Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 14/12

Christian Hendricks, Matthias Ehrhardt and Michael Günther

Integrated Forecasting of day-ahead Prices in the German Electricity Market

May 2014

http://www.math.uni-wuppertal.de

Integrated Forecasting of day-ahead Prices in the German Electricity Market

Christian Hendricks, Matthias Ehrhardt and Michael Günther

Abstract Since the start of the liberalization of energy markets the energy sector has undergone major changes. Energy companies now provide electricity at variable prices and are faced to a competitive market environment. Their trading is subject to risks and uncertainty about future price developments. In this work we introduce a regularized regression approach to forecast *Phelix Peak* prices in the German electricity market. Additionally we investigate the influence of fundamental price drivers on the forecasting accuracy. Since the problem complexity grows exponentially with the dimension of the feature space, the regression problem suffers from the curse of dimensionality. To cope with this problem we apply the combination technique. It is based on a linear combination of coarse grids to the so called sparse grid solution, which enables us to reduce the complexity while keeping a high approximation accuracy.

1 Introduction

During the last two decades the energy sector has undergone major changes. Energy companies now provide electricity at variable prices and are faced to a competitive market environment. Their trading is subject to risks and uncertainty about future price developments. These risks are mainly associated with the volatile nature of input costs, like coal and gas prices, but also other factors influence energy markets. The global concern regarding the climate change has led to the introduction of an emission trading system in the European Union (EU). Today energy suppliers have to surrender

Christian Hendricks · Matthias Ehrhardt · Michael Günther
Lehrstuhl für Angewandte Mathematik und Numerische Analysis, Fachbereich C Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany. e-mail: {hendricks, ehrhardt, guenther}@math.uni-wuppertal.de

European Emission Allowances (EUA) to offset their emission of greenhouse gases. As a conventional power plant has to burn fuel and emit CO_2 to produce electricity, these allowances can be interpreted as additional input costs. In Germany the energy transition along with the installation of renewables has accelerated. In 2012 renewable energy sources covered 22% of the total energy production in Germany [1]. In [17], [19], [13] cost savings in conventional capacity due to wind generation are analyzed. There a price decline in the German market between three to eleven € / MWh is reported. But beside a price decline also an increase of the spot price's volatility can be seen, which can be attributed to the fluctuating wind in-feed [11].

The vastly changing market environment has attracted lots of researchers to develop forecasting models for electricity markets on different time frames. They range from ARMA to neural network approaches or models known from game theory. Regarding energy markets in continental Europe mainly the Spanish and German market have been focused. [3] tried to predict prices in Spain using an autoregressive neural network model. [14] additionally included the demand for electricity into their time series model. [5] investigated if a multivariate model of electricity, carbon, fuel prices and wind forecasts can improve the forecasting accuracy compared to univariate models. In the papers considering the German market (European Energy Exchange) the price series itself has been taken as an input variable. [15], [9] tested the performance of neural networks, while [4] used ARMA models to forecast electricity spot prices.

In this work we present a non-linear approach to forecast electricity spot prices in the German electricity market. Additionally the influence of fundamental price drivers on the spot price and the forecasting accuracy shall be investigated. The arising high dimensional approximation problems will be solved with the combination technique on sparse grids [10], [16], [18], [2], [6]. This technique is an efficient way to cope with the exponentially growing complexity of high dimensional problems - the so called *curse of dimensionality*. The performance of the model will be compared to benchmark ARMA/VARMA models.

2 The Data Set

The data set consists of electricity spot prices (Phelix Peak) and time series of coal (ARA coal future [1]), gas (GASPOOL Spot), European Emission Allowances and day-ahead forecasts of wind and solar supplies[2]. All time series range from 2011 to 2012. The coal time series is quoted in USD per t and has been converted to EUR per t.

[1] front month with nearest expiry

[2] provided by www.transparency.eex.com

It is well known that electricity prices show changing price profiles during weekdays, weekends, public holidays and working days. To avoid any perturbation caused by this changing market behavior, we remove all weekends and public holidays from our time series. The time series, including only "business-as-usual" days is checked for mean and variance stationarity. With the help of the Augmented Dickey-Fuller test (ADF) we check for a unit root in each of the time series. Table 1 shows the test results: the electricity, coal, gas and emission allowance time series exhibit a unit root at the common confidence level of 5%. Therefore we differentiate the time series to eliminate the stochastic trends. To achieve variance stationarity the logarithm of all price series is taken.

	Electricity	Coal	Gas	EUA	Wind supply	Solar supply
p-value	0.1008	0.0973	0.8590	0.1370	0.0010	0.0018

Table 1 Augmented Dickey-Fuller test results

3 The Modelling Framework

In order to forecast electricity prices we apply two approaches. The first model is a regularized regression approach, which is able to capture non-linear relationships between the input variables and the desired output. Like neuronal networks, these kind of models are very versatile in approximating complex relationships, but might suffer from over-fitting. Therefore we want to compare the quality of the out-of-sample forecasting results with univariate and multivariate ARMA models. They are the standard approaches in times series analysis and work as a benchmark.

3.1 Regularized Regression

In this section we formulate the forecasting problem as a regularized least square regression. These kind of models have already proven to be useful in data mining [8], foreign exchange [7] and wind time series forecasting [12]. Compared to autoregressive models, we want to investigate if the prediction accuracy can be increased significantly in the case of further knowledge about key drivers in the market: This might be price shifts in inputs costs, such as coal and gas, as well as production forecasts of wind generators and solar modules.

Let $\Omega \subseteq \mathbb{R}^d$ denote a d dimensional feature space, then it is the goal to find a function $f : \Omega \to \mathbb{R}$, which maps the model's input $x_i \in \mathbb{R}^d$ to the desired output $y_i \in \mathbb{R}$ for $i = 1,...,n$ observations. The unknown function f belongs to some function space V, which we will specify later on. The resulting regularization problem can be written as

$$\inf_{f\in V} \left(\frac{1}{n} \sum_{i=1}^{n} L(f(x_i), y_i) + \lambda \|Pf\|^2_{L^2(\Omega)} \right), \tag{1}$$

where L is a loss function, which ensures that f is close to the output y_i. In the sequel we will consider $L(a,b) = (a-b)^2$. The second term is a penalty term for non-smooth f. The parameter $\lambda > 0$ determines the balance between the accuracy of the fitted function and its smoothness. P is a regularization operator, for example one can use $P = \nabla$ or $P = \Delta$. In the sequel we use $P = \nabla$.

In order to estimate f, a function space is needed to be specified. We will restrict ourselves to a finite dimensional space $V_m \subseteq V$ and express f with the help of basis functions $\{\phi_i(x)\}_{i=1,...s,m}$ by

$$f(x) = \sum_{i=1}^{m} \alpha_i \phi_i(x). \tag{2}$$

Plugging (2) into (1) the approximation reduces to a minimization problem, which can be rewritten as a linear equation system

$$\inf_{f\in V} \left(\lambda C + B B^T \right) \alpha = By,$$

with matrices $C_{j,k} = n\langle \nabla\phi_j, \nabla\phi_k \rangle_{L^2(\Omega)}$, $j,k = 1,...,m$, $B_{j,i} = \phi_j(x_i)$, $j = 1,...,m$, $i = 1,...,n$ and the m dimensional vector α. The vector α contains the degrees of freedom and represents a unique solution if the minimization problem is well-posed.

The dimension of the feature space is determined by the number of used variables. If additional time series are introduced to describe y, the dimension of the feature space increases. Hence the system that has to be solved grows and the curse of dimensionality shows its effects quickly. On a uniform grid with mesh size $h_N = 2^{-N}$, and level $N \in \mathbb{N}$, in each coordinate direction this would lead to $\mathcal{O}(h_N^{-d})$ degrees of freedom and an exponentially increasing complexity. To cope with this problem we use the sparse grid combination technique to reduce the complexity. [10], [16] [18] have shown that the number of grid points can be lowered to an order of $\mathcal{O}(h_N^{-1} \log(h_N^{-1}))^{d-1})$. In the following we will briefly recall the fundamentals of this technique. For a detailed introduction to sparse grids we refer to [2], [6].

The combination technique is based on linearly combining a sequence of functions. Let $\Omega := [0,1]^d$ be the d dimensional unit cube and let $\mathbf{l} = (l_1,...,l_d) \in \mathbb{N}^d$, $\mathbf{i} = (i_1,...,i_d) \in \mathbb{N}^d$ denote multi-indices, then we can define a family of grids $\{\Omega_l\}_{l\in\mathbb{N}}$ on Ω with mesh sizes $\mathbf{h}_l = (h_{l_1},...,h_{l_d}) =$

$(2^{-l_1},...,2^{-l_d})$. Each grid consists of the points $\mathbf{x}_{l,i} = (x_{l_1,i_1},...,x_{l_d,i_d})$ with $x_{l_t,i_t} = i_t h_{k_t},\ i_t = 0,...,2^{l_t},\ t = 1,...,d$. The d dimension basis functions $\phi_{l,i}$ are given by the tensor product of piecewise linear one dimensional basis functions ϕ_{l_t,i_t} for $x \in \Omega$

$$\phi_{l,i}(x) = \prod_{t=1}^{d} \phi_{l_t,i_t}(x_t).$$

The one dimensional basis function ϕ_{l_t,i_t} is given by a hat function, centered on grid point $x_t = x_{l_t,i_t}$

$$\phi_{l_t,i_t}(x_t) = \begin{cases} 1 - \left|\frac{x_t - i_t h_{l_t}}{h_{l_t}}\right| & \text{if } x_t \in [(i_t-1)h_{l_t},(i_t+1)h_{l_t}] \cap [0,1] \\ 0 & \text{otherwise.} \end{cases}$$

Since each basis function ϕ_{l_t,i_t} has support $[(i_t-1)h_{l_t},(i_t+1)h_{l_t}] \cap [0,1]$, the d dimensional function $\phi_{l,i}$ is one at grid point $x_{l,i}$ and zero at all other grid points of grid Ω_l. With the help of these basis functions the function space $V_l = \text{span}\{\phi_{l,i}(x), i_t = 1,...,2^{l_t}, t = 1,...,d\}$ on grid Ω_l can be defined. The function f_l on Ω_l is represented by

$$f_l(x) = \sum_{i_1=1}^{2_1^l} ... \sum_{i_d=1}^{2_d^l} \alpha_{l,i}\phi_{l,i}(x).$$

If we combine linearly the solution f_l from different grids Ω_l according to the formula

$$f_N(x) = \sum_{q=0}^{d-1} (-1)^q \binom{d-1}{q} \sum_{|l|_1 = N-q} f_l(x),$$

we obtain the function f_N, which lives in the sparse grid space with $\mathcal{O}(h_N^{-1}(\log(h_N^{-1}))^{d-1})$ grid points, compared to $\mathcal{O}(h_N^{-d})$ grid points of the full grid solution. Provided that f fulfills certain smoothness conditions, it is shown in [2], that the approximation error is

$$\|f - f_N\| = \mathcal{O}(h_N^2 \log(h_N^{-1})^{d-1}).$$

We have considered only the case of the d dimensional unit cube here. Please note, that this is no restriction since we can linearly transform all data to the unit cube.

3.2 Multivariate ARMA

In order to evaluate the performance quality of the fitted function of the previous section, we want to apply ARMA/VARMA models as a benchmark. In the case of no additional input variables an univariate ARMA model will be used, while for additional external inputs multivariate ARMA (Vector-ARMA) models will be tested. They can be understood as the multivariate counterpart of univariate ARMA models. The general model is of the form

$$\phi(B)y_t = \theta(B)e_t + \omega(B)u_t,$$

where B is the back shift operator. y_t and e_t are p dimensional vectors of observed output variables and unobserved residuals, Gaussian white noise, respectively. The m dimensional vector u_t contains the input variables. The coefficient functions are polynomials, which take the lag operator B as a variable. The model order is chosen with the help of the Akaike Information Criterion (AIC). The appropriate model parameters are calibrated with the help of the *R* software package *DSE*[3].

4 Forecasting Methodology and Results

In this section we want to investigate the forecasting accuracy of our function for Phelix day-ahead prices. The accuracy is quantified with the help of the following measures, where $\hat{y}_i$ is the prediction and y_i the true electricity spot price for values $i = 1,...,n$:

1. Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

2. Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

3. Root Mean Squared Error (RMSE)

$$\text{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

[3] Dynamic System Estimation (DSE): available at www.cran.r-project.org

The first two criterions measure the deviation from the forecast to the true price. The latter one is the second moment of the error and incorporates the variance of the estimator. To build up our forecasting model, we use half of our data as a training set. The first half of 2012 works as a validation set to select the model parameters; the smoothing parameter λ and the level N (see Section 3.1). Based on the selected model, we compute day-ahead forecasts for the second half of the year 2012 to test the out-of-sample behavior of our fitted models. To make both approaches as comparable as possible, we will always use the same feature space, training, validation and test set to evaluate their performance.

4.1 Forecasting Results

The easiest feature space one can think of consists of the time series itself. Since the electricity price series exhibits a unit root (see Section 2), we take the differentiated time series into considerations. With the help of the validation set, we experimentally answer the question how many delayed values should be included. Experiments turn out that two values are appropriate. Thinking in terms of autoregressive models (AR) this would correspond to an order of two. The training procedure of the regularized regression approach is therefore proposed to find a function f, which matches

$$y_t - y_{t-1} = f_1(y_{t-1} - y_{t-2}, y_{t-2} - y_{t-3})$$

for all t in the training set. In Table 2 we compare the results of the regularized regression approach and an ARMA model. In the validation set the parameters $\lambda = 0.01$ and $L = 2$ showed the best performance and have therefore been chosen. The out-of-sample accuracy is slightly worse than in the validation set. However it can beat the ARMA model in all three accuracy measures.

	$\text{RegReg}_{\text{validation}}$	$\text{RegReg}_{\text{out-of-sample}}$	ARMA
MAE	4.8914	**4.9398**	5.0255
MAPE [%]	9.8921	**9.7868**	9.8946
RMSE	6.3562	**6.9317**	7.0979

Table 2 Forecast results

We now extend the feature space by coal and gas time series. Since coal and gas fired power plants have a share of more than one half in the German energy mix [1], we want to evaluate if the forecasting results can be improved. Hence we include the last two differentiated values. The six dimensional regression problem is of the form

$$y_t - y_{t-1} = f_2(y_{t-1} - y_{t-2}, y_{t-2} - y_{t-3}, c_{t-1} - c_{t-2}, c_{t-2} - c_{t-3}, g_{t-1} - g_{t-2}, g_{t-2} - g_{t-3}),$$

where c_t is the coal price and g_t the gas price at time t in the training set. Table 3 shows the forecasting results. Compared to the previous feature space, which only consisted of lagged electricity spot prices, the introduction of both fuel price series seem to have an adverse influence on the accuracy.

	$\text{RegReg}_{\text{validation}}$	$\text{RegReg}_{\text{out-of-sample}}$	ARMA
MAE	4.9617	**5.0884**	5.0950
MAPE [%]	9.9802	**10.0623**	10.1035
RMSE	6.4023	7.1392	**7.1341**

Table 3 Forecast results including fuel prices

Since the start of the *European Union Emission Trading Scheme* (EU ETS) energy companies have to surrender EUAs to offset their emission of greenhouse gases. In order to check, whether historical prices of EUAs can enhance the accuracy, we add them to the feature space and obtain a four dimensional problem

$$y_t - y_{t-1} = f_3(y_{t-1} - y_{t-2}, y_{t-2} - y_{t-3}, e_{t-1} - e_{t-2}, e_{t-2} - e_{t-3}).$$

In Table 4 we compare the forecasting quality of both models. The regression approach slightly outperforms its benchmark. We see that an addition of EUAs to the feature space does not improve the forecasting results.

	$\text{RegReg}_{\text{validation}}$	$\text{RegReg}_{\text{out-of-sample}}$	ARMA
MAE	4.9617	**5.0884**	5.1013
MAPE [%]	9.9802	**10.0436**	10.1038
RMSE	6.4023	**7.0247**	7.1263

Table 4 Forecast results including EUAs

Along with coal and gas fired power plants, renewable energy sources play an important role in the German electricity market. In 2012 about 22% of the total electricity was produced by sustainable generators [1]. Since the preferred use of green to conventional energy is guaranteed by the *Renewable Energy Act* (EEG), there is a deep impact of production capacities provided by wind and solar generators on the spot price for electricity. The influence of wind power has already been analyzed in [17], [19], [13], [11]. Here the variable w_t denotes the wind production forecast at time t, while s_t is the solar production forecast. These are published by the *transmission system operators*[4] (TSO) and we assume that these information is available at time

[4] 50Hertz, Amprion, APG, TenneT, TransnetBW

level $t-1$. The fitting problem reads

$$y_t - y_{t-1} = f_4(y_{t-1} - y_{t-2}, y_{t-2} - y_{t-3}, w_t - w_{t-1}, s_t - s_{t-1}).$$

Table 5 shows the great improvement to the previous feature spaces. The error in terms of the MAE, MAPE and RMSE can be lowered by 28.20%, 32.07% and 27.15% compared to the first model. This increase in accuracy underlines the strong price effects of solar and wind supplies on the spot price for electricity, the so called merit order effect [17].

	$\text{RegReg}_{\text{validation}}$	$\text{RegReg}_{\text{out-of-sample}}$	ARMA
MAE	3.9290	3.5466	**5.5231**
MAPE [%]	7.5325	**6.6547**	6.9041
RMSE	5.7464	**5.0532**	5.1876

Table 5 Forecast results including wind and solar production forecasts

	f_1	f_2	f_3	f_4
N	2	0	0	4
λ	0.01	0.001	0.001	0.001

Table 6 Selected model parameters

In Table 6 the selected model parameters are stated. The level N determines the amount of grid points in our spare grid (see Section 3.1). It is clear that a finer grid has better in-sample-properties, but might suffer from overfitting if it is tested for out-of-sample data. The forecasting results underline, that a small level is sufficient and leads to stable results in the validation- and out-of-sample set.

5 Conclusion

In this paper we investigated the potential of a non-linear regression approach in the prediction of day-ahead electricity prices in Germany. The out-of-sample tests show that this model performs better than its benchmark ARMA/VARMA model. The strength of our technique lies in its ability to capture a big variety of relationships up to a high order of dimension. Within this work we considered problems up to dimension 6, but also higher dimensions are conceivable, e.g. adding further time series or more lagged values. In four different tests we evaluated the benefit of important impact factors on the prediction accuracy. The inclusion of fuel prices and CO_2 allowance prices turned out to introduce more noise. Thus leading to worse forecasting results. If wind and solar production forecasts are added to the feature space, the accuracy is greatly improved. These results underline the strong price effects of renewable energy sources in the German electricity market.

References

1. BDEW German Association of Energy and Water Industries: Erneuerbare Energien und das EEG: Zahlen, Fakten, Grafiken (in German) (2012)
2. Bungartz, H.J., Griebel, M.: Sparse grids. Acta Numerica **13**, 147–269 (2004)
3. Catalâo, J., Mariano, S., Mendes, V., Ferreira, L.: Short-term electricity prices forecasting in a competitive market: A neural network approach. Electric Power Systems Research **77**, 1297–1304 (2007)
4. Cuaresma, J., Hlouskova, J., Kossmeier, S., Obersteiner, M.: Forecasting electricity spot-prices using linear univariate time-series models. Applied Energy **77(1)**, 87–106 (2004)
5. García-Martos, C., Rodríguez, J., Sánchez, M.J.: Modelling and forecasting fossil fuels, CO_2 and electricity prices and their volatilities. Applied Energy **101**, 363–375 (2013)
6. Garcke, J.: Regression with the optimised combination technique. Proceedings of the 23rd ICML pp. 321–328 (2006)
7. Garcke, J., Gerstner, T., Griebel, M.: Intraday Foreign Exchange Rate Forecasting Using Sparse Grids. In: Sparse Grids and Applications, *Lecture Notes in Computational Science and Engineering*, vol. 88, pp. 81–105. Springer (2013)
8. Garcke, J., Griebel, M., Thess, M.: Data mining with sparse grids. Computing **67**, 225–253 (2001)
9. Gareta, R., Romeo, L.M., Gil, A.: Forecasting of electricity prices with neural networks. Energy Conversion and Management **47(13-14)**, 1770–1778 (2006)
10. Griebel, M., Schneider, M., Zenger, C.: A combination technique for the solution of sparse grid problems. Iterative Methods in Linear Algebra pp. 263–281 (1992)
11. Ketterer, J.: The Impact of Wind Power Generation on the Electricity Price in Germany. Ifo Working Paper Series Ifo Working Paper No. 143, Ifo Institute for Economic Research at the University of Munich (2012)
12. Kramer, O., Gieseke, F.: Analysis of wind energy time series with kernel methods and neural networks. In: Natural Computation (ICNC), 2011 Seventh International Conference on, vol. 4, pp. 2381–2385 (2011). DOI 10.1109/ICNC.2011.6022597
13. Neubarth, J., Woll, O., Weber, C., Gerecht, M.: Beeinflussung der Spotmarktpreise durch Windstromerzeugung. Energiewirtschaftliche Tagesfragen **56**(7), 42–45 (2006)
14. Nogales, F., Contreras, J., Conejo, A.J., Espínola, R.: Forecasting Next-Day Electrcity Prices by Time Series Models. Power Systems **17(2)**, 342–348 (2002)
15. Pao, H.T.: A Neural Network Approach to m-Daily-Ahead Electricity Price Prediction. In: J. Wang, Z. Yi, J. Zurada, B.L. Lu, H. Yin (eds.) Advances in Neural Networks - ISNN 2006, *Lecture Notes in Computer Science*, vol. 3972, pp. 1284–1289. Springer Berlin Heidelberg (2006)
16. Pflaum, C., Zhou, A.: Error analysis of the combination technique. Numerical Mathematics **84(2)**, 327–350 (1999)
17. Ray, S., Munksgaard, J., Morthorst, P., Sinner, A.F.: Wind Energy and Electricity Prices - Exploring the 'merit order effect'. A literature review by Pöyry for the European Wind Energy Association (EWEA) (2010)
18. Reisinger, C.: Numerische Methoden for hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben (in German). Ph.D. thesis, Ruprecht-Karls-Universität (2004)
19. Weigt, H.: Germany's wind energy: The potential for fossil capacity replacement and cost saving. Applied Energy **86(10)**, 1857–1863 (2009)