Applied Energy 123 (2014) 168–178

Contents lists available at ScienceDirect

Applied Energy

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

Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy

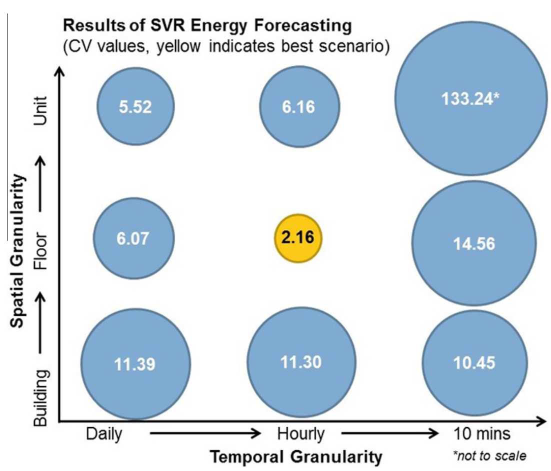
Rishee K. Jain [a,*], Kevin M. Smith [b], Patricia J. Culligan [c], John E. Taylor [d]

[a] *Center for Urban Science & Progress, New York University, 1 MetroTech Center, 19th Floor, Brooklyn, NY 11201, United States*
[b] *Department of Earth and Environmental Engineering, Columbia University, S.W. Mudd Building, 500 West 120th Street, New York, NY 10027, United States*
[c] *Institute for Data Sciences & Engineering, Department of Civil Engineering and Engineering Mechanics, Columbia University, Room 626, S.W. Mudd Building, 500 West 120th Street, New York, NY 10027, United States*
[d] *Department of Civil and Environmental Engineering, Virginia Tech, 113B Patton Hall, Blacksburg, VA 24061, United States*

HIGHLIGHTS

- We develop a building energy forecasting model using support vector regression.
- Model is applied to data from a multi-family residential building in New York City.
- We extend sensor based energy forecasting to multi-family residential buildings.
- We examine the impact temporal and spatial granularity has on model accuracy.
- Optimal granularity occurs at the *by floor* in *hourly* temporal intervals.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 16 September 2013
Received in revised form 28 December 2013
Accepted 15 February 2014
Available online 15 March 2014

Keywords:
Forecasting
Machine learning
Monitoring
Prediction
Residential building
Support vector regression

ABSTRACT

Buildings are the dominant source of energy consumption and environmental emissions in urban areas. Therefore, the ability to forecast and characterize building energy consumption is vital to implementing urban energy management and efficiency initiatives required to curb emissions. Advances in smart metering technology have enabled researchers to develop "sensor based" approaches to forecast building energy consumption that necessitate less input data than traditional methods. Sensor-based forecasting utilizes machine learning techniques to infer the complex relationships between consumption and influencing variables (e.g., weather, time of day, previous consumption). While sensor-based forecasting has been studied extensively for commercial buildings, there is a paucity of research applying this data-driven approach to the multi-family residential sector. In this paper, we build a sensor-based forecasting model using Support Vector Regression (SVR), a commonly used machine learning technique, and apply it to an empirical data-set from a multi-family residential building in New York City. We expand our study to examine the impact of temporal (i.e., *daily*, *hourly*, *10 min* intervals) and spatial (i.e., *whole building*, *by floor*, *by unit*) granularity have on the predictive power of our single-step model. Results indicate

* Corresponding author. Tel.: +1 16469970545.
E-mail addresses: rishee.jain@nyu.edu (R.K. Jain), kms2227@columbia.edu (K.M. Smith), pjc2104@columbia.edu (P.J. Culligan), jet@vt.edu (J.E. Taylor).

http://dx.doi.org/10.1016/j.apenergy.2014.02.057

that sensor based forecasting models can be extended to multi-family residential buildings and that the optimal monitoring granularity occurs at the *by floor* level in *hourly* intervals. In addition to implications for the development of residential energy forecasting models, our results have practical significance for the deployment and installation of advanced smart metering devices. Ultimately, accurate and cost effective wide-scale energy prediction is a vital step towards next-generation energy efficiency initiatives, which will require not only consideration of the methods, but the scales for which data can be distilled into meaningful information.

1. Introduction

The built environment accounts for 40% of energy consumption in the United States [1] and a similar proportion in many other parts of the world. In dense urban areas like New York City, buildings account for a staggering 94% of electricity usage and 75% of greenhouse gas (GHG) emissions [2]. Therefore, characterizing, modeling and forecasting the energy consumption of buildings is crucial if urban areas are to reduce their overall energy consumption. Accurate modeling and forecasting of building energy demand enables numerous energy management and efficiency applications such as: informing early stage design decisions [3], estimating improvements to building energy performance [4], optimizing building HVAC systems [5] and urban energy infrastructure planning [6].

Traditionally, building scale demand estimates have been performed using engineering software packages (e.g., EnergyPlus) that rely on an in-depth compilation of structural, geometric, and material building properties. The difficulty in obtaining and validating such information is a hindrance to wide-scale energy forecasting. In response, there is a growing interest in alternative "sensor based" approaches that forgo the demanding input requirements of theoretical formulations in favor of a practicable set of localized historical measurements.

In a sensor based approach, data from energy smart meters, building management systems, and weather stations are fed into a machine learning algorithm in order to infer the complex relationships between energy consumption and variables of influence such as temperature, time of day and occupancy. Prior research [7,8] has established that the prediction accuracy of sensor-based approaches is comparable, and sometimes superior, to traditional engineering based energy forecasting while requiring significantly less input data from the end user. Due to the accelerated development and proliferation of inexpensive off-the-shelf options for energy metering in recent years, sensor based approaches are becoming increasingly relevant and cost effective. A comprehensive discussion regarding the advantages of sensor based energy forecasting over traditional engineering methods can be found in [7].

While sensor based forecasting has been applied extensively in the commercial building sector since the launch of the American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE) Great Energy Predictor Shootout [9], there is a dearth of research applying this method to the residential building sector. The limited body of work [7] for the residential sector has been focused on single-family homes and has largely ignored multi-family residential units. As multi-family units are the dominant residential building stock in dense urban areas, accurate forecasting of their consumption is imperative to implementing effective energy efficiency and conservation measures on an urban scale.

In this study, we extend the sensor based energy forecasting approach to a monitored multi-family residential apartment building in New York City. Additionally, we deepen our exploration of sensor based forecasting by examining the effectiveness of our prediction algorithm at making single-step forecasts on various spatial and temporal data scales. Unlike single-family buildings, energy consumption data for multi-family residential complexes are often readily available at both the unit and building aggregate level. While utilities have traditionally reported residential energy use at the monthly scale, readily available metering technology can now report energy consumption as frequently as every minute. In general, the impact of temporal scales on sensor based energy forecasting has not been explored, and the relative utility of high-resolution monitoring remains an open question. By varying the aggregation granularity of monitoring data across several temporal (i.e., *daily*, *hourly*, *every 10 min*) and spatial (i.e., *whole building*, *by floor*, *by unit*) scales we hope to gain insight into the impact that each scale has on the forecasting accuracy of our model. In the end, our goal is to determine the optimal monitoring granularity for energy consumption that optimizes forecasting accuracy with both the financial and nonfinancial costs of installing and operating monitoring equipment.

2. Related work

2.1. Sensor based energy forecasting of commercial and residential buildings

Sensor based energy forecasting has been explored by researchers for both commercial and residential buildings. However, due to a lack of data availability for residential buildings the majority of previous work has been focused on the prediction of consumption in commercial buildings. Sensor based energy forecasting gained popularity after the first Great Energy Predictor Shootout that was hosted by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) [9]. The Great Energy Predictor Shootout was a competition in which participants were asked to develop models to predict the whole building electrical (WBE) consumption of a commercial building based on historical electrical consumption and environmental data. The overall winner of the competition [10] developed a sensor based energy forecasting model and employed an intelligent machine learning algorithm that required very little explicit domain knowledge about the actual commercial building whose energy was to be predicted. The results of the competition spurred increased attention by the research community towards the use of sensor based forecasting models.

Subsequently, several studies further explored the use of sensor based energy forecasting in commercial buildings [11–16] and to a lesser extent residential buildings [7,17]. The limited work examining residential building energy forecasting has been narrowly focused on single-family homes [7]. While energy forecasting of single-family homes is certainly valuable, in the context of dense urban areas multi-family residential units constitute a significant portion of the residential building stock and, thus, energy consumption associated with this sector. Therefore, the first contribution of this study is to investigate whether previously established sensor based forecasting models will yield acceptable levels of accuracy when applied to multi-family residential buildings. We aim to demonstrate that sensor based forecasting methods can