NTNU – Trondheim
Norwegian University of
Science and Technology

Export of Norwegian Pumped Storage

Ivar Gåsvatn

Master of Energy and Environmental Engineering
Submission date: June 2014
Supervisor: Olav B Fosso, ELKRAFT

Norwegian University of Science and Technology
Department of Electric Power Engineering

Problem description

The considerable growth of new energy sources in Europe, in particular wind and solar power, is challenging for the operation of the power system as it leads to great variability in production. One way to solve this issue is by establishing better options for energy storage. Storage of electricity in large volumes, however, is very expensive, but an alternative is storage in reservoirs. Norwegian hydropower is suitable for this purpose for several reasons.

The aim of this thesis is to optimize the scheduling of a Norwegian pumped storage power plant with a view to Norway as a future large-scale exporter of pumped storage. In this context, the German energy market is a decisive factor. The student is therefore supposed to develop a stochastic description of the German spot price, and use the model as an input for a Norwegian pumped storage power plant.

Assignment given:	Trondheim 16.01.2014
Supervisors:	Olav Bjarte Fosso
	Birger Mo

Preface

This thesis was written under the supervision of Olav Bjarte Fosso at the department of Electric Power Engineering, NTNU. It marks the end of my time as a master student in Energy and Environmental engineering at NTNU.

The thesis is divided into two main parts. The first part consists of a scientific paper, currently under evaluation for possible publication in Energy Procedia. The second part is a documentation; its purpose is to treat some of the subjects touched upon in the paper more in depth.

I would like to express gratitude towards my supervisor, Olav Bjarte Fosso, who, with his knowledge and advice, has guided me through this last year. I would also like to thank Birger Mo at SINTEF Energy Research for fruitful discussions, and Florentina Paraschiv at St. Gallen University for generously allowing me to use her program script in this thesis.

Contents

Paper

- Submitted for approval by the Scientific Committee of RERC 2014.

Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 00 (2014) 000–000

Energy Procedia

www.elsevier.com/locate/procedia

Renewable Energy Research Conference, RERC 2014

Export of Norwegian pumped storage

Ivar Gåsvatn[a]*, Olav Bjarte Fosso[a]

[a]Norwegian University of Science and Technology, O. S. Bragstads plass 2E, 7491 Trondheim, Norway

Abstract

The increased use of renewable energy sources in Continental Europe, in particular Germany, has led to a great variability in power production and prices. One proposed way to remedy the Continent's power balance – which is also economically viable in a free power market – is by the means of Norwegian pumped storage. However, the profitability in such an environment is highly dependent upon the extent of price variation in the market.

In this paper, it has therefore been sought to find a stochastic price model in which the spot price is allowed to fluctuate around a sound forecast. With historical data from the German power market as the point of departure, a deterministic price curve with seasonal and daily patterns has been obtained through linear regression. This curve has been adjusted to the market expectations contained in power futures contracts. By contrasting the updated deterministic price curve with the actual spot price, it has been possible to obtain a time series model on the basis of deviations that are mainly stochastic.

The time series model has been used to generate multiple spot price scenarios that serve as an input for a representative Norwegian pumped storage power station. Simulations show that both the power station's production planning and revenues are dependent on which scenario that is under consideration. Nonetheless, the production patterns under both the different scenarios and the real spot price are comparable, in particular with regards to the daily and weekly patterns. Similarly, the total profits depend upon the variance of the price scenarios, but all scenarios, including the actual spot price, has been shown to yield significant revenues.

Keywords: Power market; HPFC; time series; pumped storage; production planning.

* Corresponding author. Tel.: +47-977-15-933.
E-mail address: gasvatn@stud.ntnu.no

1. Introduction

New awareness of climate change and the preference for energy independency, has led many countries across Europe to adopt green energy policies. Germany has by far been the leading actor in the transition from fossil to renewable energy sources. The installed capacity of the two most prominent renewable energy sources in Germany, namely solar and wind power, now amounts to 70 MW, which is more than 40 per cent of the country's total installed capacity [1]. With the complete phase-out of nuclear power plants by 2020 and a continuation of substantial incentives for these renewables, this share is only thought to increase in the foreseeable future [2].

The major challenge with solar and wind power plants is their inherent property of intermittency; at times they can nearly cover Germany's entire energy demand, at others they produce literally nothing. Sufficient energy storage solutions is therefore vital for the transition to a green energy market. The only affordable, large scale storage solution to this date is pumped storage. However, since the geography of Germany does not allow for many such facilities, it has been proposed that Norwegian pumped storage could play a major part in Germany's "Energiewende" [3].

Pumped storage has both the property of handling vast amounts of excess energy, while at the same time being an economically self-sufficient solution. As long as the difference in price during the day, or over the week, is above a relatively minor threshold, marginal profit will be made for the plant's owner.

This paper will deal with the contribution margin in relation to export of Norwegian pumped storage; the investment costs of such a plant is beyond its scope. Furthermore, it will take German (day-ahead) spot prices, rather than Norwegian, for given. This enables us to explore the possibilities of pumped storage in an environment with stronger market connection to the European mainland.

The article will attempt to establish a stochastic model for the German day-ahead spot price that is parsimonious in its description, but still being comparable to a fundamental model in its accuracy. It tries to achieve this by first constructing a shape curve on the basis of historical data from 2009-2013, and then adjusting this curve to the market expectations given at the German futures exchange. The stochasticity of the model will be introduced by modelling the deviations between the adjusted curve and the actual spot price as an ARMA/GARCH process. This model will then be used to simulate 25 different price scenarios over three weeks in January 2014. Finally, these scenarios will serve as an input to a model of a medium sized Norwegian pumped storage power station. The article concludes by making economically as well as production planning assessments on the model's results.

Nomenclature

P_t	German day-ahead spot price at hour *t*
F_t^d	daily factor, weighting the spot price at hour *t* to the mean price of the associated day *d*
F_d^y	yearly factor, weighting the mean price of day *d* to the mean price of the associated year *y*
$H_{t,i}$,	binary explanatory variable at hour *t* and instance *i* for hours
$D_{d,i}$, $M_{d,i}$	binary explanatory variables at day *d* and instance *i* for days and months respectively
s_t	shape curve of the German day-ahead spot price at time t
α, β, γ	parameters in the regression models
ε_d, ε_t	error terms in the daily and yearly regression models respectively
f_t	hourly price forward curve (HPFC) for the German day-ahead spot price at time *t*
$F(T^S,T^E)$	price of forward contract with start date T^S and end date T^E
y_t	time series given by P_t - f_t
φ, κ, θ	parameters in the time series model
z_t	independent white noise process

2. Stochastic modeling of the German day-ahead spot price

This chapter will deal with the construction of a stochastic model of the spot price in the German day-ahead market. The first section shows how a deterministic price curve that resembles the characteristic shape of the German spot price can be established through simple linear regression in two steps. First by using the time of year and day of the week as regression variables, then by regressing on the particular hours of the day.

The second section outlines how the resulting shape curve can be adjusted through the market expectations inherent in futures contracts. The adjusted price curve will provide a basis for a stochastic time series model in the last section of this chapter.

Establishing a shape curve through linear regression

The purpose of this section is to establish a shape curve that resembles the same seasonal pattern, both yearly and daily, as the German spot price. Emphasis will therefore be placed on relative prices, rather than absolute. The section follows the same line of thought as in [4].

First, we introduce a yearly factor, F^y, that will be the response variable of its explanatory variables, i.e. weekdays and months, according to (1). To ensure that each month is approximately within the same price level, the breakdown of the year do not follow the exact calendar months. In particular, the summer and Christmas holidays are treated separately, thus giving rise to one additional month.

$$F_d^y = \frac{\frac{1}{24}\sum_{t\epsilon d} P_t}{\frac{1}{8760}\sum_{t\epsilon y} P_t} = \alpha + \sum_{i=1}^{6} \beta_i D_{d,i} + \sum_{i=1}^{12} \gamma_i M_{d,i} + \varepsilon_d \tag{1}$$

Second, a daily factor, F^d, serves as the regressand for the individual hours of the day. Since the daily price patterns change both during the seasons and over the week, each quarter of the year is assigned seven unique days. This amounts to 28 different classes, each comprised of 24 different hourly parameters, which is indicated by marking the parameters and regressors in Eq. (2) with a c:

$$F_t^d = \frac{P_t}{\frac{1}{24}\sum_{t\epsilon d} P_t} = \alpha^c + \sum_{i=1}^{23} \beta_i^c H_{t,i}^c + \varepsilon_t \tag{2}$$

The two preceding equations ascribes to the different hours of the year a specific weight, relative to the daily and yearly mean respectively. An absolute value of a particular hour can then be obtained by multiplying the two regressed factors belonging to that hour with a yearly average price, so that $s_t = F_d^y \cdot F_t^d \cdot \overline{P_t}$. The resulting values of this equation will constitute what we henceforth shall denote as the shape curve.

The shape curve constructed in this paper is based upon spot price data from EPEX SPOT in the period from 2009 through 2013 [5]. The parameters that specifies the two factors are determined by the ordinary least squares estimator, i.e. by minimizing the sum of the squared residuals in Eq. (1) and (2).

Adjusting the shape curve to market expectations

The shape curve that was established in the previous section is only based on historical data. Its accuracy is therefore dependent on whether the time interval it is describing is near its historical mean value or not. In order to make the shape curve reasonable for all hours, we will in this section explain a method that aligns the curve to the market expectations. The approach is also described in [6] and [7].

The idea of this method is that, over a delivery period, $[T^S, T^E]$, the price of a fixed forward contract should equal the mean of a function, f(t), often denoted as the hourly price forward curve (HPFC):

$$F(T^S, T^E) = \frac{1}{T^E - T^S}\int_{T^S}^{T^E} f(t)\, dt \tag{3}$$

In our case, f(t) is comprised of the shape curve, s(t), and an error term, e(t). To ensure that the HPFC upholds its seasonal pattern given from the shape curve, any adjustments on the HPFC is done on the error term. Furthermore, e(t) is defined as a polynomial spline function of the fifth order, which entails a maximum smoothness property as described in [8]. The interval of the splines are determined by the start and end dates of the settlement periods of the future contracts. The parameters of the polynomial splines are finally given by minimizing the integral of the squared, double differentiated error term. This yields the smoothest possible curve on the interval $[T^S, T^E]$ [9].

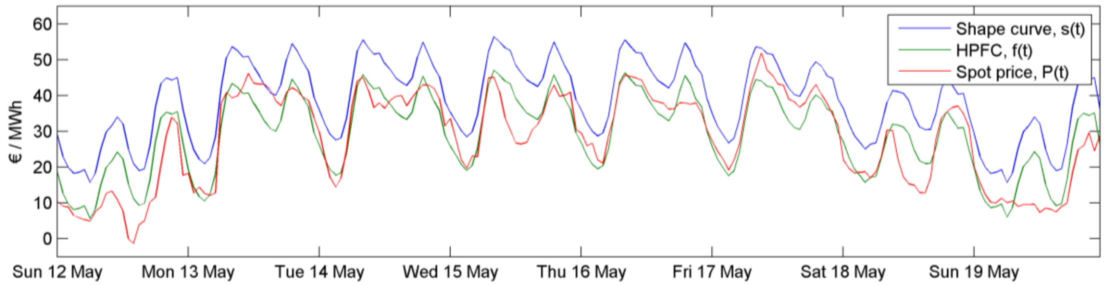


Fig. 1. Adjustment of the shape curve to futures contracts for a week in May 2013

In this paper, we have dealt with base power contracts, named Phelix Futures on the German energy exchange EEX, that were traded in 2013 [10]. Although being an exchange with a considerable turnover, the shortest time window with satisfactory liquidity is to this date weekly contracts. Therefore, data from each Friday during the year has been used to establish a new HPFC for the following week. Next, these HPFCs have been combined to give an adjusted shape curve for the whole year. When compared to the initial shape curve, the adjusted curve gave a rise in R^2 from 0,49 to 0,65. Figure 1 gives an example of an adjusted HPFC for a week in May.

Developing a stochastic time series model

Time series models are very common when describing stochastic behavior in power markets as well as in most other financial markets. Within this category, we find the ARMA class of models, which are based on the assumption that the series have an internal structure of autocorrelation. In its simplest form, it states that the time series can be properly described by a linear combination of its lagged values and lagged error terms in addition to a random error term, assumed to be independent and normally distributed [11].

Here, we are not interested in the autocorrelation of the spot price as such, but rather in the spot price's deviations from the adjusted forward curve. We will initially presuppose that $P(t) - f(t) = y_t \sim N(0, \sigma^2)$.

Our sample data is the constructed HPFC for 2013 and its associated spot price. From the autocorrelation function (ACF) in Figure 2, we can observe the series' strong autocorrelation. However, the partial autocorrelation function (PACF) indicates that the dependency is only significant for its two first lags and for the corresponding hour the previous day, thereby suggesting a multiplicative autoregressive model on the form AR(2)×(24).

It is clear from the ACF plot on the left in Figure 3 that the inferred residuals from the AR(2)×(24) model do not exhibit any significant autocorrelation, thus indicating that the model is of the right order. However, the ACF plot of the squared residuals on the right side of Figure 3 reveals strong autocorrelation. This implies that the variance of the series is not independent, but rather conditional on its previous values. We have therefore chosen to model the error term as a GARCH(1,1) process. This decision is supported by the Akaike information criterion (AIC), which decreased in value, from $5.13 \cdot 10^4$ to $4.92 \cdot 10^4$.

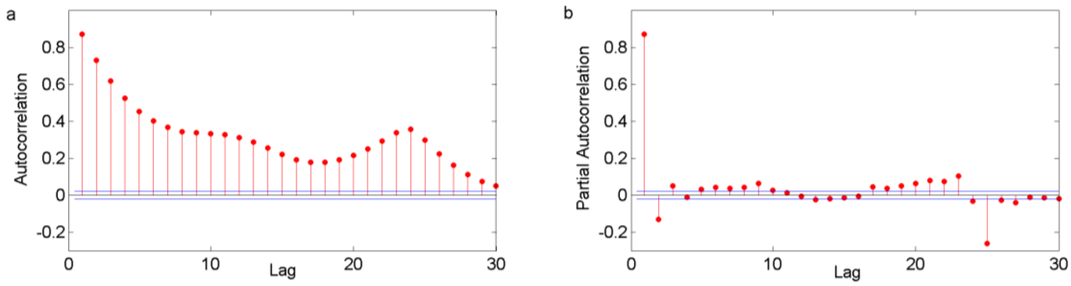


Fig. 2. (a) autocorrelation function for P(t)-f(t); (b) partial autocorrelation function for P(t)-f(t).

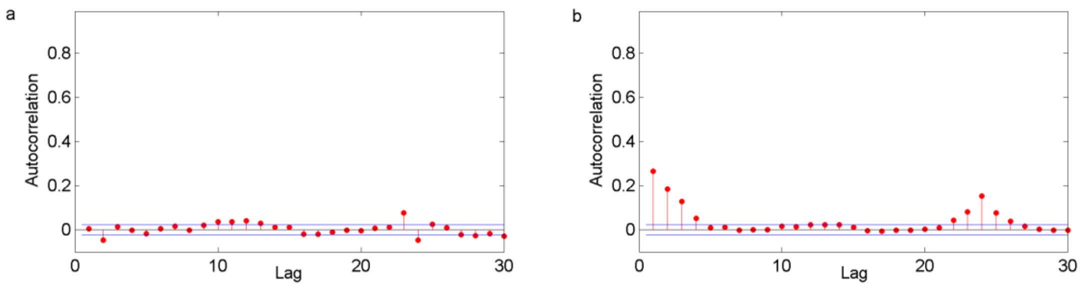


Fig. 3. (a) autocorrelation function for the residuals; (b) autocorrelation function for the squared residuals.

Upon examining the standardized residuals, we could also discern that its distribution had excess kurtosis compared to the standard normal distribution. This motivated the use of Student's t distribution for the innovation process. Here, we gained a further decrease in AIC to $4.78 \cdot 10^4$. Our final model is described in Eq. (4).

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-24} + \varepsilon_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-24} + z_t \sqrt{\kappa + \theta_1 \sigma_{t-1}^2 + \theta_2 \varepsilon_{t-1}^2} \quad (4)$$

3. A Norwegian pumped storage power plant

In developing the power station model, we have sought to make it resemble a medium sized Norwegian power plant. Its physical properties and constraints are listed in Table 1.

Table 1. Physical properties and constraints of the pumped storage power station.

Property	Value	Constraint	Lower constraint	Upper constraint
Upper reservoir initial / end level	1 000 000 m^3	Upper reservoir	0 m^3	2 000 000 m^3
Lower reservoir initial / end level	1 000 000 m^3	Lower reservoir	0 m^3	2 000 000 m^3
Plant head	100 m	Discharge capacity	0 m^3/s	100 000 m^3/h
Effiency power station / pump	0.84 / 0.90	Relift capacity	0 m^3/s	40 000 m^3/h

A natural choice for modelling production planning in this environment is through linear programming, in which the spot price times the power produced/consumed serves as the objective function. Detailed mathematical description and solving strategies of such a problem can be found in [12]. In our approach, we will simplify the analysis by disregarding start-up costs and environmental constraints, and use mean values for efficiency and head.

4. Simulations and results

In order to assess the export potential of Norwegian pumped storage in a European pricing regime, we have taken the German spot price to be a fixed input variable for the modelled power plant. Based on the constructed HPFC from the second chapter, 25 different price scenarios for a period of three weeks in January 2014 have been attained through Monte Carlo simulation. The scenarios are plotted against the HPFC and actual spot price in Figure 4.

The economic analysis reveals that every scenario yields significant revenues over these three weeks. Though ranging from 16 000 € to 29 500 €, they generate a mean net revenue of 21 400 € a week. When scheduling the production after the real spot price, however, the profits are somewhat smaller, amounting to a net income of 16 800 € a week. Similar calculation for other periods of 2014 indicates that this is a representative figure. The reason for this slight divergence is to be found in the price scenarios. It can be seen from Figure 4 that the scenarios have a greater degree of volatility than the real spot price. Since pumped storage plants takes advantage of precisely such variations, we should expect the simulated price scenarios to give greater revenues. The assessment thereby emphasizes the economic potential in a market with high volatility.

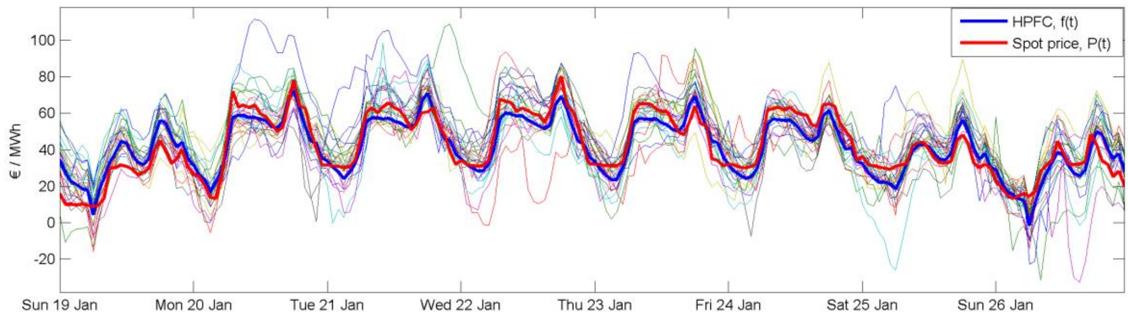


Fig. 4. Upper reservoir volume under different price scenarios. The mean scenarios and the actual spot price scenario are highlighted.

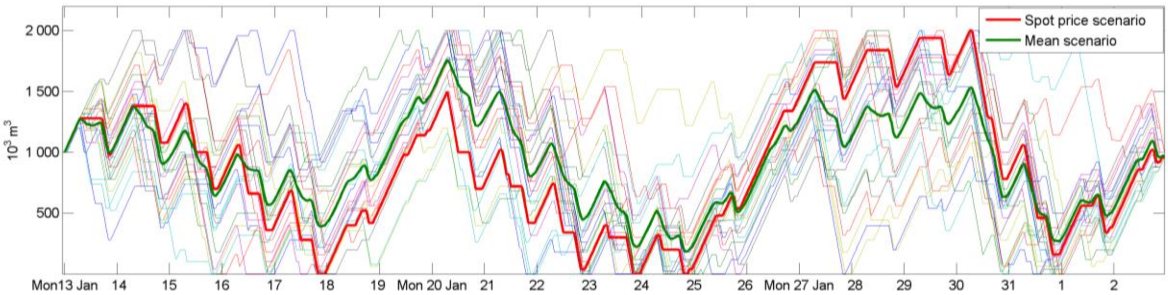


Fig. 5. Simulated price scenarios compared to HPFC and actual spot price for the second simulation week.

In the same way as the power station's bottom line is conditional on the scenarios, so is the associated production planning. Figure 5 depicts the upper reservoir's content corresponding to the different price profiles during the three weeks. As is evident, the volume levels of the reservoir varies significantly with the scenarios. However, they exhibit the same daily and weekly pattern; the reservoirs are on the whole discharged during the weekdays and refilled in the weekend, and are producing in the day and pumping during the night.

5. Conclusion

With the purpose of exploring the opportunities for Norwegian pumped storage in a European pricing regime, this paper has sought to find a simple yet accurate description of the German day-ahead spot price. This has been achieved by establishing a shape curve through linear regression and then adjusting this curve to market expectations inherent in futures contracts. Futures adjustments made for 2013 yielded a significant increase in explained variance, from 49 to 65 per cent. A stochastic model could consequently be constructed on the basis of the curve's deviations from the actual spot price. The model specification process showed that the stochasticity was well accounted for with a parsimonious ARMA/GARCH description.

The stochastic model allowed for simulation of price variations around the adjusted shape curve. These simulations have been used as an input for a modelled Norwegian pumped storage power plant. The economic analysis revealed that all scenarios generated significant revenues, including the real spot price scenario. However, the differences in profits clearly demonstrates the importance of price variations. Production planning also showed to be dependent on the scenario. Nevertheless, a common daily and weekly pattern could be identified. Furthermore, the mean scenario lay very close to the production path of the spot price. The results thus indicate that both economic and production related assessments can be made on the grounds of the developed price model.

Acknowledgements

The authors would like to thank Florentina Paraschiv from the University of St. Gallen for allowing us to use her program script to generate hourly price forward curves, and Birger Mo from SINTEF Energy Research for fruitful discussions during the work on this article.

References

[1] BDEW Bundesverband der Energie- und Wasserwirtschaft e.V. Erneuerbare Energien und das EEG: Zahlen, Fakten, Grafiken; 2014.

[2] Fraunhofer Institut für Solare Energiesysteme ISE. Stromerzeugung aus Solar- und Windenergie im Jahr 2014.

[3] Prognos AG. Bedeutung der internationalen Wasserkraft-Speicherung für die Energiewende; 2012.

[4] Blöchlinger, L. A Regime-Switching Spot/Forward Price Model with Kim Filter Estimation. PhD thesis, University of St. Gallen; 2008.

[5] EPEX SPOT SE: Auction. http://www.epexspot.com/en/market-data/auction.

[6] Benth FE, Koekebakker S, Ollmar F. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variations. The Journal of Derivatives 2007;15(1):52-66.

[7] Fleten S-E, Paraschiv F, Schürle M. Spot-Forward Model for Electricity Prices. Working Papers on Finance 1311, St. Gallen; 2013.

[8] Adams KJ, van Deventer DR. Fitting yield curves and forward rate curves with maximum smoothness. Journal of Fixed Income 1994;4:52-62.

[9] van Deventer DR, Imai K, Mesler M. Advanced Financial Risk Management. Somerset: John Wiley & Sons Inc; 2004.

[10] EEX AG: Phelix Power Futures. https://www.eex.com/en/market-data/power/derivatives-market/phelix-futures.

[11] Gershenfeld N. The Nature of Mathematical Modelling. Cambridge University Press; 1998.

[12]Fosso OB, Belsnes M, Røynstrand J, Gjengedal T, Valhovd E. Hydro power short-term scheduling in an online environment. MedPower; 2003.

Documentation

1 Introduction

Some of the topics of the paper has, as a consequence of being limited to six pages, been treated rather superficially. This second part of the thesis will therefore give a more comprehensive review of these subjects. In particular, it will deal with some of the theoretical background the paper is built upon. Also, it will give more insight into the methods and provide justification for their application.

In order to give a more natural course through the topics, the main part of this documentation is a merger of method and results. Furthermore, the structure is organized in the same way as in the paper, thus making the transition between the two main parts more seamless. The documentation ends with a chapter providing a discussion and suggestions for further work.

2 Method and results

2.1 Establishing a shape curve through linear regression

This chapter will deal with the first part of constructing a stochastic model of the German spot price, namely to make a shape curve of the price that resembles its characteristic seasonal patterns. Multiple linear regression, formulated as an ordinary least squares problem, has been the chosen method for achieving this [1].

In this class of regression problems, the dependent variable is set to be linearly dependent on its explanatory variables. In order to evalute the dependent variables seperately, the explanatory variables have been treated *nominally*, that is to say as categories. By introducing dummy variables for each category, i.e. hours, days and months, the dependent variables, in this case the spot price, can be explained effectively by the time at which it occurs. The variables and parameters that have been used in the regression analysis are summarized in table 2.1.

Table 2.1: Variables and parameters in the regression model

Dependent variables	
F_t^d	Price at time t relative to the mean of the corresponding day, d
F_d^y	Mean price at day d relative to the mean of the corresponding year, y
Independent variables	
$H_{t,i}$	Binary explanatory variables at time t and instance i for hours
$D_{t,i}$	Binary explanatory variables at time t and instance i for days
$M_{t,i}$	Binary explanatory variables at time t and instance i for months
Parameters	
α	Constant parameter
β_i, γ_i	Linear parameters for instance i
ε	Error term of the regression model

The chosen method is comprised of two steps. The first gives the shape curve its yearly seasonality by regressing on the dummy variables for days and months. Six binary variables have been used to explain the effect given by the days from Tuesday through Sunday. Since a dummy variable can not be in a linear relation the other categories, Monday is therefore not given by a separate variable, but rather accounted

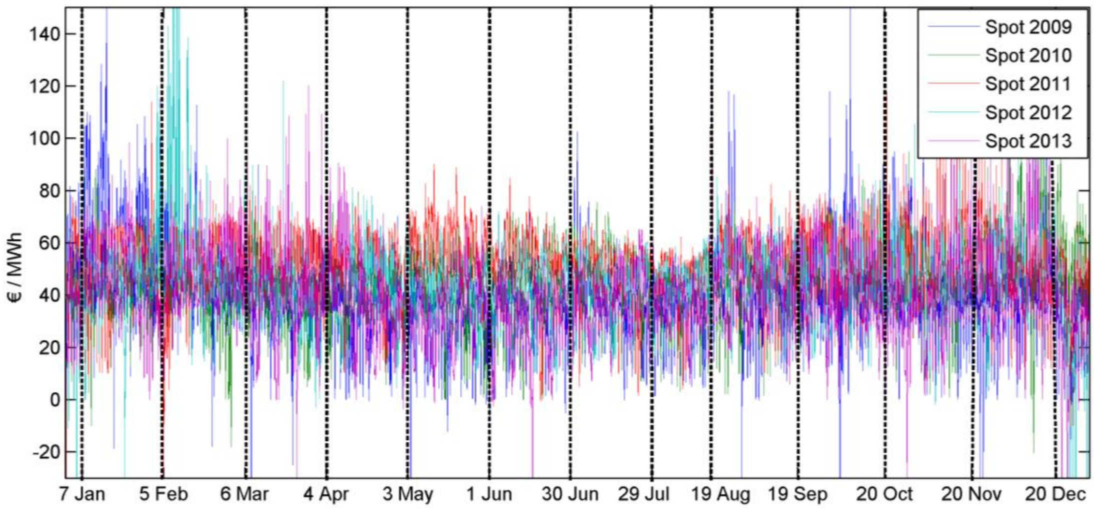


Figure 2.1: Breakdown of the year into 'monthly' periods. The spot price of the sample (from 2009 through 2013) are plotted in the background.

for by the constant parameter. Furthermore, the price characteristics of a specific day are likely to begin first at 4 am and end 24 hours later. The daily variables are therefore assigned the hours from 04^{00} through 03^{00}. This also ensures that the seasonality curve connects more smoothly, since the lowest price usually occurs at the forth hour of the day. Likewise, the monthly variables have been assigned prices for periods with similar price levels. The breakdown of the year originates with the summer and Christmas holydays being given precedence. The rest of the year are divided so that each period are of comparable length as an ordinary month. Figure 2.1 illustates how this division is done. The two holidays give rise to an additional period, yielding a total of twelve variables. Similarly as above, the first period is accounted for by the constant parameter. Table 2.2 summarizes the different categories for the first step of the shape curve regression.

Table 2.2: Categories for the variables of the yearly regression

Daily variables (each day are counted from 04^{00} that day till 03^{00} the next)			
$D_{t,1}$	Tuesday	$D_{t,4}$	Friday
$D_{t,2}$	Wednesday	$D_{t,5}$	Saturday
$D_{t,3}$	Thursday	$D_{t,6}$	Sunday
Monthly variables			
$D_{t,1}$	05-Feb – 05-Mar	$D_{t,8}$	29-Jul – 18-Aug (summer vacation)
$D_{t,2}$	06-Mar – 03-Apr	$D_{t,9}$	19-Aug – 18-Sep
$D_{t,4}$	04-Apr – 02-May	$D_{t,10}$	19-Sep – 19-Oct
$D_{t,5}$	03-May – 31-May	$D_{t,11}$	20-Oct – 19-Nov
$D_{t,6}$	01-Jun – 29-Jun	$D_{t,11}$	20-Nov – 19-Dec
$D_{t,7}$	30-Jun – 28-Jul	$D_{t,12}$	20-Dec – 06-Jan (Christmas holiday)

The regrassand for the above variables, F_d^y, is defined as the average price of the day which hour t belongs to, divided by a yearly mean. The regression model for the first part then becomes:

$$F_d^y = \frac{\frac{1}{24}\sum_{t\epsilon d} P_t}{\frac{1}{8760}\sum_{t\epsilon y} P_t} = \alpha + \sum_{i=1}^{6} \beta_i D_{t,i} + \sum_{i=1}^{12} \gamma_i M_{t,i} + \varepsilon_t \tag{2.1}$$

where P_t is the German day-ahead spot price at time t. In short, the model in 2.1 gives a particular day of the year a relative weight that is explained by the weekday and month it is ascribed to.

The second step of the method aims to provide the shape curve with a daily pattern. However, the price patterns varies both over the year and during the week. Each weekday has therefore been treated separately, while the year has been divided in four. The latter is a obviously a coarser grouping than is the case for the monthly variables above, but has been considered necessary in order to give a sufficiently large sample. Table 2.3 provides an overview of the classification.

Table 2.3: Classes for the variables of the daily regression

	Winter (Dec-Feb)	Spring (Mar-May)	Summer (Jun-Aug)	Autumn (Sep-Nov)
Monday	$H_{t.i}^{1}$	$H_{t.i}^{7}$	$H_{t.i}^{13}$	$H_{t.i}^{19}$
Tuesday	$H_{t.i}^{2}$	$H_{t.i}^{8}$	$H_{t.i}^{14}$	$H_{t.i}^{20}$
Wednesday	$H_{t.i}^{3}$	$H_{t.i}^{9}$	$H_{t.i}^{15}$	$H_{t.i}^{21}$
Thursday	$H_{t.i}^{4}$	$H_{t.i}^{10}$	$H_{t.i}^{16}$	$H_{t.i}^{22}$
Friday	$H_{t.i}^{5}$	$H_{t.i}^{11}$	$H_{t.i}^{17}$	$H_{t.i}^{23}$
Saturday	$H_{t,i}^{6}$	$H_{t,i}^{12}$	$H_{t,i}^{18}$	$H_{t,i}^{24}$
(each class, $H_{t,i}^{c}$ is comprised of 23 hourly binary variables)				

The second part of the model can thereby be formulated as follows:

$$F_t^d = \frac{P_t}{\frac{1}{24}\sum_{t\epsilon d} P_t} = \alpha^c + \sum_{i=1}^{23} \beta_i^c H_{t,i}^c + \varepsilon_t \tag{2.2}$$

Likewise as in Eq. 2.1, P_t is the German spot price at time t, while the dependent variable, F_t^d, represents the weight of the price relative to the day's mean.

The estimates for the yearly and daily factor, $\hat{F}_t^d$ and $\hat{F}_d^y$ respectively, are proportional numbers that, when multiplied, yields the shape of the price curve. To attain a shape curve, s_t, in absolute numbers, the factors must be multiplied with a mean price level. This is done according to Eq. 2.3.

$$s_t = \hat{F}_d^y \cdot \hat{F}_t^d \cdot \overline{P}_t \tag{2.3}$$

The sample data used in the regression analysis has been obtained at the European

Power Exchange (EPEX) [2], and comprises of German spot prices from 2009 through 2013. Over this time period the framework of the power exchange has largely been unchanged, including negative pricing, which was introduced in September 2008.

To illustrate how the shape curve is constructed, the estimated (in-sample) yearly and daily factors have been plotted in Figure 2.2 and 2.3, respectively. Furthermore, in Figure 2.4 the calculated shape curve has been plotted against the actual spot price for January, 2013. (The mean of the whole sample serves as the mean price level, $\overline{P}_t$.) It can be seen from the figure that much of the price's behaviour is explained by the shape curve. However, at the end of the month, one can notice a greater divergence. At this point the German power grid were fed with substantial amounts of wind power, leading to very low prices and, thus, a considerable departure from historical mean prices. This points to the need for making adjustments on the curve, as will be dealt with in the next chapter.

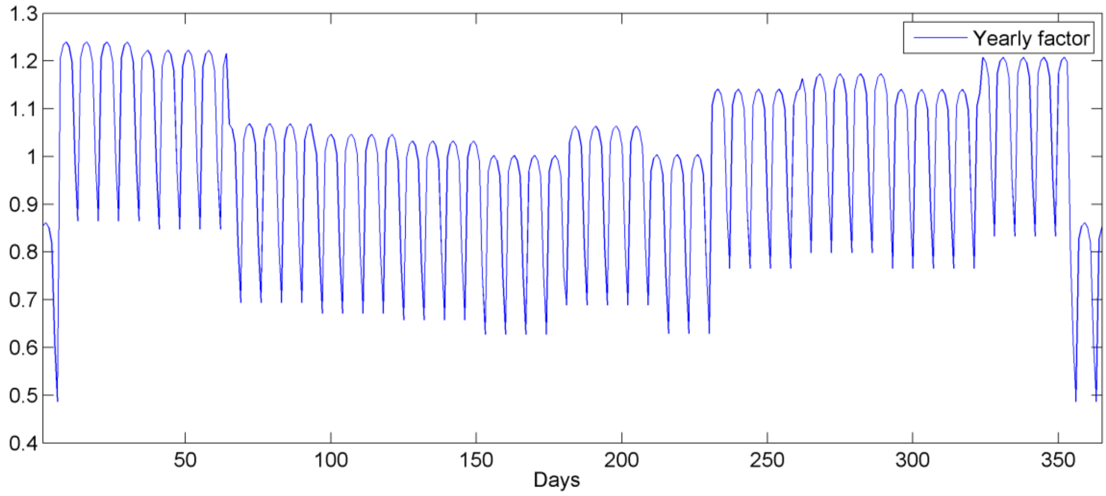


Figure 2.2: Estimated yearly factor for 2013.

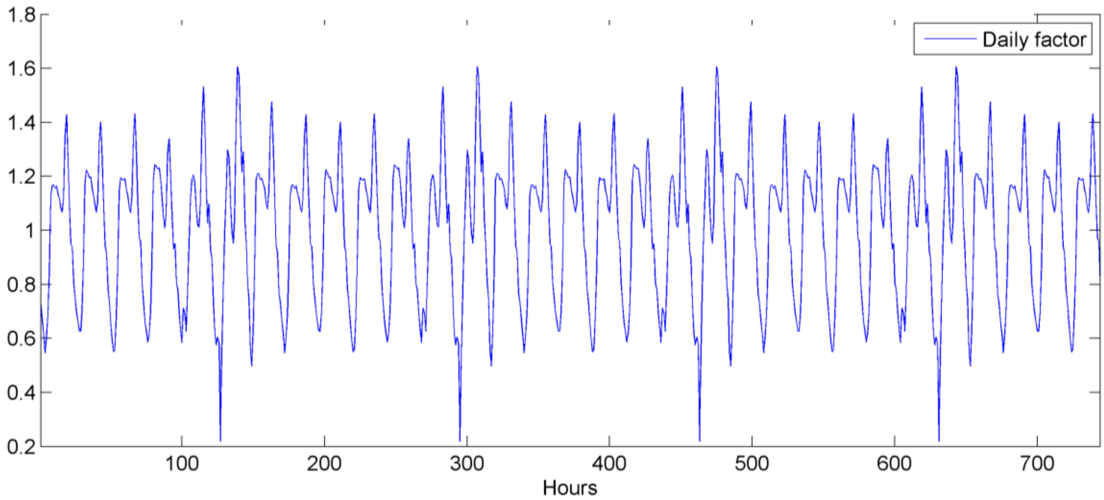


Figure 2.3: Estimated daily factor for January, 2013.

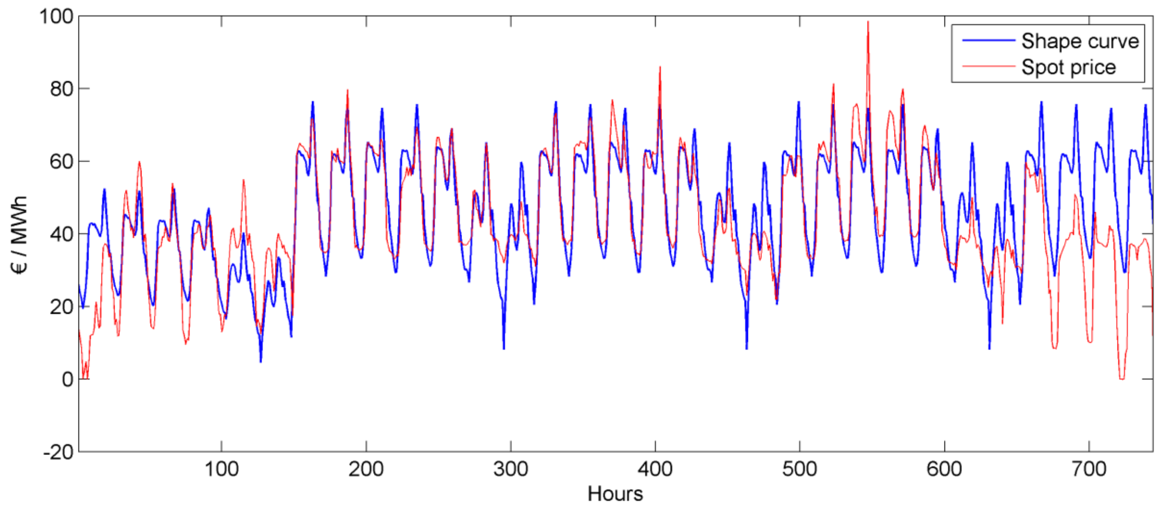


Figure 2.4: Estimated shape curve for January, 2013, plotted aginst the actual spot price.

Nonetheless, the shape curve estimated for 2013 yielded an explained variance of 49 per cent. Furthermore, as is illustrated in Figure 2.5 and 2.6, much of the initial seasonality of the spot price has been taken into account by the shape curve. From Figure 2.5 it can be seen that the price series exhibits significant autocorrelation (the concept of autocorrelation will be elaboretad upon in chapter 2.3, for now it suffice to interpret it as a measure of seasonality). The autocorrelation of the shape curve's deviation from the spot price depicted in Figure 2.6, however, makes clear that the diurnal seasonality is greatly reduced. Moreover, it can be noticed that the diurnal peaks' are levelled, thus illustrating that the regression model has succeeded in differentiating on the distinct daily patterns.

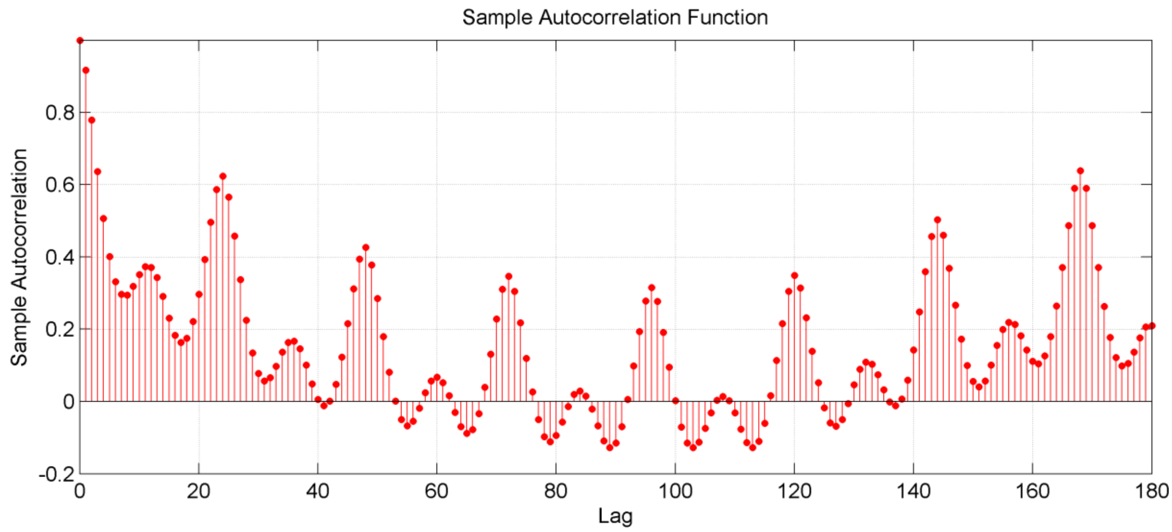


Figure 2.5: Autocorrelation of the German day-ahead spot price in 2013.

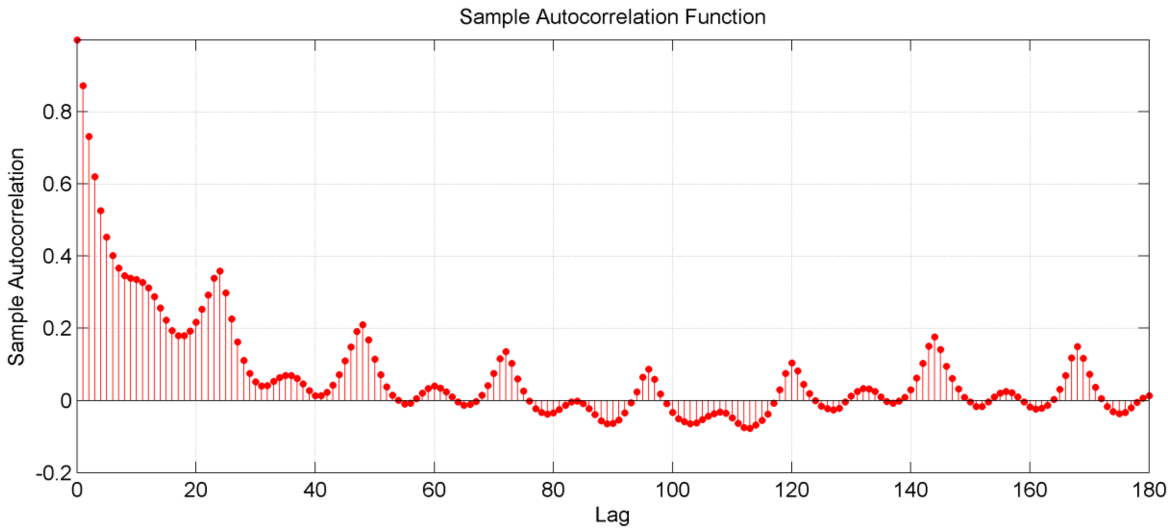


Figure 2.6: Autocorrelation of the German day-ahead spot price in 2013 subtracted by the corresponding shape curve.

2.2 Adjusting the shape curve to market expectations

The shape curve established in the previous section is only a product of the spot price's historical values. While it reproduces the seasonality of the German spot price, both the daily and yearly pattern, it will not necessarily reflect the right price level. This chapter will therefore provide a means for adjusting the shape curve such that it resembles the market expectations found in futures contracts. The advantage of this approach is that the accuracy is comparable to fundamental models, without having to analyze the production and demand in the power market. The first section will account for the theoretical background, while the second section deals with the practical application.

2.2.1 Hourly price forward curves

The method of hourly price forward curves (HPFC) originates from the concept of yield curves, in which the yields, or interst rates, gives an upward sloping relation with the corresponding contract lengths. This notion is closely related to forward curves, where it is the price of the forward, or futures, contracts that is compared to their respective maturity dates. The nature of the curve is wholly dependent on the expectations of the future development of the underlying, namely the spot price. Hourly price forward curves are a further extension of this concept. HPFCs are constructed on the basis of forward curves, but are extended to include hourly steps [3]. The idea is summarized in Eq. 2.3.

$$F(T^S, T^E) = \frac{1}{T^E - T^S} \int_{T^S}^{T^E} f(t)\, dt \tag{2.3}$$

The equation above states that the price of the financial contract with delivery period from T^S to T^E, is given by the mean of the HPFC, $f(t)$, over this period. The formula implies that contracts are settled in T^E. Furthermore, it is constructed on the basis of risk-neutral probability. This entails that there is an absence of arbitrage, i.e. the derivative's price equals the discounted expected value of the underlying.

The HPFC can be composed into two parts, a seasonal component $\mathrm{s(t)}$ and an adjustment function $\mathrm{\varepsilon(t)}$. The seasonal component will simply be the shape curve established in the section 2.1. As the name implies, the adjustments made on the HPFC will be done on $\mathrm{\varepsilon(t)}$. In this way the HPFC maintains the seasonal character given from the shape curve, while its average value equals the price of the financial contract. This is crucial since the price pattern is likely to be similar in a period with e.g. high infeed of wind power as under normal conditions, but the price level will quite certainly be different.

A maximum smoothness critera, which have been applied in earlier work done on yield curve fitting [4], will be the guideline for fitting the adjustment function:

$$min \int_{t_0}^{t_n} (\varepsilon''(t))^2 \ dt \tag{2.4}$$

In this context, the best choice for $\mathrm{\varepsilon(t)}$ is a polynomial spline of the forth order [5]. The i'th piece of the spline is defined according to Eq. 2.5.

$$\varepsilon_i(t) = a_i t^4 + b_i t^3 + c_i t^2 + d_i t + e_i \,, \qquad t_{i-1} < t < t_i \tag{2.5}$$

In a financial market, the derivatives will often have overlapping delivery periods. The approach will therefore be to divide these intervals into new sub-intervals, each accounted for by a unique polynomial as given in Eq. 2.5. The parameters of the resulting spline are determined by the qriteria in Eq. 2.4, and for that reason involves solving a quadratic programming problem.

To ensure connectivity and smoothness between the knots of the spline, the following constraints are applied for $t_1, t_2 \dots t_n$:

$$\varepsilon_{i+1}(t_i) = \varepsilon_i(t_i) \tag{2.6}$$

$$\varepsilon_{i+1}'(t_i) = \varepsilon_i'(t_i) \tag{2.7}$$

$$\varepsilon_{i+1}''(t_i) = \varepsilon_i''(t_i) \tag{2.8}$$

Additionally, the adjustment function will also encompass a risk premium given by the risk attitudes of market participants. The market price of risk will be time-varying, and short-time information such as weather conditions can influence this measure.

However, at the long end of the curve, i.e. several years ahead, the market's assessment of risk will approach a constant value, thus leading to the last constraint of the problem:

$$\varepsilon_n{}'(t_n) = 0 \tag{2.9}$$

The general minimizing problem can be formulated as

$$\min_x \ \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Hx}, \tag{2.10}$$

subjected to the equality constraint

$$\boldsymbol{A_{eq}x} = \boldsymbol{b_{eq}}. \tag{2.11}$$

The solution vector **x** contains the 5n parameters of the polynomial spline, ε(t), while **H** is a 5n x 5n matrix that accounts for the parameters coefficients. The latter has not a straightforward formulation, so its structure will be outlined in the following paragraph.

For the i'th piece of the spline, the function subject to minimization, according to Eq. 2.4, is:

$$\int_{t_{i-1}}^{t_i} (\varepsilon_1{}''(t))^2 \ dt = \int_{t_{i-1}}^{t_1} (12a_i t^2 + 6b_i t + 2c_i)^2 \, dt$$

$$= \left[\frac{144}{5} a_i^2 \tau^5 + 12b_i^2\tau^3 + 4c_i^2\tau + 2\cdot 18a_i b_i \tau + 2\cdot 8a_i c_i \tau + 2\cdot 6b_i c_i \tau\right]_{\tau=t_{i-1}}^{\tau=t_i} \tag{2.12}$$

As suggested by Eq. 2.12, **H** can be formulated symmetrically:

$$\boldsymbol{H} = \begin{bmatrix} \frac{144}{5}\Delta_1^5 & 18\Delta_1^4 & 8\Delta_1^3 & 0 & 0 & & & & & \\ 18\Delta_1^4 & 12\Delta_1^3 & 6\Delta_1^2 & 0 & 0 & \cdots & & 0 & & \\ 8\Delta_1^3 & 6\Delta_1^2 & 4\Delta_1^1 & 0 & 0 & & & & & \\ 0 & 0 & 0 & 0 & 0 & & & & & \\ 0 & 0 & 0 & 0 & 0 & & & & & \\ & \vdots & & & & \ddots & & \vdots & & \\ & & & & & & \frac{144}{5}\Delta_n^5 & 18\Delta_n^4 & 8\Delta_n^3 & 0 & 0 \\ & 0 & & & & \cdots & 18\Delta_n^4 & 12\Delta_n^3 & 6\Delta_n^2 & 0 & 0 \\ & & & & & & 8\Delta_n^3 & 6\Delta_n^2 & 4\Delta_n^1 & 0 & 0 \\ & & & & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where

$\Delta_i^k = t_i^k - t_{i-1}^k.$

The equality constraint in Eq. 2.11 are formulated with less effort. $\mathbf{A}_{eq}$ will quite simply include the constraints listed from Eq. 2.6 -2.9, in addition to the futures price contidion in Eq. 2.3. All constraints are linear in relation to x, and thus fits the description in Eq. 2.11. Furthermore, by introducing Lagrange multipliers for the constraints, Eq. 2.10 and 2.11 takes the form of a unconstrained minimazation problem:

$$\min_{x,\lambda} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{H}\boldsymbol{x} + \boldsymbol{\lambda}^T(\boldsymbol{A}_{eq}\boldsymbol{x} - \boldsymbol{b}) \tag{2.13}$$

The optimal solution can then finally be found by solving the following linearized problem formulation:

$$\begin{bmatrix} \boldsymbol{H} & \boldsymbol{A}_{eq}{}^T \\ \boldsymbol{A}_{eq} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix} \tag{2.14}$$

The program script used in this work is adopted from a restricted code developed by Florentina Paraschiv at the University of St. Gallen. However, the solution procedure follows the outlined theory above accurately. Equation 2.14 is solved in Matlab using LU (Lower Upper) factorization with partial pivoting.

2.2.2 Describing the German spot price as an HPFC

Futures contracts for the German market area are traded on the European Energy Exchange (EEX) with delivery periods stretching from days to years [6]. However, the liquidity for days and weekends are still insufficient for the purpose of establishing forward curves. Thus, the shortest time window that have been used here is contracts of weekly lengths. Moreover, futures are sold as base load, peak load and off-peak contracts. It has been natural to use the first type of contract (properly named as Phelix base future) as the input variable, since in this case it is the average price of the whole day that constitutes the underlying.

On this basis, data has therefore been gathered for base load contracts every Friday through 2013. Then, an HPFC has been calculated, starting on the following Monday and ending on the last date of the furthest yearly contract, at most seven years ahead. However, only the first week of the HPFC have been stored; the rest of curve serves as a shape curve of the next calculation. This procedure has been repeated for each week, and a HPFC for 2013 could eventually be obtained. Some pre-2013 HPFCs have also been calculated so that the shape curve's price level for 2013 could be more in

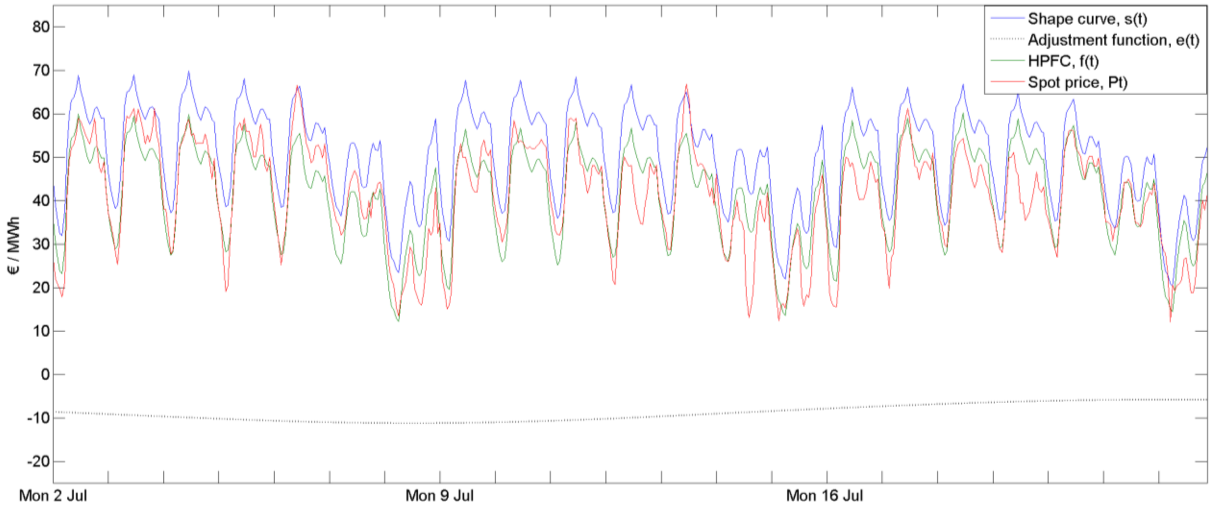


Figure 2.7: Adjusted shape curve plotted against the actual spot price for the three first weeks of July. The adjustment function indicates how the shape curve is altered.

accordance with the expectations of the market. These adjustments are more conspicuous than the weekly reshaping, since the price level of the shape curve initially is only determined by historical data. As can be seen from figure 2.7, the first adjustment, made for the first week of July 2012, makes considerable changes on the shape curve; the HPFC is clearly more in line with the actual spot price. It can also be noticed how the constraints in the above section ensures the smoothness of the adjustment function.

As the shape curve are consistently being updated with monthly and quarterly data, the long term price level will be more in agreement with the market expectations. However, adjustments still need to be done on a short term basis, so as to capture transient influences on the price such as outside temperature or wind speed. Because of their intermittency, however, their effect may only be reflected in the prices for a particular day, or even for a few hours. As the shortest delivery periods in the sample

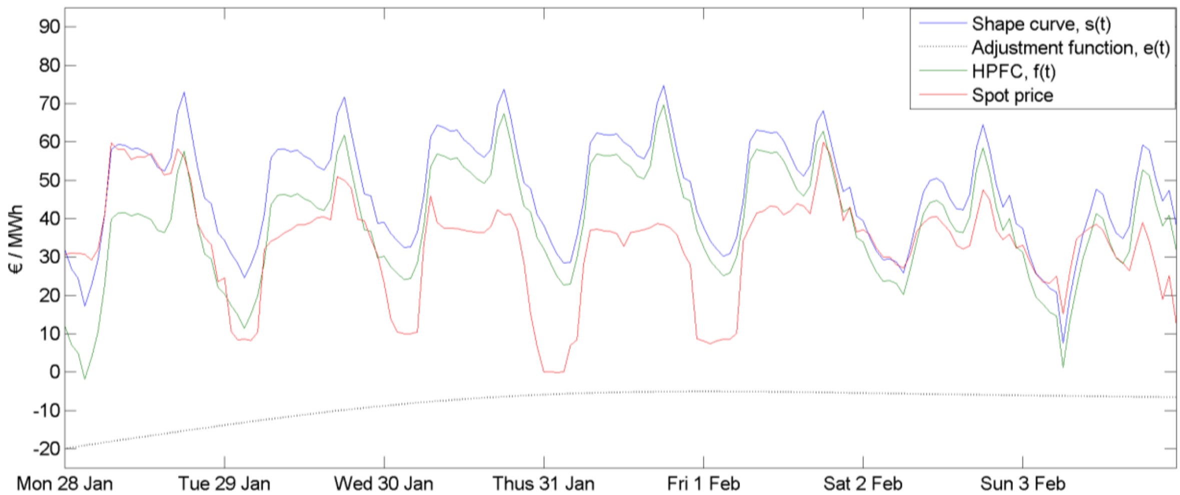


Figure 2.8: Adjusted shape curve plotted against the actual spot price for the last week of January, 2013. The adjustment function is most influential in the beginning of the period.

are weeks, the limitations of the model are most clearly seen here. Figure 2.8 illustrates this issue. It shows the HPFC for the last week in January. Expactations of considerable amounts of wind power on the grid made the futures prices for this week significantly lower than the following weeks. What the model failed to take into account, was that the excessive wind levels were mainly apparent from Tuesday to early Friday. As a consequence of its smoothing properties, the adjustment function altered the first day of the week the most, a day that were quite well described by the shape curve. The following days, on the other hand, were much less adjusted than they ought te be. However, such divergency are to be expected until a more comprehensive derivative market are in place. At any rate, the explained variance were evidently increased for the week in question.

As a whole, the hourly price forward curve for 2013, gave a considerable better fit than the initial shape curve; R^2 were improved from 0,49 to 0,65, i.e. a 16 per cent increase in explained variance. The improved fit can also be appreciated by investigating the correlation of the deviation from the spot price (less correlation points towards a better fit). Figure 2.9 depicts the autocorrelation of the divergence between the shape curve and the spot price, while Figure 2.10 compares the latter with the HPFC. It is evident from the figures that the correlation is significantly less when the adjusted shape curve is considered, especially for correlations that are more than 24 hours back in time. This last aspect indicates that while the spot price are strongly correlated within the same day, its long term correlation is not that obvious and can be accounted for by the constructed HPFC.

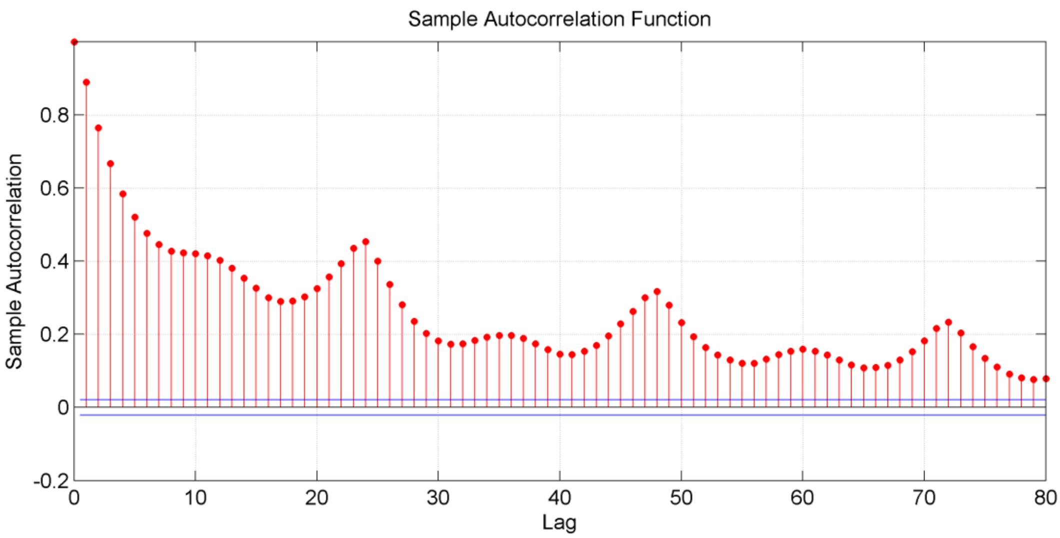


Figure 2.9: Autocorrelation of the spot price subtracted by the initial shape curve.

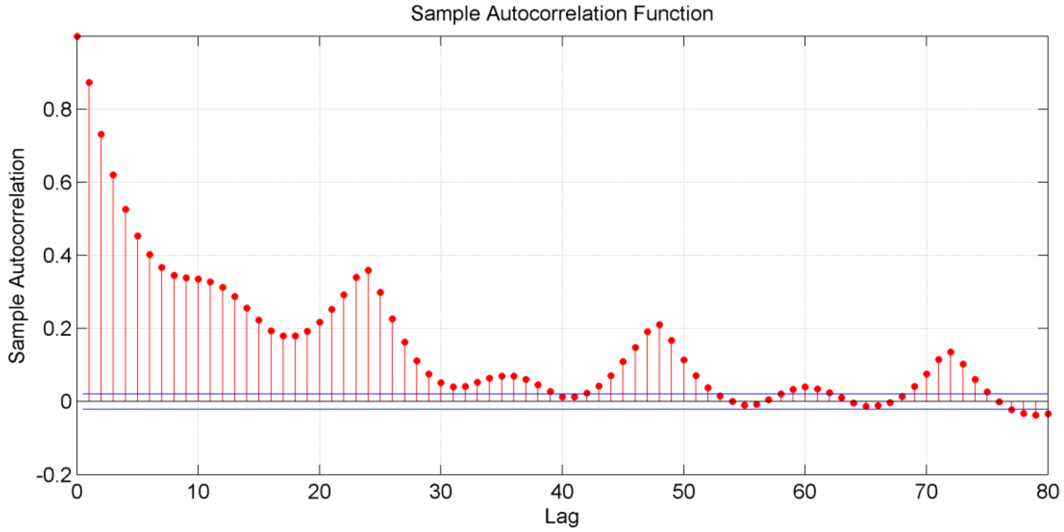


Figure 2.10: Autocorrelation of the spot price subtracted by the HPFC.

2.3 Developing a stochastic time series model

This chapter will deal with the last step in constructing a stochastic model of the German day-ahead spot price. As accounted for in the two previous sections, thus far an hourly price forward curve has been attained, constructed on the grounds of historical prices and market expectations. This curve, together with the actual spot price, are the sample data that will be used here. The chapter are organised so as to give an introduction of time series modelling in the two first parts, and then providing a motivation for the choice of a particular model in the last section.

2.3.1 Time series models

Time series are employed in a number of different fields, including the modus operandi of the marketplace. While the underlying mechanisms may be very complex, the time series they generate can be readily analysed. A typical time series contains T observations of a variable, y, observed at equal intervals over a period of time, such that

$$Y = \{y_1, \dots, y_t, \dots, y_T\} \tag{2.15}$$

Time series models are characteristic in the sense that y is sought explained by past observations. In other words, y_t has a *conditional mean* so that

$$\mu_{y_t} = \mathrm{E}(y_t | H_{t-1}) \tag{2.16}$$

where H_t represents the information known at time *t*. In addition, they also have in common a stochastic element that accounts for uncertainty, or lack of sufficient information.

Time series models that describe linear systems, fall into two categories; namely *autoregressive* (AR) and *moving average* (MA). The combination of these two models are given the acronym ARMA, and have the following mathematical description [7]:

$$y_t = c + \sum_{m=1}^{p} \varphi_m y_{t-m} + \varepsilon_t + \sum_{n=1}^{q} \theta_n \varepsilon_{t-n} \tag{2.17}$$

Following a constant term, c, the first summation in Eq. 2.17 accounts for the autoregressive part of the ARMA model. It makes y_t regress to previous values by weighting these values with an associated parameter φ_m. The third term gives the model its stochasticity; ε_t is an innovation process with mean zero and variance σ^2, assumed to be normally distributed. The last summation term provides the model with a moving average based on previous innovation terms. Its name derives from the isolated effect of producing a time series that moves around an average value. p and q indicates the order of the autoregressive and moving average part respectively. A combined model is therefore denoted as ARMA(p,q), or simply AR(p) or MA(q) if they are sepearate processes.

In addition, many time series exhibit a seasonal pattern. If this is the case, the process should be described by a seasonal multiplicative model. Such a model shares the same properties as an ordinary ARMA description, but lagged values are added at certain intervals, thus capturing effects that occur periodically. This is indicated by writing ARMA(p,q)x(P,Q), where P and Q represents the time intervals at which the correlation is apparent.

As implied earlier, the innovation process is assumed to be heteroscedastic, i.e. with constant variance. However, it is often the case that the variance changes throughout the series. This is particularly prevalent in financial markets, but also applies to markets for non-storable commodities, such as electricity. Such volatility clustering can be accounted for by introducing a *conditional variance* model, which asserts that

$$\sigma_t^2 = \mathrm{Var}(\varepsilon_t | H_{t-1}) \tag{2.18}$$

where H_t again represents the known history of the process at time *t*. It must be noted, however, that this model does not presuppose any correlation between the innovation terms; *serial correlation* is explained by the ARMA model. Rather, it states that the variance of the innovation terms are correlated and that the series thereby are *serially dependent*.

The most common conditional variance model is known as GARCH (generalized autoregressive conditional heteroscedastic). In this model the innovation process is given by

$$\varepsilon_t = \sigma^2 z_t \tag{2.19}$$

where z_t is a random variable, standardized with independent and identical distribution. In the case of normality, z_t will be drawn from the standard normal distribution. The conditional variance of a GARCH process is described in a similiar fashion as the ARMA model; it is given by a linear combination of past variances and squared innovation terms:

$$\sigma^2 = \kappa + \sum_{m=1}^{p} \alpha_m \sigma_{t-m}^2 + \sum_{n=1}^{q} \beta_n \varepsilon_{t-n}^2 \tag{2.20}$$

2.3.2 Model selection

It is not straightforward to determine the order of a time series model. First of all, the time series sample may very well be non-ideal, making it necessary to make qualitative judgements. Second, if the time series originates from a mixed process, the inclusion of a new model into an excisting description, e.g. by adding a moving average into a auto-regressive model, may infer with the original model and thus invalidate its basis. Third, overparameterizing will, in addition to the model being needlessly complicated, make the model more receptive to noise.

An important tool when conducting a qualitative analysis of the time series is the autocorrelation function (ACF), which describes the series' autocorrelation as a function of lags:

$$r(\tau) = \frac{c(\tau)}{c(0)} = \frac{\frac{1}{T}\sum_{t=\tau+1}^{T}(y_t - \bar{y})(y_{t-\tau} - \bar{y})}{\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2}, \quad \tau = 1, 2, \ldots \tag{2.21}$$

A graphical represenatation of the ACF can illuminate the characteristic properties of a sample. First, it immidately detects if the sample has any serial correlation at all. Secondly, it can also be used to detect the order of the series' process. For pure AR processes, the function will gradually tail off. This makes it hard to see whether y_t actually are linearly dependent on its lagged values, or if the correlation simply propagates backwards. For a pure MA process, however, the ACF will exhibit a characteristic cut-off for $\tau > q$, making it easy to decide its order.

A closely related concept is the partial autocorrelation function, or PACF. The function treats the time series as if it originated from an AR process, and plots the estimated *parameters* for such a model. For every lag, τ, a new model, AR(τ), is assumed and its parameters estimated. The last parameter in the model, φ_τ, is extracted, and is plotted against the lag, τ. In this way one can attain the partial correlation between the different observations in the series. The properties of the PACF are reversed from the ACF; it shows a clear cut-off if the series is autoregressive, while tailing off when examining an MA process.

If the time series originates from a mixed process, or the process in question is non-ideal, the model selection becomes more intricate. Nonetheless, a thorough use of these two functions in combination can bring to light a process' most apparent qualitative properties.

2.3.3 A time series model of the German spot price

When developing a time series model, the first thing to consider is the time series itself. In this case, the sample is to be constructed on the basis of the deviations between the expected spot price, i.e. the HPFC, and the actual spot price. In doing this, several options are possible. The first time series to be tested in this thesis were a log-normal transform of the series on the form

$$y_t = ln\frac{P_t}{f_t} \tag{2.22}$$

The choice were motivated by the anticipation of the series being normally distributed. However, this representation proved to be unfavorouble for several reasons. First of all it requires P_t and f_t to have equal sign, which is not always the case. Furthermore, since the inverse transform of the series can only yield positive values, a realistic simulation of the German spot price, which at times is negative, is unobtainable. Regardless of these minor issues, what rendered the series inadequate were the excessive price fluctuations yielded under simulation, caused by the exponential inverse transform. Figure 2.11 shows the outcome of such a simulation. The extreme variations

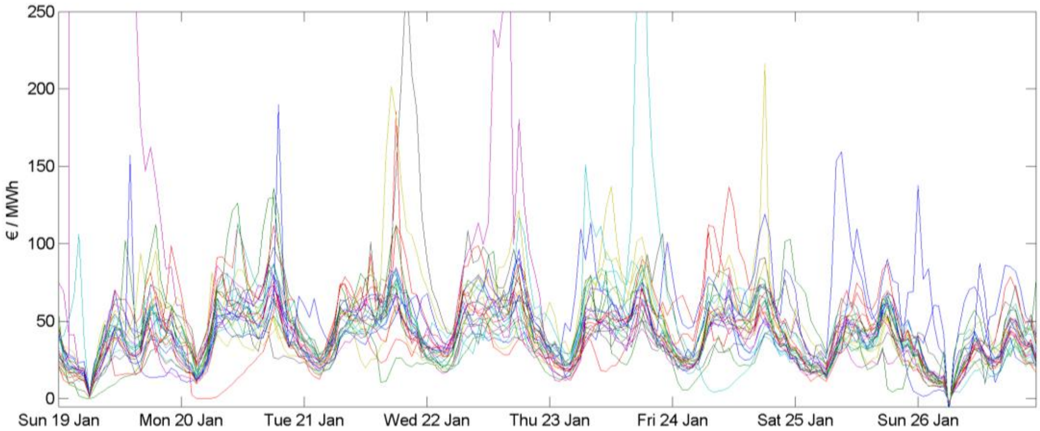


Figure 2.11: Simulated price scenarios for the third week in January. The simulations are made with a model built on the series $y_t = \ln(P_t) - \ln(f_t)$, and exhibit extensive price fluctations.

further made the power plant optimization unrealistic, especially with regards to the economic results. These concerns motivated a new time series description.

The second choice to be examined, were the same relationship as in Eq. 2.22, but without the log transform, so that

$$y_t = \frac{P_t}{f_t} \tag{2.23}$$

This option still entails the same important quality as the former, namely that price fluctuation are treated relative to the price's absolute value. A simulation of a month with a high price level, e.g. January, will then be more likely to give higher prices than a month with lower prices, e.g. July, which is a realistic characteristic. However, also this option proved to fail on the same part as the former; the simulation yielded prices that had, though not that excessive as the former, still too much fluctuations.

Since the previous alternative did not solve the problem of price variations, a new alternative was formulated; this time as a mere difference:

$$y_t = P_t - f_t \tag{2.24}$$

As suggested above, this series lacks the quality of producing simulations that varies according to the price level. However, the series showed to give much more realistic price scenarios.

With a time series in place, the next step is to choose a suitable model description. As explained in the above section, the autocorrelation function (ACF) is central at this stage. Figure 2.12 shows the ACF for the first 30 lags of the series in Eq. 2.24. The figure underlines the significant correlation within the series. Hence, a time series model is demonstrably well suited for the modelling. However, it is difficult to discern whether the series actually are serially correlated for so many lags as the figure indicates, or if the correlation merely propagates through the lags. At any rate, the series does not

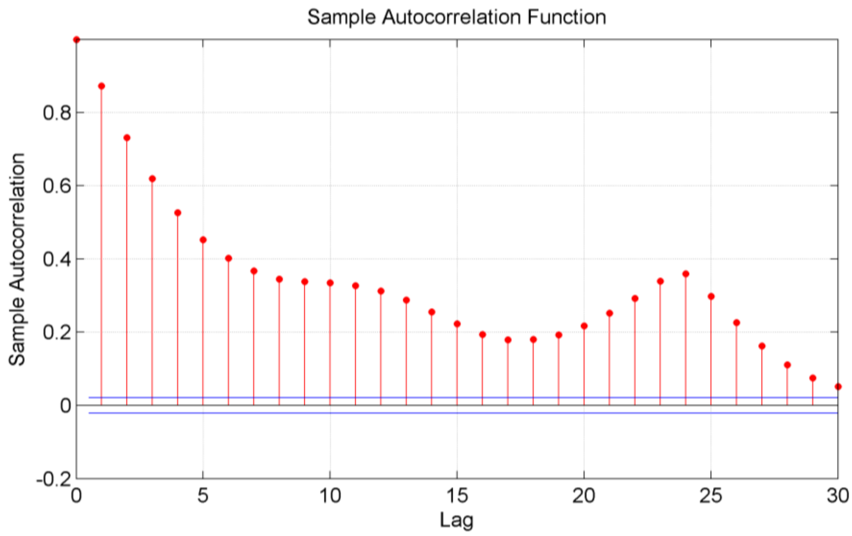


Figure 2.12: ACF plot of P_t-f_t.

come from a pure MA process, as found in many other financial applications (e.g. Brownian motion).

To better understand the nature of the process, the partial autocorrelation function (PACF) is plotted in Figure 2.13. In examining the plot, it becomes evident that the correlation is only really significant for the two first lags; the function shows a clear cut-off after the second lag. As noted in the previous section, this points towards an autoregressive process. While partial correlations for lags greater than two are not significant, correlations for 25 hours back in time are apparent. However, it is wise to model this behaviour after a more basic model is established, so as to avoid inference between the parameters [8]. For this reason an AR(2) model is constructed as the point of departure.

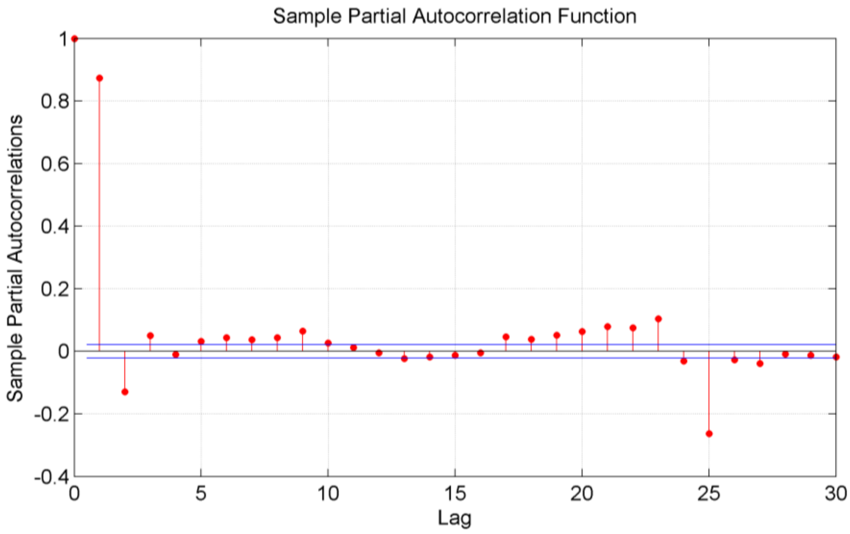


Figure 2.13: PACF plot of P_t-f_t.

Figure 2.14 shows the ACF plot of the inferred residuals of the AR(2) model. As the figure illustrates, practically all of the initial autocorrelation have been accounted for by the simple model. Indeed, most correlations lie within two standard deviations

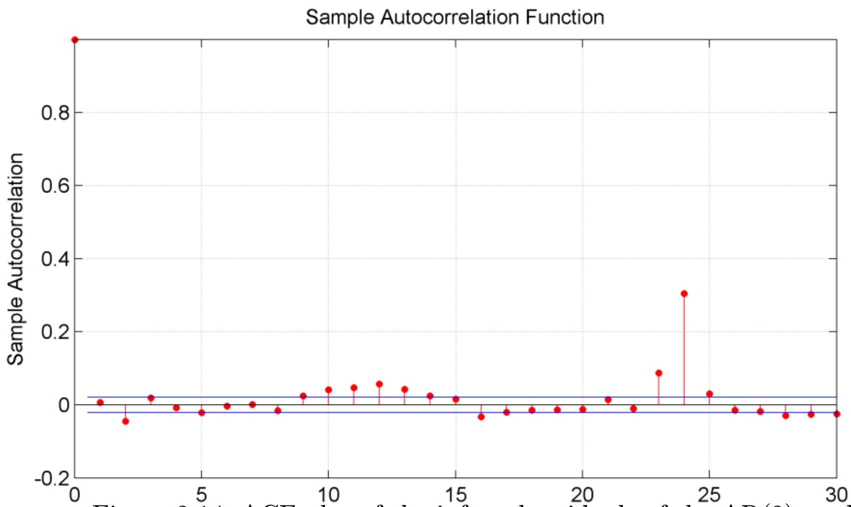


Figure 2.14: ACF plot of the inferred residuals of the AR(2) model.

of the mean (a 95 % confidence interval is marked with blue lines on the figure). Furthermore, the autocorrelation is now most significant for a lag of 24 hours. This is a more reasonable result than the 25 hour lag indicated above, and thereby underpins the principle of building a model successively. When this feature is included into an multiplicative autoregressive model on the form AR(2)x(24), very much of the series autocorrelation is accounted for. As can be seen from Figure 2.15, some diurnal correlation remains. However, including a parameter for lag 23 makes the model more vulnerable to noice, without gaining any significantly better fit.

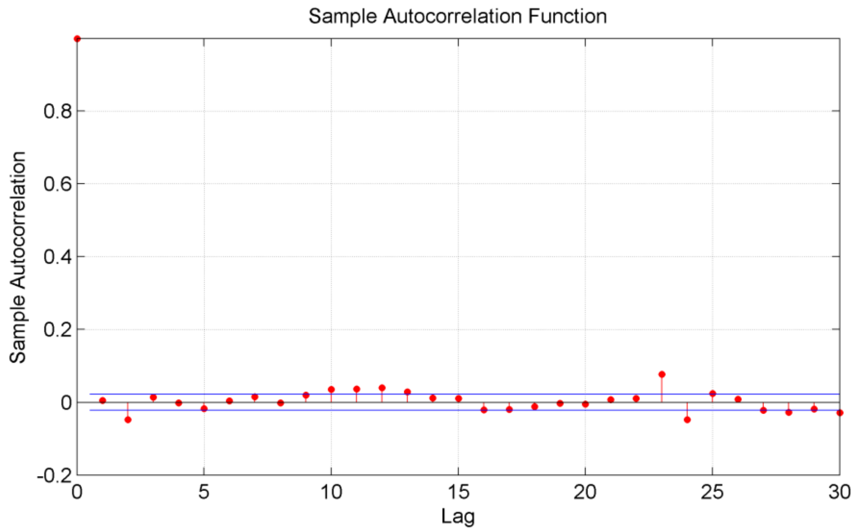


Figure 2.15: ACF plot of the inferred residuals of the AR(2)x(24) model.

While most of the serial correlation is accounted for, the series may still encompass serial dependency. In Figure 2.15, the autocorrelation of the squared residuals are depicted. If the innovation term's variance is conditional on earlier innovations, the squared residuals will exhibit autocorrelation. Provided that serial correlation allready is well described, the ACF plot will provide an unformal justification for asserting a

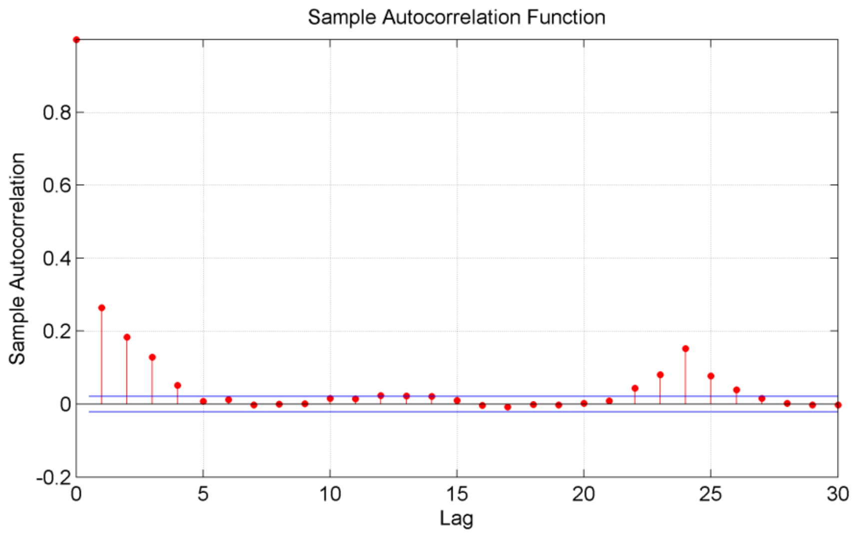


Figure 2.16: ACF plot of the squared residuals of the AR(2)x(24) model.

GARCH process for the innovation term. This is clearly indicated in the above plot, where the correlation between lagged squared residuals are significant.

A more formal test of conditional variance is the Engles ARCH test. The null hypothesis of no ARCH effects is soundly recected in favour of an ARCH(2) model. This model has the same local properties as a GARCH(1,1), and the latter is preferred because of greater flexibility. The Akaike information criterion (AIC) underpins this description; a decrease in value from $5{,}13 \cdot 10^4$ to $4{,}92 \cdot 10^4$ is obtained when a GARCH(1,1) models is compared with the pure AR model. It is more complex to make a qualitative assessment of the GARCH description than with an ARMA model, however, many studies points towards the adequacy of the GARCH(1,1) model and it is therefore retained on this basis [9].

The initial assumption of a time series model is that its innovation terms are independent and normally distributed. As became clear in the previous paragraph, the innovation process are clearly not independent. The next question that arises is if the normality assumption holds. By infering the conditional variances, modelled after a GARCH(1,1) process, the standardized residuals can be found as follows for the i'th observation:

$$st_i = \frac{r_i}{\sigma_i} \tag{2.25}$$

where r_i denotes the raw residuals inferred from the AR(2)x(24) process. A standard normal distribution would require to have 95 per cent of its innovations within an absolute value of two, and virtually all within an absolute value of three. In Figure 2.17 however, where the standardized residuals are plotted, a large number of the observations lie outside this range. This is in conflict with the assumption of normality, and thereby motivated the use of a Student's-t distribution for the innovation process.

When simulated with the Student's-t distribution, its shaping parameter, the degree of freedom (ν), was settled at 3,7, thus indicating a considerable departure from

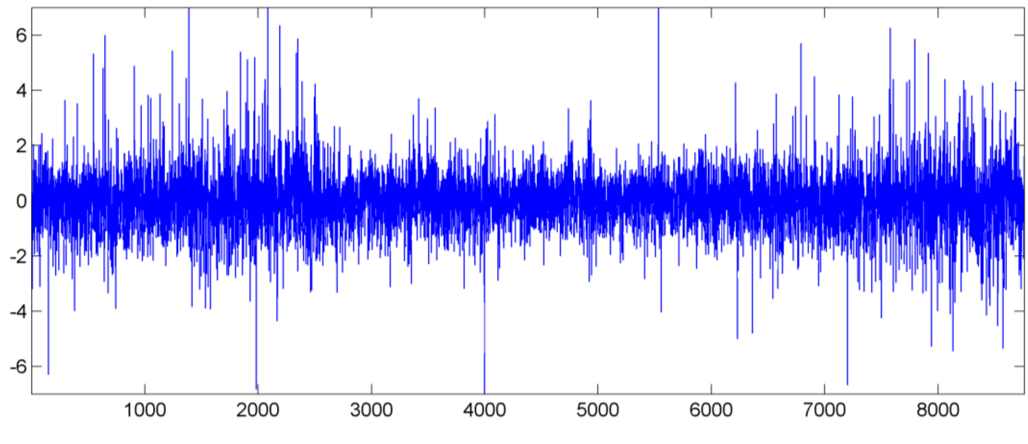


Figure 2.17: Standardized residuals inferred from the AR(2)x(24) /GARCH(1,1) model.

normality. The qualitative assessment was further supported by the AIC, which yielded an additional decrease from $4{,}92 \cdot 10^4$ to $4{,}78 \cdot 10^4$.

With this last modification, the final model of the time series could be written on the form:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-24} + z_t \sqrt{\kappa + \theta_1 \sigma_{t-1}^2 + \theta_2 \varepsilon_{t-1}^2} \tag{2.26}$$

where z_t is an i.i.d. random variable from the Student's t distribution. The parameter's value, along with their t-statistic, are listed in Table 2.4. As can be seen from the t-statistics, all parameters are statistically significant.

Table 2.4: Parameter values of the time series model

Parameters	Value	t-statistic
c	0, 1602	4, 88
φ_1	1, 0352	90, 02
φ_2	-0, 1463	-13, 12
φ_3	0, 2761	37, 59
κ	7, 67657	13, 88
α_1	0, 2596	8, 33
β_1	0, 4278	12, 29
ν	3, 6573	22, 99

The stochastic model developed in the previous sections have been used to simulate 25 different price scenarios. Three weeks in January 2014 (from January 13 to February 2) have been chosen as the simulation period. Furthermore, the last futures adjustment has been done on the preceding Friday, such that the performance of the price model is tested coincidentally. Figure 2.16 depicts the 25 simulated scenarios together with the HPFC and spot price.

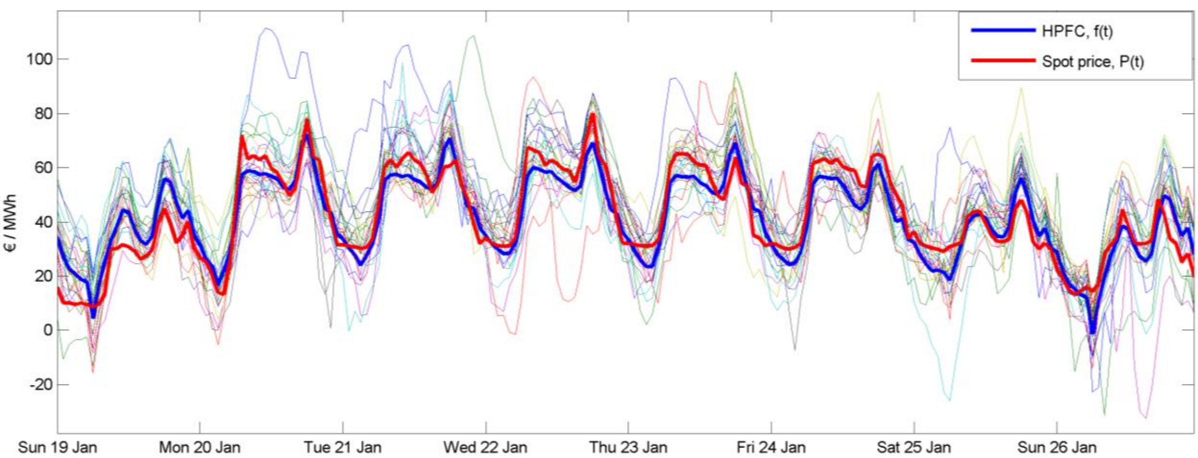


Figure 2.18: Simulated price scenarios plotted together with the HPFC and spot price.

2.4 A Norwegian pumped storage power plant

2.4.1 Building a Norwegian pumped storage power plant

This section will consider the modelling of a Norwegian pumped storage power plant. The approach draws on earlier work on hydro power scheduling [10]. However, for the purpose of lucidity, some simplifications have been made on the model used here. Constraints on production, such as start-up costs and environmental concerns, and non-linearities have not been adressed. Moreover, mean values have been used with regards to plant head and efficiency. The properties and constraints of power plant are summarized in table 2.5.

Table 2.5: Properties and constraints of the pumped storage power plant

Property	Value	
Upper reservoir initial /end level	1 000 000 m³	
Lower reservoir initial /end level	1 000 000 m³	
Plant head	100 m	
Efficiency power station /pump	0,84 /0,90	
Constraint	**Lower constraint**	**Upper constraint**
Upper reservoir	0 m³	2 000 000 m³
Lower reservoir	0 m³	2 000 000 m³
Discharge capacity	0 m³/h	100 000 m³/h
Relift capacity	0 m³/h	40 000 m³/h

The values of the properties and constraints has been chosen such that they resemble a medium sized Norwegian pumped storage power plant. In particular, the efficiency parameters are such that the combined efficiency of the cycle approximates 0,75 (B. Mo, Sintef Energy Research, conversation, May 21 2014).

Optimal scheduling of water releases can be formulated as a transportation problem [11], where water can be seen as a commodity 'sent' from one time period to the next. This enables both the reservoir levels and the upper and lower reservoir to be coupled through the time steps. The conditions can be formulated as follows:

Upper reservoir:

$$X_{upper\,res,t} = X_{upper\,res,t-1} - X_{discharge\,upper,t} + X_{relift,t} \tag{2.26}$$

Lower reservoir:

$$X_{lower\,res,t} = X_{lower\,res,t-1} + X_{discharge\,upper,t} - X_{relift,t} - X_{discharge\,lower,t} \tag{2.27}$$

Production:

$$X_{prod\,upper,t} = X_{discharge\,upper,t} \cdot \eta_{prod} \tag{2.28}$$

$$X_{prod\,lower,t} = X_{discharge\,lower,t} \cdot \eta_{prod} \tag{2.29}$$

Pumping:

$$X_{pump,t} = X_{relift,t} \cdot \eta_{pump} \tag{2.30}$$

Since the cost function and all the constraints are linear on the solution vector, the optimal allocation of the power plant can be found through linear programming:

$$max \begin{bmatrix} 0 \\ \vdots \\ \boldsymbol{P_t} \\ -\boldsymbol{P_t} \end{bmatrix}^T \boldsymbol{x}$$

$$s.t.\ \boldsymbol{Ax} = \boldsymbol{b}$$
$$and\ \boldsymbol{x} \geq \boldsymbol{0}$$

The objective function is formulated such that both electricity sold (by production) and bought (for pumping) is traded on the spot market without transaction costs. The constraints from Eq. 2.26 through 2.30 is accounted for by the A-matrix, while the b-vector is an empty matrix. Additionally, initial and end levels of the reservoir are retained by imposing upper and lower bounds on these variables. More detailed description of the optimization problem can be found in Appendix B. The problem is solved in Matlab by an interior point method.

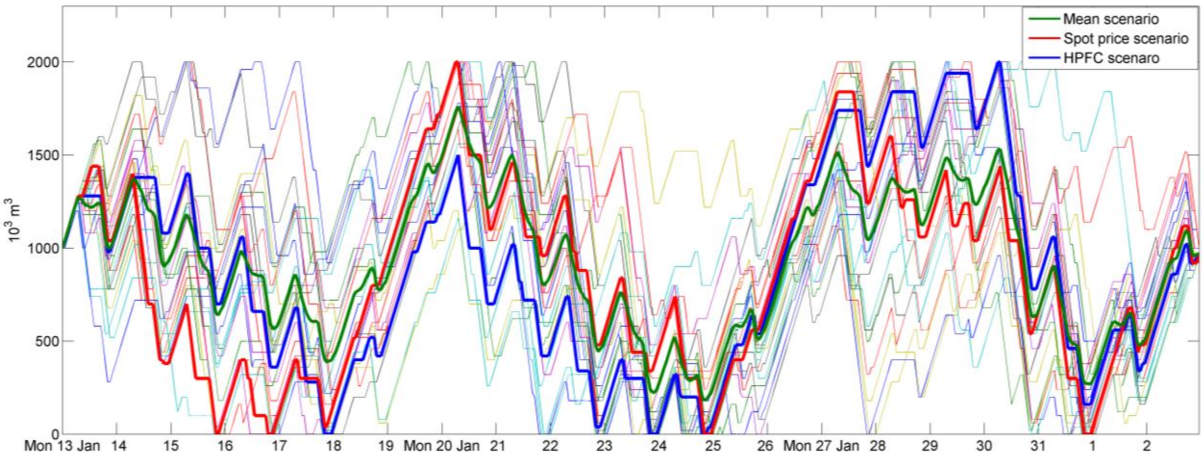
2.4.2 Scheduling in a European price regime

The simulated price scenarios from the previous chapter serves as a deterministic, fixed input for the modelled power plant. It is thereby assumed that all information regarding the price is known on beforehand. This is of course a gross simplification. Nonetheless, the assumption renders it possible to make rigorous assessments on the adequacy of the price model, and moreover, the economic potential of Norwegian pumped storage.

One of the most interesting aspect, especially from a production planning perspective, is the scheduling of water in the reservoirs. For this purpose, the upper reservoir content profiles, corresponding to each of the simulated price scenarios, are depicted in Figure 2.19. The mean of these profiles are plotted as well. Aditionally, the optimal production profile for the spot price (also assumed to be known in advance) and the HPFC scenario are illustrated in the figure.

It is clear from the figure that the content profiles corresponding to the price scenarios cover a wide solution space. However, the mean of these scenarios are in close proximity to the spot price scenario. In fact, the mean scenario actually performs better than the HPFC scenario, on which the simulations are based upon. So while the mean *price* simulations are approximately equal to the HPFC, the mean *content* scenario diverges considerably from the HPFC scenario. In other words there is a non-linear relation between the price simulations and content scenarios. This extends the usage of the price model. While the HPFC can make a good forecast of the spot price, the stochastic component contributes to a more accurate scheduling of the reservoirs.

The economic results for the simulation period shows, as suggested by Figure 2.18, a significant spread. Ranging from 16 000 € to 29 500 € a week, the simulations yielded a mean of 21 400 €. The average weekly return for the actual spot price scenario were 16 800 €, and thus clearly in the lower end. As the price scenarios are variations around the HPFC, the scenarios are expected to yield greater revenues. Hence, they clearly demonstrates the importance of price variations.



Figur 2.19: Reservoir content for the 25 price scenario plotted together with the mean scenario, and the spot price and HPFC scenario.

3 Discussion and further work

This thesis has sought to explore the opportunities for Norwegian pumped storage in a European pricing regime. Its main emphasis has therefore been on the modelling of the German spot price, so as to provide a sound basis for the power plant. The point of departure has been historical price records, and through linear regression, a mean price curve with seasonal properties could be established. Furthermore, this curve has been adjusted to give an updated curve more in align with the market expectations. The model has thus combined information from the past with expectations of the future in order to make price predictions. Moreover, the adjusted price curve has made it possible to develop a stochastic price model, which in turns, has been the basis for the power plant analysis. The following paragraphs will provide an assessment of the results from the previous chapter, as well as making suggestions for further work.

The shape curve established in chapter 2.1 is crucial for the following chapters. First of all, the curve's goodness of fit is imperative for the accuracy of the final adjusted curve; without a realistic description of the seasonality of the price, the futures adjustments will be in vain. Second, the seasonality, both daily and weekly, of the shape curve, is decisive in a pumped storage setting. Since the power plant takes advantage of differences in price during the day and over the week, the shape of the curve will be a direct contributor to the plant's profitability. Therefore, care has been taken so that the shape curve could give an accurate description of the price. Decisions regarding the scope of the explanatory variables have been taken as a trade-off between accuracy and sufficient sample size. This trade-off were important when the daily regression model were made. On the one hand, a breakdown into many classes are obviously desirable, on the other, the sample of each hour within a class could not be too small. The choice were made on four classes of the year, amounting to 65 observation per class. This led to a robust fitting, but also had as a consequence that the daily price patterns at times, especially at the beginning and end of a period, were inaccurate. This issue could be solved by dividing the year into monthly classes and compensate for the reduced sample by making a more qualitative assessment, e.g. by imposing certain rules for outliers.

With regards to the yearly regression model, a qualitative judgement were made with respect to the 'monthly' variables, especially concerning the vacation periods. However, the same holds here; en even finer breakdown of the year would clearly be advantageous. A remedy in this case could be to include a sinusoidal function in the regression model, and thus avoid the discontinuity between the periods. However, this part of the model is not that crucial, since the futures adjustment can accommodate for these inaccuracies.

The second part of the price modelling involved adjusting the shape curve to market expectations. As explained in chapter 2.2, the shortest time window with sufficient liquidity were a week. This obviously restrained the precision of the adjustment. For example, as illustrated in Figure 2.8, the shape curve might need adjustments for some days within the week, while describing the price well for others, nevertheless leading the adjustment function to alter the price level for the whole week. The validity of hourly price forward curves will unquestionably be significantly increased when futures with daily maturities are traded large-scale. Nevertheless, an increase in explained variance by 16 per cent is a considerable improvement. Hence, including futures adjustments in future work on this field, even with the same conditions as today, can clearly be recommended.

The stochastic price description were modelled on the basis of the spot price's deviations from the adjusted shape curve. This allowed for price simulations around the HPFC. As outlined in chapter 2.3, the greater part of the autocorrelation, both for the series itself and its variance, were accounted for by an ARMA/GARCH description. However, the normality assumption seemed unjustified when examining the standardized residuals. A Student's t distribution were introduced for the innovation process so that the heavy kurtosis of the residuals could better be explained. This improved the goodness of fit according to the Akaike information criterion. On the other hand, the simulations showed a fair amount of price fluctuations; when optimized in the pumped storage setting, these price scenarios yielded considerably more revenues than the actual spot price. The issue here is twofold. First, an HPFC with more detailed information would have followed the spot price more accurately. This in turn would have led to smaller residuals and, consequently, to less variance in the GARCH model. Second, the heavy tailed distribution of the residuals obliges the Student t's distribution to take a low value for its shaping parameter, thus giving rise to innovations that are unwarranted. For these reasons, the main challenge for further work will be to find an adjusted shape curve that better fits with the spot price. Furthermore, alternatives for the distribution of the innovation terms should be explored. In this context however, the variations within the price scenarios cast light upon the opportunities for Norwegian pumped storage in a more volatile pricing regime.

In developing the pumped storage power plant model, it was sought to include typical parameters found at a Norwegian hydropower plant. However, some simplifications were made, especially with regards to start-up costs. The latter leads to a high utilization of the plant, evident from the reservoir optimization in chapter 2.4. Nonetheless, the simulation showed reasonable results; the daily and weekly production patterns were especially noticeable. Furthermore, the hourly price forward curve yielded a sensible production pattern. In particular, the mean of the simulated price scenarios were very close to the optimal production plan. In this way, the stochastic price model could possibly be a simple, yet powerful tool in production planning. Further work might include simulating over different periods of the year and adding new constraints to the established power plant model.

References

[1] Blöchlinger, L. A Regime-Switching Spot/Forward Price Model with Kim Filter Estimation. PhD thesis, University of St. Gallen; 2008.

[2] EPEX SPOT SE: Auction. http://www.epexspot.com/en/market-data/auction.

[3] Fleten S-E, Paraschiv F, Schürle M. Spot-Forward Model for Electricity Prices. Working Papers on Finance 1311, St. Gallen; 2013.

[4] Adams KJ, van Deventer DR. Fitting yield curves and forward rate curves with maximum smoothness. Journal of Fixed Income 1994;4:52-62.

[5] Benth FE, Koekebakker S, Ollmar F. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variations. The Journal of Derivatives 2007;15(1):52-66.

[6] EEX AG: Phelix Power Futures. https://www.eex.com/en/market-data/power/derivatives-market/phelix-futures.

[7] Gershenfeld N. The Nature of Mathematical Modeling, Cambridge University Press; 1998.

[8] Harvey AC. Time Series Models. Oxford: Philip Allen Publishers Limited; 1981.

[9] Hansen PR, Lunde A. A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)? J. Appl. Econ. 20: 873–889; 2005.

[10]Fosso OB, Belsnes M, Røynstrand J, Gjengedal T, Valhovd E. Hydro power short-term scheduling in an online environment. MedPower; 2003.

[11]Williams HP. Model Building in Mathematical Programming. Belfast: John Wiley & Sons; 1985.

Appendix A

Source code for the second regression model.

```
hours2009_2013=zeros(24*(365+365+365+366+365),24*7*4);
hours_day=dummyvar(1:24);

date=datenum('01-Jan-2009')-1;

for i=1:(365*4+366)
    date=date+1;
    [year month]=datevec(date);
    weekday=weekday(date);

    if month==12 || month==1 || month==2

        if weekday==2
            hours2009_2013(1+24*(i-1):24*i,1:24)=hours_day;
        elseif weekday==3
            hours2009_2013(1+24*(i-1):24*i,25:48)=hours_day;
        elseif weekday==4
            hours2009_2013(1+24*(i-1):24*i,49:72)=hours_day;
        elseif weekday==5
            hours2009_2013(1+24*(i-1):24*i,73:96)=hours_day;
        elseif weekday==6
            hours2009_2013(1+24*(i-1):24*i,97:120)=hours_day;
        elseif weekday==7
            hours2009_2013(1+24*(i-1):24*i,121:144)=hours_day;
        else weekday==1;
            hours2009_2013(1+24*(i-1):24*i,145:168)=hours_day;
        end

    elseif month==3 || month==4 || month==5

        if weekday==2
            hours2009_2013(1+24*(i-1):24*i,169:192)=hours_day;
        elseif weekday==3
            hours2009_2013(1+24*(i-1):24*i,193:216)=hours_day;
        elseif weekday==4
            hours2009_2013(1+24*(i-1):24*i,217:240)=hours_day;
        elseif weekday==5
            hours2009_2013(1+24*(i-1):24*i,241:264)=hours_day;
        elseif weekday==6
            hours2009_2013(1+24*(i-1):24*i,265:288)=hours_day;
        elseif weekday==7
```

```
            hours2009_2013(1+24*(i-1):24*i,289:312)=hours_day;
        else weekday==1;
            hours2009_2013(1+24*(i-1):24*i,313:336)=hours_day;
        end

    elseif month==6 || month==7 || month==8

        if weekday==2
            hours2009_2013(1+24*(i-1):24*i,337:360)=hours_day;
        elseif weekday==3
            hours2009_2013(1+24*(i-1):24*i,361:384)=hours_day;
        elseif weekday==4
            hours2009_2013(1+24*(i-1):24*i,385:408)=hours_day;
        elseif weekday==5
            hours2009_2013(1+24*(i-1):24*i,409:432)=hours_day;
        elseif weekday==6
            hours2009_2013(1+24*(i-1):24*i,433:456)=hours_day;
        elseif weekday==7
            hours2009_2013(1+24*(i-1):24*i,457:480)=hours_day;
        else weekday==1;
            hours2009_2013(1+24*(i-1):24*i,481:504)=hours_day;
        end

    else month==9 || month==10 || month==11;

        if weekday==2
            hours2009_2013(1+24*(i-1):24*i,505:528)=hours_day;
        elseif weekday==3
            hours2009_2013(1+24*(i-1):24*i,529:552)=hours_day;
        elseif weekday==4
            hours2009_2013(1+24*(i-1):24*i,553:576)=hours_day;
        elseif weekday==5
            hours2009_2013(1+24*(i-1):24*i,577:600)=hours_day;
        elseif weekday==6
            hours2009_2013(1+24*(i-1):24*i,601:624)=hours_day;
        elseif weekday==7
            hours2009_2013(1+24*(i-1):24*i,625:648)=hours_day;
        else weekday==1;
            hours2009_2013(1+24*(i-1):24*i,649:672)=hours_day;
        end
    end
end
```

Appendix B

Source code for the pumped storage power plant model.

```
function [revenues, production, pump, reservoir_upper, reservoir_lower, sum_profit,
X] = pump_power_plant(initial_reservoir_upper, end_reservoir_upper,
initial_reservoir_lower, end_reservoir_lower, max_discharge, efficiency_prod,
max_pump, efficiency_pump, spot_price)

[hours, profiles]=size(spot_price);



%%Establishing matrices:
A_reservoir_balance1=zeros(hours,6*hours+2);
A_reservoir_balance2=zeros(hours,6*hours+2);
A_prod_balance=zeros(hours*2,6*hours+2);

head=100;
energy_eq_prod=9.81/3600*head*efficiency_prod;
energy_eq_pump=9.81/3600*head/efficiency_pump;

for i=1:hours

    A_reservoir_balance1(i,1+3*(i-1):1+3*i) = [1, -1, 1, -1];

    A_reservoir_balance2(i,2+3*(i-1):3*i) = [1, -1];
    A_reservoir_balance2(i,(1+3*hours)+i:(1+3*hours)+(i+1))=[1, -1];

    A_prod_balance(i,2+3*(i-1))=energy_eq_prod;
    A_prod_balance(i,(2+4*hours)+i)=-1;
    A_prod_balance(hours+i,3*i)=energy_eq_pump;
    A_prod_balance(hours+i,(2+5*hours)+i)=-1;

end

A=[A_reservoir_balance1;A_reservoir_balance2;A_prod_balance];

b=zeros(4*hours,1);

%%Constraints on X:
lb=zeros(6*hours+2,1);
lb(1)=initial_reservoir_upper;
lb(2+3*hours)=initial_reservoir_lower;

ub=inf(6*hours+2,1);
```

```
ub(1+3*hours)=end_reservoir_upper;
ub(2+4*hours)=end_reservoir_lower;
for i=1:hours
    ub(2+3*(i-1))=max_discharge;
    ub(3*i)=max_pump;
end

%%Linear programming:
F=zeros(6*hours+2,profiles);
X=zeros(6*hours+2,profiles);
for i=1:profiles
    F(4*hours+3:5*hours+2,i)=spot_price(:,i);
    F(5*hours+3:end,i)=-spot_price(:,i);
    X(:,i)=linprog(-F(:,i),[],[],A,b,lb,ub);
end

%%Output variables:
reservoir_upper=zeros(hours+1,profiles);
reservoir_lower=zeros(hours+1,profiles);
for i=1:profiles
    for j=1:hours+1
        reservoir_upper(j,i)=X(1+3*(j-1),i);
        reservoir_lower(j,i)=X(1+3*hours+j,i);
    end
end

revenues=zeros(hours,profiles);
for i=1:profiles
    for j=1:hours
        revenues(j,i)=(X(2+4*hours+j,i)-X(2+5*hours+j,i))*spot_price(j,i);
    end
end

production=X(4*hours+3:5*hours+2,:);

pump=X(5*hours+3:end,:);

sum_profit=sum(F.*X);

end
```