PREDICTING PEAK PRICES IN THE CURRENT DAY-AHEAD MARKET

Bernt Arild Bremdal
UiT, The Arctic University of Tromsø,
and Smart Innovation Norway
Narvik, Norway
bernt.a.bremdal@uit.no

Shayan Dadman
UiT, The Arctic University of Tromsø,
Narvik, Norway
shayan.dadman@uit.no

ABSTRACT

The work presented has studied price developments in the day-ahead market. The statistical properties inherent in the time series constituting the day-to-day prices have been investigated. It is shown that these properties are radically different in 2022 than former years and resemble more chance like properties for which make predicting future peak prices especially hard. To overcome this, a space domain approach was adopted to determine whether information about co-variant price zones could improve predictions. By using a multi-variate regression method, it is possible to accurately predict prices in NO2 using information about concurrent prices in other price zones. However, an attempt to use historic prices in multiple price zones to predict future prices gave only modest results. One reason for this is the high degree of entropy underlying price developments in the day-ahead market in 2022.

INTRODUCTION

During the recent months the energy market in Europe has been subjected to very high electricity prices due to some general market developments, combined with an extraordinary geo-political situation. Most of the European continent is affected. In Norway, specifically, the high prices and increased price volatility represent a dramatic departure from former years when prices varied from moderate in 2018 to very low in former years. The new situation causes a high impact on the economy of household consumers and businesses alike, but they also unveil challenges in the grid. Political debates and media coverage in the current situation tend to clutter principal facts that arise from the empirical data gathered over the past months. Besides, there is real need for revising present decision support tools that have been based on such empiricism. In general, there is a need to improve preparedness through better forecasting of dramatic price shifts from hour to hour. The objective behind the work presented was to determine what statistical properties in former years that are no longer valid, and which make legacy tools perform less well in the current situation. The research reported has also tried to determine on what basis prediction systems in the present market situation can be established. With a special focus on sudden price peaks and price drops we have aimed at an approach were predictions in the time domain is combined with input in the space domain, involving studies of mutual and concurrent reactions across price zones. The work reported has not looked much at the underlying market structures and deeper distribution concepts. It has taken a data centric approach and tried to determine subtle causalities through multi-variable statistics and machine learning. The principal focus for the analyses has been the recent situation in the NO2 price zone in Norway.

RELATED WORK

Forecasting electricity prices has been widely investigated in different power markets. But most of them have been designed for expected price analysis and not price spike forecasting [1] High price volatilities and price spikes are typical a result of a number of factors such as weather uncertainty, fluctuating fuel prices, transmission bottlenecks and more. This makes the task of accurate price forecasting a big challenge for the market participants. A number of models and data sources have been proposed by researchers, however, most of them find that high accuracy is often hard to achieve. [2] investigated the mechanisms that governed price spikes in the Australian energy market. Their approach succeeded in predicting price spikes on top of normal prices with a high level of confidence based on a Bayesian classification method and similarity searching techniques using the K-nearest neighbour approach. Not surprisingly the main factor affecting the market price was the demand–supply relationship. Shrivastava et al. [3] addressed the PJM day-ahead market and the whole sale electricity market in Ontario. Their principal approach applied extreme learning machines (ELM). Three models were created. A short-term classifier with a look ahead of one hour or less, a model that could support predictions of prices for the next 24 hours and a model with a forward horizon longer than a day. In their experiments they observed that the relative performance of the three models varied throughout the year. In some months, all the models performed relatively well. But during the summer months the disparity between the models was found to be high. They concluded that this could be due to high uncertainty and the volatile nature of the data for this period. Ottesen et al [4] addressed the Nordic energy market 2016 in relation to multi-market bidding and energy flexibility. Data sources included prices from Nord Pool’s Elspot day-ahead market and Statnett’s Tertiary reserves market (“RK”). They found that the value of energy flexibility improved with the more precise price predictions. However, in periods with spiky price situations proper price forecasting was difficult using their stochastic programming approach. The approach that comes closes to the work presented here, in addition to [1] is the combined decision tree and genetic algorithm method introduced by [5]. The most obvious and intuitive lesson from such historic efforts is of course that prices vary with supply and demand. But the factors that drive supply and demand have changed. Extrapolating on empiricism from former years alone, as several of the historic references, do may not offer the solution to the price forecasting problem, forward.

CHANGE IN SUPPLY AND DEMAND CHARACTERISTICS

The energy mix on the supply side is changing. The permanent downscaling and decommissioning of nuclear power plants play a role in this shift. Intermittent production based on wind and solar power is meant to replace both nuclear and coal fuelled plants. A market designed according to the marginal cost principle and a single settlement price will naturally be influenced by this. Variations in fuel prices for the remaining fossil fuelled plants add to the instability in energy prices too. Rapid temperature shifts, seasonal idiosyncrasies, the average price level itself and increased electrification of appliances and vehicles are among those factors responsible for the price increase and volatility on the demand side. And there is a pending climate issues that seemingly cast it shadows over the market. There are reasonable justified claims that European winters and summers are no longer what they used to be. Heat waves, droughts, a sudden chill with heavy showers and floods are more frequent during the summers causing more instability. In winter, people and businesses, may experience frost and very low temperatures changing to moist and mild temperatures (or even summerly degrees) in a matter of a few hours. All this accounts for a change of pace in demand from day to day and from hour to hour. Price impacts are also caused by capacity related issues in the transmission system. The increased connectivity between the different markets and price zones in Europe tend to develop complex interference patterns that can cause different types of temporary price couplings which in the next turn can be exploited for speculative purposes by market participants. Again, this influence price developments. Most recently a pandemic and the war in Ukraine have created an acute situation which has amplified instabilities stemming from these more fundamental market developments. Europe is faced with an unprecedented post-World War 2 situation, which also upsets traditional extrapolative methods, including price forecasting on history alone. The research presented here is part of an ongoing effort which attempts to offer a perspective on this new situation and to create a basis for peak price prediction without diving too much into the significant complexities that govern the market now.

METHOD ADAPTED

At the outset historic price developments were compared. The price distributions for different years were contrasted and cross entropies analysed. Focus was placed on the NO2 price zone in Norway. This price zone has experienced very high prices and rapid price shifts during 2021 and 2022, while during comparable periods in pre-pandemic years the same price zone experienced moderate to very low prices. Like many previous efforts a data scientific approach has been applied to address the problem. Both multivariate statistics and machine learning methods have been used. Different machine learning approaches were applied to determine prices in 2022 based on historic records. With the assumption that price couplings could impact the prices in NO2 efforts were set out to investigate this. This led to the hypothesis that different price zones in the Nordic countries and elsewhere in Europe could impact what [2] termed "price spikes on top of normal prices". In the reported work, a theory was developed that the prices in different zones could be used as sensors to determine the hourly prices in NO2. Prices in these zones could help predict the next day prices in NO2. Hence, this could cater for a simple time and space model. A data driven approach was adopted where price records for NordPool for multiple European prize zones, across several months were collected and studied. Each price zone addressed displayed different characteristics; some with ample and steady coal-based production, some with high-capacity connectors to other price zones, some with supply very much exposed to wind production, others with industry with high energy demands. Instead of modelling each of these characteristics individually and to avoid increased complexity we assumed that the zonal price reflected all of these characteristics and the potential interference with other zones. The question was which ones. This led to an analysis of possible co-variant developments between different zones that could shine light on peak price developments in NO2. Finally, a multi-variate regression using XGBoost was performed for the period January 1 to August 30 in 2022. This was followed by the creation of a predictor built on the platform established by the regressor.

RESULTS

*Table 1Statistical properties of the density functions for the years 2018,2019, 2022 (*Note that the prices are given in NOK/MWh and that the numbers for 2022 do not represent the full year)*

Year	Mean	Median	St.deviation
2018	415,46	368,89	110,88
2019	386,64	403,13	82,07
2022*	3102,65	3102,86	1707,51

The results generated show changes in the statistical properties in recent price developments that have not been present before. This can be shown. The probability densities for prices in three different years are shown in Figure 1. Both 2018 and 2022 are compared with year 2019. The similarities between 2018 and 2019 are quite apparent. Earlier years are not much different from 2018 and 2019. However, the discrepancy between 2019 and 2022 is very distinct. Not only does the graph for 2022

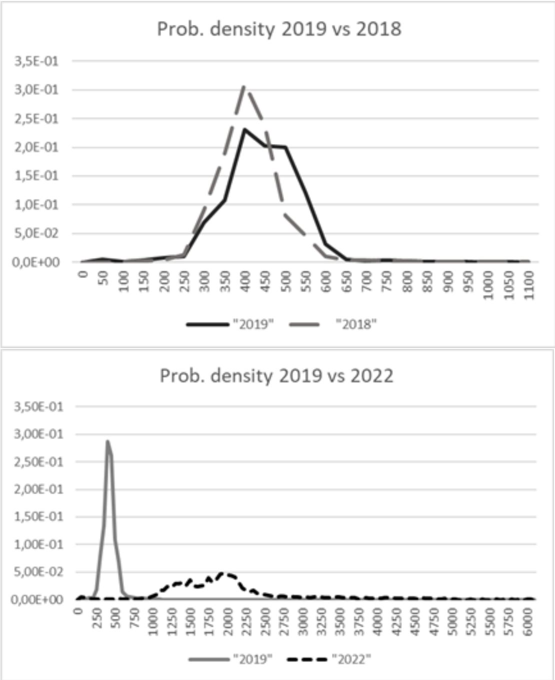

Figure 1 Probability densities for NO2 prices for three different years. Prices in 2019 compared to prices in 2018 and 2022

Table 2Table 2 Shows average results from multiple trials with different neural networks and data from both 2022 and 2022 - 2018.

Method	**MAE**	**Peak and valley timing**	**High price peaks and deep valleys**
NBEATS	1,90	Fair	Poor
TCN	1,57	Fair	Poor
LSTM	2,03	Fair	Poor
GRU	1,95	Poor	Poor
Transformer	1,89	Poor	Poor

display extreme prices compared to 2019 and 2018, but the density function is stretched out and does not show a very distinct top for the median price, which the years 2019 and 2018 do. This is reflected in the statistical properties of the distributions too (see Table 1). Two important things can be extracted from this. The 2022 prices have not gravitated around a specific price band symmetric about a distinct mean or median. The flat density function and an almost indistinctive median shows certain characteristics typical for white noise. This implies that the statistical entropy for 2022 is very high, exposing a spectrum of prices with a long, flat distribution with few occurrences per price interval. This has been established mathematically. The entropy associated with 2022 prices is close to 1. Hence, it makes sense to look outside the time domain to find other or complementary ways to produce good future predictions for peak prices in NO2. The cross-entropy between 2019 and 2022 has been determined and the KL-divergence measures 3.35. In comparison the KL-divergence between the probability distributions between

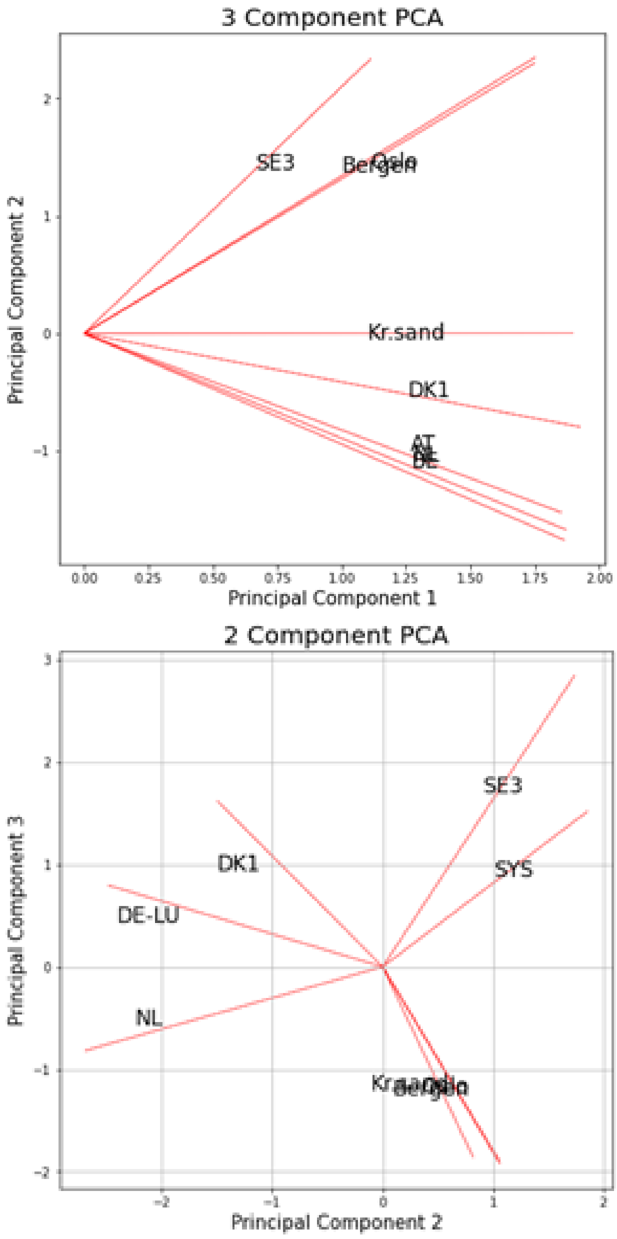

Figure 2 The PCA analysis exposed strong co-variant behaviour between different price zones. Component 1 and 2 accommodated for app. 85% of the variance. But the variance summed up for all three accounted for 93%.

years 2018 and 2019 is 0,15. Extrapolating on the price history in 2018 to predict 2019 prices are a possibility, while using historic data for the same purpose in 2022 would not be very useful. Empirical prediction models based on such is likely to falter. It also means that machine learning models alone for prediction that are dependent on so called BigData sampled over several years can be prone to errors. A suit of deep neural models were tested in multiple trials and none worked satisfactory. The following methods were tried; Neural Basis Expansion Analysis (NBEATS), Temporal Convolutional Networks (TCN), the recurrent neural network LSTM, Gated

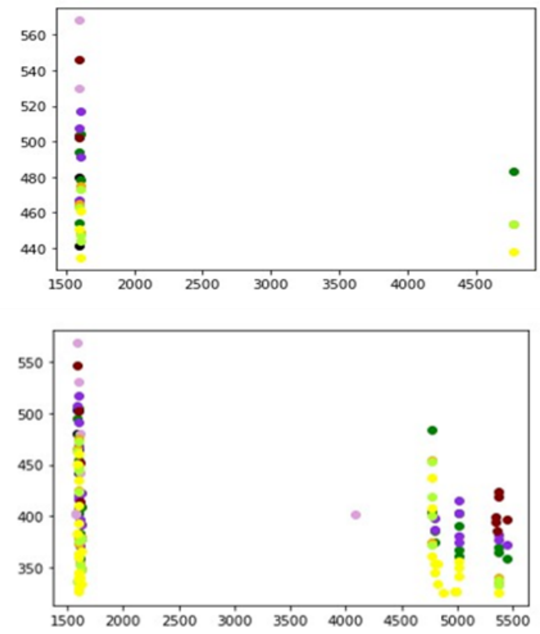

Figure 3 Price peaks in €/MWh for different European price zones plotted against a time axis in hours in 2022. The upper charts shows when peaks higher than 90% of the maximum in a zone occurs. The lower shows the peaks higher than 70% of the zone maxim

Recurrent Unit (GRU) and Transformer. Table 2 sums up the experience with these deep neural networks for predicting 2022 prices.

Non-stationarity and lack of ergodic samples for prediction that could be well be established in former years do not hold for the price situation in 2022. But by turning attention to a time-space model the project has established new insight about NO2 price developments. By looking for co-variant behaviour between different price zones it became apparent that prices in certain price zones correlated more than others. Figure 2 shows one of the results of PCA (Principal Component Analysis). Principal component 1-3 represented more than 90 % of the variance and the visual impression of the figure gives a fair perception of the co-variant nature. All price zones are more or less in sync. The system price is one reason for this. But it is also obvious that there are certain zones that are more co-variant than others. It was found that the zones south of NO2 (Kr.sand) had high impact on the prices in that zone. This can be explained in terms of the transmission capacity between southern Norway and the continent. But at the same time, it can be observed from the PC2-PC3 diagram that price developments in NO2 have certain causes that suggest a degree of independence from other zones, while NO1 (Oslo) and NO5(Bergen) and the continental ones i.e., DE, DK1, DK2, AT, BE, NL both represent more co-variant chunks that differs somewhat from NO2. By analysing peak prices, it was discovered that price peaks in NO2 happened only when similar price peaks occurred in price zones on the continent, not the other way. The top chart in Figure 3 shows when the

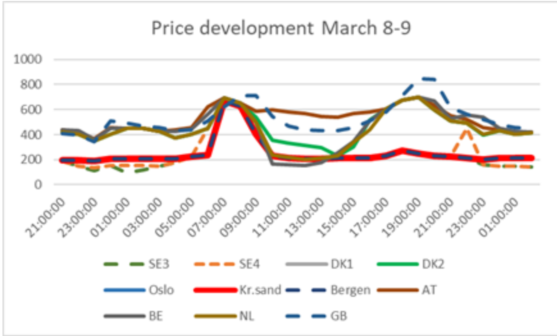

Figure 4 Peak price developments in the period March 7-9 in 2022 show how a peak in NO2 (and NO1 and N05) coincides with peaks in other zones. But the reverse never happened.

Figure 5 Shows a regression in the spatial domain where prices in 'SE3','DK1','Oslo','Bergen','NL','BE','AT','GB' for 24 hours are used to determine prices in NO2.

highest peaks in the neighbouring zones of NO2 happened. During the hours 1591 to 1602 in year 2022 several European price zones experienced a price jolt. In NO2 this occurred simultaneously with the neighbouring zones but lasted only for two hours. The top chart of Figure 3 shows prices higher than 90% of the absolute maximum price in a zone in the period between January 1 and August 1 in 2022. The lower chart shows the same for prices higher than 70% of the maximum peak price in a zone for the same period. As can be observed there is a very high degree of concurrency between peak prices in the European price zones. In fact, very high peaks in NO2 only struck when the same happened in the neighbouring zones. Figure 4 illustrates a typical incident when peaks coincide. It appears like a ripple effect, spreading from the continent and northwards.

Based on these findings it appeared intuitive to apply a time-space approach for predicting prices in NO2. A decision tree based XGBoost regressor with depth 6 and a very low learning rate was used. With this regression, prices for neighbouring zones were applied as input in order to estimate the price in NO2 for the same hour (see figure 5). Figure 6 shows a very good fit, suggesting that the space domain offers a good, additional basis for supporting time-oriented regressions. However, when combined with the time domain (Figure 7 and Figure 8) we

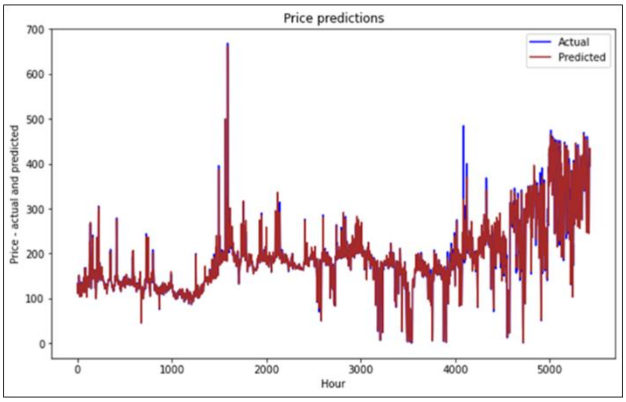

Figure 6 Predicting peak prices at hour, t, in NO2 based on prices in 'SE3','DK1','Oslo','Bergen','NL','BE','AT','GB' for the same hour.

Figure 7 Scatter diagram showing the result of a combined time-space prediction model where input is a record set from 'SE3', 'DK1', 'Oslo', 'Bergen', 'NL', 'BE', 'AT', 'GB' at time t-24 (hours) and output is prediction in NO2 at time t (hour).

can observe that the fit is quite poor. The peaks and valleys are not very well predicted.

CONCLUSIVE DISCUSSION

We can conclude that the co-variant aspects between price zones offer a good way to determine prices in NO2 based on information on the market state of neighbouring price zones. However, the challenge lies with predictions in the time domain. Adding information on historic prices in neighbouring price zones gave modest results when predicting future prices in NO2. Like for [2],[3] and [4] the problem lies with the radical upwards and downwards price shifts. The analysis of the cross entropy and statistical properties inherent in the time series studied provides at least a partial explanation for this. The time-based affinity between NO2 prices from day to day is poor and resemble a stochastic process close to white noise. However, the work done has pointed to different idiosyncrasies between price zones, which also seem to

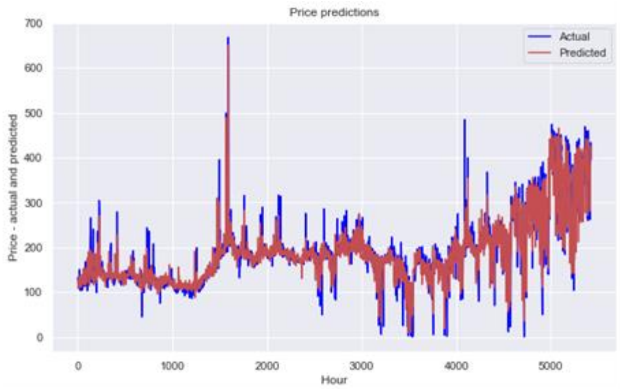

Figure 8 Using t-24 input for the same price zones as in Figure 6 to predict prices in NO2 at time t (hour).

appear in the latent space between principal component 2 and 3. This will be exploited in our continued work. However, this also suggests diving deeper into the underlying mechanisms that determine demand and supply in the different price zones.

ACKNOWLEDGEMENTS

The work in the INTHydro project is supported by the Norwegian Research Council.

REFERENCES

[1] J. H. Zhao, Z. Y. Dong, X. Li and K. P. Wong, "A Framework for Electricity Price Spike Analysis with Advanced Data Mining Methods," in IEEE Transactions on Power Systems, vol. 22, (1), 2007, doi: 10.1109/TPWRS.2006.889139.

[2] Xin Lu, Zhao Yang Dong, Xue Li, "Electricity market price spike forecast with data mining techniques," Electric Power Systems Research, Vol. 73,2005, doi.org/10.1016/j.epsr.2004.06.002.

{3] Shrivastava, N.A., Panigrahi, B.K. & Lim, MH. Electricity price classification using extreme learning machines. Neural Comput & Applic 27, (2016). https://doi.org/10.1007/s00521-013-1537-1

[4] Ødegaard Ottesen,S. Tomasgard, A. Fleten, S-E, Energy, 2018, vol. 149, issue C, DOI: 10.1016/j.energy.2018.01.187

[5] Shrivastava, N.A.,Singh, D., Ajay, S.P., Maini, T. A Novel Feature Selection and Short-Term Price Forecasting Based on a Decision Tree (J48) Model, Energies 2019, 12(19), doi.org/10.3390/en12193665