A Model for Forecasting short-term Electricity Prices for Electric Utilities

Ingela Tietze[1], Charlotte Newiadomsky[2] and Georg Vossen[3]

Abstract: Structural changes in the electricity market led to a decrease in the accuracy of electricity price forecasts. An extended model taking into account not only historical spot price data but also exogenous parameters achieves significantly higher forecast accuracy. Our analysis shows that considering only one additional parameter (predicted residual load) achieves similar results as considering three additional parameters. The developed model is coded in MATLAB with a modular structure. The modular structure allows the model to be highly comprehensive. It can be integrated easily into the existing software architecture used by electric utilities, the model is easy to use, and the required input data is typically available to these companies.

Keywords: Electricity Price Modelling, Price Forecasting, Hourly Price Forward Curve, HPFC, Prototype

1   Introduction

A sound spot price forecast is key to the competitiveness of energy suppliers. A common standard in spot price forecasting is the so called Hourly Price Forward Curve (HPFC), which depicts the expected hourly prices of future energy purchases or sales. Several important daily decisions are based on the forecasts of the hourly prices for the next days, for example purchases at the wholesale market and the dispatch of generation capacities. In the longer term HPFCs are used for the pricing of customers.

Due to the importance of the HPFC bigger electric utilities usually run their own HPFC models. These models are often very sophisticated. They are developed and maintained in modelling departments of electric utilities and are adapted to the specific requirements of each electric utility. Smaller electric utilities without modelling departments either use commercial software tools to calculate HPFCs or buy the HPFC itself (which means they normally receive the calculated price forecast each morning before trading starts). In these cases the electric utilities have no insight into the mechanisms of price forecasting and thus, critical reflection on the forecasts is barely possible.

---

[1] Pforzheim University, Institute for Industrial Ecology, Tiefenbronner Str. 65, D-75175 Pforzheim, ingela.tietze@hs-pforzheim.de
[2] Hochschule Niederrhein, SWK-Energiezentrum E , Reinarzstraße 49, D-47805 Krefeld, charlotte.newiadomsky@hs-niederrhein.de
[3] Hochschule Niederrhein, Faculty of Mechanical and Process Engineering & Institute of Modelling and High-Performance Computing, Reinarzstraße 49, D-47805 Krefeld, georg.vossen@hs-niederrhein.de

The missing information on the calculation of the forecasts has become more important over the last years, as the performance of the commercial HPFCs in many cases decreased. In order to fill this gap, we aimed for a model for the calculation of an HPFC that is not only comprehensible, but also meets the following criteria:

- runs with data typically available at electric utilities,

- can be easily integrated into the existing software architecture of electric utilities,

- needs little effort from the user, and

- copes with negative prices[4] and the effects from increasing renewable energy generation.

2   Background

The number of market participants as well as the trading volumes at the spot markets (short-term markets, contract fulfilment up to three days ahead) and forward markets (long-term markets, contract fulfilment up to several years ahead) increased since the liberalisation of the electricity markets [St16]. To cope with the new requirements regarding electricity trading, existing models describing price dynamics, for instance in the oil market, have been adapted. For the application to the electricity markets these models had to be extended, e.g. by taking into account different seasonalities and jumps in prices.

In general, the main input data for the calculation of the HPFC are the historical market prices. Relevant parameters influencing the spot electricity prices are for example weather, season, daytime, hourly demand, and the feed-in from renewable energies [Se10]. Until 2013, nevertheless, HPFCs available on the market considered historical spot prices only [SSH13].

Two major reasons are given for the increasing difference between the calculated HPFC and real spot prices:

- Introduction of negative prices at the European Energy Exchange in September 2008 (see e.g. [An10]), and

- Increased feed-in from renewable energies, especially from photovoltaics (see e.g. [Ba13] and [Bö15]).

The introduction of negative prices constitutes a relevant change for HPFC models. Many of the models applied so far were only able to give positive results. The integration of negative results would require different modelling approaches.

---

[4] Until 2008 the major electricity exchange for Germany only allowed prices greater equal zero. In September 2008, prices less than zero were introduced at the exchange.

The increased feed-in from renewable energies, especially from photovoltaics, leads to different daily price patterns. Due to high amounts of renewable electricity at noon and during the afternoon, German electricity prices at the spot market decrease in this time (see e.g. [Ha14] and [DDE13]). This effect only applies to days with high renewable feed-in, so that two different daily price patterns can be observed: one with increasing or more or less stable prices in the early afternoon and the other with decreasing prices in the early afternoon. Most HPFC modelling approaches use averaged daily patterns of the last couple of days for the forecasting of the future daily patterns. Under the new circumstances, this approach needs a correction to account for the two different patterns.

3   Methodological Approach

Usually, HPFCs are saved in central databases so that the different departments requiring the information (sales, trading, power plant dispatch) can get access easily. Typically, these databases are provided in formats suitable for integration into MATLAB and/or Excel. Thereby, exporting data sets into Excel and/or MATLAB as well as importing data sets as basis for the electricity price forecast is possible.

Weron gives an extensive review on modelling approaches for electricity price modelling and differentiates between the following approaches: multi-agent models, fundamental models, reduced-form-models, statistical models and computational intelligence models [We14]. We decided to develop a statistical model and excluded the other modelling approaches due to the following reasons:

- Multi-agent models give qualitative results and not the required quantitative results. The integration into the existing software architecture would not be feasible, as neither the strategies of competitors are available as data sets to electric utility companies nor could the existing software process the qualitative results.

- Fundamental models are based on huge data sets (see e.g. [SW07]), data collection occurs over a longer time period. The forecasting period mostly applies to mid- or long-term.

- Reduced-form models are used to give a forecast about the daily characteristics (e.g. hour with lowest or highest price) and not to provide the hourly prices itself.

- Computational intelligence models are used for short-term electricity price forecasting but are not applied in our case. Even though the statistic models are complex and sophisticated, they are more comprehensible to practitioners in electric utilities than the approaches of computational intelligence as statistical models are used for various applications in electric utilities. In addition, in practical applications the performances of the two modelling approaches are comparable [We14].

Statistical models use historical prices and/or historical or current values of exogenous parameters such as demand or weather data. The most popular statistical approaches are

additive models, in which the predicted price is the sum of different components [We14]. Thus, we developed our model as additive statistical model.

4   Model Description

In the following, we describe the model framework before we present our additive modelling approach consisting of two major parts: First, a forecast using historical spot price data and secondly, a forecast using exogenous parameters. In addition, we provide information about parameter fitting and the model architecture as well as the methodology to determine the forecast accuracy of our model.

4.1   Model Framework

The model is written in MATLAB since, on the one hand, this allows a convenient realisation of our model structure consisting of several sub functions and, on the other hand, it contains several implemented methods such as regression and convolution which are useful for our model. The average duration for calculation amounts to about 5 minutes. The model imports the necessary data sets from the electric utility's database and writes the results back into the database. The decision makers in different departments can read the HPFC directly from the database and do not need an introduction to MATLAB. Electric utilities using databases without MATLAB integration can apply the model using EXCEL as intermediate programme. The calculation time increases in this case.

The main forecasting period is the next day, but the model provides data for the next 30 days (with decreasing forecast accuracy over time).

4.2   Model Component 1: Forecasting using historical Spot Price Data

Mean reversion approaches are used extensively in energy commodity forecasting. The models using mean reversion approaches achieve sound results and thus we decided for a mean reversion approach even though the approach is being discussed controversially.

The forecasting using historical spot price data sums up the following components:

- Ornstein-Uhlenbeck process following Schwartz' approach [Sc97] to describe the mean reversion of electricity prices,

- Arithmetic Brownian motion to describe the deviation from the long-term average following Schwartz and Smith [SS00],

- Weekly seasonality to describe the typical hourly demand pattern during one week using the moving average.

Initially, the historical spot price data are corrected for discontinuities in terms of jumps and spikes to calculate the forecast components without unexpected extreme values [KUW13]:

- To identify discontinuities the long-term average is subtracted from each historical price. In case two subsequent results are higher than the threshold level of three standard deviations, the historical prices will be corrected.

- Identified jumps are corrected to the value of the long-term average, identified spikes are corrected to the long-term average plus or minus three times the standard deviation (depending on the spike's direction).

In literature, the Ornstein-Uhlenbeck process is often presented as "one-factor-model" [Sc97] to describe the non-systematic part of short-term electricity price motion. Schwartz' approach from 1997 [Sc97] is often used as a basis to describe the development of commodity prices. The logarithmized commodity price $\ln P_t^r$ is modelled as

$$\ln P_t^r = S_t \tag{1}$$

with

$$dS_t = \kappa_S(\nu_s - S_t)dt + \sigma_s dW_t^s \tag{2}$$

κ_S describes the mean reversion rate, ν_s the long-term average, σ_s the volatility, and W_t^s a Wiener process. κ_S, ν_s, and σ_s are nonnegative constants [Ge05]. To allow also for negative prices, we consider P_t^r instead of $\ln P_t^r$. The main characteristic of the Ornstein-Uhlenbeck process is the mean reversion which is determined by the mean reversion rate κ_S. Higher κ_S leads to a quicker reversion towards the long-term average.

An arithmetic Brownian motion is used to model the long-term uncertainty of electricity price motion. The arithmetic Brownian motion describes the permanent deviation from the long-term average of the short-term motion described above. The "one-factor-model" is transferred into a "two-factor-model" consisting of the Ornstein-Uhlenbeck process and the Brownian motion [SS00]:

$$lnP_t^r = S_t + L_t \tag{3}$$

with dS$_t$ as described above and

$$dL_t = \mu dt + \sigma_L dW_t^L \tag{4}$$

μ describes the Brownian motion with drift, σ_L the volatility, and W_t^L a second Wiener process.

In order to describe the typical seasonality of electricity prices, furthermore a deterministic component is integrated into the model. In order to calculate the weekly seasonality, data from the last 31 days are taken as a basis. These data are processed with the weighted moving average method, where an average value is calculated for each hour of each day. The weighted average is defined as the convolution of the datum points with a fixed weighting function. This means that it has multiplying factors to give different weights to data at different positions in the sample window (31 days). After an analysis of the seasonality data, we found out that public holidays have a similar deviation as Sundays. Subsequently, we replaced the seasonality of public holidays by the seasonality from the Sunday previous to this public holiday. The calculated deterministic component is finally added to the Ornstein-Uhlenbeck process and the Arithmetic Brownian motion resulting in the spot price forecast.

4.3   Model Component 2: Forecasting using exogenous Parameters

Due to the importance for trading, various data is available at electric utilities. As the temporal resolution is hourly or less, huge numbers of data items need to be processed at electric utilities. In order to keep the processing of our model easy, the model comprehensible, and the duration for calculation low, a major development target has been the improvement of the forecast accuracy with as few exogenous parameters as possible.

For the development of model component 2, we first applied univariate and multivariate analyses in historical data to find correlations between the following parameters:

- spot prices and electricity input of wind power plants (whole year, seasons, specific blocks of hours),

- spot prices and electricity input of photovoltaic power plants (whole year, seasons, specific blocks of hours),

- spot prices and combined electricity inputs of wind power plants and photovoltaic power plants,

- spot prices, load and combined electricity inputs of wind and photovoltaic power plants,

- spot prices and standardised load profile (whole year, hours and week days),

- spot prices and predicted residual load (predicted demand minus predicted feed-in from wind and photovoltaics),

- spot prices and predicted temperature, and

- spot prices and standardised load profiles.

In the investigated data, there was no correlation between the exogenous parameters and the spot price which was valid for arbitrary times simultaneously. However, when

considering each hour and each weekday separately, a strong correlation in the data could be found. The parameters predicted residual load, predicted temperature and standardised load profiles have the biggest correlation to the spot price and hence, we decided to consider these parameters in a first implementation. Due to a strong correlation a second implementation only takes the parameter predicted residual load into account. We found out that a model component based on a linear regression approach with regression coefficients for each hour separately provided very good results.

4.4   Parameter Setting

Parameter fitting by using an optimisation algorithm provides the following:

- the optimal weighting of the long-term price level resulting from the Ornstein-Uhlenbeck process and the short-term price level resulting from the arithmetic Brownian motion,

- the optimal weighting of the forecast using historical spot prices and the forecast using exogenous parameters, and

- the optimal fitting period (amount of historical data considered).

4.5   Model Structure

Fig. 1 depicts the simplified model structure. The MATLAB code contains at the first sub-level a module to calculate the HPFC, a module to export the results to the database and a module to plot the results.

The calculation module 1.1 is structured into four functions: Function 1.1.1 imports the data sets from the database and function 1.1.2 subsequently processes and corrects the data for discontinuities. Functions 1.1.3 and 1.1.4 calculate the price forecast itself, using the results from the second function. For this purpose, function 1.1.3 uses historical spot price data only, whereas function 1.1.4 uses exogenous parameters.

The three components of function 1.1.3 are each structured as sub-functions since each of them comprises a significant part of the complete implementation. At the end of function 1.1.3, the forecast is determined as the sum of all sub-function results.

Module 1.1 calculates the final HPFC considering the results from functions 1.1.3 and 1.1.4.

Module 1.2 exports the results to the electric utility's database. Module 1.3 optionally displays the results in a graph.
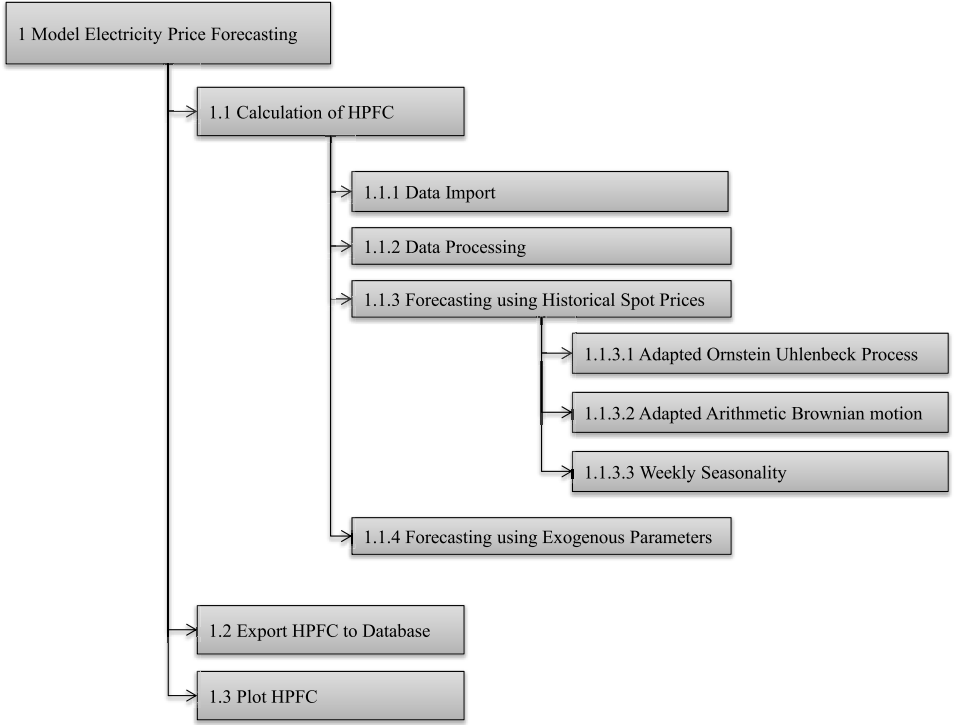
Fig. 1: Structure of the model to forecast short-term electricity prices

4.6   Selected computation results

Fig. 2 depicts exemplarily the resulting HPFC from our model and the spot price from the day-ahead auction at European Power Exchange EPEX SPOT SE. In general, the calculated HPFC on the basis of historical spot data and predicted exogenous data describes the development of the spot price.

To determine the forecast accuracy we average the Manhattan norm: The sum of the absolute differences between spot price $\hat{x}$ and predicted spot price x is divided by the number of comparisons N.

$$\|\hat{x} - x\| = \frac{\|\hat{x} - x\|_1}{N} = \sum_{n=1}^{N} \frac{|\widehat{x_n} - x_n|}{N} \tag{5}$$

Note that due to the averaging, we do not work with a monotone p-norm. However, we chose this norm for a better comparability when testing our algorithm with different time horizons.
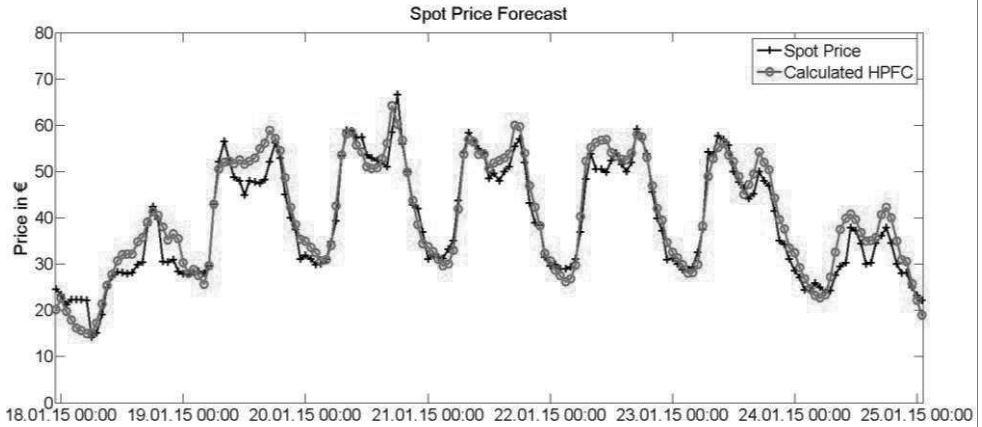
Fig. 2: Calculated HPFC and real spot prices for the period 18.01.15 to 25.01.15

The input data from the years 2012 to 2014 were used for the parameter fitting and were thus not only a forecasting period but also the training period of our model. 2015 is the first forecasting only period. We calculated a one-day-ahead forecast for each hour of the years 2012 to 2015 taking into account 26 weeks of historical data. Table 1 depicts the forecast accuracy i.e. the average deviation of model results from the real spot prices for the hours of the years 2012 to 2015.

The numbers in column 1 show the forecast accuracy, measured by (5), using only model component 1 (forecasting using historical spot price data[5]). Columns 2 and 3 provide the forecast accuracy using only model component 2 (forecasting using exogenous parameters) where column 2 considers only the predicted residual load and column 3 all exogenous parameters, i.e. predicted residual load, predicted temperature and standardised load profile. Columns 4 and 5 show the forecast accuracy for the additive model taking into account both model components, where column 4 considers only the residual load as exogenous parameter and column 5 considers all three parameters.

The consideration of exogenous parameters leads to a higher forecast accuracy for all years analysed. Already the sole integration of the predicted residual load into the model leads to the improvement of the forecast. The additional consideration of the predicted temperature and the standardised load profile only leads to similar results. Considering more parameters thus does increase the complexity but does not improve the forecast accuracy.

---

[5] This still represents the approach of many commercially available HPFCs today.

	1 historical spot prices	2 predicted residual load	3 predicted - residual load - temperature - standardised load profiles	4 combination of 1 and 2	5 combination of 1 and 3
2012	7,27	9,98	10,01	6,87	6,82
2013	7,66	12,78	14,91	7,32	7,38
2014	5,66	11,36	13,31	5,25	5,31
2015	10,10	5,23	5,36	4,60	4,62

Tab 1: Forecast accuracy of different model components and their combinations for a fitting period of 26 weeks

5   Conclusion and Outlook

The developed statistical model fulfils the requirements regarding comprehensibleness, data availability, integration into existing software architecture, low user effort and forecast accuracy (taking into account negative prices and predicted electricity generation from wind and photovoltaics). The acceptance of our statistical model coded in MATLAB at the electric utilities is assured by the following:

• well acknowledged calculation methods for the component on the basis of historical spot prices are used,

• a comprehensible method for the component on the basis of exogenous parameters is integrated, and

• the weighing of the major components on the first sub-level allows for users with little MATLAB modelling knowledge, to adapt the weighing parameters in case of structural changes in the market.

The developed model assists the competitiveness of electric utilities as it improves the forecast accuracy for a HPFC by considering exogenous factors. The forecast accuracy increased significantly by the consideration of predicted residual load, predicted temperature and standardised load profiles. This applies especially to years in which the forecasting using historical spot prices have lower forecast accuracy.

Big data is an important issue in spot price prognosis since the data which can be used to fit the model is very comprehensive. In addition to the proposed exogenous parameters (predicted residual load, predicted temperature and standardised load profile), we tested several other types of data. We found out that the three parameters used have the biggest correlation to the spot price and hence, decided to consider these parameters. However, the results for multivariate fitting with residual load and only one other parameter as well

as the univariate model on basis of the residual load solely achieved nearly the same forecast accuracy. A fit on the basis of the data neglecting the residual load provided in all cases distinctly worse accuracy. This indicates a strong influence of the residual load parameter on the spot price.

For the years under consideration our analysis shows no advantage in taking into considerations further parameters other than the predicted residual load. A quicker and more comprehensible model considering only the parameter predicted residual load achieves a similar forecast accuracy as a model considering further parameters.

A comparison of our results with a commercially available HPFC shows an increase in forecast accuracy by about 30 %. This increase cannot only be explained by the consideration of exogenous factors but also by the prediction of negative prices and the intensive parameter fitting undertaken.

Compared to a HPFC taking into account the historical spot price data only the biggest improvement in forecast accuracy could be achieved for 2012 and 2015. An explanation for the big improvements in 2012 could be, that in 2011 and 2012 especially the German electricity sector has been subject to major structural changes in the form of a massive increase in generation from photovoltaics [Bd16]. This led to different daily price patterns depending on the feed in from photovoltaics. Models not considering exogenous parameters can barely account for this. For 2015 the explanation could be that it is the first year out of the training period for the parameter setting. Additional analysis shall be undertaken to verify or falsify these hypotheses in order to further the understanding on improvements to short-term electricity price forecasting.

As the parameter fitting assists significantly to the performance of the model, the integration of the parameter fitting into the model would be an interesting further development. On the basis of the recent forecast accuracy, an additional optimisation module determines the best weighting parameters and the best fitting period to maintain and improve the forecast accuracy continuously.

Funding Information

References

[An10]   Andor, M.; Flinkerbusch, K.; Janssen, M.; Liebau, B.; Wobben, M.: Negative
            Strompreise und der Vorrang Erneuerbarer Energien. Zeitschrift für Energiewirtschaft
            34/2010, 91–99. 2010.

[Ba13]   Bader, A: Sind erneuerbare Energien die treibende Kraft am Intraday-Strommarkt in
            Deutschland?. Energiewirtschaftliche Tagesfragen 63/2013, 48–50, 2013.

[Bd16]   bdew: Erneuerbare Energien und das EEG: Zahlen, Fakten, Grafiken. 2016.

[Bö15]   Böschen, H.; Knaut, A.; Obermüller, F.; Tode, C.: Preisvorhersage im Intraday-Markt
            auf Basis von Prognoseänderungen der Einspeisung aus Wind und Sonne.
            Energiewirtschaftliche Tagesfragen 65/2015, 31–33,2015.

[DDE13]   De Jong, C.; van Dijken, H.; Enev, E.: Wie erneuerbare Energien die (Preis-) Zukunft
            gestalten. Energiewirtschaftliche Tagesfragen 2013/63, 18-20, 2013.

[Ge05]   Geman, H.: Commodities and commodity derivatives: Modelling and pricing for
            agriculturals, metals, and energy. John Wiley & Sons, West Sussex, 2005.

[HA14]   Hasenbeck, M.: Die Price-Forward-Kurve wird grün; Einflüsse von Wind- und
            Solarstrom auf das Pricing. e|m|w Zeitschrift für Energie, Markt, Wettbewerb 2014, 32–
            35, 2014.

[KUW13]   Krohmer, A.; Utz, S.; Wagner, A.: Modellkalibrierung: Ein oft unterschätzter Faktor für
            die Modellgüte. Energiewirtschaftliche Tagesfragen 2013/63, 45-48, 2013.

[Sc97]   Schwartz, E.: The Stochastic Behavior of Commodity Prices: Implications for Valuation
            and Hedging. The Journal of Finance 52/1997, 923–973, 1997.

[Se10]   Seifert, J.: Preismodellierung und Derivatebewertung im Strommarkt; Theorie und
            Empirie. KIT Scientific Publishing, Karlsruhe, 2010.

[SS00]   Schwartz, E.; Smith, J.: Short-Term Variations and Long-Term Dynamics in
            Commodity Prices. Management Science 46/2000, 893–911, 2000.

[SSH13]   Syben, O.; Spang, E.; Hatakka. T: Einfluss der HPFC auf die mittelfristige Bewertung
            von Erzeugungseinheiten; Was sind Kraftwerke morgen wert?. BWK - Das Energie-
            Fachmagazin 65/2013, 24–26,2013.

[St16]   Trading Volumes am EEX Spot- und Terminmarkt für Strom in den Jahren 2002 bis
            2014   (in   Terrawattstunden).   http://de.statista.com/statistik/daten/studie/12486/
            umfrage/entwicklung-der-eex-handelsvolumina, 23.03.2016

[SW07]   Swider, D.; Weber, C.: The Costs of Wind's Intermittency in Germany: Application of
            a Stochastic Electricity Market Model. European Transactions on Electrical Power
            17/2007, 151-172, 2007.

[We14]   Weron, R.: Electricity price forecasting: A review of the state-of-the-art with a look into
            the future. International Journal of Forecasting 30/2014, 1030–1081, 2014.