Applied Energy 172 (2016) 251–263

Contents lists available at ScienceDirect

Applied Energy

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

Short-term building energy model recommendation system: A meta-learning approach

Can Cui [a], Teresa Wu [a,*], Mengqi Hu [b], Jeffery D. Weir [c], Xiwang Li [d]

[a] School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, 699 S. Mill Ave., Tempe, AZ 85281, USA
[b] Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, 842 W. Taylor St., Chicago, IL 60607, USA
[c] Department of Operational Sciences, Air Force Institute of Technology, 2950 Hobson Way, Wright-Patterson Afb, OH 45433, USA
[d] Center for Green Buildings and Cities, Graduate School of Design, Harvard University, 20 Sumner Rd, Cambridge, MA 02138, USA

HIGHLIGHTS

- A Building Energy Model Recommendation (BEMR) system is proposed.
- BEMR has excellent performance on regular and extrapolation forecasting.
- Building data and physical characteristics features are both considered.
- BEMR is validated on real building, proving its generalizability.

ARTICLE INFO

Article history:
Received 13 December 2015
Received in revised form 16 March 2016
Accepted 28 March 2016

Keywords:
Building energy consumption
Time series forecasting
Recommendation system
Machine learning
Meta-learning
Feature reduction

ABSTRACT

High-fidelity and computationally efficient energy forecasting models for building systems are needed to ensure optimal automatic operation, reduce energy consumption, and improve the building's resilience capability to power disturbances. Various models have been developed to forecast building energy consumption. However, given buildings have different characteristics and operating conditions, model performance varies. Existing research has mainly taken a trial-and-error approach by developing multiple models and identifying the best performer for a specific building, or presumed one universal model form which is applied on different building cases. To the best of our knowledge, there does not exist a generalized system framework which can recommend appropriate models to forecast the building energy profiles based on building characteristics. To bridge this research gap, we propose a meta-learning based framework, termed Building Energy Model Recommendation System (BEMR). Based on the building's physical features as well as statistical and time series meta-features extracted from the operational data and energy consumption data, BEMR is able to identify the most appropriate load forecasting model for each unique building. Three sets of experiments on 48 test buildings and one real building were conducted. The first experiment was to test the accuracy of BEMR when the training data and testing data cover the same condition. BEMR correctly identified the best model on 90% of the buildings. The second experiment was to test the robustness of the BEMR when the testing data is only partially covered by the training data. BEMR correctly identified the best model on 83% of the buildings. The third experiment uses a real building case to validate the proposed framework and the result shows promising applicability and extensibility. The experimental results show that BEMR is capable of adapting to a wide variety of building types ranging from a restaurant to a large office, and gives excellent performance in terms of both modeling accuracy and computational efficiency.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

According to the U.S. Energy Information Administration (EIA), buildings consume nearly half (48%) of the total energy and produce almost 45% of CO_2 emissions in the United States [1]. This drives the need to develop high-fidelity and computationally efficient energy forecasting models for building systems to ensure

* Corresponding author at: Room 308, 699 S. Mill Ave., Tempe, AZ 85281, USA.
E-mail addresses: ccan1@asu.edu (C. Cui), Teresa.Wu@asu.edu (T. Wu), mhu@uic.edu (M. Hu), Jeffery.Weir@afit.edu (J.D. Weir), xiwang_li@gsd.harvard.edu (X. Li).

http://dx.doi.org/10.1016/j.apenergy.2016.03.112
0306-2619/© 2016 Elsevier Ltd. All rights reserved.

optimal automatic operation, reduce energy consumption, and improve the building's resilience capability to power grid disturbances [2]. Existing building energy models are in general categorized as: physics-based models, hybrid models and data-driven models [3]. Physics-based models employ the physical concepts and knowledge of the low level devices and aggregate the mathematical expressions to model the building system. It heavily relies on domain expertise and often is computationally prohibitive [4]. Hybrid models use simplified physical descriptions combined with parameter identification algorithms to predict energy consumption. Nevertheless, without a description of the building geometry and materials, it is difficult to estimate the model parameters. In contrast, the emerging technology advancements in the energy industry make it possible to collect massive amounts of data from sensors and meters, which enable data-driven modeling to unfold the underlying knowledge [5]. As most industrial, institutional, and commercial buildings built after 2000 include a building automation systems (BAS), there is a growing interest to mine valuable information and derive additional insights from data collected. The data-driven approach motivates and drives the building energy research in various aspects including estimation of energy consumption [6–8], real-time performance validation and energy usage analysis [9], and energy saving operational control [3,10,11]. A significant advantage of the data driven approach lies in that it considerably reduces the design cycle iteration time for building design and operations, which includes not only simulation, but also analysis of results and optimization of actions based on these results. It allows for fast realizations of the design and operation tasks for any building scenario in an industrial context. Based on the updating cycle and horizon, the load forecast models can also be categorized into short term load forecasting (STLF), medium term load forecasting (MTLF), and long term load forecasting (LTLF) [12].STLF focuses on the load forecasting on daily basis and/or weekly basis, and MTLF and LTLF are based on monthly and yearly collected data for transmission and distribution (T&D) planning [13], and financial planning, which assist with medium to long term energy management, decision making on the utilities project and revenue management. STLF is important for real-time energy operations and maintenance. For daily operations, system operators can make switching and operational decisions, and schedule maintenance based on the patterns obtained during the load forecasting process [14]. To better assist the operations and control strategies development, this study develops a novel STLF methodology for buildings, which provides accurate load forecasts for daily and weekly based energy system management. The model, however, could be viably transformed into MTLF or LTLF, by adding features of economy and land use, and extrapolating the model to longer horizons.

Various data-driven methods have been studied and implemented for building load forecasting including (1) statistical methods such as autoregressive, moving average, exponential smoothing [15], state space [16,17], polynomial regression [18], and (2) machine learning methods such as neural networks [19] and support vector regression [8,20]. Statistical regression models simply build the correlation between the energy consumption and the simplified influential features such as weather parameters. These empirical models are developed from historical performance data to train the models. Machine learning models are good at building non-linear models and are especially effective on complex applications.

A regression-based approach was tested on the peak and hourly load forecasts of the next 24 h using Pacific Gas and Electric Company's (PG&E) data [21]. The regression model was thoroughly tested and concluded to be superior to the existing system load forecasting algorithms used at PG&E. In another study, five methods (autoregressive integrated moving average (ARIMA) modeling; periodic AR modeling, an extension for double seasonality of Holt-Winters exponential smoothing; an alternative exponential smoothing formulation; and a principle component analysis (PCA) based method) were compared on 10 load series from 10 European countries on an hourly interval and 24-h horizon [22]. They concluded that the double seasonal Holt-Winters exponential smoothing method outperformed the others. Another interesting study by Ahmed et al. [23] explored machine learning methods. Eight machine learning models for time series forecasting on the monthly M3 time series competition data (around a thousand time series) were investigated. These eight are multilayer perceptron, Bayesian neural networks, radial basis functions, generalized regression neural networks, K-nearest neighbor regression, CART regression trees, support vector regression, and Gaussian processes. They concluded that the best two methods turned out to be the multilayer perceptron and the Gaussian process regression. Chirarattananon and Taveekun [24] developed a model for building energy consumption forecasting based on overall thermal transfer value and concluded that the model does not present good generalizability on some types of buildings, especially on hotels and hospitals. Yik et al. [25] predicted the energy consumption for a group of different types of buildings using a number of physical parameters such as air conditioning system type, year the building was built and geographical information. The resulting model showed high correlation to the detailed simulation model. One novel data-characteristic-driven modeling methodology for nuclear energy consumption was proposed in [26], in which two steps, data analysis and forecasting modeling, were involved in formulating an appropriate forecasting model in terms of the sample data's own data characteristics. Experimental results showed that "data-characteristic-driven modeling" significantly improves prediction performance compared to all other benchmark models without consideration of data characteristics. However, only time series data characteristics and univariate forecasting models were explored in this study. One observation from these extensive studies is model performance varies and is highly dependent on the characteristics of the building systems, which leads the researchers come to inconsistent conclusions regarding the performance of various forecasting models. This concurs with what was found by [27]: he thoroughly reviewed twenty-five years of research and concluded that no algorithm is best for all load forecasting tasks. He suggested that the identification of which methods should be chosen with respect to the situations should be done via experimental studies.

Noting that a building system is stochastic, nonlinear and complex [28], research so far has mainly focused on an approach of trial-and-error or one-size-fits-all. In the cases where little prior knowledge of the building systems is available, previous studies either develop multiple models and identify the outperformer among them, which is computationally expensive and impractical for real-time building energy management and operations, or subjectively presume one model fits any type of building, suffering from high-bias modeling. In short term building load forecasting, the main goal is to minimize the forecasting error with computationally-efficient solutions. Building management control tasks can range from real-time load forecasting and user behavior analysis to predictive building control. For these tasks, the meter data are usually generated at a rate ranging from per minute to per hour. Due to the dynamics of building energy systems and for real-time supervisory purposes, the control and operations should be updated dynamically by analyzing the time series data. This impedes the trial-and-error modeling approach in that the computational complexity for constructing multiple models is unaffordable, especially in the case where data volume is large.

Download English Version:

https://daneshyari.com/en/article/6683244

Download Persian Version:

https://daneshyari.com/article/6683244

Daneshyari.com