International Journal of Forecasting 34 (2018) 147–162

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Probabilistic forecasting of industrial electricity load with regime switching behavior

K. Berk [a,*], A. Hoffmann [b], A. Müller [a]

[a] *Department of Mathematics, University of Siegen, Walter-Flex-Str. 3, 57072 Siegen, Germany*
[b] *statmath GmbH, Spandauer Str. 2, 57072 Siegen, Germany*

ARTICLE INFO

Keywords:
Electricity load forecasting
Probabilistic forecasting
Time series models
Seasonality
Inhomogeneous Markov switching model
Regime-switching models

ABSTRACT

This paper suggests a novel inhomogeneous Markov switching approach for the probabilistic forecasting of industrial companies' electricity loads, for which the load switches at random times between production and standby regimes. The model that we propose describes the transitions between the regimes using a hidden Markov chain with time-varying transition probabilities that depend on calendar variables. We model the demand during the production regime using an autoregressive moving-average (ARMA) process with seasonal patterns, whereas we use a much simpler model for the standby regime in order to reduce the complexity. The maximum likelihood estimation of the parameters is implemented using a differential evolution algorithm. Using the continuous ranked probability score (CRPS) to evaluate the goodness-of-fit of our model for probabilistic forecasting, it is shown that this model often outperforms classical additive time series models, as well as homogeneous Markov switching models. We also propose a simple procedure for classifying load profiles into those with and without regime-switching behaviors.

1. Introduction

The modeling and forecasting of electricity loads is an important issue for all market participants, typically with a focus on short-term point forecasts. However, recent years have seen an increase in the volatility of the supply, demand and prices of electricity, making probabilistic forecasting increasingly relevant. It is important both for the planning of the availability of production facilities and for the risk management purposes of buyers, producers and retailers. One needs not only short-term but also at least medium-term probabilistic forecasts, for horizons of a few months to a few years.

The aim of this paper is to contribute to the literature on the medium-term probabilistic forecasting of the electricity load of industrial companies. These often have a complicated demand pattern, as a small number of machines may be responsible for most of the electricity demand, and thus, their electricity load depends heavily on whether or not these machines are in use. In some cases, the operating hours of these machines are quite predictable. However, in other cases one may observe quite random fluctuations between operating hours and some kind of standby regime with only a very low electricity load. In this paper we want to suggest a stochastic model for the electricity load for dealing with exactly this type of consumption pattern, with two regimes and random switches between the regimes.

There already exists a vast body of literature on different load forecasting techniques, ranging from parametric models to artificial intelligence approaches; an overview is provided by Weron (2006). In the past, the literature has focused largely on short-term point forecasts, but in

* Corresponding author.
E-mail addresses: berk@mathematik.uni-siegen.de (K. Berk), hoffmann@statmath.de (A. Hoffmann), mueller@mathematik.uni-siegen.de (A. Müller).

https://doi.org/10.1016/j.ijforecast.2017.09.006

recent years, the research focus has shifted more and more towards probabilistic forecasting. Much of the more recent research has been fostered by the recent Global Energy Forecasting Competition (GEFCom2014). Hong et al. (2016) provide an overview of the results of that competition. The top performing teams use a variety of approaches, one of which was a hybrid of kernel density estimation and quantile regression (Haben & Giasemidis, 2016). In the course of a later probabilistic load forecasting competition organized by Tao Hong in Fall 2015, it turned out that the lasso estimation proposed by Ziel and Liu (2016) also provided very good results in the GEFCom2014 load forecasting track. A comprehensive overview of the topic of probabilistic electric load forecasting is provided by Hong and Fan (2016). Further important recent contributions, especially on forecasting for longer time horizons, include those by Hong, Wilson, and Xie (2014) and Hyndman and Fan (2014). Liu, Nowotarski, Hong, and Weron (2017) and Nowotarski, Liu, Weron, and Hong (2016) discuss methods of averaging sister forecasts to improve the probabilistic forecasting of the electricity load.

This paper builds on the approach to the modeling of industrial companies' electricity demands that was described by Berk (2015) and Berk and Müller (2016). Like these authors, we assume that the only information that we can use for forecasting is that given by individual companies' historical smart metering data. Their idea of the models is that of conventional additive time series models. The electricity demand follows a certain seasonal pattern that is supplemented by a residual process which describes the random deviations from that pattern. Often, these deviations are assumed to follow autoregressive moving-average (ARMA) type processes. They may also depend on exogeneous variables like the temperature, but this is not considered here. However, there are load profiles for which these approaches perform badly because of stochastic changes in demand that do not exhibit seasonalities. A common characteristic of these approaches is sudden drops in demand that appear to happen at random times, and also typically have random durations.

The economic effects of such regime changes are vast, both for customers and for suppliers. If they are not modeled properly, the risk of the specific load profile is very likely to be underestimated, which implies high imbalance costs. This paper proposes a Markov switching model with time-varying transition probabilities. To be more precise, the model consists of an underlying inhomogeneous Markov chain that drives the binary process which determines the regime, and a seasonal component combined with an ARMA process in the production regime.

An important reference for regime-switching models is the paper by Hamilton (2008); however, in that case homogeneous Markov chains are used. Regime-switching models with time-varying transition probabilities are considered by Diebold, Lee, and Weinbach (1994), who suggest an expectation–maximization (EM) algorithm for estimating the parameters and describe an application to exchange rate and business cycle modelling. An inhomogeneous model for business cycle phases is also proposed by Filardo (1994), who uses stock indices as exogeneous variables that drive the Markov chain. In the context of energy markets, Benschop and Cabrera (2014) develop a Markov switching generalized autoregressive conditional heteroscedasticity (GARCH) model for CO_2 spot prices. An approach for the simulation of electricity prices is provided by Paraschiv, Fleten, and Schürle (2015).

The remainder of this paper is organized as follows. Section 2 illustrates what we see as the main characteristics that distinguish between 'normal' load profiles and those which we refer to as regime-switching ones. We use these special characteristics to develop an automated test procedure, based on Hartigan's dip test for unimodality, for identifying regime-switching load profiles. Section 3 introduces the inhomogeneous Markov switching model that we suggest for the probabilistic forecasting of electricity loads that exhibit regime-switches. Section 4 describes a database of load profiles that we use for an empirical study. We develop an algorithm for the automatic identification of company holiday periods and include such periods in the model. The procedure of parameter estimation via differential evolution is described, together with the methods of evaluating the forecast accuracies of probabilistic models. Here, we use the continuous ranked probability score (CRPS), which is a proper scoring rule that is useful for this purpose (see Gneiting and Katzfuss, 2014 and Gneiting and Ranjan, 2011). This idea is motivated by the GEFCom2014, where the closely related so-called pinball loss score is used to compare different models for the probabilistic forecasting of electricity prices and loads, as well as wind and solar production. The empirical results of the modeling procedure are given in Section 5. Section 6 concludes.

2. Classification of regime-switching load profiles

This section describes a method for classifying load profiles into one of two categories, which we refer to as regime-switching and non-regime-switching. More precisely, one should say that we want to detect stochastic regime switching behaviors, as almost all load profiles show some kind of deterministic regime-switching behavior, if one considers easy-to-predict seasonalities like different loads during the day and at night as regime-switching. Many companies have electricity demands that are quite homogeneous, except for seasonality and some noise. As an example, let us consider a customer from the retail industry. They have fixed opening hours during which the electricity demand reaches a certain level, then during closing hours the consumption drops to what we refer to as the standby level. Three weeks of a typical load profile are shown in the top plot of Fig. 1. The very homogeneous structure during the weekdays and the low demand on Sundays are both clear. The picture also includes special events, since the first weekend shown is the Easter weekend, and the demand is lower on Good Friday and Easter Monday. However, these different regimes during opening and closing hours are easy to predict. Thus, we can deseasonalize the data in order to get something quite homogeneous. Note that all electricity loads plotted in this paper are measured in kW/h, but the absolute values are not shown for the sake of privacy. However, the scale of the load is not relevant for our methods.