (p) for efficiency, and x_{iO} shows the probability at different period p for the centroid characterization. Once the 'k-means process of clustering' is terminated, the ability modification is executed to establish that area limit is exquisite. A number of the profile in a subordinate cluster is inequivalent to the area space, the subsequent amendment, known as the "*C* rule" is adopted according to the extent among the different centroid and different profiles groups Fig. 6 [129].

(1) (b_i) can be explained capability of the ith area. Meanwhile, it is limited than the number of profile I_i, $b_i < I_i$. b_i of profiles nearest point to the centroid which is allocated to the area of the zone, while the remaining profiles are selected for the different area, where centroids are nearest to the centroid of the ith area. This practice is recognized as *C1* in this review paper.

(2) The ability or capacity of the ith area of density b_i is bigger as compared the profile number I_i, $b_i > I_i$, all characterizations are allocated to this area, while the additional openings are permeated by the consumer profiles which are approaching to the contemporary area of the ith zone of the centroid. This edict is recognized as *C2* in this review.

(3) R_g and R_l practices are executed iteratively unless every area capacity b_i is according to its characterization amount I_i [129]. Table 6 depicts the summery of different statistical and regression-based data driven models.

2.3. ANN based data-driven models

Connectionist systems or artificial neural networks are a computational model used in 'machine learning practice', and other related experimentation developments. ANNs penitentiary is molded from precedents, preferably than explicitly particularized. They are numerous commonly used in ointments that are too complicated to communicate in a conventional microcomputer algorithm that enlists familiar, programming based on rules. ANNs provide a variety of solutions for a large assortment of responsibilities, including speech recognition, computer vision, medical diagnostics and machine translation.

2.3.1. Forecasting prediction models

Prediction of building energy has obtained acceptance because of the catapulting analysis in the building efficiency field. A wide diversity of building power models becomes recognized for the short-term, medium-term and long-term power forecasting [130,131]. The commonly used patterns in current times are comprises 'machine learning and statistical based models' that precisely determine the power prediction and usage correlating to beforehand designated data. Latest studies furnish specified recitals on available prediction models and their organizational [132–134] of the majority of these models, however, are comprised of the base of ANNs and their advancements [135–138]. Ayidinalp et al. [139] applied ANN for forecsting the power utilization of apparatus, space cooling, lighting, in the domestic area. The application of artificial neural networks demonstrated exceptional prediction capabilities when contrasted against a building simulation model. After two years, Aydiinalp et al. [140] utilized the artificial neural network to model space water power utilization and heating. Zamarre et al. [141] practiced feedback artificial NN to prognosticate short-term power load. The developed feedback pattern or model was part of the PhD exposition of 'Schenker' [142].

This arrangement incorporates the section of the yield, which is fed back as an input value as well as the forecasting failure relative to the contained outputs that are then used to train the network. Karatasou et al., artificial neural networks were utilized by Aazadeh et al. in reference [143] to determine long-term power utilization in authority concentrated making enterprises in Republic of Iran. They also conducted the similar study examining measures for actively projecting monthly power usage [144]. Yokoyamaa et al. [145] employed back propagation artificial neural network to foretell the cooling load requirement for an office building. They introduced a novel method called global optimization is known as a modal trimming method to detect the parameters of models [146]. The 'global optimization method' can estimate the different influence neurons in the different (hidden layers) and inputs on the exactness of the prognostication.

Existing applications and advantages of ANN are appropriately admitted in an amount of restrictions, particularly in nonlinear modelling. The principal component in artificial neural networks is the energy as mentioned earlier neurons that lined up in stripes and attached to neurons in different zones into sections [147]. These connections are apprehended as the synaptic type weights-could, and the intentions of the process of model training are to extract certain substances. Activation of the neurons is circumscribed by the aggregated of the model inputs parameters and can be (mathematically) indicated as in Eq. (21):

$$O = f\left(\sum (w_{ij} x_j)\right) \tag{21}$$

Here, O is the neuron output, x_j is the neuron input, w_{ij} presents the connection weight of the input, and f is the substitution function. This function employed in ANNs is commonly the

Table 6
Summary of different statistical and regression-based data-driven models.

Model category	Statistical and regression based data-driven models
Forecasting prediction models	
Prediction model	ARMA [76], Adaptive Model [73], HWT method [74], Time series (univariate) model [75], Multiplicative autoregressive models [75], Kalman filtering [75], Univariate neural network Models [74], HWT method [74], Suboptimal seasonal autoregressive Models [74], Nonparametric regression [75], Structural models [75], Curve-fitting procedures [75], Genetic Algorithms (RSVMG) [76], Time-series modeling [79],
Software	Matalb, EnergyPlus, Trnsys, etc.
Application and usage	Modelling of intraday electricity prices, Traffic management and call centre staff scheduling, Data pre-processing
Advantages	Greater forecasting performance and efficiency, Can be appropriate for contemptible models, Convergence secured, Computationally efficient by plan and design, Not model dependent
Benchmarking models	
Prediction model	Semiparametric time series models [84], Wavelet models [85], ARMA [84], Dynamic regression model [86], ARX models [84], Time Series Models [86], Regime-switching regressions [84], Stochastic linear programming model [88]
Software	MFE Toolbox, SCA, Matlab, EnergyPlus
Application and usage	Management of the risk, Prediction total consumed energy spot values and prices, Intervals or density predictions, Daubechies wavelets
Advantages	Can be computationally effective, Not different kind model or algorithm dependent, processes are stationary, Secured to concentrate on optimal enzyme inadequate number of variables, computationally highly efficient and effective, Frequency can be considered high
Energy mapping models	
Prediction model	Genetic algorithms [94], Simple mathematical model [99], Regression model [94], Probabilistic energy model [100], Conditional demand analysis [94], Statistical modelling [102], Regression techniques [94], Top-down modeling approach [105], BREDEM [94], Autoregressive distributed lag model [94], Error correction model [97], Energy simulation model [98], Bottom-up models [94], GIS-based statistical downscaling approach [107]
Software	ESP-r, ERAD, EnergyPlus, eQuest, DOE-2.1E, eQUEST/DOE-2.2, BREDEM, HOT 2000, EPIQR, ECCABS, VBA, Sketch-up
Application and usage	Supply analysis based on long term electricity projections, Electricity requirement by estimating for historic data response, Electricity requirement augmentation of the different end practices including of behavioural perspectives, Cooling and heating loads, Use of energy; CO_2 emissions
Advantages	Energy forecasting in Long-term perceptive, Contains occupant response and performance, Algorithm with advancement technologies, Embodiment of socioeconomic and macroeconomic impacts, Ascertainment of appropriate energy supplying as well as their use in the building environment, Simple and manageable input knowledge, Encompasses energy inclinations, Practices the data of billing simplistic survey erudition
Energy profiling models	
Prediction model	Stochastic occupancy models [115], Learning-Based Model [127], Derivative-free optimization algorithm [126], Agent-based model [117], Stochastic process method [117], Stochastic Model Predictive Control [124], Generalised stochastic model [118], Spectral partitioning algorithm [123], SHOCC [118], Poisson model [119], Autoregressive-moving-average (ARMA) algorithm [129], K-means algorithm [129], MPC control algorithm [125]
Software	DOE, EnergyPlus, ESP-r, DeST, TRNSYS, MATLAB, SUNtool, SER-Res
Application and usage	Indoor environment, Electricity usage, In security Pertinence, Control of Lighting, HVAC system electricity usage when the flow of air linked with occupancy control and building
Advantages	Models can produce the accurate location of every occupant of the building and the zone level, Can be measure occupancy rate estimation for the entire building

'sigmoidal function', which contains the resulting structure (22):

$$S(t) = \frac{1}{1+e^{-t}} \tag{22}$$

The neuron output is transferred to the next-layer into output of this neuron and these weights in the subsequent film might be denoted, and simplest form is shown in Eq. (23), wherever, O_l is the neuron output for the next layer the and w_{jl} is the 'weight' bonding the earlier neuron:

$$O_l = f\left(\sum (w_{jl} f(w_{ij} x_j))\right) \tag{23}$$

The objective of the preparation of the model training procedure is to depreciate the square error among the measured outputs and the predicted outputs. E is diminished from the (gradient descent model), which includes estimating the influenced derivative of E relative to the specific substance in the interface.

$$E = \sum^{\frac{1}{2}} (O_p - O_m).^2 \tag{24}$$

Here, E is the sum of error refer to total, O_p is the estimated output, and O_m is designated output.

2.3.2. Benchmarking prediction models

Energy benchmarking is imperative for calculating the use of building energy and contrasting it to related buildings in like environments. The benchmarking results added calculates could be used to decrease power usage when the case building has been evaluated to utilize immense power levels compared with the other similar premises. An ANN based benchmarking model is exhibited as a profoundly useful approach. This technique focuses particularly on forecasting a weighted power use index by exercising within account specific variables, before-mentioned as lighting type, plug load density and times of the process, 'air conditioning equipment type and efficiency' etc. Also, the artificial neural network benchmark paradigm for prognosticating latent electricity conservations from the different retrofit schemes was estimated.

Benchmarking input parameters were altered to indicate inherent power conserving by a retrofit scheme, and the set of new the inputs was simulated with the artificial neural network design. Matson and Piette [148] have reviewed power studies for office and commercial structures of the building. A linear regression technique proposed by Sharp [236,150] is various usually applied in different benchmarking type buildings. Massive details of the power star (for building energy) can be developed from the research by [151]. In a study by [152], a model based 'benchmarking method' particularly relevant for lab room was formed. Supplementary analysis of lab benchmarking and pragmatic attentiveness and conditions can be comprehended from reference [153,154]. Table 7 describes the different aspects of ANN-based data driven models.

2.4. SVM based data-driven models

An SVM is a discriminative classifier identified by an isolating hyperplane. Put another way, given designated the training data,

Table 7
Summary of ANN-based data-driven models.

Model category	ANN based data-driven models
Forecasting prediction models	
Prediction model	Feed forward neural networks [137], Radial basis function (RBF) [147], Numerical models [232], Multiple Linear Regression with Interactions [131], AI-based approaches [131], Hybrid models [134], Mathematical models [232], Causal models [130], Adaptive models [135], Engineering model [138], Pseudo-steady state model [233], Simplified ANN model [230], Grey model [145], ANOVA model [143]
Software	Matlab, DOE-2, EnergyPlus, BLAST, ESP-r, EnergyPlus
Application and usage	Building electricity usage prognostication, energy consumption data pre-processing, controlling and predicting the peak energy demand, Flow of air distribution inside the room, Wind speed, Indoor air-temperature, heating ventilation and refrigeration system dissection, forecasting of electricity usage, prediction of the energy usage, Solar radiation
Advantages	Higher class efficiency, Easy to practice and use, Forecasting can be implemented in the entire building section, Immeasurable at determining different non-linear difficulties, Comprehensive amount of zones, Accurate characterization of the flow of fluid occurring inside the building sector, A consistent volume of data in form of training with the essential vector data
Benchmarking prediction models	
Prediction model	Multilayer neural networks [149], CART [149], Classification Tree [149], CHAID [149], Linear regression model [152], CART and Negative Binominal Regression (NBR) [149], Empirical model [152], Simple appraisal methods [152], Logistic regression [149], Classification tree analysis [149]
Software	FLUENT, COMSOL, MIT-CFD, PHOENICS-CFD, TrnSys, EnergyPlus, IDA-ICE, ESP-r, Clim2000, BSim and BUILDOPT-VIE
Application and usage	Evolution of the energy demand, Natural ventilation, Passive solar, Air conditioning system, Climate environment, occupants behaviour, Indoor thermal comfort, Artificial and natural ventilation, Evolution of time of the world electricity usage, Space-averaged indoor temperature forecasting
Advantages	High efficiency and performance in forecasting the substantive electricity usage, Time distribution and Spatial of confined state parameters (concentration, temperature, airflow, pressure) in a high volume, Implementation is quite easier, Versatile zone of the buildings; Computation time is reasonable, Function representing the system; Efficacious modeling and optimization model

the model outputs of an optimal-hyperplane which classifies additional illustrate. In increasing to delivering linear group, support vector machine can effectively conduct a nonlinear analysis known as the 'kernel trick', precisely outlining their facts within high dimensional characteristic periods. During the energy usage data are not labelled, controlled training is impossible, and an unsupervised approach to learning is essential. This kind of procedure acts to locate spontaneous 'clustering' of the data to combinations, and later maps current data to certainly established assortments. The type of clustering model enhances the SVM. This is activity is apprehended as support vector (SV) clustering and is frequently applied in industrial assiduousness, unless during data are not labelled or meanwhile unique some of the data are labelled as a preprocessing for a distributed approach.

2.4.1. Forecasting prediction models

The electric market, rendering the suitable quantity of power at the appropriate time, including the best invitation demand is crucial for utility companies to amplify their profits. The mid-term power market clearing price prediction has shifted paramount for sources' maintenance scheduling, relocation, budgeting and planning and bilateral contracting. While there are various available approaches for the short-term power market clearing price prediction, the amount of mid-term power market clearing price forecasting is insignificant. Currently, most electricity markets clearing price forecasting studies pertain to short-term power market clearing price prediction, which is typically associated as the 24:00 h one day ahead energy price prediction. Enough small research has been accompanied by forecasting power market clearing price by mid-term [155–160].

Recently, support vector machine, a novel training approach comprises on the fundamental uncertainty minimization, possesses achieved expanded awareness regarding power price prediction [161–164]. The primary benefits of support vector over the artificial neural network or any prediction techniques are that support vector machine can dodge challenging issues that include the likes of overfitting data and unpredictably high out of sample in energy consumption data erroneousness, also corresponding time delivering immeasurable outcomes. Support vector machine is also a quite formidable prediction technique. Notwithstanding of the fundamental advantages, support vector machine usage will be viewed consistently as one that delivers results. Additionally, the SVM has the added value of less flexible variables correlated to artificial neural network and, consequently, it is considered less involved in the control variable assortment. A conventional support vector machine can perform approximately 3% [165] improved achievement contrasted to a typical artificial neural network as it pertains to the short-term power market price forecasting prediction.

Many models are employed to optimize the training of support vector machine to enhance SVM prediction precision. These models combine genetic algorithm models [166–169], swarm artificial fish algorithm [170], component analysis algorithm [171,172], and imprecise positions models [173,174]. An improved support vector machine, which is known as the (least squares support vector machine) was also formed to increase the veracity of the inventive support vector machine [175,176]. While particular classification has determined demonstrated partial increases, the operation accuracy.

An SVM is a novel machine learning approach comprise on fundamental uncertainty minimization. At its infancy, support vector machine was used in different classification purpose [177]. After on, the 'nonlinear function of regression computation' was enumerated by determining a curved 'quadratic optimization' predicament. The different subject class can be acknowledged as computation for regression with reoccurring the amount of threshold. Assume $\{(X_t, y_t)\}$ for $(t = 1)$ to (N) is a presented position of different data wherever $X_t = (-x_{t1}, x_{t2}, \ldots, x_{tk})$ is the vector of inputs at time t with k components and y_t is the analogous at different data price at time t that might be interpreted as

$$y_t = f(X_t) = \langle W, \emptyset(Xt) \rangle + b \tag{25}$$

wherever $h; i$ express the product of dot, W shows the vector for weight, b represents the bias, and $\emptyset$ depicts the function of mapping exchanges of the input vector X_t toward a substantial greater dimensional characteristic period. The same optimization difficulty is, then:

$$\begin{aligned} & minimize\ \frac{1}{2}\|W\|^2 + C\sum_{t=1}^{N}(\zeta_t + \zeta_t^*) \\ & subject\ to \begin{bmatrix} y_t - \langle W, \emptyset(Xt)\rangle - b \le \varepsilon + \zeta_t \\ \langle W, \emptyset(Xt)\rangle + b - y_t \le \varepsilon + \zeta_t \\ \zeta_t \zeta_t^* > 0 \end{bmatrix} \end{aligned} \tag{26}$$

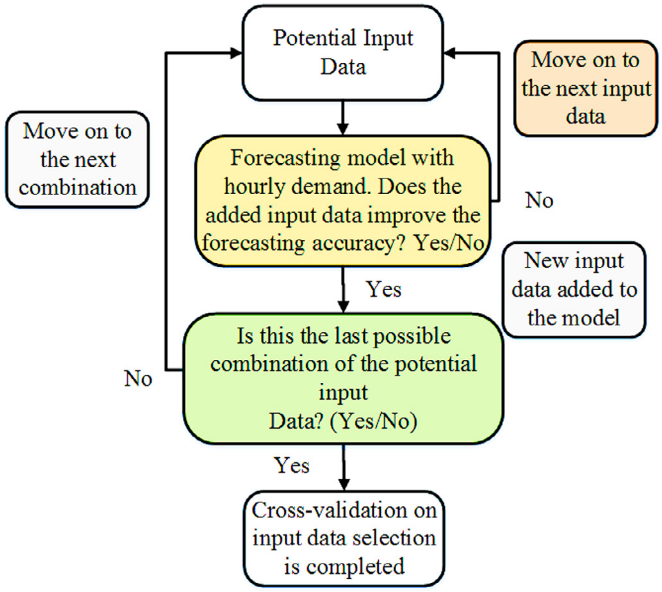

Fig. 7. Data selection process [177].

Where C is the constant of regularization relative to the segment price of transgressions. ζ_t and ζ_t^* are the slack parameters that estimate the expense of the errors two before and subsequently the objective amount in the training circumstances.

The 'e-insensitive function of loss' with 2ε bandwidth is presents as:

$$|y_t - f(X_t)|\varepsilon = \begin{bmatrix} 0, \; if \; |y_t - f(X_t)| \le \varepsilon \\ |y_t - f(X_t)| - \varepsilon, \; otherwise \end{bmatrix} \tag{27}$$

The dual problem, then, is:

$$aximize - \frac{1}{2}\sum_{t,l=1}^{N}(\alpha_t - \alpha_t^*)\left(\alpha_l - \alpha_l^*\right)\langle \emptyset(X_t), \emptyset(X_l)\rangle$$
$$-\ \varepsilon\sum_{t=1}^{N}(\alpha_t - \alpha_t^*) + \sum_{t=1}^{N}y_t(\alpha_t - \alpha_t^*) \tag{28}$$

$$subject\ to \begin{bmatrix} \sum_{t=1}^{N}(\alpha_t - \alpha_t^*) = 0 \\ \alpha_t, \alpha_t^* \in [0, C] \end{bmatrix} \tag{29}$$

$$subject\ to \begin{bmatrix} \sum_{t=1}^{N}(\alpha_t - \alpha_t^*) = 0 \\ \alpha_t, \alpha_t^* \in [0, C] \end{bmatrix} \tag{30}$$

Where α_t, α_t^*, α_l^* and α_l are the 'Lagrange multipliers'. The concluding support vector machine for nonlinear functions can be formulated as

$$y_t = f(X_t) = \sum_{t=1}^{N}(\alpha_t - \alpha_t^*)K(X, X_t) + b \tag{31}$$

Where $K(X, X_t) = \langle \emptyset(X), \emptyset(X_t)\rangle$ depicts 'kernel function'. Radial basis Gaussian function (32) is the several redoubtable in nonlinear purpose estimation.

$$K(X, X_t) = \exp\left(-\frac{\|X - X_t\|^2}{2\sigma^2}\right) \tag{32}$$

By executing so, the selection of the data has shown Fig. 7, will ensure testing not only each particular component but each desirable sequence of variables that will increase to prediction the output results [177]. The ultimate input elements that become been acquired toward description for the suggested multiple support vector machine prediction pattern at every hour t are confined to energy hourly requirement at the period t, one day maximum electricity requirement, average demand of electricity, natural gas prices on daily basis, from last year's electricity prices number (1 to 4 depicting the rate zones: high, medium and low peak) at t, at the month (1 to 12) and as well as hour of the day (amount 1–24). Additionally, the data used for the target of an hour t and the prediction of price module is the power marketing clearing price at the time t.

2.4.2. Benching marking models

Energy benchmarking permits the energy achievement of the building sector to be weighed against the goal of efficient power control as well as the design of the building. Power benchmarking can elicit systematize consumption of power notices and determine the baseline of power performance in an assigned market. Management of energy is used to achieve or enhance the 'benchmarked level', and some benchmarking strategies that models utilize could be expanded to accommodate the control of the energy of the building environment, as described by Wen and Li [178]. The conception of power benchmarking for the building sector is not a novel one with its methodologies, ratings and even labelling processes evolving as time goes on. Lombaird et al. [179] have rendered a general consideration in identifying the fundamental variables that are critical for benchmarking of energy for buildings.

These findings can prove helpful for comprehending the basic idea of benchmarking models. The research pattern in introducing the benchmark can be distorted, alike 'normalization', or in further refined procedures, like engineering and machine modelling [180,181]. ML techniques in the form of support vector machine applied to the different kind of regression problems. Support vector machine is remarkably useful for interpreting nonlinear predicaments when using a little sample intensity for the training data [182]. The engineering modelling method is not as a standard for the activities of benchmarking. Nevertheless, the deployment of this system is completely examined and described by Torceillini et al. [183] and Fumio et al. [184], also the process is moreover interpreted from Yaan et al. [185]. The accessories for the building energy simulations of the different patterns are immediately accessible, Crawley et al. [186] has rendered a thoroughgoing investigation of certain accessories. Energy modelling, database tools are prepared by the department of energy, United State of America home page [187]. The apparatus picked from this research is (EnergyPlus) language or software, authorized and developed by the department of energy [188].

An assortment of power 'benchmarking' practices has been described and is circumscribed by necessitates and outcomes. The restraints can be found as the approachability of electricity usage data, the size of the data as well as coverage of data [155]. From this study, utilizes used three techniques in energy forecasting, which are: (a) machine learning; (b) engineering modelling; and (c) statistical modelling. These approaches can be classified as bottom-up a strategy where separated data are utilized. The bottom-up approaches signals comprehensive sets of evidence, along with refined data retrieval [189]. The decided benchmarking techniques are not exclusive to the energy consumption sets of data. Different other approaches can be used. For illustration, the principal component regression technique can be utilized to compensate resonance influence in the selected dataset [190]. Nevertheless, the ways expanded in this research illustrate low-cost classifications employing 'OLS' to the highly assiduous benchmarking method of engineering modelling.

Sample set can be defined as $\{(X_i, Y_i)\}\cdot i_1^N$ wherever N total sample, X_i describes the power usage of the ith specimen and Y_i is the normalised power consumption of the ith unit. The association of

X_i to Y_i is given by:

$$Y = W.\emptyset(X) + b \tag{33}$$

The support vector machine depreciates the distinction among the exact value Y and the forecasting value $f(X)$. This is accomplished by depreciating the regularized function as:

$$minimize; \left(\frac{1}{2}\right)\|w\|^2 + \left(\frac{C}{N}\right)\sum_{i=1}^{N} L_\epsilon(Y_i.f(X_i)) \tag{34}$$

The second-term ascertains the error of the prognostication and it is practiced the values $L_\epsilon(Y_i.f(X_i) = 0$ if $|Y_i - f(X_i)| \le \in$ and $L\in(Y_i, f(X_i)) = |Y_i - f(X_i) - \in|$ for others. $\in$ is referred to as the radius. The values for W and b are received by employing the 'Lagrangian multiplier' approach. Hence,

$$Y = \sum_{i=1}^{N}(\alpha - \bar{\alpha})^T K(X_i, X) + b \tag{35}$$

$K(X_i, X) = \emptyset(X_i)^T\emptyset(X)$ depicts the function of kernel. In this review, uses libsvm, account 2.60 in 'R' programming [191].

2.4.3. Energy mapping models

A novel strategy to assemble as well as train ANNs was recently disclosed, which is void of such limitations. The modern channels are denominated support vector machine [192]. SVM energy mapping models enlist a special training algorithm that augments the dividing margin between two classes of data. Sensor-based energy forecasting has been researched for both commercial and residential buildings. However, due to the lack of information regarding residential buildings, a large share of prior work has been focused on energy consumption predictions in commercial buildings. Machine learning algorithms form the foundation for sensor-based energy mapping models are presented in reference [193–197]. SVR refers to the latest version of the support vector machine for 'regression' evaluation. It was introduced in [34]. Prior work [198,199] has demonstrated that SVR outperforms ANN methods when predicting building energy loads associated with cooling.

In this part, depicts an extensive overview of the general model employed for the selection and evaluation of energy mapping model selection, as shown in Fig. 8 [200]. The algorithm is general in that it is invariant to the ephemeral (spatial) granularity of the input dataset, D, to preclude errors resulting from inconsistent procedures across scales [200]. In this figure, the flowchart is usual in the knowledge which is not changed to the spatial and ephemeral of the data set by input D, to impede mistakes or errors originated into irregular methods beyond computations. Substantially, in this review resembled the measurements regarding D to relevant time-scale (every 10 min, hourly, daily, i.e.,), and through determining the pair of different houses and families inside the house $f = B$ that would be employed to forecast the total house power consumption. Based on the unit-scale, every scale-unit formed like an individualized home as well as floor measurement, and every house floor is constructed or modeled such as individual groups, whereas at the entire home measure the response is displayed as a one substance $f = B$. Further, the selection of variable D is conducted individually for any illustrated house (family) f. The model to the variables assortment is rendered in Fig. 8. Subsequent selection of variables is established for house consists of a family f; methods are commonly evaluated and prepared during the process of bootstrapping that provides limitations in model confidence efficiency and performance as estimated by the coefficient of variation amount. To every house, the sets of the data are begun N intervals. Every bootstrap started D of I is divided into a different validation set V and training set T that are estimated 30% and 70% of D respectively.

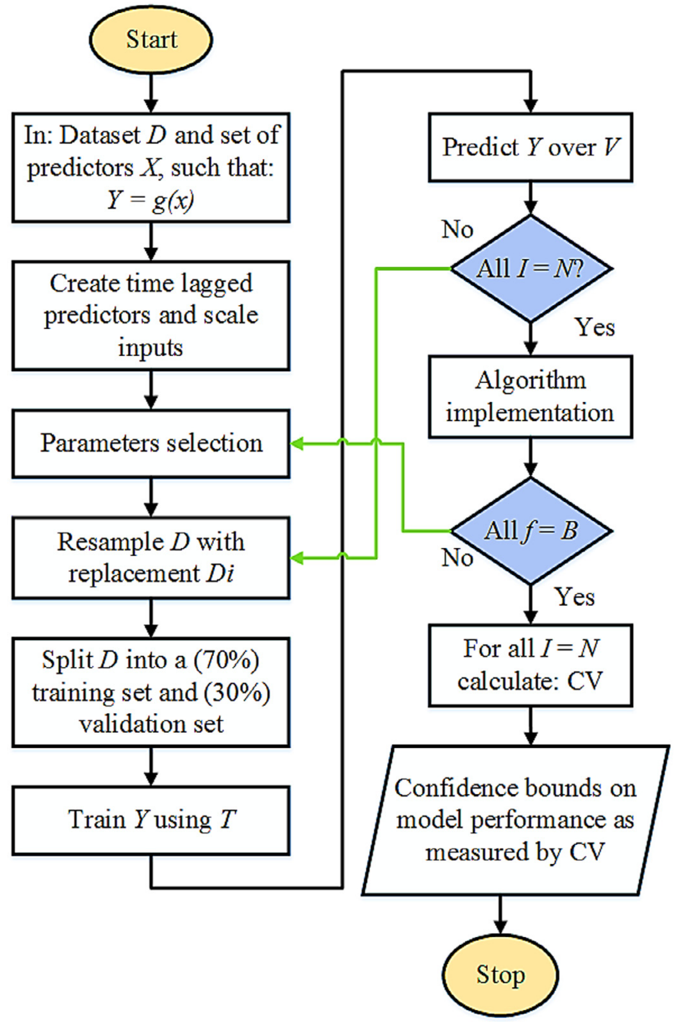

Fig. 8. A general algorithm used for model selection and validation [200].

We define two sets of model inputs: M_1 for validation of the 'Great Energy Forecasting Shootout' data and M_2 for testing on our empirical data set. M_1 is consistent with inputs presented in [201]and is defined as the following:

$$M1 = \overrightarrow{x(t)} = [y(t-1.1), y(t-2.1), T(t)sh, s, ch] \tag{36}$$

Wherever $y(t{-}1.1)$ and $y(t{-}2.1)$ describe identified power usage amounts for the former two-time steps, $T(t)$ represents present temperature, $S(t)$ shows the flux of solar, s is a symbol parameter that denotes weekend/holiday or weekday, sh is the sine of the present hour (time) and ch is the cosine of the present hour. Based on previous work [202,203] and a lack of solar flux data, we modify the model inputs as $M2$ for testing on the Watt Hall data set as:

$$M2 = \overrightarrow{x(t)} = [y(t-1.1), y(t-2.1), T(t), sh, s, ch] \tag{37}$$

Table 8 demonstrates the brief discussion of discussion of SVM based data driven models.

3. Building sector energy forecasting using large scale based approaches

Grey-box based models, White-box based models and black-box based models can direct various large scale residential power associated predicaments, e.g., forecasting and prediction, mapping urban energy consumption, profiling, benchmarking, as well as first

Table 8
Summary of SVM based data-driven models.

Model category	SVM based data-driven models
Forecasting prediction models	
Prediction model	Self-organising-map (SOM) network [176], SVM models [164], Autoregressive integrated moving average (ARIMA) [176], Independent component analysis-support vector machine [170], Time series [164], Markov models (IOHMM) [176], Architecture of the hybrid network [176], Genetic Algorithm based optimization, termed GA LS-SVM [168], Time-series analysis [176], Wavelet transform [168], Box-Cox mathematical [176], Regression Modeling of Support Vector Machine [170], Machine-based price-forecasting model [165], Engineering modelling [155]
Software	Matlab, CFD, TrnSys, EnergyPlus
Application and usage	Function approximation and regression estimation, Future energy demand estimation, Industrial and commercial load demand prediction,
Advantages	Eliminating redundant information and other areas to reduce training data of SVM, Processing of large-scale data, Overcome the shortage of SVM algorithm
Benching marking models	
Prediction model	Bottom-up building stock models [189], Piecewise Linear Regression [182], Principal Component Regression [190], Multi-level assessment method [185], Econometric top-down models [189], Bottom-up approach [189], Residual-clustering approach [190], Multiple linear regression analysis [190], Daylight prediction models [186], Hybrid approach [185], Degree-day method [184], Radial basis function neural network (RBFNN) [180], Frequency domain spectral density analysis [178]
Software	EnergyPlus, Matlab, Trnsys,
Application and usage	Building energy simulation, Used particularly in old buildings where the data availability is often problematic
Advantages	Extensive range of electricity usage features with an appropriate variation and selection of the energy variables, Adopts the structure risk minimization, Employed by conventional neural networks, Seeks to decrease an lower and upper bound by the error, Valid and effective or approaching overall target regression and classification problems, Show a linear regression and non-linear mapping, Reduce air infiltration and rate of heat loss
Energy mapping models	
Prediction model	Box-Jenkins [196], ANN model [194], Support vector machines for regression [193], Radial Basis Function [194], Empirical risk minimization (ERM) principle method[193], Sparse model [194], Chaotic GASA algorithm [196], SVR models [194], Transfer function method [198], ARIMA [196], Knowledge based expert system (KBES) approach [196], Bayesian estimation model [196], Chaos-based searching algorithms [196], Vapnik-Chernoverkis (VC) theory [199], Genetic algorithm (GA) [196], Support Vector Regression (SVR) [200], Machine learning [201], Single-step forecasting model [200], Fuzzy C-Means [201], Adaptive Neuro Fuzzy Inference System (ANFIS) [201], Non-linear regression models [202], Least Squares Support Vector Machine (LS-SVM) [201], Data-driven method [203], Genetic algorithm-adaptive network-based fuzzy inference system (GA-ANFIS) [203]
Software	EnergyPlus, IDA-ICE, ESP-r , Clim2000, Trnsys, Matlab
Application and usage	Non-linear efficiency of weather data and building energy performance, SVM is achievable and appropriate in prognostication of periodically landlord disco's bills, Employed by conventional neural network, Building heating and cooling load demand prediction
Advantages	Minimization of structure risk, SVM endeavours to decrease an lower and upper bound of the error, Useful for directing general target regression and classification problems, Ground-up electricity calculation, Very simple information by inputs, Embodiment of socioeconomic and macroeconomic impacts, Ascertainment of every end use power usage by variety, rating, etc., Compasses inclinations

clustering interpretation, all practised variables' determination in which to develop prototypes. The research reviews of the models are furnished below.

3.1. Energy forecasting white box data mining based approaches

The improvement of white-box techniques for an intact planned residential building asset would demand a substantial measure of time to simulate the individual's building behavior. Besides, it would require extensive data collection to ascertain that the calculate the building energy response are adequately detailed. The city areas power building responsibility is specified as a white box approaches to simulate the building energy pattern, described archetypes; these are produced later an explicit and precise association of the various general physical properties of diverse assortments of homes that exhibit commonalities [204–206].

This scheme investigates a smaller part of the behavior of building energy stock, providing building energy management in briefly, which furthermore maintains the capability to discriminate the building energy efficiency of the intact portfolio satisfactorily. Further to the point, the order from building energy stock with characteristic erections recognized for the development of correct (benchmarking algorithms) at a district or local based level. The different benefit of this proposition lies in the probability to evaluate mysterious retrofit electricity saving measures and associating, potential situations for the remaining city [207].

3.2. Grey box data mining approaches

Grey box methods or models [208–211] integrate past knowledge with physical approach information derived from the energy consumption data sources. Grey box algorithms are typically comprised of a hybrid structure assimilating data-driven approaches as well as first-principle physics. They include advantages as well as shortcomings of the black and white and box methods. In the greater number of big or large scale techniques, the residential building property is epitomized according to an analogy of an energy path, where a decreased succession 'Resistance-Capacitance (RC)' circuit is capable of establishing and specifying the power performance of the building sector [212].

3.3. Black box data mining based approaches

An extensive amount of black box approaches applied building level energy consumption investigation predicts the electricity usage behavior at a for the LS of buildings instead single building. And although black box approaches obtain extensive use of power consumption, according to the assortment of hierarchically significant data inputs [213–215], different examples of the applicability of these strategies in large-scale presented in reference [216,217]. The conventional widely used the black-box methods for forecasting and prediction at building sector are [218]: MLR, the model of simple regression, ANN, Decision Trees (DT), and support vector machine. Data driven based approaches are dependent upon the accessibility of previous electricity usage data to predict the energy efficiency and performance.

Based on this challenge, it is essential to develop a pattern to train the different methods. Another issue arises, however, whereby privacy of the data policy and financial affairs result in difficulty with the collection process, often diminishing the quality of decisions. Geographic erudition arrangement is emerging as an

Table 9
Summary of building sector energy forecasting using large scale based approaches.

Model category	Building sector energy forecasting using large scale based approaches
Energy forecasting white box data mining based approaches, Grey box data mining based approaches, Black box data mining based approaches	
Prediction model	Archetypal model [206], Back Propagation Neural Networks(BPNN) [215], Grid-based methods [214], Thermal model [208], Simplified model [212], Data driven (black-box) method [210], Simplified engineering [215], Regression models [206], FCM models [214], Linear model [210], Hierarchical clustering method [214], The mathematical models [208], NARX models [217], Parametric archetypal model [206], Conventional GMDH method [215], Ant colony optimization algorithm [214], Steady-state models [212], Explicit Euler methods [208], lumped parameter models [210], CDA method [218], Fourier series Methods [208], Empirical approach [224], Geographic Information System (GIS) techniques [220], City-oriented simulation tools [212], Elaborate engineering methods [215], K-means algorithm [214], Support vector machines (SVM) [213], Two-node model [208], Model-based methods [214], Nodal network model [212], Bottom-up models [218], Conditional demand analysis [218], Thermal network models [211], Moore's method [212], The physical model [208], ARX-models [210], Linear ARX model [217], Two conditional parametric models [210], Nonlinear ARX model [217], Full-knowledge physical model [212], Linear and time-invariant system [212], Empirical risk minimization (ERM) principle [213], Engineering-based approaches [222], Fuzzy c-means(FCM) [214], Hybrid method [215], Hierarchical methods [214], Partitioning methods [214], Statistical approaches [224], Density-based methods [214], Iterative refinement clustering [214], Hybrid method of Group Method of Data Handling (GMDH) [215], Least Square Support Vector Machine (LSSVM) [215], Black-box approaches [217], Artificial Intelligence Methods [215], Nonlinear Autoregressive Model with Exogenous Inputs) [217], Data mining algorithms [215], GA model [218], GIS-based statistical downscaling approach [222], Integration of system dynamics models [224]
Software	TRNSYS, eQuest, IES and ESP-r, Energy Plus, DOE-2.1E, ESP-r, MATLAB, SPSS, R
Application and usage	Building energy simulation applications, Predict energy demand, Application in forecasting analysis provides promising results for forecasting building electricity usage, Uses simple survey information and billing data, Encompasses trends
Advantages	Simple, efficient and scalable, Easy to implement, Strong ability of antinomies, Can recognize the numerous vital features including self-stability, Can be implemented without any problem, Network can execute the task that a linear program cannot, Can be executed in any application, Large scale energy forecasting models can learn and no need to be programmed, Long-term prediction in the inadequacy of each dis-continuity, Comprises inhabitant response, Addition of socioeconomic and macroeconomic impacts, Inclusion of macroeconomic and socioeconomic effects, Determination of typical end-use energy contribution, Simple input information, Ascertainment of qualities end-use comprises on simulation, Ascertainment of every consumer's electricity usage by raging, type, etc.,

increasingly significant source in which produce large scale energy methods. This is because their capabilities of and visualizing and allocating as presented in references [219–222].

Although challenges continue here as well since a limited number of GIS databases include appropriate for grasping the energy consumption achievement of a town or city. Other pertinent data sets comprise: census [223], national reserves [224], standardizing [224], local and national surveys [225], inquiries [219] and environmental data. Latterly, different novel erudition gathering techniques like as masses energy data sourcing become emerged for developing and populating entire-city databases [226–229]. New and different knowledge need to be concentrated, however, which is dependent upon the energy prediction methodologies. Here the most relevant elements are: development limit, employment (surface/volume), floors quantity, orientation, surface/glazing factor, solar amount of different floors, solar shading, orientation, windows area electricity usage at aggregated level or building level, electricity measures erudition, and environmental data. Table 9 shows the summary of building sector energy forecasting using large scale based approaches.

4. Conclusion

Numerous attempts have been enlisted to compose large scale building power usage approaches at varying levels of functionality and granularity. A host of various approaches has been created with this intention. Data driven based approaches and large scale building energy prediction models to offer an adequate equivalence among decreasing the time it takes to develop the method at the time of sustaining an acceptable accuracy level. In this review, building energy demand prediction models has been reviewed in detail where it was discovered that each country is engaged in comprehensive electricity preparation and planning for its sustainable advancement. The illustrated methods in this review paper are widely employed in the building sector, but their applicability at large scale yields fewer results. The study, which is summarized in two main categories, is:

1. Data-driven based approaches.
2. Large scale based building energy forecasting approaches.

The data driven based approach is further divided into four types: (a) artificial neural networks; (b) statistical and machine learning models; (c) artificial neural networks; (d) support vector machine; (e) clustering equally to base models. Each type includes four sub-categories: (i) energy forecasting; (ii) energy mapping; (iii) benchmarking; (iv) energy profiling. It has been discovered that aggregates of several approaches are applied at generous scale for forecasting the future electricity usage and to compensate for the scarcity of existing data in the form of energy consumption and climatic parameters. The evaluation approaches revealed that some famously observed data driven based approaches, such as support vector machine and artificial neural network are quiet not employed at large scale which could otherwise present comprehensive knowledge on the consumption of building power stock. This is understandable acknowledging the tremendous level of specification of particular building data which is necessary to create these patterns.

From the earlier analysis and description, it's understood that a substantial amount of estimation and calculations are required to estimate energy efficiency and performance in the building environment, from building level to sub-system level and also national and regional level. From Individually algorithms, contains its benefits in several circumstances of assiduousness. The large-scale energy production models show significant variations. Various other remunerations can be associated with expanding this algorithm. It is observed that data-driven based approaches which are appropriate for precise and accurate estimates. In contradiction, by assuming significant policies, access to improve while sustaining the model accuracy. Artificial neural networks and support vector machine are exquisite at determining the non-linear obstacles, executing them very appropriate to forecasting energy in the building environment. They can provide the profoundly realistic prognostication, and selection of model with variable setting is adequately implemented. Support vector machine shows, still further, higher prediction performance than ANN.

At district level, data driven structure necessitates a substantial infiltration of electricity metering arrangements and circumstances in which to analyze single energy consumption data of the entire building assets; certain contingencies are however not readily available. It is established that the approaches associated power, economics and climate conditions for sustainable planning the future energy utilization. It is anticipated that before-mentioned approaches will accommodate electricity planners to accurately and precisely outline for the future aspects. Additional research on the plausibilities of data driven, large scale based approaches is needed in the direction to explain such significant issues as (i) whereby is it conceivable to produce accurate erudition for buildings described by various patterns of accessible energy consumption data?

Acknowledgements

The author appreciatively acknowledges the support of 'National Natural Science Foundation of China' (Grant 51576074 and 51328602).

References

[1] W. Zeyu, S.S. Ravi, A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models, Renew. Sustain. Energy Rev. 75 (2017) 796–808 75.
[2] T. Catalina, V. Iordache, B. Caracaleanu, Multiple regression model for fast prediction of the heating energy demand, Energy Build. 57 (2013) 302–312.
[3] S.S. Kwok, R.K. Yuen, E.W. Lee, An intelligent approach to assessing the effect of building occupancy on building cooling load prediction, Build. Environ. 46 (8) (2011) 1681–1690.
[4] D. Chirag, Z. Fan, Y. Junjing, E.L. Siew, W.S. Kwok, A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models, Renew. Sustain. Energy Rev. 78 (2017) 796–808.
[5] B. Yildiz, J.I. Bilbao, A.B. Sproul, A review and analysis of regression and machine learning models on commercial building electricity load forecasting, Renew. Sustain. Energy Rev. 73 (2017) 1104–1122.
[6] J. Phil, P. Joanne, L. Simon, Modelling the built environment at an urban scale-Energy and health impacts in relation to housing, Landscape Urban Plann. (2007) 39–49.
[7] A. Foucquier, S. Robert, F. Suard, L. Stephan, A. Jay, State of the art in building modelling and energy performances prediction: a review, Renew. Sustain. Energy Rev. 23 (2013) 272–288.
[8] L. Suganthi, A.A. Samuel, Energy models for demand forecasting-a review, Renew. Sustain. Energy Rev. 16 (2012) 1223–1240.
[9] V.S.K.V. Harish, A. Kumar, A review on modeling and simulation of building energy systems, Renew. Sustain. Energy Rev. 56 (2016) 1272–1292.
[10] A.S. Ahmad, M.Y. Hassan, M.P. Abdullah, H.A. Rahman, F. Hussin, H. Abdullah, et al., A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (2014) 102–109.
[11] A. Foucquier, S. Robert, F. Suard, L. Stéphan, A. Jay, State of the art in building modelling and energy performances prediction: a review, Renew. Sustain. Energy Rev. 23 (2013) 272–288.
[12] C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, MarilynA. Brown, RamM. Pendyala, Machine learning approaches for estimating commercial building energy consumption, Appl. Energy 08 (2017) 889–904.
[13] Z. Li, G. Huang, Re-evaluation of building cooling load prediction models for use in humid subtropical area, Energy Build. 62 (2013) 442–449.
[14] Y. Zhang, Z. Neill, B. Dong, G. Augenbroe, Comparisons of inverse modelling approaches for predicting building energy performance, Build. Environ. 86 (2015) 177–190.
[15] M. Leung, N.C. Tse, L. Lai, T. Chow, The use of occupancy space electrical power demand in building cooling load prediction, Energy Build. 55 (2012) 151–163.
[16] J.S. Chou, D.K. Bui, Modeling heating and cooling loads by artificial intelligence for energy-efficient building design, Energy Build. (2014) 437–447.
[17] R.E. Edwards, J. New, L.E. Parker, Predicting future hourly residential electrical consumption: a machine learning case study, Energy Build. (2012) 591–603.
[18] R.Z. Jovanović, A.A. Sretenović, B.D. Živković, Ensemble of various neural networks for prediction of heating energy consumption, Energy Build. 91 (2015) 189–199.
[19] J. Massana, C. Pous, L. Burgas, J. Melendez, J. Colomer, Short-term load forecasting in a non-residential building contrasting models and attributes, Energy Build. 92 (2015) 322–330.
[20] D. Monfet, M. Corsi, D. Choiniere, E. Arkhipova, Development of an energy prediction tool for commercial buildings using case-based reasoning, Energy Build. 81 (2014) 152–160.
[21] R. Platon, V.R. Dehkordi, J. Martel, Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis, Energy Build. 92 (2015) 10–18.
[22] R. Mena, F. Rodríguez, M. Castilla, M. Arahal, A prediction model based on neural networks for the energy consumption of a bioclimatic building, Energy Build. 82 (2014) 142–155.
[23] X. Lv, T. Lu, C.J. Kibert, M. Viljanen, A novel dynamic modeling approach for predicting building energy performance, Appl. Energy 114 (2014) 91–103.
[24] E.G. Escrivá, C.B. Álvarez, C.B. Roldán, O.M. Alcázar, New artificial neural network prediction method for electrical consumption forecasting based on building end-uses, Energy Build. 43 (2011) 3112–3119.
[25] J. Wu, J. Wang, H. Lu, Y. Dong, X. Lu, Short term load forecasting technique based on the seasonal exponential adjustment method and the regression model, Energy Conversion Manag. 70 (2013) 1–9.
[26] F.W.H. Yik, J. Burnett, I. Prescott, Predicting air-conditioning energy consumption of a group of buildings using different heat rejection methods, Energy Build. 33 (2) (2001) 151–166.
[27] B. Howard, L. Parshall, J. Thompson, S. Hammer, J. Dickinson, V. Modi, Spatial distribution of urban building energy consumption by end use, Energy Build. 45 (2012) 141–151.
[28] C. Yang, S. Létourneau, H. Guo, Developing data-driven models to predict BEMs energy consumption for demand response systems, Lecture Notes Comput. Sci. 8481 (1) (2014) 188–197.
[29] T. Hong, L. Yang, D. Hill, W. Feng, Data and analytics to inform energy retrofit of high performance buildings, Appl. Energy 126 (2014) 90–106.
[30] M. Yalcintas, O.U. Aytun, An energy benchmarking model based on artificial neural network method utilizing US commercial buildings energy consumption survey (CBECS) database, Int. J. Energy Res. 31 (4) (2007) 412–421.
[31] M. Yalcintas, An energy benchmarking model based on artificial neural network method with a case example for tropical climates, Int. J. Energy Res. 30 (14) (2006) 1158–1174.
[32] C. Fan, F. Xiao, S. Wang, Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques, Appl. Energy 127 (2014) 1–10.
[33] EU directive 95/46/EC, 2012/148/EU2014/724/EU: Commission Recommendation of 10 October 2014 on the Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems.
[34] H. Drucker, C.J. Burges, L. Kaufman, A. Smola, V. Vapnik, Support vector regression machines, Adv. Neural Inf. Process Syst. (1997) 155–161.
[35] P. Franses, R. Paap, Periodic Time Series Models, Oxford University Press, 2003.
[36] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, Singapore, 2006.
[37] T.G. Nikolaou, D.S. Kolokotsa, G.S. Stavrakakis, I.D. Skias, On the application of clustering techniques for office buildings' energy and thermal comfort classification, smart grid, IEEE Trans. 3 (4) (2012) 2196–2210.
[38] X. Gao, A. Malkawi, A new methodology for building energy performance benchmarking: an approach based on intelligent clustering algorithm, Energy Build. 84 (2014) 607–616.
[39] L. Arambula, P. Romagnoli, A. Gasparella, Selection of representative buildings through preliminary cluster analysis, International High-Performance Buildings Conference, 2014.
[40] M. Heidarinejad, M. Dahlhausen, S. McMahon, C. Pyke, J. Srebric, Cluster analysis of simulated energy use for LEED certified U.S. office buildings, Energy Build. 85 (2014) 86–97.
[41] N. Gaitani, C. Lehmann, M. Santamouris, G. Mihalakakou, P. Patargias, Using principal component and cluster analysis in the heating evaluation of the school building sector, Applied Energy 87 (6) (2010) 2079–2086.
[42] D. Hsu, Characterizing energy use in Ney York City commercial and multifamily buildings, ACEEE Summer Study on Energy Efficiency in Buildings, 2012.
[43] S. Banihashemi, G. Ding, J. Wang, Developing a hybrid model of prediction and classification algorithms for building energy consumption, Energy Proc. 110 (2017) 371–376.
[44] K. Fatai, L. Oxley, F.G. Scrimgeour, Modelingand forecasting the demand for electricity in New Zealand: a comparison of alternative approaches, Energy J. 24 (1) (2003) 75–102.
[45] C. Hamzac, Forecasting of Turkey's net electricity energy consumption on sector bases, Energy Policy 35 (3) (2007) 2009–2016.
[46] K.J. Baker, R.M. Rylatt, Improving the prediction of UK domestic energy-demand using annual consumption-data, Appl. Energy 85 (6) (2008) 475–482.
[47] S. Siitonen, M. Tuomaala, M. Suominen, P. Ahtila, Implications of process energy efficiency improvements for primary energy consumption and CO_2 emissions at the national level, Appl. Energy 87 (9) (2010) 2928–2937.
[48] M. Espinoza, C. Joye, R. Belmans, Short-term load forecasting, profile identification and customer segmentation: A methodology based on periodic time series, IEEE Trans. Power Syst. 20 (3) (2005) 1622–1630.
[49] I.P. Panapakidis, T.A. Papadopoulos, G.C. Christoforidis, G.K. Papagiannis, Pattern recognition algorithms for electricity load curve analysis of buildings, Energy Build. 73 (2014) 137–145.
[50] J.W. Taylor, Short-term load forecasting with exponentially weighted methods, IEEE Trans. Power Syst. 27 (1) (2012) 673–679.
[51] D.W. Bunn, E.D. Farmer, Comparative Models For Electrical Load Forecasting, Wiley, 1985.
[52] W. Mingwu, W. Dongfang, L. Jian, J. Hui, J. Juliang, A novel clustering model based on set pair analysis for the energy consumption forecast in China, Math. Problems Eng. 8 (2014) 191–242.
[53] M.W. Wang, J.L. Jin, Y.L. Zhou, Set Pair Analysis Based Coupling Methods and Applications, Science Press, Beijing, China, 2014.
[54] Dong Wang, AlistairG. Borthwick, Handan He, Yuankun Wang, Jieyu Zhu, Yuan Lu, Pengcheng Xu, Xiankui Zeng, Jichun Wu, Lachun Wang, Xinqing Zou, Jiufu Liu, Ying Zou, Ruimin He, A hybrid wavelet de-noising and Rank-Set Pair Analysis approach for forecasting hydro-meteorological time series, Environ. Res. 160 (2018) 269–281.

[55] European Union, Directive 2002/91/Ec of the European Parliament and of the Council on Energy Performance of Buildings, December (2002), 2002.
[56] Z.H. Zhan, J. Zhang, Y. Li, H.S.H. Chung, Adaptive particle swarm optimization, IEEE Trans. Syst. Man Cybernet. Part B 39 (Dec. 2009) 1362–1381.
[57] Jones, et al., Planning for a sustainable city: an energy and environmental prediction model, J. Environ. Plann. Manag. 43 (6) (2000) 855–872.
[58] ODPM, Reforming the Home Buyingselling Process in England and Wales Contents of the Home Information Pack a Consultation, HMSO, London, 2003.
[59] H. Zhao, F. Magoulès, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16 (2012) 3586–3592.
[60] Welsh Assembly Government, The Welsh Housing Quality Standard-Guidance for Local Authorities on the Assessment Process and Achievement of the Standard, Welsh Assembly Government, 2002.
[61] M. Munson, G. Jaskulski, C. Thomas, If these walls could think, IEEE PowerEnergy Mag. 9 (2011) 50–55.
[62] J.L. Mathieu, P.N. Price, S. Kiliccote, M.A. Piette, Quantifying changes in build-ing electricity use, with application to demand response, IEEE Trans. Smart Grid 2 (2011) 507–518.
[63] O.M. Masoso, L.G. Grobler, The dark side of occupants' behaviour on building energy use, Energy Build. 42 (2010) 173–177.
[64] G.J. Tsekouras, N.D. Hatziargyriou, E.N. Dialynas, Two-stage pattern recognition of load curves for classification of electricity customers, IEEE Trans. Power Syst. 22 (2007) 1120–1128.
[65] I.P. Panapakidis, M.C. Alexiadis, G.K. Papagiannis, Load profiling in the deregulated electricity markets: a review of the applications, in: Proc. of the 9th International Conference on the European Energy Market, 2012, pp. 1–8.
[66] G. Chicco, R. Napoli, F. Piglione, Comparisons among clustering techniques for electricity customer classification, IEEE Trans. Power Syst. 21 (2006) 933–940.
[67] J.J. López, J.A. Aguado, F. Martín, F. Munoz, A. Rodríguez, J.E. Ruiz, Hopfield––K-Means clustering algorithm: a proposal for the segmentation of electricity customers, Electric Power Syst. Res. 81 (2011) 716–722.
[68] A. Mutanen, M. Ruska, S. Repo, P. Järventausta, Customer classificationload profiling method for distribution systems, IEEE Trans. Power Delivery 26 (2011) 1755–1763.
[69] G. Chicco, I.S. Ilie, Support vector clustering of electrical load pattern data, IEEE Trans. Power Syst. 24 (2009) 1619–1628.
[70] C. Cortes, V.N. Vapnik, Support-vector networks, Machine Learn. 20 (1995) 273–297.
[71] P. Domingos, M. Pazzani, On the optimality of the simple Bayesian classifier under zero-one loss, Machine Learn. 29 (1997) 103–137.
[72] E. Carpaneto, G. Chicco, R. Napoli, M. Scutariu, Electricity customer classification using frequency-domain load pattern data, Int. J. Electric. Power Energy Syst. 28 (2006) 13–20.
[73] K. Moslehi, R. Kumar, A reliability perspective of the smart grid, Smart Grid, IEEE Trans. 57 (1) (2010) 64.
[74] J.W. Taylor, Triple seasonal methods for short-term electricity demand forecasting, Eur. J. Oper. Res. 204 (1) (2010) 139–152.
[75] H.S. Hippert, C.E. Pedreira, R.C. Souza, Neural networks for short-term load forecasting: a review and evaluation, power systems, IEEE Trans. 16 (1) (2001) 44–55.
[76] N. Sapankevych, R. Sankar, Time series prediction using support vector machines: a survey, Comput. Intell. Mag. 4 (2) (2009) 24–38.
[77] Y. Ji, J. Hao, N. Reyhani, A. Lendasse, Direct and recursive prediction of time series using mutual information selection, in: Computational Intelligence and Bioinspired Systems, Springer, 2005, pp. 1010–1017.
[78] D.C. Sansom, T.K. Saha, Energy constrained generation dispatch based on price forecasts including expected values and risk, in: Power Engineering Society General Meeting, 2004, pp. 261–266.
[79] B.J. Chen, M.W. Chang, et al., Load forecasting using support vector machines: a study on EUNITE competition, Power Syst. IEEE Trans. 19 (4) (2004) 1821–1830.
[80] K.W. Tri, H. Samuel, V. Matteo, A. Karl, Residential electricity load forecasting: evaluation of individual and aggregate forecasts.
[81] D. Wu, Y. Ke, J. Yu, P. Yu, L. Chen, Detecting leaders from correlated time series, in: H. Kitagawa, Y. Ishikawa, Q. Li, C. Watanabe (Eds.), Database Systems for Advanced Applications of Lecture Notes in Computer Science, Springer, Berlin Heidelberg, 2010, pp. 352–367.
[82] P.J.M., Empirical Analysis of Demand Response Baseline Methods. (April 2011) White Paper.
[83] The demand response baseline, White Paper (2011).
[84] R. Weron, A. Misiorek, Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models, Int. J. Forecast 24 (2008) 744–763.
[85] A.J. Conejo, J. Contreras, R. Espìnola, M. Plazas, Forecasting electricity prices for a day-ahead pool-based electric energy market, Int. J. Forecast 21 (2005) 435–462.
[86] F.J. Nogales, J. Contreras, A.J. Conejo, R. Espfnola, Forecasting next-day electricity prices by time series models, IEEE Trans. Power Syst. 17 (2) (2002) 342–348.
[87] K. Skytte, The regulating power market on the Nordic power exchange Nord Pool: an econometric analysis, Energy Econ. 21 (4) (1999) 295-08.
[88] S.E. Fleten, E. Pettersen, Constructing bidding curves for a price-taking retailer in the Norwegian electricity market, IEEE Trans. Power Syst. 20 (2) (2005) 701-08.
[89] M. Olsson, L. Söder, Modeling real-time balancing power market prices using combined SARIMAMarkov processes, IEEE Trans. Power Syst. 23 (2) (2008) 443–450.
[90] S. Jaehnert, H. Farahmand, G.L. Doorman, Modelling of prices using the volume in the Norwegian regulating power market, IEEE PowerTech, Bucharest, 2009.
[91] M.O. Brolin, L. Söder, Modeling Swedish real-time balancing power prices using nonlinear time series models, in: IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, 2010, pp. 358–363.
[92] T.K. Boomsma, N. Juul, S.E. Fleten, Bidding in sequential electricity markets: the Nordic case, No.6, Stochastic Programming E-Print Series, Institute fur Mathematic, 2013.
[93] K. Gro, L. Anders, S.E.F. Eriksrud, Benchmarking time series based forecasting models for electricity balancing market prices, Energy Syst. 6 (2015) 43–61.
[94] L. Swan, V. Ugursal, Modeling of end-use energy consumption in the residential sector: a review of modelling techniques, Renew. Sustain. Energy Rev. 13 (2009) 1819–1835.
[95] E. Hirst, A model of residential energy use, Simulation 30 (3) (1978) 69–74.
[96] G.P. Saha, J. Stephenson, A model of residential energy use in New Zealand, Energy 5 (2) (1980) 167–175.
[97] J. Bentzen, T. Engsted, A revival of the autoregressive distributed lag model in estimating energy demand relationships, Energy 26 (1) (2001) 45–55.
[98] E. Mata, A.S. Kalagasidis, F. Johnsson, A modelling strategy for energy, car-bon, and cost assessments of building stocks, Energy Build. 56 (2013) 100-08.
[99] G.V. Fracastoro, M. Serraino, A methodology for assessing the energy performance of large scale building stocks and possible applications, Energy Build. 43 (2011) 844–852.
[100] W. Tian, R. Choudari, A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London, Energy Build. 54 (2012) 01–11.
[101] S. Guerra, L. Itard, H. Visscher, The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock, Energy Build. 41 (2009) 1223–1232.
[102] K.J. Baker, R.M. Rylatt, Improving the prediction of UK domestic energy-demand using annual consumption-data, Appl. Energy 85 (2008) 475–482.
[103] P. Caputo, G. Costa, S. Ferrari, A supporting method for defining energy strategies in the building sector at urban scale, Energy Policy 55 (2013) 261–270.
[104] I. Theodoridou, M. Karteris, G. Mallinis, A.M. Papadopoulos, M. Hegger, Assessment of retrofitting measures and solar systems potential in urban areas using geographical information systems: application to a Mediterranean city, Renew. Sustain. Energy Rev. 16 (2012) 6239–6261.
[105] S. Heiple, D.J. Sailor, Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles, Energy Build. 40 (2008) 1426–1436.
[106] B. Howard, L. Parshall, C. Thompson, S. Hammer, J. Dickinson, V. Modi, Spatial distribution of urban building energy consumption by end use, Energy Build. 45 (2012) 141–151.
[107] A.M. Rotterdam, B. Olivier, S. Francesca, L. Ulrich, Estimating energy savings for the residential building stock of an entire city: a GIS-based statistical downscaling approach applied to Rotterdam, Energy Build. 75 (2014) 358–367.
[108] DOE, Building energy data book, 2011 Available at http://buildingsdatabook.eren.doe.gov/. Accessed 07/10, 2013.
[109] W. Detlef, K. Scott, Energy consumption characteristics of commercial build-ing HVAC systems, in: Volume I: Chillers Refrigerant Compressors, and Heating Systems, 20, 2001, pp. 33745–33750.
[110] W. Detlef, K. Scott, Energy consumption characteristics of commercial build-ing HVAC systems, Volume Ii: Thermal Distribution, Auxiliary Equipment, and Ventilation, 1999.
[111] W. Grondzik, R. Furst, Hvac Components and Systems, Vital Signs Project, 2000.
[112] A. Standard, Standard for the design of high-performance green buildings except low-rise residential buildings, in: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, 2009, p. 189.
[113] ASHRAE, ASHRAE Handbook: Fundamentals, ASHRAE, Atlanta, GA, 2009.
[114] U.S. Energy Information Administration, Commercial buildings energy consumption survey - overview of commercial buildings characteristics, 2003 Available at http://www.eia.doe.gov/emeu/cbecs/cbecs2003/introduction.html. Accessed Dec, (2013).
[115] C. Duarte, D.K. Van, C.Rieger Wymelenberg, Revealing occupancy patterns in an office building through the use of occupancy sensor data, Energy Build. 67 (2013) 587–595.
[116] A. Standard, Energy standard for buildings except low-rise residential buildings, Am. Soc. Heating Refrigerating Air-Conditioning Engineers 90 (1) (2007).
[117] D.Yan C.Wang, Y. Jiang, A Novel Approach For Building Occupancy Simulation Building Simulation, 4, Tsinghua Press, Beijing, China, 2011.
[118] J. Page, D. Robinson, N. Morel, J. Scartezzini, A generalised stochastic model for the simulation of occupant presence, Energy Build. 40 (2) (2008) 83–98.
[119] D. Wang, C.C. Federspiel, F. Rubinstein, Modelling occupancy in single person offices, Energy Build. 37 (2005) 121–126.
[120] L.J. Lo, A. Novoselac, Localized air-conditioning with occupancy control in an open office, Energy Build. 42 (7) (2010) 1108–1120.
[121] Y. Zhu, M. Liu, T. Batten, H. Noboa, D.E. Claridge, W.D. Turner, Optimization of control strategies for HVAC terminal boxes, in: Proceedings of 12th Sym-

posium on Improving Building Systems in Hot and Humid Climates, San Antonio, 2000, pp. 278–284.
[122] G.P. Henze, C. Felsmann, G. Knabe, Evaluation of optimal control for active and passive building thermal storage, Int. J. Thermal Sci. 43 (2) (2004) 173–183.
[123] K. Deng, P. Barooah, G.P. Mehta, S.P. Meyn, Building thermal model reduction disaggregation of states, in: American Control Conference (ACC), Baltimore, MD, IEEE, 2010, pp. 5118–5123.
[124] F. Oldewurtel, A. Parisio, N.C. Jones, et al., Energy efficient building climate control using stochastic model predictive control and weather predictions, in: American Control Conference (ACC), IEEE, 2010, pp. 5100–5105.
[125] Y. Ma, G. Anderson, F. Borrelli, A distributed predictive control approach to building temperature regulation, in: American Control Conference (ACC), San Francisco, IEEE, 2011, pp. 2089–2094. (2011).
[126] X.N. Truong, G.J. Pappas, Receding-horizon supervisory control of green buildings, in: American Control Conference (ACC), San Francisco, IEEE, 2011, pp. 4416–4421.
[127] A. Aswani, N. Master, J. Taneja, D. Culler, C. Tomlin, Reducing transient and steady state electricity consumption in HVAC using learning-based model –predictive control, Proc. IEEE 100 (1) (2012) 240–253.
[128] G. Gao, K. Whitehouse, The self-programming thermostat: optimizing setback schedules based on home occupancy patterns, in: Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, Berkeley, CA, ACM, 2009, pp. 67–72. November 04–06.
[129] Y. Zheng, B.G. Burcin, The coupled effects of personalized occupancy profile based HVAC schedules and room reassignment on building energy use, Energy Build. 78 (2014) 113–122.
[130] H.S. Hippert, C.E. Pedreira, R.C. Souza, Neural networks for short-term load forecasting: a review and evaluation, IEEE Trans. Power Syst. 16 (2001) 44–55.
[131] M.E. Baran, H.L. Willis, Modelling and forecasting hourly electric load by multiple linear regression with interactions, in: IEEE PES Gen. Meet., IEEE, 2010, pp. 1–8.
[132] L.P. Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy Build. 40 (2008) 394–398.
[133] E. Delzendeh, S. Wu, A. Lee, Y. Zhou, The impact of occupants' behaviours on building energy analysis: A research review, Renew. Sustain. Energy Rev. 80 (2017) 1061–1071.
[134] A. Foucquier, S. Robert, F. Suard, L.A. Stéphan, State of the art in building modelling and energy performances prediction: a review, Renew. Sustain. Energy Rev. 23 (2013) 272–288.
[135] J. Yang, H. Rivard, R. Zmeureanu, On-line building energy prediction using adaptive artificial neural networks, Energy Build. 37 (2005) 1250–1259.
[136] A.H. Neto, F.A.S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, Energy Build. 40 (2008) 2169–2176.
[137] S. Karatasou, M. Santamouris, V. Geros, Modeling and predicting building's energy use with artificial neural networks: methods and results, Energy Build. 38 (2006) 949–958.
[138] A.S. Ahmad, M.Y. Hassan, M.P. Abdullah, H.A. Rahman, F. Hussin, H. Abdullah, et al., A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (2014) 102–109.
[139] V. Aydinalp, U. Ismet, A.S. Fung, Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks, Appl. Energy 71 (2002) 87–110.
[140] M. Aydinalp, I.V. Ugursal, A.S. Fung, Modelling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks, Appl. Energy 79 (2004) 159–178.
[141] P.A. González, J.M. Zamarre, Prediction of hourly energy consumption in buildings based on a feedback artificial neural network, Energy Build. 37 (2005) 595-01.
[142] G. Schenker, Prediction Control Using Feedback Neural Networks and Partial Models, Swiss Federal Institute of Technology, 1996.
[143] A. Azadeh, S.F. Ghaderi, S. Sohrabkhani, Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors, Energy Conversion Manag. 49 (2008) 2272–2278.
[144] A. Azadeh, S.F. Ghaderi, S. Sohrabkhani, Forecasting electrical consumption by integration of Neural Network, time series and ANOVA, Appl. Math. Comput. 186 (2007) 1753–1761.
[145] R. Yokoyama, T. Wakui, R. Satake, Prediction of energy demands using neural network with model identification by global optimization, Energy Conversion Manag. 50 (2009) 319–327.
[146] R. Yokoyama, Capability of global search and improvement in modal trimming method for global optimization, JSME Int. J. Ser. C 48 (2005) 730–737.
[147] D. Chirag, S.E. Lee, Y. Junjing, S. Mattheos, forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks, Energy Build. 121 (2016) 284–297.
[148] N. Matson, M.A. Piette, Review of California and national methods for energy-performance benchmarking for commercial buildings, in: California Energy Commission, Public Interest Energy Research Program, LBNL, 2005, p. 57364.
[149] Yuan Y Liu, Min Yang, Malcolm Ramsay, Xiao S. Li, Jeremy W. Coid, A comparison of logistic regression, classification and regression tree, and neural networks models in predicting violent re-offending (2011) 27:547–573.
[150] T Sharp, Benchmarking energy use in schools, in: Proceedings of the ACEEE 1998 Summer Study on Energy Efficiency in Building, 3, 1988, pp. 305–316.
[151] T.W. Hicks, B.N. Von, An evaluation of America's first ENERGY STAR buildings: the class of 1999, in: Proceedings of the ACEEE 2003 Summer Study on Energy Efficiency in Buildings, 4, 2000, pp. 177–185.
[152] C. Federspiel, Q. Zhang, E. Arens, Model-based benchmarking with applications to laboratory buildings, Energy Build. 34 (3) (2002) 203–214.
[153] P. Mathew, D. Sartor, G.O. Van, S. Reilly, Rating energy efficiency and sustainability in laboratories: results and lessons from the Labs21 program, in: Proceedings of the ACEEE 2003 Summer Study on Energy Efficiency in Buildings, 4, 2003, pp. 321–329.
[154] Y. Melek, An energy benchmarking model based on artificial neural network method with a case example for tropical climates, Int. J. Energy Res. 30 (2006) 1158–1174.
[155] V. Shabunko, C.M. Lim, S. Brahim, Developing building benchmarking for Brunei Darussalam, Energy Build. 85 (2014) 79–85.
[156] S.S. Torbaghan, A. Motamedi, H. Zareipour, L.A. Tuan, Medium-term electricity price forecasting, in: Proceedings of North American power symposium, 2012, pp. 1–08.
[157] X. Yan, Electricity market clearing price forecasting in a deregulated electricity Market, Department of Electrical Engineering University, Saskatchewan, 2009.
[158] X. Yan, N.A. Chowdhury, Electricity market clearing price forecasting in a deregulated electricity market, in: Proceedings of the 11th international conference on probabilistic methods applied to power systems Singapore, 2010, pp. 36–41.
[159] X. Yan, N.A. Chowdhury, Electricity Market Clearing Price Forecasting in a Deregulated market: a Neural Network Approach, VDM Verlag, Germany, 2010.
[160] S.S. Torghaban, H. Zareipour, L.A. Tuan, Medium-term electricity market price forecasting: a data-driven approach, in: Proceedings of North American power symposium, 2010, pp. 1–7.
[161] W. Sun, J.C. Lu, M. Meng, Application of time series based SVM model on next day electricity price forecasting under deregulated power market, in: Proc International Conference Mach Learn Cyber, 2006, pp. 2373–2388.
[162] D.M. Zhou, F. Gao, X.H. Guan, Application of accurate online support vector regression in energy price forecast, in: Proceedings of the 5th World congress on intelligent control and automation, 2004, pp. 1838–1842.
[163] L.L. Hu, G. Taylor, H.B. Wan, M. Irving, A review of short-term electricity price forecasting techniques in deregulated electricity markets, in: Proceedings of the 44th international universities power engineering conference, 2009, pp. 1–5.
[164] R.A. Swief, Y.G. Hegazy, T.S. Abdel-Salam, M.A. Bader, Support vector machines (SVM) based short term electricity load–price forecasting, in: Proceedings of the IEEE Bucharest PowerTech, 2009, pp. 1–5.
[165] L.M. Saini, S.K. Aggarwal, A. Kumar, Parameter optimization using genetic algorithm for support vector machine-based price-forecasting model in National electricity market, IET Generation Trans. Distrib. 4 (2010) 36–49.
[166] W. Sun, J. Zhang, Forecasting day ahead spot electricity prices based on GASVM, in: Proc International Conference Internet Computing Sci Engineering, 2008, pp. 73–88.
[167] Y.G. Chen, G.W. Ma, Electricity price forecasting based on support vector machine trained by genetic algorithm, in: Proceedings of the 3rd international symposium on intelligent information technology application, 2, 2009, pp. 292–305.
[168] M.J Mahjoob, M. Abdollahzade, R. Zarringhalam, GA based optimized LS-SVM forecasting of short term electricity price in competitive power markets, in: Proceedings of the 3rd IEEE conference on industrial electronics and applications, 2008, pp. 73–88.
[169] D.S. Gong, J.X. Che, J.Z. Wang, J.Z. Liang, Short-term electricity price forecasting based on novel SVM using artificial fish swarm algorithm under deregulated power, in: Proceedings of the 2nd international symposium on intelligent information technology application, 2008, pp. 85–99.
[170] H. Zheng, L.Z. Zhang, L. Xie, X. Li, SVM model of system marginal price based on developed independent component analysis, in: Proc International Conference Power Systems Technology, 2004, pp. 1437–1440.
[171] Y. Wang, S.Q. Yu, Price forecasting by ICA-SVM in the competitive electricity market, in: Proceedings of the 3rd IEEE Conference on Industrial Electronics and Applications, 2008, pp. 314–319.
[172] J.Y. Tian, Y. Lin, Short-term Electricity price forecasting based on rough sets and improved SVM, in: Proceedings of the 2nd International Workshop on Knowledge Discovery and Data Mining, 2009, pp. 68–71.
[173] T. Wang, L.J. Qin, Application of SVM based on rough set in electricity prices forecasting, in: Proceedings of the 2010 International Conference on Environmental Science and Information Application Technology, 2010, pp. 317–320.
[174] L. Xie, H. Zheng, L.Z. Zhang, Electricity price forecasting by clustering-LSSVM, in: Proceeding International Power Engineering Conference, 2007, pp. 697–702.
[175] S. Fan, J.R. Liao, K. Kaneko, L.N. Chen, An integrated machine learning model for day-ahead electricity price forecasting, in: Proceeding Power System Conference Expos, 2006, pp. 1643–1649.
[176] S. Fan, C. Mao, L. Chen, Next-day electricity-price forecasting using a hybrid network, IET Generation Trans. Distrib. 1 (2007) 176–182.
[177] Y. Xing, A.C. Nurul, Mid-term electricity market clearing price forecasting: a multiple SVM approach, electrical power and energy systems, Electrical Power Energy Syst. 58 (2014) 206–214.
[178] X. Li, J. Wen, Building energy consumption on-line forecasting using physics based system identification, Energy Build. 82 (2014) 1–12.
[179] L.L. Prez, J. Ortiz, R. Gonzlez, I.R. Maestre, A review of benchmarking rating

and labelling concepts within the framework of building energy certification schemes, Energy Build. 41 (3) (2009) 272–278.
[180] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks, Energy Conversion Manag. 50 (1) (2009) 90–96.
[181] C. Filippin, Benchmarking the energy efficiency and greenhouse gases emissions of school buildings in central Argentina, J. Build. Environ. 35 (5) (2000) 407–414.
[182] R.E.E. Edwards, J. New, L.E. Parker, Predicting future hourly residential electrical consumption: a machine learning case study, Energy Build. 49 (2012) 591–603.
[183] P. Torcellini, M. Deru, M. Griffith, K. Benne, D.B. Crawley, DOE Commercial Building Benchmark Models, National Renewable Energy Laboratory, 2008.
[184] N. Fumo, P. Mago, R. Luck, Methodology to estimate building energy consumption using EnergyPlus benchmark models, Energy Build. 42 (12) (2010) 2331–2337.
[185] C. Yan, S. Wang, F. Xiao, A simplified energy performance assessment method for existing buildings based on energy bill disaggregation, Energy Build. 55 (2012) 563–574.
[186] B. Crawley, W. Hand, T. Kummert, Contrasting the capabilities of building energy performance simulation programs, Building Environ. 43 (2008) 661–673.
[187] USA Department of Energy, DOE-2, Available Simulation Tools, 2014, Available from: http://apps1.eere.energy.gov/buildings/toolsdirectory/subjects.cfm/pagename=subjects/pagenamemenu=wholebuildinganalysis/pagen mesubmenu=energy simulation (accessed 09.02.14).
[188] USA Department of Energy, EnergyPlus: energy simulation software, 2012 Available from http://apps1.eere.energy.gov/buildings/energyplus. (accessed 01.05.12).
[189] M.M. Kavgic, A.A. Mavrogiani, D.D. Mumovic, A.J. Summerfield, Z.Z. Stevanovic, Djurovic-M. Petrovic, A review of bottom-up building stock models for energy consumption in the residential sector, Building Environ. 45 (2010) 1683–1697.
[190] E. Wang, Z. Shen, K. Grosskop, Benchmarking energy performance of building envelopes through a selective residual-clustering approach using high dimensional dataset, Energy Build. 75 (2014) 10–22.
[191] D. Meyer, Support Vector Machines. The Interface to libsvm in Package e1071, FH Technikum Wien, Austria, 2014, Available from:https://cran.rproject.org/web/packages/e1071/vignettes/svmdoc.pdf (accessed 04.04.14).
[192] V. Vapnic, in: Statistical Learning Theory, Springer, New York, 1995, pp. 133–157.
[193] B. Dong, C. Cao, S.E. Lee, Applying support vector machines to predict building energy consumption in tropical region, Energy Build. 37 (2005) 545–553.
[194] K. Kavaklioglu, Modeling and prediction of Turkey's electricity consumption using support vector regression, Appl. Energy 88 (2011) 368–375.
[195] H.X. Zhao, F. Magoulès, Feature selection for predicting building energy consumption based on statistical learning method, J. Algor Comput. Technol. 6 (2012) 59–78.
[196] W.Y. Zhang, W.C. Hong, Y. Dong, G. Tsai, J.T. Sung, G.F. Fan, Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting, Energy 45 (2012) 850–858.
[197] M. Espinoza, J.A.K. Suykens, B.D. Moor, Fixed-size least squares support vector machines: a large scale application in electrical load forecasting, CMS 3 (2006) 113–129.
[198] K.W. Mui, L.T. Wong, Cooling load calculations in subtropical climate, Building Environ. 42 (7) (2007) 2498-7.
[199] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Applying support vector machine to predict hourly cooling load in the building, Appl. Energy 86 (2009) 2249–2256.
[200] K. Rishee, M.K. Smith, J.P. Culligan, E.J. Taylor, 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 123 (2014) 168–178.
[201] R.E. Edwards, J. New, L.E. Parker, Predicting future hourly residential electrical consumption: a machine learning case study, Energy Build. 49 (2012) 591-03.
[202] S. Karatasou, M. Santamouris, V. Geros, Modeling and predicting building's energy use with artificial neural networks: methods and results, Energy and Building 38 (2006) 949–958.
[203] K. Li, H. Su, J. Chu, Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: a comparative study, Energy Build. 43 (2011) 2893–2899.
[204] I. Ballarini, S.P. Corgnati, V. Corrado, N. Talà, Improving energy modeling of large building stock through the development of archetype buildings (2011) 2874–81.
[205] I. Ballarini, S.P. Corgnati, V. Corrado, Use of reference buildings to assess the energy saving potentials of the residential building stock: the experience of TABULA project, Energy Policy 68 (2014) 273–284.
[206] I. Korolija, L.H Marjanovic, Y. Zhang, V.I. Hanby, UK office buildings archetypal model as methodological approach in development of regression models for predicting building energy consumption from heating and cooling demands, Energy Build. 60 (2013) 152–162.
[207] G. Mustafaraj, D. Marini, A. Costa, M. Keane, Model calibration for building energy efficiency simulation, Appl. Energy 130 (2014) 72–85.
[208] J.H. Kämpf, D. Robinson, A simplified thermal model to support analysis of urban resource flows, Energy Build. 39 (4) (2007) 445–453.
[209] D.F.H. Robinson, J. Kämpf, P. Leroux, D. Perez, A. Rasheed, U. Wilke, CITYSIM: comprehensive micro-simulation of resource flows for sustainable urban planning, Proc Build. Simul. (2009).
[210] H.A. Nielsen, H. Madsen, Modelling the heat consumption in district heating systems using a grey-box approach, Energy Build. 38 (1) (2006) 63–71.
[211] M. Lauster, J. Teichmann, M. Fuchs, R. Streblow, D. Mueller, Low order thermal network models for dynamic simulations of buildings on city district scale, Build. Environ. 73 (2014) 223–231.
[212] E.J. Kim, G. Plessis, J.L. Hubert, J.J. Roux, Urban energy simulation: simplification and reduction of building envelope models, Energy Build. 84 (2014) 193–202.
[213] B. Dong, C. Cao, S.E. Lee, Applying support vector machines to predict building energy consumption in tropical region, Energy Build. 37 (5) (2005) 545–553.
[214] K.I. Zhou, S.I. Yang, C. Shen, A review of electric load classification in smart grid environment, Renew. Sustain. Energy Rev. 24 (2013) 103–110.
[215] A.S. Ahmad, M.Y. Hassan, M.P. Abdullah, H.A. Rahman, F. Hussin, H. Abdullah, A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (2014) 102–109.
[216] S. Humeau, T.K. Wijaya, M. Vasirani, K. Aberer, Electricity load forecasting for residential customers: exploiting aggregation and correlation between households, in: 2013 Sustainable Internet and ICT for Sustainability - Conference Proceeding, 2013, pp. 1–6.
[217] K.M. Powell, A. Sriprasad, W.J. Cole, T.F. Edgar, Heating, cooling, and electrical load forecasting for a large-scale district energy system, Energy 74 (2014) 877–885.
[218] L.G. Swan, V.I. Ugursal, Modeling of end-use energy consumption in the residential sector: a review of modeling techniques, Renew. Sustain. Energy Rev. 13 (8) (2009) 1819–1835.
[219] F. Ascione, R.F.M. De, F. Rossi, R. Fistola, M. Sasso, G.P. Vanoli, Analysis and diagnosis of the energy performance of buildings and districts: methodology, validation and development of Urban Energy Maps, Cities 35 (2013) 270–283.
[220] F. Ascione, M. Canelli, M. Sasso, G.P. Vanoli, Combined cooling, heating and power for small urban districts: an Italian case study, Appl. Thermal Eng. 71 (2) (2014) 705–713.
[221] S. Taylor, D. Fan, M. Rylatt, Enabling urban-scale energy modelling: a new spatial approach, Build. Res. Inf. 42 (1) (2013) 4–16.
[222] A. Mastrucci, O. Baume, F. Stazi, U. Leopold, Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam, Energy Build. 75 (2014) 358–367.
[223] P. Caputo, G. Costa, S. Ferrari, A supporting method for defining energy strategies in the building sector at urban scale, Energy Policy 55 (2013) 261–270.
[224] F. Ascione, M. Canelli, R.F. Masi, M. Sasso, G.P. Vanoli, Combined cooling, heating and power for small urban districts: an Italian case study, Appl. Thermal Eng. 71 (2) (2014) 705–713.
[225] K. Orehounig, G. Mavromatidis, R. Evins, V. Dorer, predicting energy performance consumption of a neighbourhood using building performance simulation, Energy (2011).
[226] GlasgowEnergyApp: http://www.iesve.com/news/article/New-Energy-App-Helps-Residents-Cut-Bills/4332?modal=1.
[227] T. Giovanni, K. Ruth, O. Mike, O. James, F. Donal, Data driven approaches for prediction of building energy, consumption at urban level, in: 6th International Building Physics Conference, IBPC, 2015, pp. 3378–3383. Energy Procedia 78 (2015).
[228] T. Ahmad, Non-technical loss analysisprevention using smart meters, Renew. Sustain. Energy Rev. 75 (2017) 573–589.
[229] T. Ahmad, H. Chen, J. Wang, Y. Guo, Review of various modeling techniques for the detection of electricity theft in smart grid environment, Renew. Sustain. Energy Rev. Part 3 82 (Feb. 2018) 2916–2933.
[230] D.B. Jani, M. Mishra, P.K. Sahoo, Application of artificial neural network for predicting performance of solid desiccant cooling systems – A review, Renew. Sustain. Energy Rev. 80 (2017) 352–366.
[231] K. Kulkarni, P.K. Sahoo, M. Mishra, Optimization of cooling load for a lecture theater in a composite climate in India, Energy Build. 43 (2011) 1573–1579.
[232] D.B. Jani, M. Mishra, P.K. Sahoo, Performance prediction of rotary solid desiccant dehumidifier in hybrid air-conditioning system using artificial neural network, Appl. Thermal Eng. 98 (2016) 1091-03.
[233] D.B. Jani, M. Mishra, P.K. Sahoo, Performance prediction of solid desiccant - vapor compression hybrid air-conditioning system using artificial neural network, Energy 103 (2016) 618–629.
[234] D.B. Jani, M. Mishra, P.K. Sahoo, Performance studies of hybrid solid desiccant–vapor compression air-conditioning system for hot and humid climates, Energy Build. 102 (2015) 284–292.
[235] V. Ok, A procedure for calculating cooling load due to solar radiation: the shading effect from adjacent or nearby buildings, Energy Build. 19 (1) (1992) 11–20.
[236] T. Sharp, Energy benchmarking in commercial office buildings, in: Proceedings of the ACEEE 1996 Summer Study on Energy Efficiency in Buildings, 4, 1996, pp. 321–329.