DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

Modeling and forecasting electricity price jumps in the Nord Pool power market

Oskar Knapik

CREATES Research Paper 2017-7

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Modeling and forecasting electricity price jumps in the Nord Pool power market ☆

O. Knapik[a,*]

[a]*Aarhus University, Department of Economics and Business Economics, CREATES - Center for Research in Econometric Analysis of Time Series, Fuglesangs Alle 4, 8210 Aarhus V, Denmark*

Abstract

For risk management traders in the electricity market are mainly interested in the risk of negative (drops) or of positive (spikes) price jumps, i.e. the sellers face the risk of negative price jumps while the buyers face the risk of positive price jumps. Understanding the mechanism that drive extreme prices and forecasting of the price jumps is crucial for risk management and market design. In this paper, we consider the problem of the impact of fundamental price drivers on forecasting of price jumps in NordPool intraday market. We develop categorical time series models which take into account i) price drivers, ii) persistence, iii) seasonality of electricity prices. The models are shown to outperform commonly-used benchmark. The paper shows how crucial for price jumps forecasting is to incorporate additional knowledge on price drivers like loads, temperature and water reservoir level as well as take into account the persistence in the jumps occurrence process.

Keywords: autoregressive order probit model, categorical time series, seasonality, electricity prices, Nord Pool power market, forecasting, autoregressive multinomial model, fundamental price drivers

JEL: C1, C5, C53, Q4

1. Introduction

Electricity prices often tend to temporarily jump to extreme levels, a phenomenon which is usually associated with with the non-storability of electricity, unexpected increases in demand,

☆Acknowledgment: The author appreciates support from Center for Research in Econometric Analysis of Time Series, CREATES funded by the Danish National Research Foundation with grant number DNRF78.

*Corresponding author

Email address: oknapik@econ.au.dk (O. Knapik)

unexpected supply shortages or over production from wind turbines or the failures of the transmission infrastructure (see Geman and Roncoroni, 2006; Christensen et al., 2012; Eydeland and Wolyniec, 2003; Harris, 2006, among others). The spikes may occur due to the fact that the central dispatch process sometimes needs to rely on the bids of the high marginal cost of production generators in order to satisfy demand. It is important to remember though, that the severity of spikes cannot be explained only by the dispatch of units with a high marginal cost of production but it is mainly driven by gaming or speculation (see Weron, 2006). Electricity price jumps are visible for intra-day or day-ahead prices on a half-hourly, hourly or daily time grid, but not for forward prices. Extreme price events are particularly hazardous for power market participants. Many traders on the electricity market are mainly concerned with either the risk of negative or positive price jumps, i.e., buyers face the risk of positive, while sellers face the risk of negative price jumps (see Hellstrom et al., 2012). Consequently, improving our understanding of the factors which contribute to the occurrence of extreme prices as well as accurate forecasting of these events, is crucial for effective risk management in the energy sector. It is the forecasting problem of price jumps which is the central concern of this paper.

Models for electricity prices typically fall into three categories: traditional autoregressive time series models, nonlinear time series models (with a particular emphasis on Markov-switching models) and continuous-time diffusion or jump-diffusion models (see Christensen et al., 2012). All of these models aim to characterize the trajectory of the spot price or return across time. A review of the models with special attention to forecasting of electricity prices can be found in (Weron, 2014). Taken at face value, these models appear to be different in their identification and treatment of price jumps (see Janczura et al., 2013, for the review).

There is a scant research focusing directly on price jumps modeling and forecasting. Lu et al. (2005) explore the reasons for price spikes using Bayesian classification and similarity searching. Their approach uses the measurement of the proposed composite indexes reflecting the relationship among electricity demand, supply and reserve capacity using Queensland half hourly prices. An important research in this area is the work of Jong (2006) representing the class of regime-switching models for electricity price models called Independent Spikes models (see Lindstrom et al., 2015). Mount et al. (2006) show that a stochastic regime-switching model with two regimes and the two transition probabilities being functions of the load and/or the implicit reserve margin can reflect the volatile behavior of wholesale electricity prices associated with price spikes. Becker et al. (2007) claim that the use of a time-varying probability regime-switching model with transition probabilities modeled with logit transformation of variables capturing demand (daily average and maximum load) and weather (daily average and maximum temperature, dew-point) can help to predict price spikes for Queensland data. Amjady and Keynia (2011) propose a data mining selection technique for prediction of both likelihood and severity of price spikes. Cartea et al. (2012) implement "tight market conditions" in capacity constraints in the form of a threshold variable to a regime-switching model for England and Wales power markets. They reveal that 85% of spikes occur when the demand-to-capacity ratio is in the interval [0.908; 0.960]. Christensen et al. (2012) consider the time series of price spikes for half-hourly data from the Australian market and introduce a nonlinear variant of the autoregressive conditional hazard (ACH) model. Clements et al. (2013) propose a semi-parametric approach based on a framework developed for forecasting realized volatility for forecasting spikes in the Australian market. Maryniak and Weron (2014) analyze forward looking data that is available to all participants in the UK power market and showed that the reserve margin has a huge potential for explaining the spike probability.

The majority of models for electricity prices treat price spikes as a memoryless process. However, evidence suggests there is a significant persistence component and helps to explain the intensity of the jumping process (see Christensen et al., 2009, 2012; Clements et al., 2013, among others). Most of the models consider also the case of two states: normal price and spike price. Here, we consider categorical time series model of three electricity price states (normal price and positive/negative spike) taking into account the persistence.

The focus of majority of the papers in the literature is on Australia, UK and US power markets. There are only few papers that refer to the Nord Pool power market, when the trends are on extreme price events. Hellstrom et al. (2012) explore the possible reasons behind electricity price jumps in the Nordic electricity market by the use of a mixed GARCH-EARJI[1] jump model. Voronin and Partanen (2013) propose data mining and time series techniques for prediction of both normal prices and price spikes in the Finish Nord Pool Spot day-ahead power

[1] $EARJI(r, s)$ is an exponential autoregressive jump intensity model

market. (Voronin et al., 2014) propose a hybrid forecasting model for the Finnish electricity spot market and show that hybridization of the normal range price and price spikes forecasts may provide comprehensive and valuable information for electricity market participants. Lindstrom et al. (2015) extend the Independent Spike Model used to model the electricity price and find out that consumption can be used to forecast extreme events in the Nord Pool power market.

This paper studies electricity prices from the Nord Pool power market. In the Nordic countries, more than 80% of the hourly consumed electricity is traded on the Elspot market, the day-ahead electricity market. Since security of supply is very important and forecasts for demand and/or supply a day ahead are not very accurate, several other spot markets have been put in place. One of them is the Elbas market where up to an hour before delivery producers and retail suppliers can update the quantity of power traded. The Nord Pool market plays a key role in the development of intraday power trading in Europe and it is becoming increasingly important due to a visions share of wind power production. Future prospects indicate exponential growth, reaching 1.900 GW installed wind capacity worldwide in 2020 (Source: World Wind Energy Association). This type of market can be crucial in increasing the share of renewable energy in the energy mix. On the other hand, Elbas is also almost not explored. Therefore the main focus of this research is on Elbas market.

This paper makes several contributions to the existing state of knowledge in the modeling of extreme price event occurrences in Nord Pool power markets. We extensively study the fundamental price drivers for the Nord Pool power market. We propose new categorical time series models to properly model the persistence in the electricity price jumps process. One important characteristic of the proposed econometric models which are considered is that they embed the information content of previous jumps and relate it to explanatory variables modeled in three ways with the use of a first-order Markov chain model with a time-varying transition matrix, an autoregressive ordered probit model, and an autoregressive conditional multinomial model. The paper is organized as follows. Section 2 describes the Nord Pool power market structure, focusing on hypothetical causes of price jumps. Section 3 describes the data. Section 4 introduces the main price drivers. Section 5 presents the models and statistical inference. Sections 6 and 7 provide the empirical and forecasting results. Section 8 concludes.

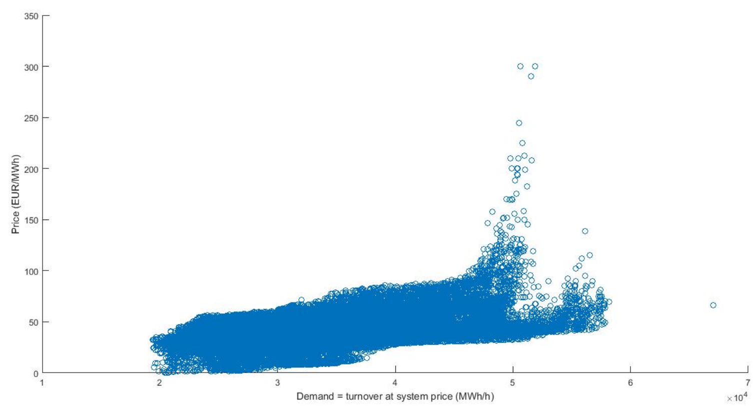
2. The Nord Pool power market and hypothesised causes of price jumps

The Nordic power market was fully established in 2000 when the regional electricity markets of Sweden, Norway, Finland and Denmark merged. Nowadays, the Nord Pool power market also includes the Baltic countries of Estonia, Latvia and Lithuania. Nord Pool operates as one market in which supply to a region is aggregated and generators are dispatched in order to satisfy the demand as cost-effectively as possible. If in one of the regions the local demand is higher than the local supply or the electricity in a neighboring region is cheap enough to guarantee transmission, the electricity is imported and exported in between, subject to the physical constraints of the transmission infrastructure. Hydroelectric production (which supplies around 57% of the Nord Pool capacity and nuclear generators (which supply around 18% of the Nord Pool capacity) have relatively high start-up costs and low marginal costs of production. Gas turbines and oil-fired plants together supply around 10% of the market, and only take about 20 min to start power generation, but have a comparatively high marginal cost of production, used typically for peak periods only. Wind power production supplies around 9% of electricity demand with an increasing share. This type of renewable energy is less predictable than more traditional sources of energy and may lead to price drops.

The majority of the volume handled by Nord Pool Spot is traded on the day-ahead market called Elspot. To a large extent, the balance between supply and demand is secured there. However, incidents may happen between the closing of Elspot at noon CET and delivery on the following day. A nuclear power plant can suddenly stop operating or strong winds may cause higher wind power generation than expected. Therefore, Nord Pool Spot's intraday trading system Elbas has been introduced. The importance of the intraday market is growing as more wind power enters the grid. Wind power production is unpredictable by nature and fluctuates in relation to day-ahead contracts. Therefore produced volume often needs to be offset. Elbas will play a key role in the development of intraday power trading. Covering the Nordic and Baltic regions as well as Germany and, recently, also the UK, Elbas supplements Nord Pool Spot's day-ahead market and helps to secure the necessary balance in real time between supply and demand in the power market for Northern Europe.

Figure 1 displays a scatter plot of the dependence between price and load (approximated as turnover at system price) in the Nord Pool Elspot power market. We might observe that price spikes appear when the load is high. On the other hand, negative jumps/drops can be observed for low load. Therefore, load should be considered an important explanatory variable.[2]

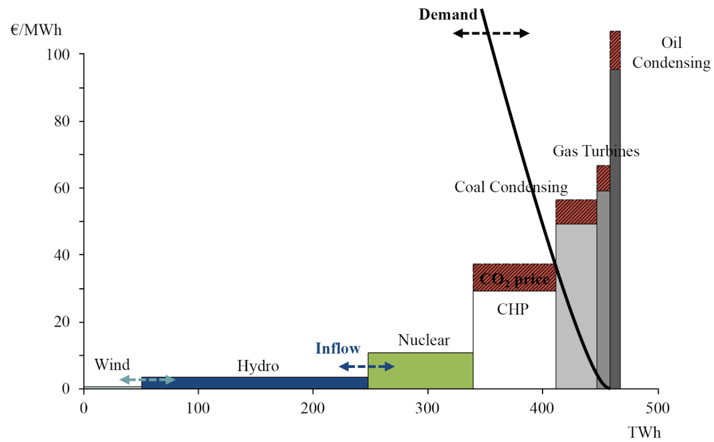
Figure 1: Price and demand in Nord Pool Elspot power market



The reasons behind the occurrence of price jumps relate to the interaction between system demand and supply (see Barlow, 2002; Geman and Roncoroni, 2006, among others). The demand for electricity is very inelastic, as the demand side is protected from pool price fluctuations by retailers who buy electricity at the spot price and sell the electricity at fixed rates. Electricity during normal conditions is provided by traditional low-cost generators (coal-fired and hydroelectric generators). If the system is affected by increases in demand and/or reductions in supply, the spot price jumps above the threshold, where it is cost-effective for generators with higher costs of production (gas-fired and diesel generators) to compete with low-cost generators. On the other side, if the system is affected by decreases in demand and/or increases in supply, the spot price jumps below the threshold, where paying the buyers for taking the additional energy (rather than shutting down the power plant) happens in the most extreme cases.

[2]In this paper we do not consider transmission capacity constraints as we work with system prices and volume-weighted prices for the entire Nord Pool. However, those constraints might be important explanatory variables for the electricity prices for single areas within the grid.

Figure 2: Power production price curve in the Nordic electricity market



Source: `http://www.nordpoolspot.com`

Figure 2 shows the power production price chart in the Nordic electricity market, where the production cost is on the y-axis and the annual total production is on the x-axis. The blocks in the figure represent different means of power production. The width of the blocks reflects the generation capacity and the height represents the marginal production costs. The red-striped areas illustrate the price increases caused by the EU ETS CO2 emissions allowances. The annual power demand in the region is illustrated by the black demand curve in the picture. The chart doesn't include wind- and biomass-based power generation, as they represent a relatively small share of the total production capacity. Furthermore, as hydropower represents around half of the total power generation in the market and its marginal production cost is almost zero, fluctuations in the hydropower supply (marked with the blue-dashed arrows) shift the other means of production along the x-axis. If the nuclear power generation remains stable, and does not balance the hydropower generation fluctuations so that power demand could be satisfied with hydropower and nuclear only, the next means of generation along the x-axis are CHP (combined heat and power) and coal condensing, with both using coal as raw material input. Thus, if demand for electricity in the Nordic market exceeds the combined production capacity of hydropower and nuclear, the marginal production methods are coal-fired methods of production and, therefore, the marginal price in the market should be linked to coal prices. Another important factor affecting the cost of coal-fired power is the price of the emission allowances set by the European Union that have a huge impact on coal-based power production costs. Kara et al. (2008) estimated that the average electricity spot price would increase by 0.74 EUR/MWh for every 1 EUR per tonne CO2 allowance price in the Nordic area between 2008 and 2012. If electricity demand in the market cannot be satisfied by the above-mentioned means of power production, other power production means are gas turbines and oil condensing. As natural gas prices are highly correlated with oil price (see, for example Krichene (2002) and Villar and Joutz (2006)), the cost of these means of production is highly dependent on oil price.

3. Data

We study hourly volume-weighted average prices from the Elbas power market. The data covers the period from 14 September 2009 to 31 December 2013. A further sample length of 24 months from 1 January 2014 to 31 December 2015 is reserved for assessing the out-of-sample performance of the models.

The estimation period is chosen to cover a few very interesting price peak events in the Nord Pool power market. Firstly, prices peaked during the winter of 2009-2010. For Sweden, Finland, Eastern Denmark as well as Mid and Northern Norway there were three very high price peaks that occurred on 17 December 2009 at 17-18 in the evening and on 8 January 2010 at 8-9 in the morning. For these areas, the prices during the three peaks were 1400, 1000 and 1400 EUR/MWh, respectively. The system price remained at 300 EUR/MWh or below. At the same time the price in Southern Norway and Western Denmark was relatively low, at about 65 EUR/MWh. Secondly, the estimation period covers the price spike event on 24 August 2010, when spot prices in pan-Nordic power exchange Nord Pool Spot's Estonia market area reached an exceptional 2000 EUR/MWh following the loss of 160MW of generation capacity in the Baltic state. Prices for Tuesday delivery between 8-11 and 12-14 reached the maximum technical level[3] that can be recorded under Elspot day-ahead market rules.

For the purpose of this paper, a price jump will be formally defined as a situation where the spot electricity price exceeds a particular threshold value that is chosen to lie outside the normal range of daily fluctuations. To define price jump events, some method to identify jumps

[3] At this time. Currently, the upper price limit is set to be 3000 EUR/MWh.

must be chosen. A variety of such methods can be found in the work of Janczura et al. (2013). Following the works of Becker et al. (2007) and Christensen et al. (2009), Christensen et al. (2012), Clements et al. (2013) we use static thresholds for spike/drop identification. Whilst the actual threshold used is market-specific, the reason for using a threshold to define extreme events is generic (see Kanamura and Ohashi (2007), Mount et al. (2006)). This logic comes from the fact that applications such as demand-side management do not require precise values of future spot prices but use specific price thresholds for making scheduling decisions Weron (2014). Additional justification for using static thresholds is that the sample size is small enough that market conditions should not change. Using a static threshold is also justified by the fact that average marginal costs of producing power with different sources do not change. It is also more informative than dynamic thresholds for market participants.

Figure 3 presents the quantiles of Elbas hourly average electricity prices and Elspot hourly spot system electricity prices. It can be seen that the electricity spot price fluctuates between 20 EUR/MWh and 60 EUR/MWh under ”normal” conditions. The threshold chosen for defining price spike events in Nord Pool is that of 80 EUR/MWh, which lies above the 90% spot prices for each hour of the day. This threshold is also slightly larger than the average marginal cost of gas turbine generation capacity bought online during periods of market stress (around 77 EUR/MWh). The threshold chosen for defining price drop events is chosen that of 10 EUR/MWh, which lies below the 10% spot prices for each hour of the day and is around the average marginal cost of production based on nuclear power. It should be noted that the price limits imposed by the market rules in force during the sample period are -500 EUR/MWh for the lower limit and 3000 EUR/MWh for the upper limit. The sensitivity analysis shows that the results are relatively robust for the choice of the exact values. Based on the analysis we can define the same thresholds for both considered power markets. The choice of the threshold is market specific and requires additional analysis in order to establish it for different power market. However, one can consider application of some outlier detection methods as considered in the literature (see Janczura et al., 2013; Chen and Liu, 1993a,b; Johansen and Nielsen, 2016, among others).

Figure 3: Quantiles of Elbas average prices by hour

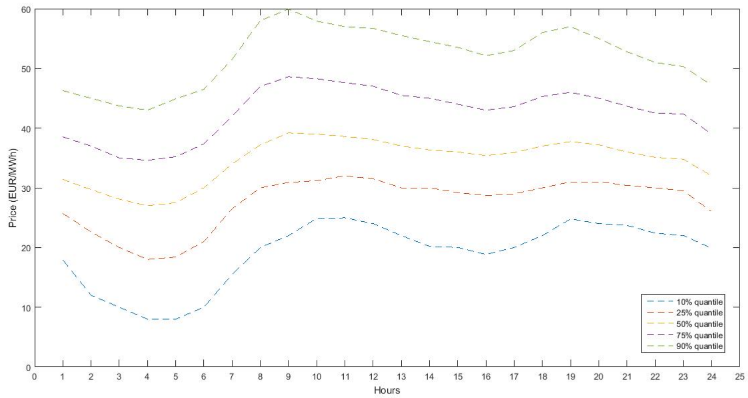


Figure 4: Quantiles of Elspot spot prices by hour

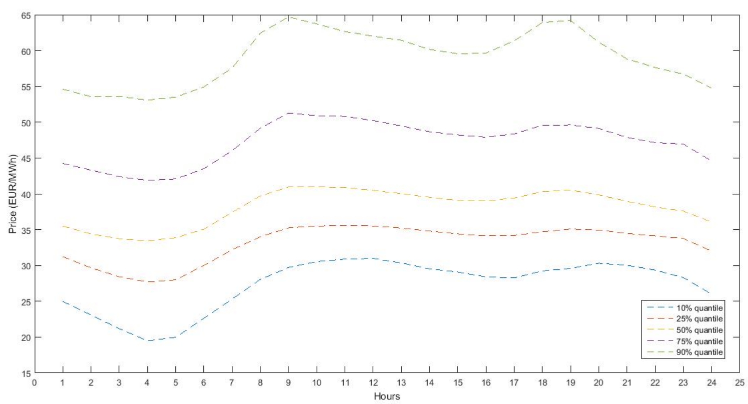


Figure 5 and 6 present hourly Elspot system electricity prices and Elbas volume-weighted average prices by hours and days in the considered time period. We can observe that there is a periodic dependency in the occurrence of price spikes and jumps.

Figure 5: Elbas average electricity prices by hours and days

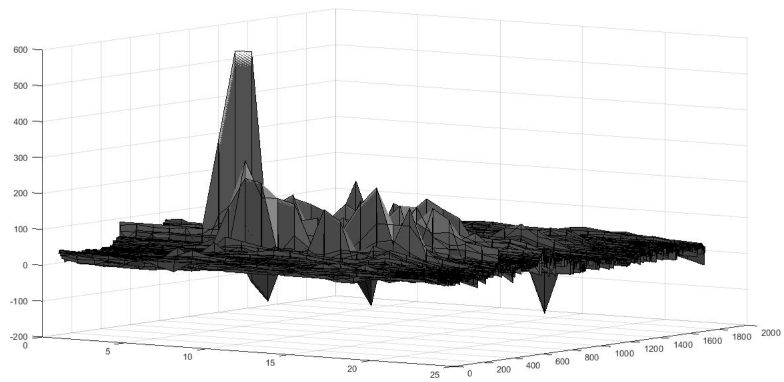


Figure 6: Elspot spot electricity prices by hours and days

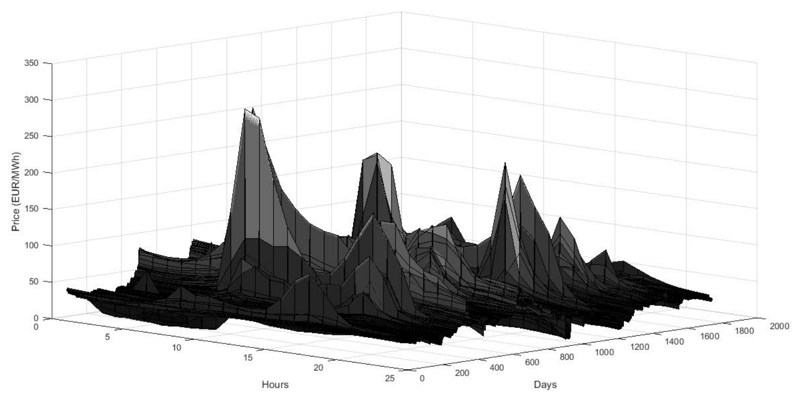


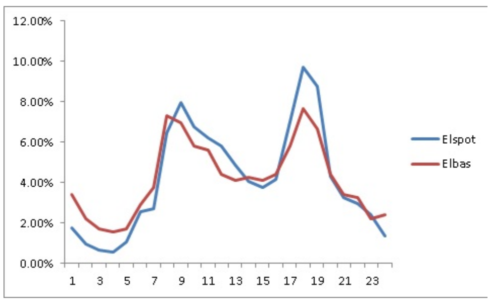
Table 1 shows the structure of electricity prices in the Elbas power market. 92.5% of all prices are within the range of 10 EUR/MWh to 80 EUR/MWh. Negative jumps and positive jumps are respectively 6.06% and 1.44% of the total number of observations.

Table 1: Elbas electricity prices: negative jumps, normal prices and positive jumps

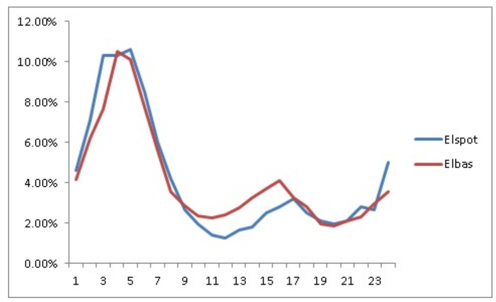
	Negative jumps	Normal prices	Positive jumps
Number	2284	34855	541
Frequency	6.06%	92.5%	1.44%

Figure 7a-9b display the respective number of price spikes (exceedance above 80 EUR/MWh) and price drops (exceedance below 10 EUR/MWh) on hourly, daily and monthly bases, and also provide casual empirical evidence to support temporal dependency and seasonality in the Nord

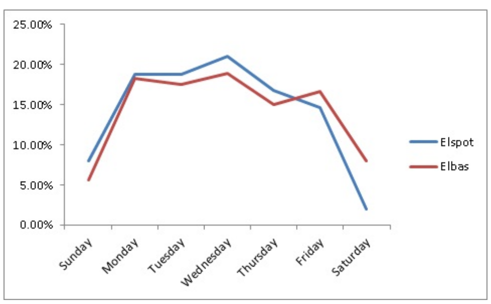
Pool Elspot and Elbas power markets. The percentage of spikes and drops is computed with respect to the total number of spikes and drops, respectively. Price spikes are more likely to appear early in the morning (8:00-11:00) and late afternoon (17:00-19:00), on working days (Monday-Friday) and during the winter. The figures illustrate also that price drops are more often during the night and early morning (24:00-7:00), from Wednesday to Sunday, and in July and October. This suggests that jumps are very much driven by the demand.



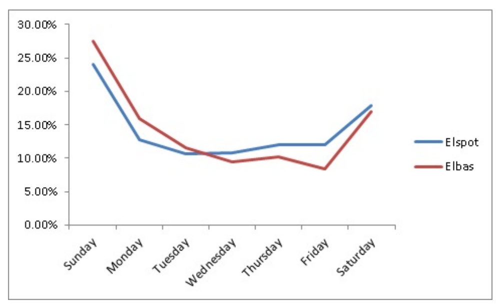
(a) 7a Hourly seasonality of spikes



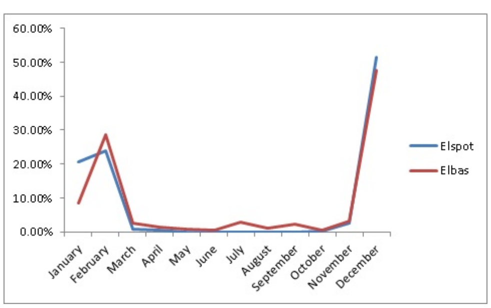
(b) 7b Hourly seasonality of drops



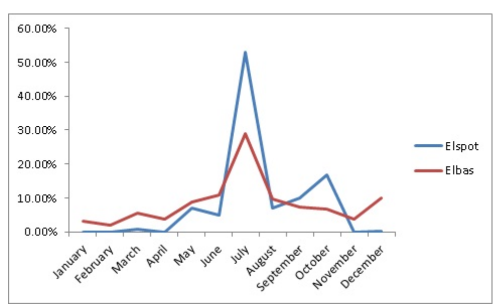
(a) 8a Daily seasonality of spikes



(b) 8b Daily seasonality of drops

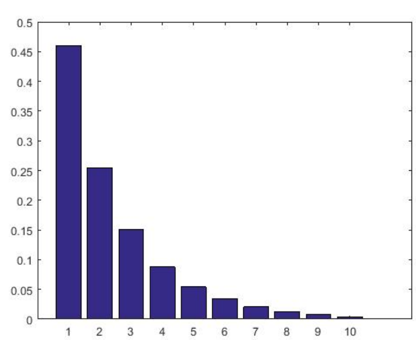


(a) 9a Monthly seasonality of spikes

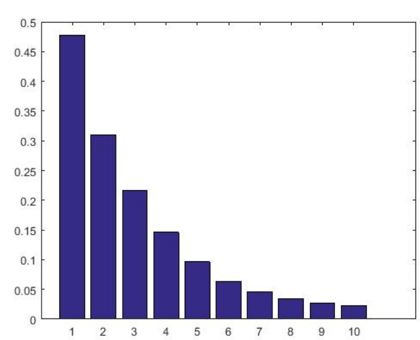


(b) 9b Monthly seasonality of drops

Figures 10a, 10b, 11a and 11b present proportions of jumps (spikes and drops separately) that appeared in next 10 hours, assuming that a jump occurred at a given hour in the total number of jumps (spikes and drops)[4]. The counting is done as follows: in the loop we check if for a given hour the price is a spike or drop; if it is, we check whether the prices in the next 10 hours are also in the same price state, and increase the numbers corresponding to each of the next 10 hours if needed. The whole number for each following hour is divided by the total number of jumps (spikes or drops) observed in the system. We might observe strong persistence in each case, i.e. jump clustering may be observed. The potential reasons behind abnormal price events, consistent with the clustering of price jumps (shown in Figures 8a to 9b), might lie in the persistence of system stress (see Becker et al., 2007; Christensen et al., 2009, 2012, among others). This is supported by Figures 8a to 9b , which show that a lot of jumps occur in successive hours, meaning that the probability that a jump occurs in the next hour, given that one has occurred in the previous hour (hours), is higher than otherwise, excluding other temporal effects. For example, the proportion of spikes that occur one hour after another spike has occurred is around 45%. The proportion of prices being in a spike regime for the next five hours is 5%. This finding will also be supported by the estimated transition matrix of a first-order Markov chain. This reasonably mild finding has important implications for model choice, as it renders many of the current methods of capturing jumps inadequate.



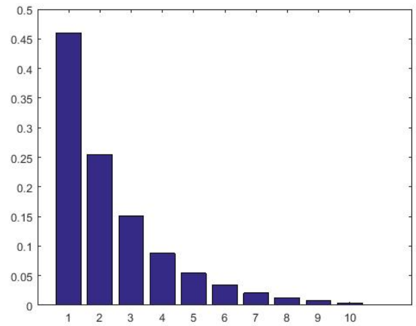
(a) 10a: Persistence of spikes: Elbas



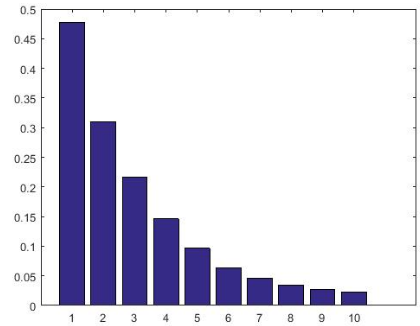
(b) 10b: Persistence of drops: Elbas

[4]Technically speaking the method is based on computations of conditional probabilities of the price being in given state given price(s) in the past were in the same state, e.g. for spikes

$$P\left(Y_t = 2|Y_{t-1} = 2, \ldots, Y_{t-10} = 2\right), P\left(Y_t = 2|Y_{t-1} = 2, ..., Y_{t-9} = 2\right), ..., P\left(Y_t = 2|Y_{t-1} = 2\right)$$



(a) 11a: Persistence of spikes: Elspot



(b) 11b: Persistence of drops: Elspot

Section 5 provides a more analytical approach supporting the observation that there is persistence of exceedance in electricity prices, which shows temporal dependence. The formal framework of categorical time series can be used to determine the statistical significance of possible factors driving the occurrence of price jumps. The total number of abnormal price events arising per day may be treated as a counting measure, which measures the stress acting on the system. Here we will look at one of the measures of persistence, which is a first-order Markov chain, as the estimated transition matrix parameters.

Table 2 and 3 contain the maximum likelihood estimates of the transition matrix of the first-order Markov chain for Elbas and Elspot power markets and their standard errors. The transition matrix describes the probabilities of moving between different price states in time t, and conditionally on being in a given state in time $t-1$. For instance, if the price in the Elbas market is in a negative drop state in time $t-1$, we have around a 48% chance that it will remain in the same state in time t. On the other hand, if the price in time $t-1$ is in a normal price state, then it will be in the same state with around a 98% chance. However, if the price in time $t-1$ is in a spike state, then we have around a 50% chance that it will be in a normal state and a 45% chance that it will remain in the spike state in time t.

However, looking at the seasonal character of price jump occurrence, it is reasonable to suspect that those transition matrices might vary over time and can be explained with some covariates. Kanamura and Ohashi (2007) show that in the case of the U.S. electricity market PJM the transition probabilities of electricity prices cannot be constant, and depend on both the current demand level relative to the supply capacity and the trends of demand fluctuation. Therefore, it is important to have a closer look at the main price drivers in the Nord Pool power

Table 2: First-order Markov chain: MLE estimates transition matrix for Elbas

				t
		Drop	Normal price	Spike
t-1	Drop	0.4778	0.4966	0.0256
		(0.0202)	(0.0206)	(0.0047)
	Normal price	0.0140	0.9798	0.0062
		(0.0006)	(0.0048)	(0.0004)
	Spike	0.0393	0.5009	0.4560
		(0.0086)	(0.0306)	(0.0293)

Table 3: First-order Markov chain: MLE estimates transition matrix for Elspot

				t
		Drop	Normal price	Spike
t-1	Drop	0.8551	0.1448	0.0000
		(0.0345)	(0.0142)	(0.0000)
	Normal price	0.0025	0.9947	0.0028
		(0.0002)	(0.0048)	(0.0002)
	Spike	0.0000	0.1604	0.8396
		(0.0000)	(0.0147)	(0.0336)

market.

4. The main price drivers

In order to understand the reasons behind price jumps, Geman and Roncoroni (2006) and Mount et al. (2006) analyze the generator bid curves, particularly focusing on the transition in bids from low-cost high-supply generators to high-cost low-supply generators. Systematic changes in demand due to weather or business demands, reductions in supply due to scheduled infrastructure maintenance, and non-systematic reductions in supply due to generator or network failure are some of the factors that can shift the demand and supply curves (see Christensen et al. (2012)). The empirical literature has found that the occurrence of price jumps varies across time and supports the potential reasoning behind abnormal price events. For example, Escribano et al. (2002) and Knittel and Roberts (2005) show that for some markets, the intensity parameter of the spiking process in electricity prices demonstrates seasonal dependence. Kanamura and Ohashi (2007) also observe a seasonal dependency in the transition probabilities between jump and non-jump regimes due to systematic changes in demand.

The ability of the Nordic power system to store energy in hydro-reservoirs causes less variation in the Nordic price structure than that of, for example, Germany. Inflow during summer and in periods with low demand can be used in the winter. Furthermore, hydrologic forecasts have an impact on prices, since more or less energy than expected will cause the prices to go up or down. Hydropower reservoir levels are collected on a weekly basis from the beginning to the end of the tested period. Reservoirs are taken as a percentage of the total hydropower capacity available in the Nord Pool area. The reservoir levels and capacity data are from Norwegian Water Resources and Energy Directorate (NVE), Svensk Energi (Swedenergy AB), and the Finnish Environment Institute (SYKE). Reservoirs taken into account from Sweden and Finland are those after their integration in the Nord Pool market. Denmark is not included because its power production resources do not include any hydropower reservoir plant. In the Nord Pool electricity market, about 53% of power production is generated by hydropower reservoirs. The influence of reservoir levels on electricity futures prices at Nord Pool has been studied by Gjolberg and Johnsen (2001), Botterud et al. (2002), Forsund and Hoel (2004) and Fehr et al. (2005). The researchers conclude that hydropower reservoir levels are an important factor that explains futures and spot prices. The seasonality of reservoir levels has a highly important influence on electricity spot prices. From a storage theory perspective, inventory seasonals generate seasonals in the marginal convenience yield – and in the basis (see Fama and French (1987), p. 56). Taking reservoir levels as inventories of electricity, the effect of demand and supply shocks on spot electricity prices will depend on reservoir levels and how they are managed. Thus, demand or supply shock is easily offset when reservoirs are high. Reservoirs being low, a demand or supply change, are more difficult to balance and will be persistent, allowing spot prices to rise. In order to better understand the impact of reservoir levels on prices, two extreme cases can be studied in a hydropower generation market: very high reservoir levels and very low reservoir levels. In the case of an overflow the potential gains of producers will be reduced. This causes a negative convenience yield, that is, producers would prefer to sell power at a lower price than to allow overflows. As the main idea of hydropower management is to distribute water during the periods when reservoirs are nearly full, spot prices will be lower than usual and futures prices will be above spot prices. On the other hand, when reservoirs are very low, the convenience yield will be positive and might include big values. In this case, spot prices will be higher than short-term futures prices. If reservoir levels are not enough to satisfy demand, electricity prices will probably increase together with power imports. Furthermore, precipitation can be an important factor in order to obtain an estimation of the water inflow to hydroelectric reservoirs, and the expectation of a dry or rainy period will be clearly influenced by its values.

The behavior of weather variables can also produce some predictable seasonal patterns in spot prices. The relationship between weather variables and electricity load and price has been studied by many researchers. Temperature is the main price driver in Nordic countries. Cold temperatures increase heat demand, since electricity is very much used for heating in Nordic countries. Colder temperatures usually increase prices because of higher power demand. However, in special cases, e.g. combined heat and power plants where heat is the primary product, the demand for the heat could trigger secondary electricity production and cause the prices to decrease. Li and Sailor (1995), and Munoz et al. (1998) show in a few US states that temperature is the most significant weather variable explaining electricity and gas demand. The influence of air temperature has also been described by other authors who obtained a significant explicative power in their modeling; see, for example, Peirson and Henley (1994), Peirson and Henley (1998) and Pardo et al. (2002). Heating degree day (HDD) is a variable that shows the demand for energy needed for heating. It is taken from measurements of outside air temperature. The heating requirements for a specific structure in a specific place tend to be directly proportional to the number of HDDs in that location. In this study we will consider average temperature measured on a daily basis in 13 Nordic cities (Oslo, Bergen, Trondheim, Tromsø, Helsinki, Sodankyla, Vaasa, Tampere, Stockholm, Göteborg, Östersund, Luleå and Copenhagen). The data on heating degree days comes from `DegreeDays.net`.

Due to the fact that there is no fuel cost for production and unpredictability, additional wind energy can lead to a price decrease. This type of energy may, in some cases, cause even negative prices in hours with low demand and additional supply. On the other hand, when wind production falls short of expected values, it can trigger high prices in both day-ahead and intra-day markets.

Finally, electricity system prices from Elspot might contain important information about Elbas electricity prices. The same situation can be seen in futures markets, where the basis is the difference between the futures price and the underlying spot price. Figure 12 and 13 present the dependence between Elspot and Elbas electricity prices. Within a range of price fluctuation (20-60 EUR/MWh) we can see a very strong correlation between Elspot and Elbas prices. In general, Elspot electricity prices are higher than Elbas average electricity prices. The reason for that could be the additional supply caused by wind power in the Elbas market. As the Elspot electricity prices are formulated in a day-ahead market, they might be treated as an important explanatory variable for the Elbas average electricity prices and, consequently, as price jump occurrence factor.

Figure 12: Dependence between Elspot and Elbas electricity prices

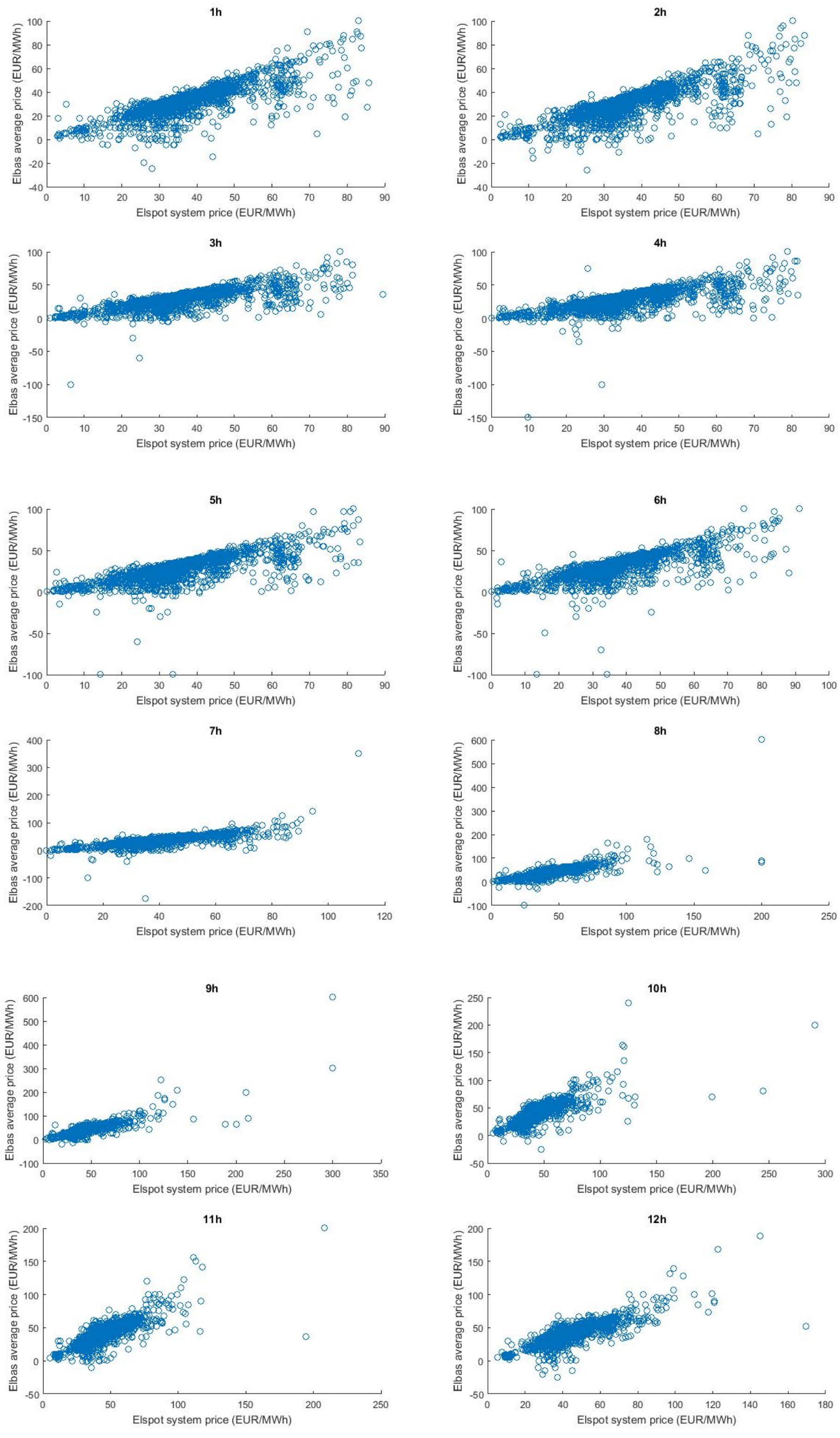
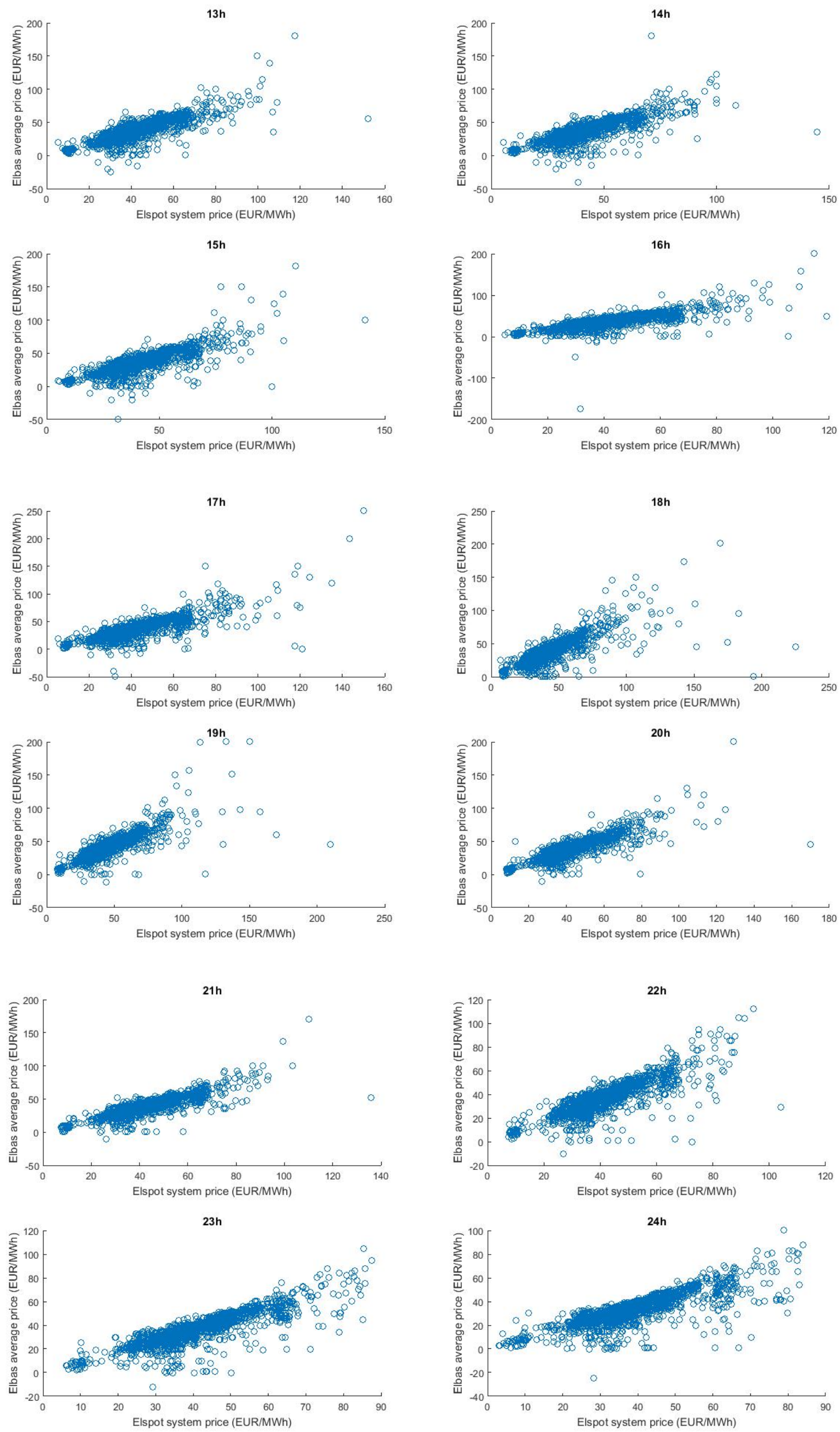


Figure 13: Dependence between Elspot and Elbas electricity prices



5. Models for electricity price jumps

Let P_t be the electricity spot price and consider the three states

$$Y_t = \begin{cases} 0, \ if \ -500 \leq P_t \leq 10 \ (\text{negative jump, drop}) \\ 1, \ \ if \ 10 < P_t < 80 \ (\text{normal price}) \\ 2, \ if \ 80 \leq P_t \leq 3000 \ (\text{positive jump, spike}) \end{cases} \tag{1}$$

Y_t is then a categorical time series of electricity prices. There are several approaches to modeling that type of processes. Here, we will consider three models: an autoregressive ordered probit (AOP) model, a first-order Markov model with a time-varying transition matrix, and an autoregressive conditional multinomial model. Each model provides different insights into the dynamics of the data-generating mechanism.

5.1. Autoregressive ordered probit model

Let Y_t^* denote a latent (unobserved) variable. The autoregressive ordered probit model assumes the autoregressive effect in the response variable and it is built around an auxiliary/latent regression model in the following way:

$$Y_t^* = X_t\beta + \rho Y_{t-1} + \varepsilon_t, \tag{2}$$

Based on Y_t^* the autoregressive ordered probit model defines the mechanism leading to the price being in a specific state as follows:

$$Y_t = \begin{cases} 0 \ if \ Y_t^* \leq \mu_0 \\ 1 \ \ if \, \mu_0 < Y_t^* \leq \mu_1 \\ 2 \ if \ \mu_1 < Y_t^*, \end{cases} \tag{3}$$

We assume that ε_t is normally distributed with zero mean and unit variance. The probability that electricity price Y_t at time t is in a specific state, conditional on explanatory variables X_t

and previous state Y_{t-1}, is then given by

$$
\begin{aligned}
P(Y_t = 0|X_t, Y_{t-1}) &= \Phi\left(-X_t\beta - \rho Y_{t-1}\right)\\
P(Y_t = 1|X_t, Y_{t-1}) &= \Phi\left(\mu_1 - X_t\beta - \rho Y_{t-1}\right) - \Phi\left(-X_t\beta - \rho Y_{t-1}\right)\\
P(Y_t = 2|X_t, Y_{t-1}) &= 1 - \Phi\left(\mu_1 - X_t\beta - \rho Y_{t-1}\right)
\end{aligned} \tag{4}
$$

where $\Phi()$ denotes the cumulative distribution function of the standard normal distribution, and $\theta = (\beta, \rho, \mu_0, \mu_1)$ is the vector of unknown parameters, including β vector, ρ scalar and $\mu = (\mu_0, \mu_1)$ parameters called cutpoints or threshold parameters.

The cutpoint parameters are estimated by the data and help to match the probabilities associated with each discrete outcome. These are subject to the obvious restriction $\mu_0 < \mu_1$. We assume that explanatory variables X_t and lagged response variable Y_{t-1} are observed, but that latent selection variable Y_t^* is not.

For the three probabilities, the partial (marginal) effects of changes in the explanatory variables and lagged response variable are

$$
\begin{aligned}
\tfrac{\partial P(Y_t=0|X_t)}{\partial X_t} &= -\phi\left(-X_t\beta\right)\beta\\
\tfrac{\partial P(Y_t=1|X_t)}{\partial X_t} &= \left[\phi\left(\mu_1 - X_t\beta\right) - \phi\left(-X_t\beta\right)\right]\beta\\
\tfrac{\partial P(Y_t=2|X_t)}{\partial X_t} &= \phi\left(\mu_1 - X_t\beta\right)\beta\\
\tfrac{\partial P(Y_t=0|Y_{t-1})}{\partial Y_{t-1}} &= -\phi\left(-Y_{t-1}\rho\right)\rho\\
\tfrac{\partial P(Y_t=1|Y_{t-1})}{\partial Y_{t-1}} &= \left[\phi\left(\mu_1 - Y_{t-1}\rho\right) - \phi\left(-Y_{t-1}\rho\right)\right]\rho\\
\tfrac{\partial P(Y_t=2|Y_{t-1})}{\partial Y_{t-1}} &= \phi\left(\mu_1 - Y_{t-1}\rho\right)\rho
\end{aligned} \tag{5}
$$

where $\phi()$ is the density function of the standard normal distribution.

We can evaluate these as sample means, or take a sample average of the marginal effects. Unlike the probit, the signs of the "interior" marginal effects are unknown and not completely determined by the signs of β and ρ. We can, however, sign the effects of the lowest and highest categories based on β, ρ. The others, however, cannot be known by the reader simply by looking at a table of point estimates.

Model parameters $\theta = (\beta, \rho, \mu)$ are estimated by partial maximum likelihood. In order to describe the estimation procedure we should note that the t'th observation of categorical time series Y_t expressed by vector $\mathbf{Y_t} = (Y_{t1}, Y_{t2})'$ of length $q = v - 1 = 2$ with elements

$$Y_{tj} = \begin{cases} 1, \text{ if the } j\text{th category is observed at time } t \\ 0, \text{ otherwise} \end{cases}$$

for $t = 2, \ldots, N$ and $j = 1, 2$. Denote by $\pi_t = (\pi_{t1}, \pi_{t2})'$ the vector of conditional probabilities given F_{t-1}

$$\pi_{tj} = E\left[Y_{tj}|F_{t-1}\right] = P\left(Y_{tj} = 1|F_{t-1}\right), \; j = 1, 2$$

for every $t = 2, \ldots, N$. At times we refer to the π_{tj} as "transition probabilities". As before, σ-field F_{t-1} stands for the whole information up to and including time t. Define

$$Y_{tm} = 1 - \sum_{j=1}^{q} Y_{tj}$$

and

$$\pi_{tm} = 1 - \sum_{j=1}^{q} \pi_{tj}.$$

In addition, put $\{W_{t-1} = (X_t, Y_{t-1}), t = 1, \ldots, N-1\}$ for the $p \times q$ matrix that represents a covariate process. Each response Y_{tj} corresponds to a vector of length p of random time-dependent covariates which forms the j'th column of W_{t-1}. The covariate matrix may contain lagged values of the response process and of any other auxiliary process as discussed earlier (Fokianos and Kedem, 2003, see). In estimation, X_t contains Elspot electricity system prices and turnover at a system price, wind power production, heating degree days, and water reservoir level.

We assume that the vector of transition probabilities - that is, conditional expectation of response vector Y_t - is linked to the covariate process through the equation

$$\pi_t(\theta) = h\left(W'_{t-1}\theta\right), \tag{6}$$

with θ being a p-dimensional vector of time-invariant parameters. The equation (6) gives the general form. In our case the $\pi_t(\theta)$ is given by

$$\pi_t(\theta) = \begin{pmatrix} \pi_{t1}(\theta) \\ \pi_{t2}(\theta) \end{pmatrix} = \begin{pmatrix} P(Y_t = 0|X_t, Y_{t-1}) \\ P(Y_t = 1|X_t, Y_{t-1}) \end{pmatrix} \tag{7}$$

where the probabilities $P()$ are defined as in equation (4).

Furthermore, we can obtain

$$\begin{aligned} Var[Y_t] =& \Sigma_t(\theta) = \\ =& \begin{pmatrix} \pi_{t1}(\theta)(1-\pi_{t1}(\theta)) & -\pi_{t1}(\theta)\pi_{t2}(\theta) \\ -\pi_{t1}(\theta)\pi_{t2}(\theta) & \pi_{t2}(\theta)(1-\pi_{t2}(\theta)) \end{pmatrix} \end{aligned}$$

The partial likelihood is a product of the multinomial probabilities

$$\prod_{j=1}^{3} \pi_{tj}^{Y_{tj}}(\theta)$$

that is,

$$PL = \prod_{t=2}^{N} \prod_{j=1}^{3} \pi_{tj}^{Y_{tj}}(\theta).$$

Thus, the partial log-likelihood is given by

$$l(\theta) = \log PL(\theta) = \sum_{t=2}^{N} \sum_{j=1}^{3} Y_{tj} \log \pi_{tj}(\theta).$$

The maximization of the log-likelihood function has to be done numerically. More details can be found in Fokianos and Kedem (2002).

5.2. Markov model

The second model is the first-order non-homogeneous Markov model for electricity price Y_t with $\upsilon = 3$ regimes whose transition matrix depends on some covariates. The transition matrix is parametrized by two $\upsilon \times 1$ vectors: π_t and ϕ_t. The π_t parameters control the probability of moving into each of the states unconditionally on the state from which it is moving, conditional on a move occurring. This single vector is not sufficient for modeling all of the dependence among the data. In particular, the probability of staying in a state given that the previous state is the same, is likely to be understated. The ϕ_t vector of probabilities is added to increase the chance of remaining in the same state. Each element of ϕ_t corresponds to the probability of moving to another state, although the new state may be the same as the old state. Thus, there are two mechanisms by which the new state can be the same as the old state: there can be no jump, or there can be a jump but to the same state. One appealing feature of this model is that it is flexible enough to accommodate states that form frequent small patches (i.e. a consecutive sequence of a state) as well as those that form occasional large patches with the same overall proportions. The parametrization based on the work of Foster et al. (2009) leads to transition probabilities $p_{t,ji}$ from state y_i to state y_j from observation number t to observation number $t+1$.

$$p_{t,ij} = P\left(Y_{t+1} = y_j | Y_t = y_i\right) = \begin{cases} (1-\phi_{ti}) + \phi_{ti}\pi_{tj}, & \text{if } i = j \\ \phi_{ti}\pi_{tj}, & \text{if } i \neq j \end{cases} \tag{8}$$

This can be represented in the form of the following time-varying transition matrix:

$$P_t = diag\left(1-\phi_t\right) + \pi_t \phi_t^T \tag{9}$$

The explanatory variables enter the model via logistic and additive logistic transformations of linear combination of the covariates (Aitchison (1982), Billheimer et al. (2002)). That is to say,

$$\phi_{ti} = \frac{\exp\left(x_{ti}^T\gamma_i\right)}{1+\exp\left(x_{ti}^T\gamma_i\right)} \tag{10}$$

$$\pi_{ti} = \begin{cases} \dfrac{\exp\left(u_t^T\beta_i\right)}{1+\sum\limits_{j=1}^{\upsilon-1}\exp\left(u_t^T\beta_j\right)} & \text{if } 1 \leq i \leq \upsilon-1 \\ 1-\sum\limits_{j=1}^{\upsilon-1}\pi_{tj} & \text{if } i = \upsilon \end{cases} \tag{11}$$

where x_{ti} $(1 \leq i \leq \upsilon)$ and u_t are the vectors of covariate values associated with the γ_i and β_j $(1 \leq j \leq \upsilon-1)$ parameter vectors respectively. We denote the entire set of parameters through two sets of vectors, namely γ and β with vector elements $|\gamma| = \upsilon$ and $|\beta| = \upsilon-1$, respectively. It is possible that the individual ϕ_{ti} are affected by different explanatory variables by specifying different x_{ti} in the logistic transformation (10). This is not sensible for the individual π_{ti} as they are all dependent on all parameters β_i ($1 \leq i \leq \upsilon - 1$) in the additive logistic transformation (11).

Estimation of the sets of parameters γ and β is carried out via direct maximization of the likelihood. Let us assume that a categorical time series or sample from chain $y_1, \ldots, y_N$ is observed. The joint probability of this realization is given by

$$P(Y_1 = y_1, \ldots, Y_N = y_N) = P(Y_1 = y_1) \prod_{t=2}^{N} P(Y_{t+1} = y_{t+1} | Y_t = y_t)$$

The log-likelihood is defined as

$$l(\gamma, \beta) = \log(P(Y_1 = y_1)) + \sum_{t=2}^{N} \log(P(Y_{t+1} = y_{t+1} | Y_t = y_t))$$

The log-likelihood is maximized using a Quasi-Newton algorithm.

5.3. Autoregressive conditional multinomial

The autoregressive conditional multinomial model was introduced by Russell and Engle (2005) and Russel (1996). The presentation of the model follows the work of Russell and Engle. Let us consider the categorical electricity price series where each Y_t is a random variable which may take one of υ states at time t, and consider a first-order Markov chain for the price change between different states.

The conditional distribution of Y_t is characterized by

$$\mathbf{\Pi_t} = \mathbf{P Y_{t-1}}$$

where $\mathbf{\Pi_t}$ denote a vector of conditional probabilities that the jth element of $\mathbf{Y_t}$ takes the value of 1, $\mathbf{P}$ is the $\upsilon \times \upsilon$ transition matrix, and $\mathbf{Y_t}$ is the jth column of the identity matrix if the jth state of Y_t occurred.

Transition matrix $\mathbf{P}$ must satisfy that all elements are nonnegative and all columns must sum to unity. In general settings, $\mathbf{P}$ might vary over time. However, the restrictions on $\mathbf{P}$ are directly satisfied by simple estimators when the transition matrix is constant, and imposing those restrictions for the time-varying case is not an easy task. Therefore, Russell and Engle suggested a different approach.

Let one state of the Y_t variable, state v, be chosen as the base state. The logs of odds ratios of variable Y_t taking the mth state against the vth state are defined as

$$\log\left(\frac{\pi_{tm}}{\pi_{tv}}\right) = \log\left(\sum_{j=1}^{v} P_{mj}\mathbf{Y}_{(t-1)j}\right) - \log\left(\sum_{j=1}^{v} P_{vj}\mathbf{Y}_{(t-1)j}\right) = \sum_{j=1}^{v-1} P^*_{mj}\tilde{\mathbf{Y}}_{(t-1)j} + c_m$$

where π_{ij} is the j-th element of Π_i and P_{ij} is the j-th element of the i-th row of the $\mathbf{P}$ matrix. Here, c_m is a scalar constant and P^*_{mj} denotes ratio $\log\left(\frac{P_{mj}}{P_{vj}}\right)$; $\tilde{\mathbf{Y}}$ is now the $v-1$ dimensional vector.

$(v-1)\times 1$ elements π_{mj}/π_{vj} are collected in matrix P^* of dimension $(v-1)\times(v-1)$. The probabilities of Y_t variable taking states $1,\ldots,v$ may be computed directly from the matrix. The probability the variable Y_t taking state v follows the restriction that the probabilities sum to 1. Moving from direct modeling of the transition matrix towards modeling logs of odds ratios allows for the following dynamic specification of the odds ratio equation

$$h(\pi_t) = \sum_{j=1}^{p} A_j\left(\mathbf{Y}_{t-j} - \mathbf{\Pi}_{t-j}\right) + \sum_{j=1}^{q} B_j h(\pi_{t-j}) + \gamma z_t$$

where $h()$ is the inverse logistic function, A_j and B_j denote the jth $(k-1)\times(k-1)$ parameter matrices, z_t contains exogenous variables and the constant, $\mathbf{\Pi_t}$ is the vector of transition probabilities at time t, and $\mathbf{Y_t}$ is the jth column of the identity matrix if the jth state of Y_t occurred.

The conditional probabilities of variable Y_t taking the kth state can be computed as follows:

$$\pi_{tk} = \frac{\exp\left(\sum_{j=1}^{p} A_{jk}\left(\mathbf{Y}_{t-j} - \mathbf{\Pi}_{t-j}\right) + \sum_{j=1}^{q} B_{jk}\log\left(\frac{\pi_{t-j,k}}{\pi_{t-j,v}}\right) + \gamma z_t\right)}{1 + \sum_{m=1}^{v-1}\exp\left(\sum_{j=1}^{p} A_{jm}\left(\mathbf{Y}_{t-j} - \mathbf{\Pi}_{t-j}\right) + \sum_{j=1}^{q} B_{jm}\log\left(\frac{\pi_{t-j,m}}{\pi_{t-j,v}}\right) + \gamma z_t\right)}$$

The model is relatively easy to interpret when $p = q = 1$, that is, for the $ACM(1,1)$ case. Matrix A determines the impact of previous periods on current transition probabilities and the eigenvalues of matrix B indicate how quickly the impact is weakening. If all transition probabilities are to be non-zero, the condition that all solutions of the equation in z, $|I - B_1 z - B_2 z^2 - \ldots - B_q z^q| = 0$ are smaller than 1 in absolute value must be satisfied.

The model is estimated with the maximum likelihood method. The likelihood can be represented as the product of the conditional densities. Letting π_{tj} denote the jth element of $\pi_{\mathtt{t}}$, the log-likelihood can be established as follows:

$$L = \sum_{t=1}^{N} \sum_{j=1}^{K} (Y_{tj} \log(\pi_{tj})) = \sum_{t=1}^{N} Y_t' \log(\pi_t) \tag{12}$$

The optimization of the likelihood function has to be done numerically with a proper optimization algorithm, e.g. the BHHH algorithm (see Berndt et al. (1974) for the details). The ML estimates of the parameters are consistent and asymptotically normal under common regularity conditions.

6. Estimation and forecasting results

This section contains estimation results based on the models described in Section 5 and the data considered in Section 3. We study the influence of the explanatory variables on extreme price event occurrence in Elbas electricity prices. As is discussed in Section 4, it is anticipated that the variables relating to the load (the turnover at system price), temperature, water reservoir and wind power production ought to influence the occurrence of extreme price events. Therefore, we use the load, Elspot prices, heating degree days, water reservoir level and wind power production as explanatory variables.

6.1. Estimation results

6.1.1. Model 1

The results of the maximum likelihood estimation of the autoregressive ordered probit model parameters for Elbas electricity prices are exhibited in Table 4, where standard errors are again calculated using the typical sandwich form. All of the explanatory variables are statistically significant. In accordance with the results of other studies for different power markets (see Kanamura and Ohashi, 2007; Mount et al., 2006; Christensen et al., 2009, 2012, among others), the coefficients of the load are positive and significant, indicating that higher values of load are associated with higher states of electricity prices (spike regime) and lower values might lead to negative electricity jumps. The heating degree days also have a statistically significant effect on Elbas prices. Wind power, as it was expected, has statistically significant negative effect on Elbas electricity price. The Elbas price is more likely to be in a higher state (regime) with a higher load, lower water reservoir level, lower wind production and higher heating degree day. The Elspot electricity price also has a statistically significant positive effect on Elbas electricity prices. Statistically significant persistence has been confirmed: Elbas price from the previous hour has a statistically significant positive effect on the current Elbas price. The thresholds/intercept parameters are significantly different from each other, so the three price states should not be combined into one and suggested thresholds for price drops and spikes seem to be reasonable.

Each of the price drivers is statistically significant and might be useful when forecasting the occurrence of price jumps. Concerning the measure of goodness of fit, the categorical R^2 has been computed as well as the corrected categorical R^2. Their values are 0.9899 and 0.8648 respectively. They confirm that the model fits to the data very well.

Table 4: Maximum likelihood estimates of the autoregressive ordered probit model parameters

Coefficients	Estimate	Std. Error	z value	$Pr(>\|z\|)$
$y_t = 1$ (Elspot)	1.5108	0.0875	17.281	2.00E-16
$y_t = 2$ (Elspot)	2.9003	0.1127	25.726	2.00E-16
Load	0.3360	0.0290	11.606	2.00E-16
HDD	0.0095	0.0026	3.682	0.0002
water reservoir	-0.0020	0.0001	-2.906	0.0036
wind power	-0.0026	0.0001	-17.983	2.00E-16
$y_{t-1} = 1$ (Elbas)	2.2000	0.0353	62.329	2.00E-16
$y_{t-1} = 2$ (Elbas)	4.0941	0.0830	49.283	2.00E-16
$0\|1$	2.4358	0.1187	20.52	2.00E-16
$1\|2$	7.5855	0.1321	57.4	2.00E-16

Table 5 contains marginal effects of the autoregressive ordered probit model for Elbas electricity prices. Marginal effects are calculated as the mean of the independent variables. The results are in accordance with the expectations. The highest impact on the changes in probabilities associated with Elbas prices being in different states has a previous state of Elbas prices. If the past Elbas price state was in a normal state, then we will have, on average, a 46.7% decrease in the probability of the current Elbas price being in drop regime, a 46.5% increase in the probability of being in a normal price state, and a 0.2% increase in the probability of being in a spike state. If the past Elbas price was in a spike state, then we will have, on average, a 0.21% decrease in the probability of the current Elbas price being in a drop state, a 81.3% decrease in the probability of being in a normal price regime, and a 83.4% increase in the probability of being in a spike price state. If load (the turnover at system prices) increases by 10,000 MWh, then we will have, on average, a 1.5% decrease in the probability of the Elbas price being in a drop state, a 1.4% increase in the probability of being in a normal price regime, and a 0.1% increase in the probability of being in a spike regime. If the Elspot price is in a normal regime then we will have on average 24.7% decrease in the probability of Elbas price being in a drop regime, a 24.6% increase in the probability of being in a normal price state, and a 0.1% increase in the probability of being in a spike state. If the Elspot price state is in a spike state, then we will have, on average, a 2.1% decrease in the probability of Elbas price being in a drop state, a 39.1% increase in the probability of being in a normal price regime and a 41.3% increase in the probability of being in a spike price state. Changes in three of the explanatory variables - heating degree days, water reservoir level and wind power production - (although statistically significant) turn out to have a close-to-zero average effect on the probability of the price being in any of the considered states.

Table 5: Marginal effects

	$y_t = 0$	$y_t = 1$	$y_t = 2$
$y_t = 1$ (Elspot)	-0.247	0.246	0.001
$y_t = 2$ (Elspot)	-0.021	-0.391	0.413
Load	-0.015	0.014	0.001
HDD	0.000	0.000	0.000
water reservoir	0.000	0.000	0.000
wind power	0.000	0.000	0.000
$y_{t-1} = 1$ (Elbas)	-0.467	0.465	0.002
$y_{t-1} = 2$ (Elbas)	-0.021	-0.813	0.834

6.1.2. Model 2

Table 6 contains maximum likelihood estimates of parameters of the π_t and ϕ_t matrices. Three explanatory variables are used: Elspot prices, load, and the heating degree days. All explanatory variables turn out to be statistically significant, but their influence on the transition probabilities is different. π_t parameters control the probability of moving into each of the states unconditionally on the state from which it is moving, conditional on a move occurring. ϕ_t corresponds to the probability of moving to another state, although the new state may be the same as the old state.

Table 6: Maximum likelihood estimates of the parameters of ϕ_t and π_t matrices

Probability	Parameters	Explanatory variable	Estimate	St Error	z value
ϕ_1	γ_{11}	y_t (Elspot)	0.7237	0.3201	2.2605
	γ_{12}	Load	-0.4885	0.2704	-1.8063
	γ_{13}	HDD	0.0152	0.0233	0.652
ϕ_2	γ_{21}	y_t (Elspot)	0.6454	0.2685	2.4038
	γ_{22}	Load	-1.7767	0.1963	-9.0522
	γ_{23}	HDD	0.2409	0.0147	16.3883
ϕ_3	γ_{31}	y_t (Elspot)	-1.0336	0.222	-4.6554
	γ_{32}	Load	0.5409	0.1954	2.7686
	γ_{33}	HDD	-0.0195	0.0252	-0.7736
π_1	β_{11}	y_t (Elspot)	1.2018	0.3185	3.7733
	β_{12}	Load	-1.0118	0.2577	-3.9262
	β_{13}	HDD	0.2161	0.0191	11.3294
π_2	β_{21}	y_t (Elspot)	1.1486	0.2969	3.8683
	β_{22}	Load	-2.7392	0.242	-11.3211
	β_{23}	HDD	0.3695	0.0209	17.6475

6.1.3. Model 3

Table 7 contains the maximum likelihood estimates of the parameters of $ACM(1,1)$ with explanatory variables. The eigenvalues of matrix B indicate how quickly the impact fades. The B parameter matrix has the important property that all of its eigenvalues are less than one in absolute value. This means that the model is correctly specified in terms of the properties of the state change process, as the estimated process should assume every possible state infinitely many times in the future. Moreover, parameter matrix B implies that the price regime change process is of relatively high persistence, i.e. if some event increased the probability of a price regime change at time t, then at time $t+1$ the effect will still be present. The difference between the estimates of b_1 and b_2 is statistically significant. Matrix A determines the impact of previous periods on current transition probabilities. The estimates of the parameters on the diagonal of A, i.e. a_{11} and a_{22}, are both positive, which means that after each price regime increases the probability of another price regime increase. Positive value a_{12} implies that after the price regime changes, the probability of an opposite change increases. The parameters of explanatory variables are all statistically significant. Moreover, the positive values of g_{12} and g_{22} imply that the probability of being in a normal or spike regime increases with the increase of turnover at system price. On the other hand, the negative values of g_{15} and g_{25} suggest that the probability of the price being in a normal or spike regime decreases when wind power production increases.

Table 7: Maximum likelihood estimates of $ACM(1,1)$ model parameters

	Estimate	Std Error	z value	$P(>\|z\|)$	Explanatory variable
c_1	-1.43	0.13	-11.00	2.1188E-27	
a_{11}	3.19	0.08	40.91	0	
a_{12}	1.23	0.27	4.51	1.4959E-05	
b_1	0.77	0.01	59.46	0	
g_{111}	1.48	0.11	13.12	1.5857E-38	$y_t = 1$ (Elspot)
g_{112}	1.20	0.16	7.44	3.7828E-13	$y_t = 2$ (Elspot)
g_{12}	0.35	0.04	10.11	2.4379E-23	Load
g_{13}	-0.01	0.00	-4.00	0.0001	HDD
g_{14}	-0.01	0.01	-1.50	0.1295	water reservoir
g_{15}	-0.02	0.00	-10.00	7.6946E-23	wind power
c_2	-12.03	18.56	-0.65	0.3234	
a_{21}	3.93	0.60	6.60	1.3563E-10	
a_{22}	4.49	0.64	7.07	5.7368E-12	
b_2	0.32	0.07	4.30	3.8985E-05	
g_{211}	4.53	18.58	0.24	0.3873	$y_t = 1$ (Elspot)
g_{212}	6.43	18.58	0.35	0.3757	$y_t = 2$ (Elspot)
g_{22}	1.30	0.24	5.39	1.9770E-07	Load
g_{23}	0.16	0.02	7.09	4.8143E-12	HDD
g_{24}	-0.21	0.06	-3.25	0.0020	water reservoir
g_{25}	-0.08	0.02	-5.47	1.2930E-07	wind power

Note:

$$h(\pi_t) = c + A \cdot (Y_{t-1} - \pi_{t-1}) + B \cdot h(\pi_{t-1}) + \gamma z_t$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & 0 \\ 0 & b_{22} \end{bmatrix}$$

$$\gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ ... & ... \\ \gamma_{71} & \gamma_{72} \end{bmatrix}$$

7. Forecasting results

To assess the forecasting performance of the considered models, the following procedure is adopted. The parameters of the models are estimated using the sample period data (14 September 2009 to 31 December 2013). The observations from 1 January 2014 to 31 December 2015 are used for the forecast performance evaluation. For each model, one-step-ahead forecasts from an extended window are computed. The model parameters are reestimated once a new observation is included and then one-step-ahead forecasts are made. We do not attempt to forecast the explanatory variables. As the explanatory variables forecast, the actual values are taken. The prediction in the Markov model and an autoregressive conditional multinomial model is based on the mode of the conditional distribution of Y_{t+1}, given that $Y_t = y_j$ is y_j, the last realization of the categorical time series. The choice of hour-ahead forecasts is driven by the fact that the Elbas operates as a continuous-trading market for each hour interval. In the event of a price jump being forecasted for the next hour interval, effective risk management requires an immediate action on the part of retailers in order to mitigate the effects of the potential price jump. Retailers can reduce their reliance on the pool to meet their demand by activating the standby capacity that they may have available. Both hydroelectric and gasfired peaking plants can be brought from an idle state to full capacity in an hour. If this physical response is not available, an alternative strategy is to use the futures market, which trades in real time.

Table 8 presents the results of the forecasting exercise based on the data from 1 January 2014 to 31 December 2015. The out-of-sample time series does not contain any observations above 80 EUR/MWh. Therefore, we will focus only on the correct predicting of price drops and the full 3×3 confusion matrix can not be constructed. Table 8 contains information on the number of correct predictions of price jump occurrence based on four models. As a benchmark, two simple models are considered: the Markov model, in which transition matrix is constant in time, and the ordered probit model with seasonal dummies as explanatory variables. The Markov model was not able to correctly predict any of the price jump events. The ordered probit model with seasonal dummies was able to predict 2037 jump occurrences. This shows that it is necessary to incorporate some additional information in explanatory variables in order to be able to predict jump occurrences more accurately. The best forecasts of price jump occurrences are based on the ACM model. The model was able to correctly predict 2856 jump events from a total of 3527. The second model is the autoregressive ordered probit model with 2797 correct predictions. Slightly worse is the non-homogeneous Markov model with 2755 correct predictions of price jump occurrences. There is also a difference between the models with and without explanatory variables. We show that it is crucial for the price jumps prediction to include the explanatory variables in the model.

Table 8: Forecasting of price jump occurrences

Jump events forecasting		
autoregressive ordered probit (with explanatory variables)	Correct	2797
	False	730
autoregressive ordered probit (without explanatory variables)	Correct	1200
	False	2327
non-homogeneous Markov model (with explanatory variables)	Correct	2755
	False	772
non-homogeneous Markov model (without explanatory variables)	Correct	1158
	False	2369
autoregressive conditional multinomial (with explanatory variables)	Correct	2856
	False	671
autoregressive conditional multinomial (without explanatory variables)	Correct	1353
	False	2174
homogeneous Markov model (benchmark)	Correct	0
	False	3527
ordered probit model with seasonal dummies (benchmark)	Correct	2037
	False	1490

8. Conclusions

The accurate forecast of extreme price events is of high importance for risk management in the electricity sector. However, most of the electricity-pricing models are an adaptation of models for prices or returns from financial econometrics with a different degree of success. Opposed to the majority of contributions to this field, the current paper focuses on the forecasting of extreme price events, the occurrence of which is treated as a realization of categorical time series. An autoregressive ordered probit, a Markov model, and an autoregressive conditional multinomial model were used to analyze the drivers of the process and to forecast extreme price events. We show that it is crucial for correct jump occurrence prediction to build a statistical model which takes into account both persistence in the series and external information on price drivers. High loads were found to have a significant impact on the probability of a price spike and low loads were found to increase the probability of a price drop. Essentially, the persistence of the categorical price process was also confirmed to be significant in explaining the occurrence of extreme price events. The considered models provide hour-ahead forecasts of price jumps that are superior to the forecasts made by a memoryless model and an ordered probit model with seasonal dummies as explanatory variables. The best forecasts of the extreme price events are obtained based on the $ACM(1,1)$ model.

J. Aitchison. *The Statistical Analysis of Compositional Data*, volume 44. Journal of the Royal Statistical Society, 1982.

N. Amjady and E. Keynia. A new prediction strategy for price spike forecasting of day-ahead electricity markets. *Applied Soft Computing*, 11:4246–4256, 2011. ISSN 1568-4946. doi: http://dx.doi.org/10.1016/j.asoc.2011.03.024.

M. Barlow. A diffusion model for electricity prices. *Mathematical Finance*, 12:287–298, 2002. doi: 10.1007/978-4-431-53947-6_25.

R. Becker, S. Hurn, and V. Pavlov. Modelling Spikes in Electricity Prices. *The Economic Record*, 83:371–382, 2007.

E. Berndt, B. Hall, R. Hall, and J. Hausman. Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3:653–665, 1974.

A. Botterud, A.K. Bhattacharyya, and M. Ilic. Futures and spot prices–an analysis of the scandinavian electricity market. In *Proceedings of North American Power Symposium*, pages 1–8, 2002.

A. Cartea, M.G. Figueroa, and H. Geman. Modelling Electricity Prices with Forward Looking Capacity Constraints. *Applied Mathematical Finance*, 16:103–122, 2012.

C. Chen and L.M. Liu. Forecasting time series with outliers. *Journal of Forecasting*, 12(1): 13–35, 1993a.

C. Chen and L.M. Liu. Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association*, 88(421):284–297, 1993b.

T. Christensen, S. Hurn, and K. Lindsay. It never rains but it pours: modeling the persistence of spikes in electricity prices. *The Energy Journal*, 207:543–556, 2009.

T.M. Christensen, A.S. Hurn, and K.A. Lindsay. Forecasting spikes in electricity prices. *International Journal of Forecasting*, 28:400–411, 2012. ISSN 0169-2070.

A. Clements, J. Fuller, and S. Hurn. Semi-parametric forecasting of spikes in electricity prices. *Economic Record*, 89:508–521, 2013. ISSN 1475-4932. doi: 10.1111/1475-4932.12072.

A. Escribano, J.I. Pena, and P. Villaplana. *Modeling electricity prices: international evidence*. PhD thesis, Oxford Bulletin of Economics and Statistics, 2002.

A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. Wiley Finance, Hoboken, N.J, 2003. ISBN 978-0-471-10400-1.

E.F. Fama and K.R. French. Dividend yields and expected stock returns. *Journal of Financial Economics*, 22:3–25, 1987.

N. Von Der Fehr, E. Amundsen, and L. Bergman. The nordic market: signs of stress? *The Energy Journal Issue on European Electricity Liberalization*, pages 71–98, 2005.

K. Fokianos and B. Kedem. *Regression Models for Time Series Analysis*. Wiley Finance, 2002.

K. Fokianos and B. Kedem. Regression Theory for Categorical Time Series. *Statistical Science*, 18:357–376, 2003.

F. Forsund and M. Hoel. Properties of non-competitive electricity market dominated by hydro-electric power. *SSRN Electronic Journal*, 30:1116–1157, 2004.

S.D. Foster, M.V. Bravington, A. Williams, F. Althaus, G.M. Laslett, and R.J. Kloser. Analysis and prediction of faunal distributions from video and multi-beam sonar data using markov models. *Environmetrics*, 20:541–560, 2009. doi: 10.1002/env.952.

H. Geman and A. Roncoroni. Understanding the fine structure od electricity prices. *The Journal of Business*, 79:1225–1262, 8 2006.

O. Gjolberg and T. Johnsen. Electricity futures: Inventories and price relationships at nord pool. Technical report, Norwegian School of Economics and Business Administration, Sept 2001. Working paper, Norwegian School of Economics and Business Administration.

C. Harris. *Electricity Markets: Pricing, Structures and Economics*. Wiley, 2006.

J. Hellstrom, J. Lundgren, and H. Yu. Why do electricity prices jump? empirical evidence from the nordic electricity market. *Energy Economics*, 34:1774–1781, 2012. ISSN 0140-9883.

J. Janczura, S. Trueck, R. Weron, and R. Wolff. Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics*, 38:96–110, 2013.

S. Johansen and B. Nielsen. Asymptotic Theory of Outlier Detection Algorithms for Linear Time Series Regression Models. *Scandinavian Journal of Statistics*, 43(2):321–348, 2016.

C. De Jong. The nature of power spikes: A regime-switch approach. *Nonlinear Analysis of Electricity Prices*, 10:27–46, 2006.

T. Kanamura and K. Ohashi. A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation. *Energy Economics*, 2007.

C.R. Knittel and M.R. Roberts. An empirical examination of restructured electricity prices. *Energy Economics*, 27:791–817, 2005.

N. Krichene. World crude oil and natural gas: a demand and supply model. *Energy Economics*, 24:557–576, 2002.

X. Li and D.J. Sailor. Electricity use sensitivity to climate and climate change. *World Resource Review*, 1995.

E. Lindstrom, V. Noren, and H. Madsen. Consumption management in the nord pool region: A stability analysis. *Applied Energy*, 146:239–246, 2015. ISSN 0306-2619. doi: http://dx.doi.org/10.1016/j.apenergy.2015.01.113.

X. Lu, Z.Y. Dong, and X. Li. Electricity market price spike forecast with data mining techniques. *Electric Power Systems Research*, 73(1):19–29, 2005. ISSN 0378-7796. doi: http://dx.doi.org/10.1016/j.epsr.2004.06.002.

P. Maryniak and R. Weron. Forecasting the occurence of electricity price spikes in the UK power market. *HSC Research Reports*, 14, 2014.

T.D. Mount, Y. Ning, and X. Cai. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics*, 28:62–80, 2006. ISSN 0140-9883.

J. R. Munoz, J. Rosen, and D.J. Sailor. Natural gas consumption and climate: a comprehensive set of predictive state-level models for the united states. *Energy the International Journal*, 23:91–103, 1998.

J.R. Pardo, M. Ridal, D. Murtagh, and J. Cernicharo. Microwave temperature and pressure measurements with the odin satellite: I. observational method. *Can. J. Phys.*, 80:443–454, 2002.

J. Peirson and A. Henley. Electricity load and temperature: Issues in dynamic specification. *Energy Economics*, 16:235–243, 1994.

J. Peirson and A. Henley. Residential energy demand and the interaction of price and temperature: British experimental evidence. *Energy Economics*, 20:157–171, 1998.

J.R. Russel. *Econometric Analysis of Irregularly-Spaced TransactionData Using a New Class of Accelerated Failrue Time Models with Applicationsto Financial Transaction Data*. PhD thesis, University of California, San Diego, 1996.

J. R. Russell and R.F. Engle. A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial: Autoregressive conditional duration model. *Journal of Business & Economic Statistics*, 23(2):166–180, 2005. doi: 10.2307/27638809.

J.A. Villar and F.L. Joutz. The relationship between crude oil and natural gas prices. Eia manuscript, Energy Information Administration, 2006.

S. Voronin and J. Partanen. Price Forecasting in the Day-Ahead Energy Market by an Iterative Method with Separate Normal Price and Price Spike Frameworks. *Energies*, 6:5897–5920, 2013.

S. Voronin, J. Partanen, and T. Kauranne. A hybrid electricity price forecasting model for the nordic electricity spot market. *International Transactions on Electrical Energy Systems*, 24: 736–760, 2014. ISSN 2050-7038.

R. Weron. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, Chichester, 2006.

R. Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081, 2014. ISSN 0169-2070.

Research Papers 2016

CREATES
Center for Research in Econometric Analysis of Time Series

2016-23: Yunus Emre Ergemen and Carlos Vladimir Rodríguez-Caballero: A Dynamic Multi-Level Factor Model with Long-Range Dependence

2016-24: Shin Kanaya: Convergence rates of sums of α-mixing triangular arrays: with an application to non-parametric drift function

2016-25: Gustavo Fruet Dias, Marcelo Fernandes and Cristina M. Scherrer: Improving on daily measures of price discovery

2016-26: Martin M. Andreasen, Tom Engsted, Stig V. Møller and Magnus Sander: Bond Market Asymmetries across Recessions and Expansions: New Evidence on Risk Premia

2016-27: Kim Christensen, Ulrich Hounyo and Mark Podolskij: Testing for heteroscedasticity in jumpy and noisy high-frequency data: A resampling approach

2016-28: Kim Christensen, Roel Oomen and Roberto Renò: The Drift Burst Hypothesis

2016-29: Hossein Asgharian, Charlotte Christiansen, Rangan Gupta and Ai Jun Hou: Effects of Economic Policy Uncertainty Shocks on the Long-Run US-UK Stock Market Correlation

2016-30: Morten Ørregaard Nielsen and Sergei S. Shibaev: Forecasting daily political opinion polls using the fractionally cointegrated VAR model

2016-31: Carlos Vladimir Rodríguez-Caballero: Panel Data with Cross-Sectional Dependence Characterized by a Multi-Level Factor Structure

2016-32: Lasse Bork, Stig V. Møller and Thomas Q. Pedersen: A New Index of Housing Sentiment

2016-33: Joachim Lebovits and Mark Podolskij: Estimation of the global regularity of a multifractional Brownian motion

2017-01: Nektarios Aslanidis, Charlotte Christiansen and Andrea Cipollini: Predicting Bond Betas using Macro-Finance Variables

2017-02: Giuseppe Cavaliere, Morten Ørregaard Nielsen and Robert Taylor: Quasi-Maximum Likelihood Estimation and Bootstrap Inference in Fractional Time Series Models with Heteroskedasticity of Unknown Form

2017-03: Peter Exterkate and Oskar Knapik: A regime-switching stochastic volatility model for forecasting electricity prices

2017-04: Timo Teräsvirta: Sir Clive Granger¤s contributions to nonlinear time series and econometrics

2017-05: Matthew T. Holt and Timo Teräsvirta: Global Hemispheric Temperatures and Co–Shifting: A Vector Shifting–Mean Autoregressive Analysis

2017-06: Tobias Basse, Robinson Kruse and Christoph Wegener: The Walking Debt Crisis

2017-07: Oskar Knapik: Modeling and forecasting electricity price jumps in the Nord Pool power market