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

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
0378-7788/

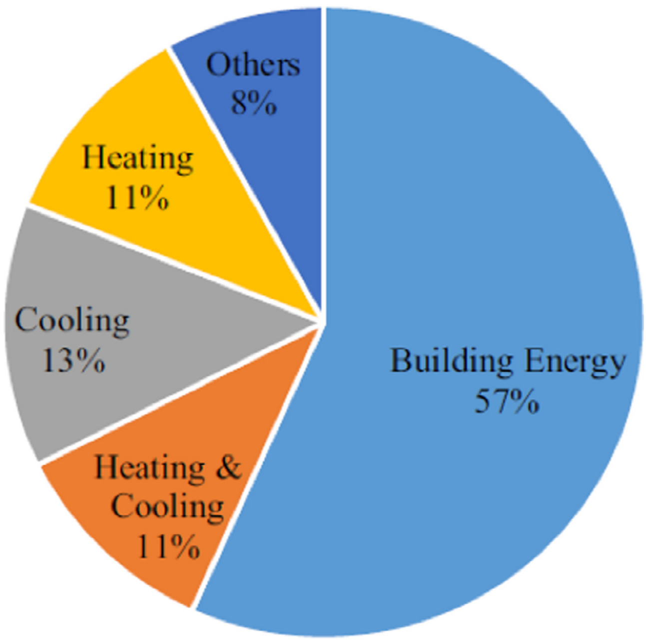

Fig. 1. The composition of energy type [1].

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