Electricity price prediction using time-series models

Jaka Konda
jaka.konda@student.tugraz.at
Graz University of Technology
Graz, Austria

ABSTRACT

Electricity market is becoming more and more open, available for different market players and wider audience to participate in the trading. Due to the nature of electricity, hard to store and limitation of the infrastructure available for transfer between the regions, the market behaves very volatile. In this paper we train and compare ARIMA and LTSM neural network models for forecasting Austrian electricity prices and compare it against a baseline of two hour lagged value. We achieved a substantial accuracy gain against a baseline, although room for large improvement remains.

CCS CONCEPTS

• **Mathematics of computing** → **Time series analysis**; • **Applied computing** → **Forecasting**.

KEYWORDS

datasets, time-series prediction, electricity price prediction

1 INTRODUCTION

In the last decades, worldwide electricity went over extensive and numerous changes. It started as a closed system where all activities (generation, transmission and power distribution) were regulated by the government and slowly transitioned into an open market. The main driving force was the fact that competition could result in more efficient utilisation of resources, higher reliability and consequently lower prices for end consumers [6].

With deregulation, the people now have a choice to pick their own electricity supplier, transforming the market into customer oriented, driven by the supply and demand relationship. In the new market several different actors emerged, generators, investors, trades and load serving bodies, each trying to maximise their profit margin. Their interests are most often in conflict with each other. If we take, for example, distributors and suppliers, distributors try to maximise their profit by buying low from suppliers and selling high, while suppliers try to sell to distributors as high as possible introducing a conflict of interests.

Trading is therefore done in a range of market clearing price (MCP) also known as price equilibrium, illustrated in figure 1. Buyers must submit their offers for at least 24 hours ahead of time without the knowledge of competitor prices and same is done by suppliers or more, up to a month in case of trading with futures. MCP is afterwards determined by intersecting the aggregated curves of supply and demand. Any deviations might result in an offer not being accepted by the other party. The addressed problem depends on many factors that are constantly evolving through time. Some of those have long term evolving impact such as the transition from fossil fuels to renewable energy source, slow adoption of electric vehicles and the transition to heating pumps in consumer homes. On the other hand, there are constantly changing factors such as weather impact, affecting wind turbines, rainfall having an influence to hydro generation and other connected markets such as prices of fossil fuels.

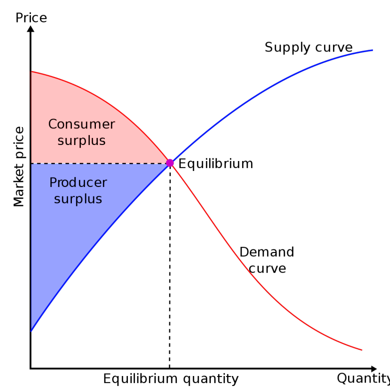

Figure 1: Price equilibrium, dictated by supply and demand.

2 DATA

The data is taken from the ENTSO-E (European Network of Transmission System Operators for Electricity) [2] FTP server. It was established in December 2008 with a goal to unify until then separate organisation and start combined operation with 2009. It currently represents 43 electricity transmission system operators (TSOs) from 36 countries across Europe. Its goal is to support implementation Europe's energy which is responsible for:

- climate policy - integration of renewable energy sources such as wind and solar power,
- single energy market - provide better collaboration between member states to improve affordability, sustainability and security of energy supply,

- single system - a single focal point for all technical, market and policy issue relating to European energy network for all participating players.

In 2013, a new regulation was accepted, which mandates all member states to submit fundamental information related to electricity generation and type, load, transmission and balancing starting with January 2015. Data is therefore publicly available since then for participating countries. Prior to that date, available data varies by country and the amount of attributes available.

Trading data is due to its public nature available in aggregated forms by the hour or quarter of an hour. It varies across the countries. High resolution trading data, is available for a fee on European Energy Exchange (EEX) [1].

Provided data is distributed in CSV files, one for each month since the beginning of January 2015. Inside the file, data is unsorted and split by different bidding regions. We used data for Austrian region, where trading data is reported every hour without missing values.

Exploration

In order to get visual overview of the data, we first aggregated the data by year and week number, extracted from the timestamp. Result is shown in Figure 2.

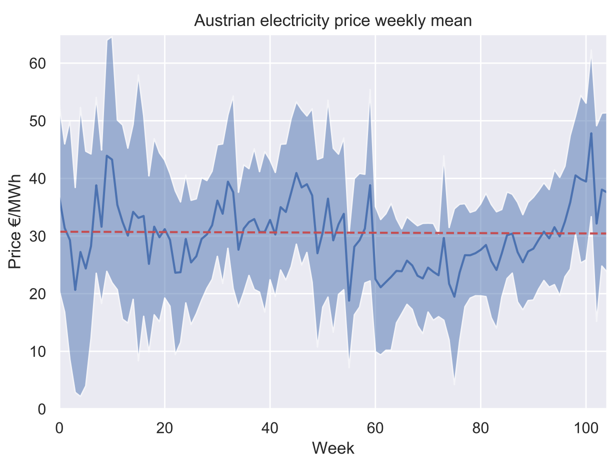

Figure 2: Austrian electricity prices in the second week of May 2016.

As there is no clear pattern observed from the broad overview, we visualised randomly selected week without any aggregation, shown in Figure 3. We can observe two spikes in the price of electricity in the working days, one in the morning and the other in the afternoon. Such price spikes normally indicate larger consumption, which is correlated with people going or returning from work. However this pattern breaks on weekends and holidays, where the price is substantially smaller, sometimes even reaching negative values.

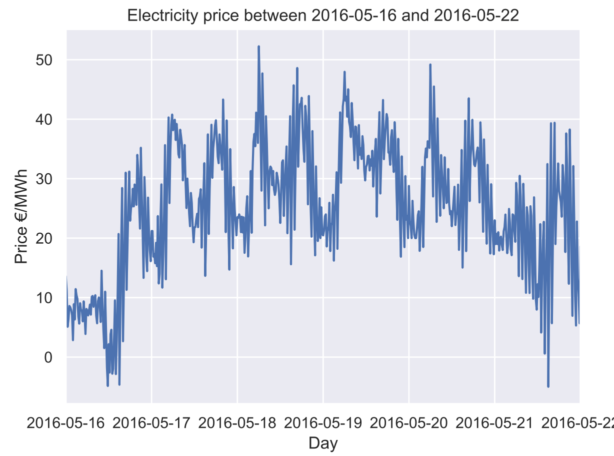

Figure 3: Weekly mean and std. dev. for electricity between January 2015 and December 2016. Red line represents a trend line.

Stationarity Test

As we can observe, the trend is very slowly downwards, meaning that the series is most probably not stationary, but look can often be deceiving. To eliminate bias and human error we performed augmented Dickey-Fuller (ADF) [4] test. The test's null hypothesis checks for presence of a unit root, which indicates a random walk with drift. If the unit root is present, the time-series is not stationary.

The test showed we can reject the null hypothesis since the test statistic is much smaller than the 1% critical value, meaning the time-series is stationary. The calculated ADF values are in shown table 1.

Table 1: Values obtained with augmented Dickey-Fuller test.

Parameter	Values
test statistic	-12.237
p-values	5.18×10^{-23}
critical value (10%)	-2.567
critical value (5%)	-2.862
critical value (1%)	-3.430

3 METHODS

In this section we describe used methods for forecasting comparison.

Baseline

For the baseline we used a simple lagged value approach. We tested all lagged values between 2 and 168 (1 week). Forecasting lag 0 and lag 1 values are not available due to one hour publishing delay by the ENTSO-E.

Auto Regressive Integrated Moving Average

Auto Regressive Integrated Moving Average (ARIMA) [5] is a statistical analysis model that uses time series data for better understating current or forecasting future trends.

It is combined out of an Auto Regressive (AR) [3] model that uses lagged values as predictor variables. It has a single p parameter named order that denotes how many lagged values are taken into an account.

With integrated, consecutive data points are differenced in order to make the time-series stationary by removing trends and seasonal structures that might cause a drift in the model parameters and decrease forecasting accuracy. It is controlled by parameter d, the degree of differencing.

The last part is Moving-Average model (MA), not to be mistaken by a rolling mean. Similar as AR uses lagged values and parameter q specifies the number of those values to be averaged. Contrary to the AR models where they use the values for prediction, MA uses errors from the previous forecasts to predict the current forecast.

Parameter search. Parameters were selected using two options. First manual approach with the help of autocorrelation plot (ACF) and partial autocorrelation (PACF) from the series, visible in Figure 4, and the second grid search to verify the results.

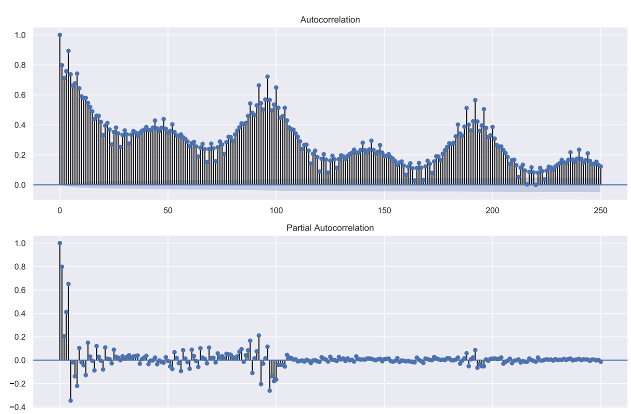

Figure 4: Auto correlation and partial auto correlation plots.

Due the the large amount of lagged values causing auto correlation and series drift, we differenced the series by setting the d parameter to 1. AR terms were set to 5 based on partial auto correlation plot and moving average terms experimentally set to 3. The resulting fit is shown in Figure 5.

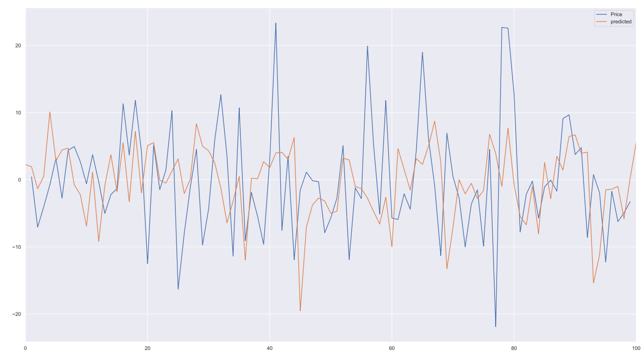

Figure 5: ARIMA(5, 1, 3) fit on the validation dataset with differenced values.

In order to test parameter neighbourhood, we performed an additional grid search. Parameters (2, 1, 2) performed almost as good as manually selected with faster execution, therefore we picked those as they represent a simpler model without sacrificing any accuracy. Those were also recommend by the autoarima function.

Long Short-Term Memory Neural Network

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning dependence in sequence prediction problems related to time-series such as stock prices or speech recognition.

The main difference between feed-forward neural networks is that they are able to find the trained patterns (long-term memory) in the past lagged values (short-term memory) adding a powerful option for models to train on.

Model. For this problem, we used a neural with two hidden layers, structure can be seen in Table 2. Having more than two output neurons helped increase the model accuracy. Different architectures with varied parameters such as number of input or output neurons, number of hidden layers and neurons per layer were also tested. Best performing model on the validation set was used for final test. The second output neuron was always used for evaluation since this is the relevant for predicting the next price.

Training. The model optimization method used during training is ADAM [7]. It was stopped once the validation loss increased twice consequently, indicating that the model stopped learning and might start to over-fit on the training data. Best model on the validation set was stored.

Table 2: Model of the used LTSM network.

Layer	Neurons
Input	12
1^{st}	70
2^{nd}	30
Output	10

4 EVALUATION

The European electricity market is split into different bidding zones. We picked Austrian region for price prediction using time-series models. The train, validation and final test sets were split by dates so that years 2015 and 2016 were used for training, 2017 for parameter fitting or validating dataset and final test year 2018.

For calculating the goodness of a fit, we used Root-mean-squared error (RMSE) and Symmetrical mean absolute percentage error (sMAPE) [9]. RMSE(1). First is the standard deviation of the residuals (forecasting errors) with the same unit as the forecasting data. Is the measure most commonly used as a loss function in machine learning. Value zero represents perfect match between actual and forecasted values and is always non-negative.

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n} (y_i - \hat{y}_i)^2}{N}} \tag{1}$$

sMAPE (2) is a measure of prediction accuracy used for forecasting methods in statistics. It is commonly used due to simple interpretation, as the resulting value represents percentage error.

$$sMAPE = \frac{100}{N} \sum_{i=1}^{N} \frac{|y_i| - \hat{y}_i}{(|y_i| + |\hat{y}_i|)/2} \tag{2}$$

5 RESULTS

In the table 3 we can see the results from the selected methods. Both LTSM and ARIMA showed an improvement from the baseline, however neither was able to capture the full complexity so the produced results are not suitable at all for practical use.

Table 3: Final results of tested models.

Method	RMSE	sMAPE (%)
Baseline (-2)	14.43	29.21
ARIMA (2, 1, 2)	13.13	26.13
LTSM	11.16	22.49

6 CONCLUSION

In this paper we explored two methods, ARIMA and LTSM neural network on a time-series forecasting problem for predicting Austrian electricity prices. We compared both methods using a naive approach of the delayed value by 2 hours. The chosen problem is perhaps not as interesting as some other problems which include long-term predictions, however our results showed that the problem is difficult due to the extremely volatile nature of the market generating the time-series.

LTSM offered a substantial improvement over a baseline and performed the best of all tested model, sMAPE value should be at least additional 8% lower to be perhaps considered feasible.

ARIMA model showed some improvement over baseline, however it was only minor. After visualizing the forecast, it can be quickly noticed that it captures the ups and down, however those are not accurate at all and the model later with time averages into a flat line. These might indicate that either parameters were chosen incorrectly, but grid search contradicts these. The second, more likely option is that the model itself is not appropriate for the selected problem which we suspect.

Future Work

SARIMA model might be more appropriate, since it is more flexible by capturing also the seasonal data which we expect to perform a better fit to the actual data, lowering the error.

With LTSM neural network we explored only one of the most basic architectures. Current newer method include combining LTSM neural networks with convolutional layers and deeper architectures, to capture more complex patterns although at the expense of a required much larger training set.

Problem of predicting future prices will always be trending due financial background and most likely never solved due to unpredictability and the amount of players in the market. There is still a lot of improvement as the study from Lago et al. [8] shows, where WARIMA model performed 4% better based on the sMAPE metric and Deep neural network achieved sMAPE of 12.34%.

7 PERSONAL NOTE

I learned a great deal about the time-series. Both methods were new to me and it was fun and interesting exploring them. ARIMA model was the one causing me the most issues and my first encounter with it surprised me with extremely slow execution compared to other learning models. The model itself was chosen without any prior knowledge and I found during the work that it was most likely not suitable at all for the task or the data should first be simplified which would defeat the purpose. While the result are not particularity good, I am still satisfied with the progress made especially due too many difficulties tackling the topic for the first time.

REFERENCES

[1] 2002. European Energy Exchange (EEX). Retrieved 2019-05-27 from www.eex.com

[2] 2008. European Network of Transmission System Operators for Electricity (ENTSO-E). Retrieved 2019-05-27 from www.entsoe.eu

[3] Hirotugu Akaike. 1969. Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics* 21, 1 (1969), 243–247.

[4] Yin-Wong Cheung and Kon S Lai. 1995. Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business & Economic Statistics* 13, 3 (1995), 277–280.

[5] Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. 2003. ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems* 18, 3 (2003), 1014–1020.

[6] V. P. Gountis and A. G. Bakirtzis. 2004. Bidding strategies for electricity producers in a competitive electricity marketplace. *IEEE Transactions on Power Systems* 19, 1 (Feb 2004), 356–365.

[7] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[8] Jesus Lago, Fjo De Ridder, and Bart De Schutter. 2018. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy* 221 (2018), 386–405.

[9] Spyros Makridakis. 1993. Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting* 9, 4 (1993), 527–529.