Omega 59 (2016) 47–59

Contents lists available at ScienceDirect

Omega

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

Forecasting electricity smart meter data using conditional kernel density estimation

Siddharth Arora [1], James W. Taylor *,[1]

Saïd Business School, University of Oxford, Park End Street, Oxford OX1 1HP, UK

ARTICLE INFO

Article history:
Received 8 January 2014
Accepted 28 August 2014
Available online 4 December 2014

Keywords:
Forecasting
Electricity demand
Nonparametric density estimation
Smart meter

ABSTRACT

The recent advent of smart meters has led to large micro-level datasets. For the first time, the electricity consumption at individual sites is available on a near real-time basis. Efficient management of energy resources, electric utilities, and transmission grids, can be greatly facilitated by harnessing the potential of this data. The aim of this study is to generate probability density estimates for consumption recorded by individual smart meters. Such estimates can assist decision making by helping consumers identify and minimize their excess electricity usage, especially during peak times. For suppliers, these estimates can be used to devise innovative time-of-use pricing strategies aimed at their target consumers. We consider methods based on conditional kernel density (CKD) estimation with the incorporation of a decay parameter. The methods capture the seasonality in consumption, and enable a nonparametric estimation of its conditional density. Using 8 months of half-hourly data for 1000 meters we evaluate point and density forecasts, for lead times ranging from one half-hour up to a week ahead. We find that the kernel-based methods outperform a simple benchmark method that does not account for seasonality, and compare well with an exponential smoothing method that we use as a sophisticated benchmark. To gauge the financial impact, we use density estimates of consumption to derive prediction intervals of electricity cost for different time-of-use tariffs. We show that a simple strategy of switching between different tariffs, based on a comparison of cost densities, delivers significant cost savings for the great majority of consumers.

1. Introduction

A smart meter is an electronic device that measures electricity consumption at the installed facility, and transmits this information to the consumer and the energy supplier/operator on a near real-time basis. It is anticipated that based on smart meter information, significant energy and financial savings can be achieved through detailed consumer feedback and tariffs designed for facilitating energy savings [1,2]. The large scale installations of smart meters will generate massive amounts of data, offering unique insight into the consumption behavior of different consumers. Over the coming years, smart meters are scheduled to replace the existing electronic meters. It is estimated that by 2019, approximately 60 million meters will be installed and operable in the United States [3]. In the European Union, all member states must have smart meters installed for at least 80% of consumers by 2020, with full deployment by 2022 [4]. It has been estimated that the cost of investment in smart electricity grids in the European Union will be around €51 billion [5].

Unlike conventional meters, smart meters provide site-specific information regarding electricity consumption throughout the day. This information can potentially change the landscape of energy markets, by allowing suppliers to make highly data-dependent decisions to develop innovative dynamic pricing strategies for their target consumers. Smart meters, along with different time-of-use (TOU) tariffs, can help consumers shift their consumption away from peak hours, which can result in significant savings [6]. With the liberalization of electricity markets, market participants rely on accurate forecasts to make informed energy transactions [7]. Also, smart meters can assist electricity scheduling, thereby facilitating safe and efficient operation of the power system.

In a recent trial involving electricity smart meter installations in Ireland, the deployment of TOU tariffs and information stimuli, such as bi-monthly billing and an in-home display device, resulted in an overall reduction in electricity usage by 3.2% and peak usage by 11.3% [1,2]. Given the potential for smart meters in enabling the efficient use of energy, and its financial implications for energy markets, it is imperative to develop accurate methods for modeling electricity smart meter data. This is the focus of this paper.

* Corresponding author.
E-mail addresses: Siddharth.Arora@sbs.ox.ac.uk (S. Arora), James.Taylor@sbs.ox.ac.uk (J.W. Taylor).
[1] Tel.: +44 1865 288800.

http://dx.doi.org/10.1016/j.omega.2014.08.008
0305-0483/

Electricity consumption data from individual smart meters exhibits seasonality, comprising both intraday and intraweek cycles, and in comparison with total national demand, consumption recorded from individual smart meters exhibits much higher variability. It, therefore, seems appropriate to provide a forecast of the probability density function for consumption recorded by a smart meter, rather than just producing a point forecast. Focusing on the density forecasts of electricity consumption from smart meters can help (a) ensure that the risk associated with complex decision making based on such forecasts is adequately assessed, and, (b) generate accurate consumption estimates at varying aggregation levels, potentially resulting in improved demand side management. The literature on modeling electricity smart meter data is small. There are some recent papers on short-term point forecasting of smart meter data [8,9], and forecast error measures [10]. However, we are not aware of existing studies on modeling the density of electricity consumption data from individual smart meters.

The non-Gaussian and highly variable nature of individual smart meter electricity consumption data motivates nonparametric probability density estimation methods. In this study, we propose methods based on kernel density (KD) and conditional kernel density (CKD) estimation (see, for example, [11,12]). The conditioning in our CKD implementations aims to capture seasonality. Although this seasonality is usually far less clear than the seasonality in a series of the total electricity demand for a country, both types of data tend to exhibit intraday and intraweek seasonal cycles. This prompts us to consider forms of KD and CKD that are inspired by the structure of models presented in the literature for total national consumption, namely the exponential smoothing and autoregressive moving average (ARMA) models presented by Taylor [13,14] and Gould et al. [15]. The CKD method involves kernel weighting over the conditioning variable. We consider different implementations of the CKD estimator, where we condition consumption on the period of week, period of day, and lagged consumption. We chose methods based on KD and CKD estimation for modeling smart meter data, because these methods (a) model the full density function, (b) can accommodate seasonality in the time series, and, (c) make no distributional assumption for the shape of the density, allowing the estimated density to be, for example, multi-modal, fat-tailed or skewed. The incorporation of a decay factor within the CKD estimation helps model temporal evolution in the relationship between consumption and the conditioning variables.

This paper employs 8 months of half-hourly data, recorded from 1000 smart meters, to evaluate the KD and CKD methods, in terms of accuracy of their density, quantile, and point forecasts, for lead times ranging from one half-hour up to 1 week ahead.

Although weather variables are often used in energy modeling [16], the KD and CKD methods that we propose use only the historical consumption observations. We felt that we could not assume the availability and affordability of weather predictions for a location reasonably close to each smart meter. Furthermore, the use of weather data in a large-scale online prediction system raises issues of robustness [17].

Recent advances in smart metering technology may pave the way for easy switching between suppliers, and between different payment schemes from the same supplier [18]. In this study, we use our density estimates of electricity consumption to derive prediction intervals for electricity cost, for different TOU tariffs. We compare different costs that would potentially be incurred in the future, for each available tariff, and select the tariff that would result in the greatest cost savings. In a case study of a warehouse environment, Sanders and Graman [19] emphasize the importance of evaluating the impact of forecast errors on organizational cost.

In Section 2, we describe the smart meter data. Section 3 presents different KD and CKD methods, along with an exponential smoothing benchmark method. Empirical results regarding forecast accuracy are provided in Section 4. Section 5 derives prediction intervals for electricity cost. Section 6 summarizes and concludes the paper.

2. Smart meter data

We used 8 months of half-hourly smart meter data for electricity consumption from 2 January to 31 August 2010. We used the first 7 months, comprising 10,128 observations (*in-sample*), for optimizing method parameters, while the final month, constituting 1488 observations (*post-sample*), was used to evaluate forecast accuracy. The final month of the in-sample period was used for cross-validation. Using a moving window of 6 months, we generated a sequence of density forecasts, for lead-times ranging from one half-hour up to 1 week ahead, by using as forecast origin each midnight in the post-sample data. The data was recorded for 800 residential consumers and 200 small to medium-sized enterprises (SMEs). The data was obtained from the Commission for Energy Regulation (CER) based in Ireland [1,2]. Figs. 1 and 2 show

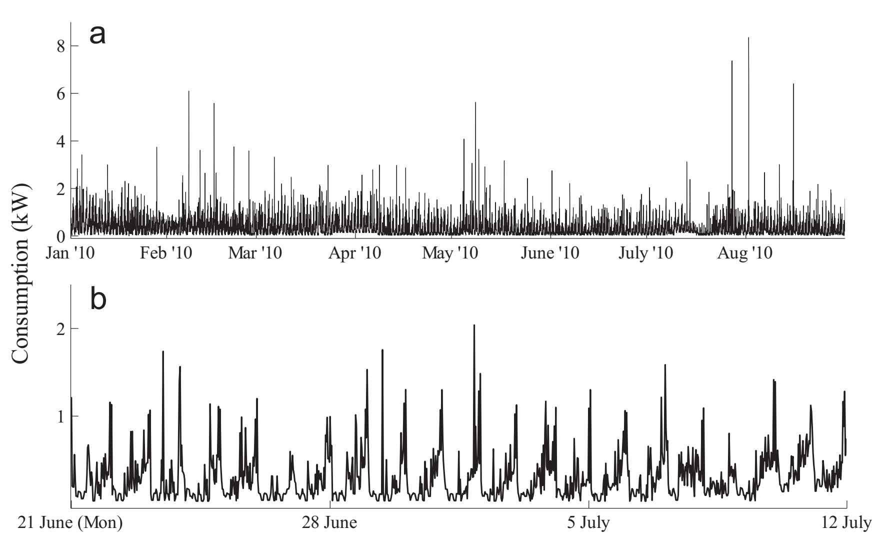

Fig. 1. Consumption for a residential consumer for (a) the full 8 months and (b) a typical three-week period.