NTNU
Norwegian University of
Science and Technology

Faculty of Social Sciences
and Technology Management
Department of Industrial Economics
and Technology Management

MASTER THESIS

for

STUD.TECHN CAMILLA WESTAD KOLSRUD

Field of study	**Financial Engineering** Investering, finans og økonomistyring
Start date	15.01.2010
Title	**Reservoir Hydropower: The Value of Flexibility** Vannkraftverk med magasin: Betydning av fleksibilitet
Purpose	Provide insights in how storage flexibility impacts the expected revenues of hydropower plants. Analyse the importance of the complexity of price models.

Main contents:

1. Conduct an empirical analysis of the different factors affecting flexibility and how they impact the ability of hydropower producers to exploit high prices.

2. Apply insights from empirical analysis in a general valuation model for hydropower plants. Analyse the impact of flexibility and the level of complexity of price models on expected revenues.

Monica Rolfsen
Deputy Head of Department

Stein-Erik Fleten
Supervisor

MASTER THESIS

for

STUD.TECHN MARTIN PROKOSCH

Field of study	**Financial Engineering** Investering, finans og økonomistyring
Start date	15.01.2010
Title	**Reservoir Hydropower: The Value of Flexibility** Vannkraftverk med magasin: Betydning av fleksibilitet
Purpose	Provide insights in how storage flexibility impacts the expected revenues of hydropower plants. Analyse the importance of the complexity of price models.

Main contents:

1. Conduct an empirical analysis of the different factors affecting flexibility and how they impact the ability of hydropower producers to exploit high prices.

2. Apply insights from empirical analysis in a general valuation model for hydropower plants. Analyse the impact of flexibility and the level of complexity of price models on expected revenues.

Monica Rolfsen
Deputy Head of Department

Stein-Erik Fleten
Supervisor

DECLARATION

Stud.techn.   CAMILLA KOLSRUD

Institute of Industrial Economics and Technology Management

I hereby declare that I have written the above mentioned
thesis without any kind of illegal assistance.

TRONDHEIM   04.06.2010

Place   Date

Signature

DECLARATION

Stud.techn.   **MARTIN PROKOSCH**

Institute of Industrial Economics and Technology Management

I hereby declare that I have written the above mentioned
thesis without any kind of illegal assistance.

TRONDHEIM   04.06.2010
Place   Date

Signature

Preface

This master thesis was conducted at the Norwegian University of Science and Technology (NTNU), Department of Industrial Economics and Technology Management. Our work falls within the Group of Financial Engineering.

We would specially like to thank our supervisor, Professor Stein-Erik Fleten for helpful assistance and valuable discussions. Last but not least, we thank the fourteen anonymous Norwegian hydropower producers who have contributed with data. Without their contribution this thesis would not have been possible.

Trondheim, June 8, 2010

_____   _____

Camilla Westad Kolsrud   Martin Prokosch

Reservoir Hydropower:
The value of flexibility

Camilla Westad Kolsrud
Martin Prokosch

Master thesis TIØ4900
Norwegian University of Science and Technology

June 4, 2010

Abstract

The main goal of this study is to quantify the impact of long term storage flexibility on potential revenues for hydropower stations. Using empirical data from 14 Norwegian producers, we find that the value of storage flexibility on average accounts for 22 % of the actual revenues, ranging as high as 40 % for power stations with the highest relative regulation. With the motivation to quantify the impact of factors driving flexibility (relative regulation, degree of inflow seasonality and capacity factor) on revenues, we build a complete stochastic model for hydropower stations with reservoir. This model includes a stochastic inflow model, taking into account autocorrelation and seasonality, and an application on of the geometric multi-factor spot price model proposed by Benth et al. [2008] on Nord Pool data. The correlation between water inflow and spot prices is incorporated by letting the mean level of the base component of the spot price depend on the deviation from normal aggregated reservoir content in Norway. We use a parameterization of the water value function proposed by Näsäkkälä and Keppo [2008] as one of two methods to simulate operator's dispatch strategy. By comparing the simulated results with the empirical results, we evaluate the overall performance of the stochastic model to be sufficient for our analysis.

By conducting a sensitivity analysis with regards to relative regulation, degree of inflow seasonality and capacity factor, we find that the value of storage flexibility increases with increasing degree of inflow seasonality. Given high degree of inflow seasonality, increasing relative regulation from 0.1 to 1.0 increases the revenue potential with as much as 35 %, while such an increase has no effect on revenues under low inflow seasonality. Further, an increase of production capacity, will reduce the need for (value of) storage flexibility.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The main focus of this study is to understand how the of degree flexibility impacts revenues for hydropower stations with reservoir. According to IHA [2010], 16 % of the global electricity supply is generated from hydropower, making it the biggest, and according to REN21 [2009] the fastest growing, source of renewable energy in the world. The ability to store electricity in form of water in reservoirs, and the capability to quickly adjust generation output, makes hydropower a versatile source of electricity for both base and peak load. In this way hydroelectricity from power stations with reservoir differs substantially from the two other fastest growing renewable energy sources, wind and solar (REN21 [2009]). For both these sources, the generation output is directly dependent on uncontrollable exogenous conditions; wind speed and sunlight. Hence, with the increase of electricity generation from these renewable sources, the flexibility of hydropower will become even more important in order to ensure security of supply in the years to come. According to the Nordic power exchange Nord Pool [2010], 57 % of the total electricity in the Nordic countries is generated from hydropower. In Norway, which is one of the largest hydropower producers in the world, 99 % of the generated electricity (around 120 [TWh] yearly) is produced by hydropower stations. In this study we use data from 14 different Norwegian hydropower producers, consisting of inflow, reservoir and production time series from April 2000 to April 2009, to study the value of the ability to store water from the viewpoint of individual price taking power stations. To our knowledge there has not been conducted any similar empirical studies of hydropower with such an extensive dataset.

To understand the basics behind the hydropower production and the Nordic power market, a short introduction follows. Hydropower producers generate revenues by dispatching water from one or multiple reservoirs under the uncertainty of prices and future inflow. The flexibility of a power station depends on the size of this reservoir, the amount and distribution of inflow flowing into the reservoir and the production capacity. If the size of the reservoir is small relative to the annual inflow, the producer will be less flexible as it has to produce more often to avoid spillage. The amount of electricity generated by a hydropower plant is determined by the energy coefficient of the particular plant, which depends on the head (the difference in height between reservoir level and the turbine). The larger the energy coefficient (or equivalently the larger the head), the more energy per volume of water is generated. Every day the producers submit price dependent bids to the spot market on Nord Pool (Elspot) for the hourly production the following day. The demand and supply bids are aggregated to determine the equilibrium price, called the system price. Further, power derivatives are traded on Nord Pool financial market on working days, enabling the power producers to hedge against price uncertainty.

When analyzing the flexibility, it is necessary to distinguish between short term and long term flexibility. Having short term flexibility enables power stations to profit from hourly changes in prices by changing the dispatch of water from hour to hour. Long term flexibility, however, enables the producers to obtain higher prices by storing water throughout the season. We will focus on long term flexibility in this paper, with a time horizon of one season. The power stations profit from seasonal variations in demand (and hence prices) primarily driven by the need for heating during cold winter months in the Nordic region. Having negligible start-up costs and fast response times, the flexibility of hydropower stations is mainly limited by water availability and production capacity. This flexibility, defined as the degree of freedom available when establishing their dispatch plans, affects the revenue potential as it enables producers to benefit from volatile electricity prices by producing when the price is high and save water when the price is low. Using an analogy from financial theory, hydropower stations provide the operator with a series of American put options with very low strike prices (production costs). The number of options and their time to maturity depends on the inflow distribution, reservoir size and production capacity. This study focuses on how these factors influence the revenues for hydropower stations, the main goal being to isolate the value of storage flexibility.

There exists extensive literature analyzing the dynamics of storage assets. This ranges back to the classical "warehouse problem" introduced by Cahn [1948], who analyzes the optimal pattern of operation under seasonal price and cost variations. In more recent work, Secomandi [2010] studies the optimal management of commodity storage assets, with an application on natural gas storage. Oil and natural gas storage is the topic of several recent studies, as in papers by Boogert and de Jong [2008], Wu et al. [2010], Ludkovski and Carmona [2009], Thompson et al. [2009] and Bjerksund et al. [2008]. Depending on the region, gas prices can show seasonal trends similar to electricity prices, due to for example varying heating demand. Boogert and de Jong [2008] use a Monte Carlo method incorporating gas price dynamics and flexibility constraints, to quantify the value of natural gas storage assets. Wu et al. [2010] focus on the limited flexibility due to storage capacity and the maximum injection and withdrawal rates, and show how optimal revenues depend on these limits to flexibility in a three-period model. Thompson et al. [2009] apply a real option approach to derive differential equations for the value of natural gas storage assets, opening for the possibility of incorporating complex (continuous) spot price dynamics. In an earlier study, Thompson et al. [2004] use the same approach to value a pumped-storage hydropower facility, sharing many similarities with natural gas storage. Ludkovski and Carmona [2009] also show how storage flexibility affects the value function of natural gas storage, and how the same model can be applied to a pumped-storage asset. The main focus of papers on natural gas and hydropower pumped-storage is when to purchase and when to sell in order to profit from short and long term price variation. Traditional reservoir hydropower, being the focus of this study, deviates from pumped-storage in the way that the producers have no possibility to pump (inject) water (gas). The inflow into the reservoir is stochastic, which changes the problem dynamics.

The main focus of literature on traditional hydropower is on how to optimally dispatch the water available, taking into account the uncertainty of future inflow. The optimal production scheduling under these conditions has been studied by Fosso et al. [1999], which give an overview of the production scheduling problem and possible tools

in the long-, medium- and short term. The study of Näsäkkälä and Keppo [2008] also focuses on medium- and long term planning, proposing a parameterization for the optimal production strategy. For a more comprehensive review of the economic theory of hydropower scheduling, see Foersund [2007]. Fleten et al. [2009] conduct an empirical study related to the optimal dispatch on water, showing how financial information impacts of on the production strategy of producers based on empirical data. In this study we use findings from both literature on storage assets and literature on traditional hydropower scheduling to study the impacts on revenues of factors limiting flexibility. Following an approach similar to Wu et al. [2010], we define several flexibility cases describing different levels of storage flexibility and information availability. Using detailed empirical data, we provide insights on how the revenues vary with different degrees of flexibility. Looking at the power station's technical characteristics gives a preliminary picture on how reservoir size, inflow distribution and production capacity impact the revenues. Further, we isolate the value of storage flexibility, finding that on average this accounts for 22 % of the actual revenues for the power stations in our sample. To further quantify the impact of the factors, we build a comprehensive stochastic model for hydropower stations, enabling the arbitrarily change of any factor. This model consist of several components; an inflow model taking into account different forms of seasonality in the level and variance, a detailed model for Nord Pool spot prices using a multi-factor model, as described by Benth et al. [2008], with the estimation approach of Meyer-Brandis and Tankov [2008], and two production strategies for the dispatch of water. One of these dispatch models is an application of the threshold function method described by Näsäkkälä and Keppo [2008], expanded to incorporate findings from the project thesis by Kolsrud and Prokosch [2009]. The model components are estimated based on daily empirical data from the 14 power stations, Nord Pool electricity prices and aggregated reservoir data for Norway. The results from Monte Carlo simulations of the model dynamics are compared with the empirical results to review overall model performance. By running sensitivities, we quantify the impact of reservoir size, inflow distribution and production capacity on the revenue potential and the value of storage flexibility.

The layout of this study is as follows: In Chapter 2, we present the empirical data used in the preliminary analysis and estimation of the stochastic model. In Chapter 3, we introduce the framework with the different flexibility cases and define the value of storage flexibility. The components for the stochastic model , and how these are estimated, is presented in Chapter 4. In Chapter 5, we compare the model with empirical data and quantify the impact of different factor on revenues through sensitivity analysis. Chapter 6 concludes this thesis.

Chapter 2

Data description

The empirical analysis and estimations in this study are based on inflow, production and reservoir data for 14 Norwegian hydro power stations, electricity spot and future/forward prices and aggregated reservoir levels in Norway.

2.1 Power producer data

When selecting power stations to include in the analysis, we follow a number of selection criteria. As the underlying problem of this paper is to analyze the impact of flexibility of power stations, we want power stations that have a certain level of freedom in choosing when and how much to produce. The selection criteria are summarized below:

- **Independent:** The power station should be as independent of other stations as possible, meaning no coupling to other stations upstream of the reservoir or downstream of the generator. This is to avoid having to take into account interdependencies between power stations in the analysis and modeling. We also include only one power station per producer to further ensure independence.

- **Sufficient reservoir capacity:** Due to limited reservoir capacity, or lack thereof, run-of-river plants are significantly less flexible than power stations with dedicated reservoirs. In the analysis in Chapter 3 we argue that the run-of-river equivalent represents the value of having no long term flexibility, and use this equivalent as a basis to value flexibility. Hence no pure run-of-river plants are included in our sample.

- **Price taker:** We want to assume that the producer's production decisions have no influence on spot or derivatives prices. By excluding the largest producer in Norway, Statkraft, we ensure that none of the producers in the sample have more than 1.8 % market share, which limits their potential market power. The market share for the top 10 producers in Norway is listed in Table A.2 in the Appendix. By doing this we also ensure that changes in the reservoir levels of our power stations have small impact on the aggregated reservoir content in Norway.

Table 2.1 gives an overview of the power stations included in the sample with selected technical characteristics. The calculated values for relative regulation and capacity factor both affect the amount of flexibility of each producer. For each power station,

Table 2.1: Selected technical characteristics for each power stations in the sample. The average energy coefficient is defined as the average energy generated [kWh] per volume of water [m³]. Relative regulation is defined as reservoir capacity divided by average annual inflow. Capacity factor is average annual inflow divided by production capacity, here given in percentage of one year.

Power station	Price area	Energy coefficient	Production capacity	Reservoir capacity	Average inflow	Relative regulation	Capacity factor
		[kWh/m³]	[MW]	[MWh]	[MWh/yr]	[yr]	[%]
A	NO4	1.24	128	222 063	605 586	0.367	54
B	NO4	1.69	120	801 152	488 719	1.639	46
C	NO5	1.23	31	50 226	92 338	0.544	34
D	NO3	1.27	33	51 816	138 606	0.374	48
E	NO5	0.67	28	118 900	87 831	1.354	36
F	NO2	0.16	23	14 016	153 029	0.092	76
G	NO1	1.25	68	255 000	271 250	0.940	46
H	NO1	1.09	167	272 500	451 661	0.603	31
I	NO4	1.50	62	142 000	230 643	0.616	42
J	NO3	0.95	41	42 599	81 337	0.524	23
K	NO3	1.36	153	380 800	662 862	0.574	49
L	NO4	1.55	182	940 075	916 284	1.026	57
M	NO4	0.66	26	89 562	124 440	0.720	55
N	NO3	1.66	90	172 972	371 680	0.465	47

the factors are defined as follows:

$$Relative\ regulation = \frac{R^{max}}{\overline{I}^{tot}} \tag{2.1}$$

$$Capacity\ factor = \frac{\overline{I}^{tot}}{P^{max}} \tag{2.2}$$

where R^{max} is the power station's reservoir capacity, $\overline{I}^{tot}$ the average annual inflow and P^{max} the production capacity. Relative regulation should be interpreted as the amount of time it takes to fill up the reservoir with the average annual inflow, while the capacity factor is the share of time the producer needs to run at maximum capacity to produce all inflow in an average year. Note that a high production capacity corresponds to a low capacity factor.

Table 2.1 shows that the sample is well diversified both in terms of head (directly influencing the energy coefficient), production capacity, reservoir size and average annual inflow. The power stations are also geographically well spread, here seen by the Nord Pool price area. Power station F stands out with the smallest reservoir in the sample. This low reservoir size combined with a a large drainage basin causing high inflow, reduces F's relative regulation. In this way F is close to being a run-of-river plant, and hence has the lowest degree of storage flexibility in our sample. Regarding the rest of the power stations, B, E, G and L stand out with the highest relative regulation, and are expected to be able to obtain additional value by storing inflow.

For the inflow, reservoir and production data series we have data from season 2000/01 to season 2008/09, where each season is defined to start in week 17 and last throughout week 16 in the following year. Hence the sample consist of 9 years, 470 weeks and 3

290 days, starting from April 24th 2000 and the last day being April 19th 2009. The choice of using week 17 as start of the season is discussed in the reservoir data section below. Both production, inflow and reservoir data are here given in their energy equivalents [MWh], where inflow and reservoir volumes are converted using the average energy coefficient[1]. For power stations with multiple reservoirs, inflow and reservoir data are aggregated. We lack data from some seasons for a few power stations, as pointed out below.

2.1.1 Inflow data

The inflow data series show high measurement uncertainty, as they contain negative daily and/or hourly values for several power stations. Our hypothesis is that producers correct measuring errors during the last days/hours by registering negative inflow. To account for the uncertainty, we reduce the number of negative data points by aggregating daily/hourly inflow to weekly values. Even after doing so, some weeks still have negative inflow, and hence these are in our dataset set equal to zero. This method leads to a higher value for total inflow than when summing the original data series. However, the difference between the total seasonal inflow obtained by the aggregated data and the original data is relatively low for all producers; power station J has the largest average annual deviation of 7 %. Daily data points are obtained by evenly distributing the weekly inflow throughout the week. We will lose the daily variations in inflow by performing the weekly aggregation, but do on the other hand obtain more consistent values.

Table 2.2: Descriptive statistics for weekly inflow for each power station. Note: Negative inflow values were set equal to zero.

Power station	Min	Max	Weekly stdv	Seasonal stdv
	[MWh]	[MWh]	[%]	[%]
A	0	74 882	105	15
B	0	63 533	119	9
C	0	7 539	81	18
D	0	12 065	81	15
E	0	8 643	98	21
F	0	34 496	115	43
G	48	52 465	140	17
H	93	81 076	141	19
I	0	32 790	127	12
J	0	12 224	116	13
K	0	71 400	113	13
L	0	104 463	116	10
M	0	22 658	137	12
N	0	52 680	103	15

Table 2.2 shows statistics for weekly inflow for each power station. Not surprisingly the variation in weekly inflow is relatively high for all power stations, here shown by the standard deviation. The variation across seasons are relatively low for the majority of

---

[1]The calculation of average energy coefficient is described in Table 2.1.
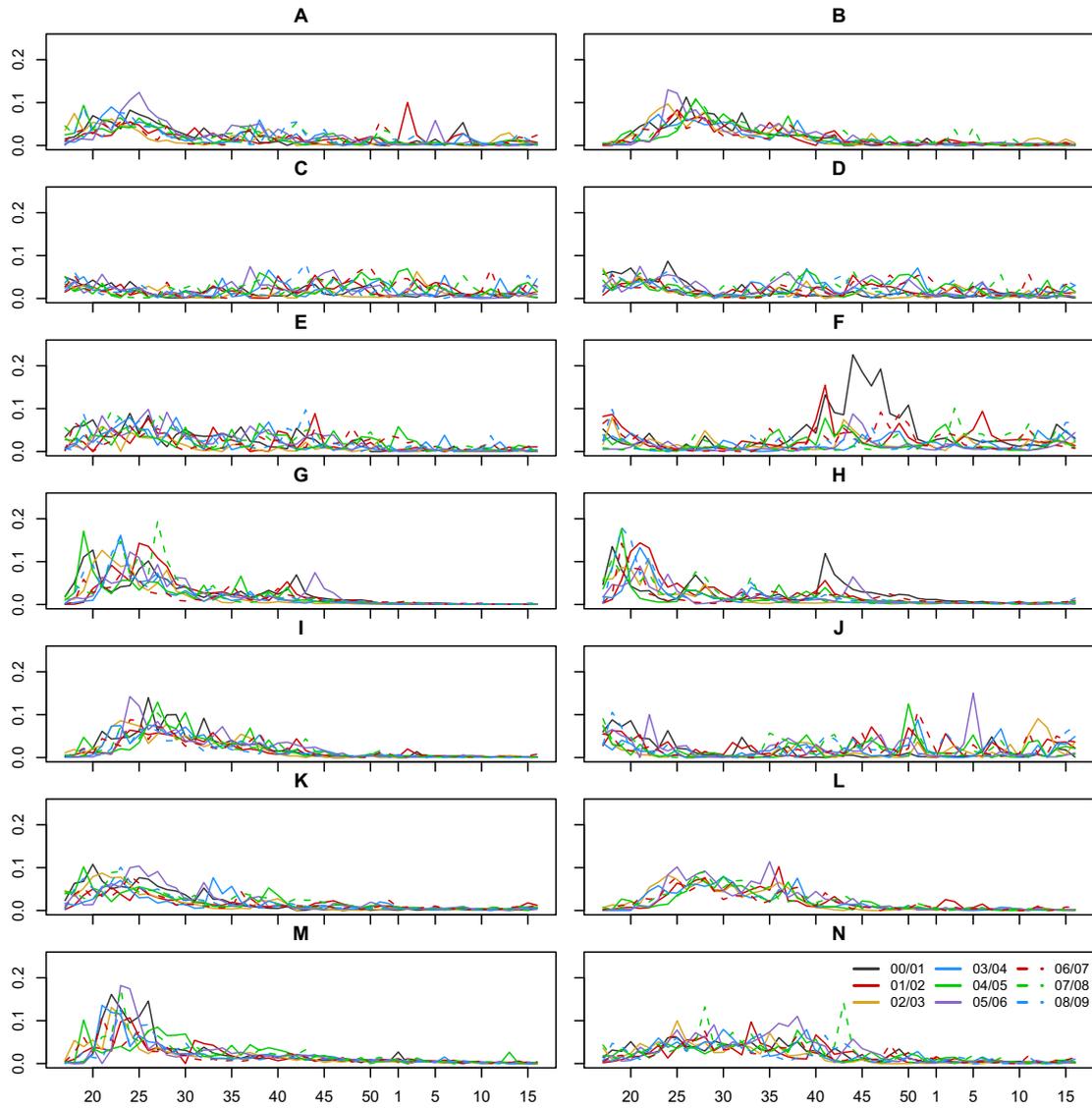
Figure 2.1: Weekly inflow relative to the average seasonal total. The horizontal axis represents week number
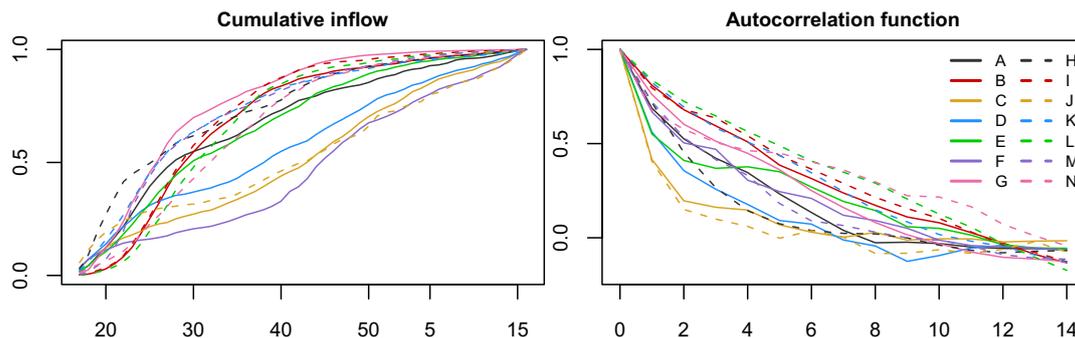
Figure 2.2: Left: Average cumulative weekly inflow for each power station, given in percentage of total seasonal inflow, with week number on the horizontal axis. The vertical axis shows the share of total seasonal inflow expected to occur up to a specific week. Right: Autocorrelation function of weekly inflow for each power station, with maximum number of lags equal to 14 weeks on the horizontal axis.

power stations. An exception is power station F, which has a seasonal standard deviation of 43 %. This is to a large degree caused by extremely high inflow in the season 2000/01, as seen in Figure 2.1, which shows weekly inflow per season. A majority of the power stations experience high seasonal variation; often the largest share of the inflow occurs in the first part of the season. To investigate the distribution of inflow throughout the year, cumulative average inflow (counting from the beginning of the season) is plotted left in Figure 2.2. We see power stations C, D, F and J have relatively evenly distributed inflow, while the inflow of the other power stations is more concentrated. How the inflow is distributed depends on geographical placing of the reservoir; areas in the north and high above sea level have a larger share of inflow coming snow during the winter, which results in low inflow in winter but high inflow during the melting period in the spring. Power stations along the coast can expect more evenly distributed inflow. The right plot of Figure 2.2 shows the autocorrelation function up to a lag of 14 weeks. The autocorrelation seems to decrease exponentially for most power stations, but with varying rate of decay. Power stations A, D, E, H and J all have below 50 % correlation between weeks further than 1 week apart, while power stations B, I, K and L have the slowest decaying autocorrelation function with over 50 % correlation remaining after 4 weeks.

Figure 2.3 shows total seasonal inflow relative to the historical average. For most of the power stations inflow in seasons 2002/03 and 2006/07 where relatively low. Note that inflow data from season 2008/09 is missing for power station B, and from seasons 2000/01 and 2008/09 for L.

2.1.2   Reservoir data

Reservoir content is for some power stations measured each hour, some once a day and for some on weekly basis. We use daily reservoir series in our analysis, and hence the hourly and weekly series are converted accordingly. We use the hourly reservoir level at midnight as the daily level the following day, while weekly reservoir data are interpolated where missing. Daily reservoir levels for each power station are plotted in Figure 2.4. The data shows a clear seasonal trend; reservoir levels are at their lowest in the spring and are filled up during the summer and fall. This seasonality is a result of the relationship
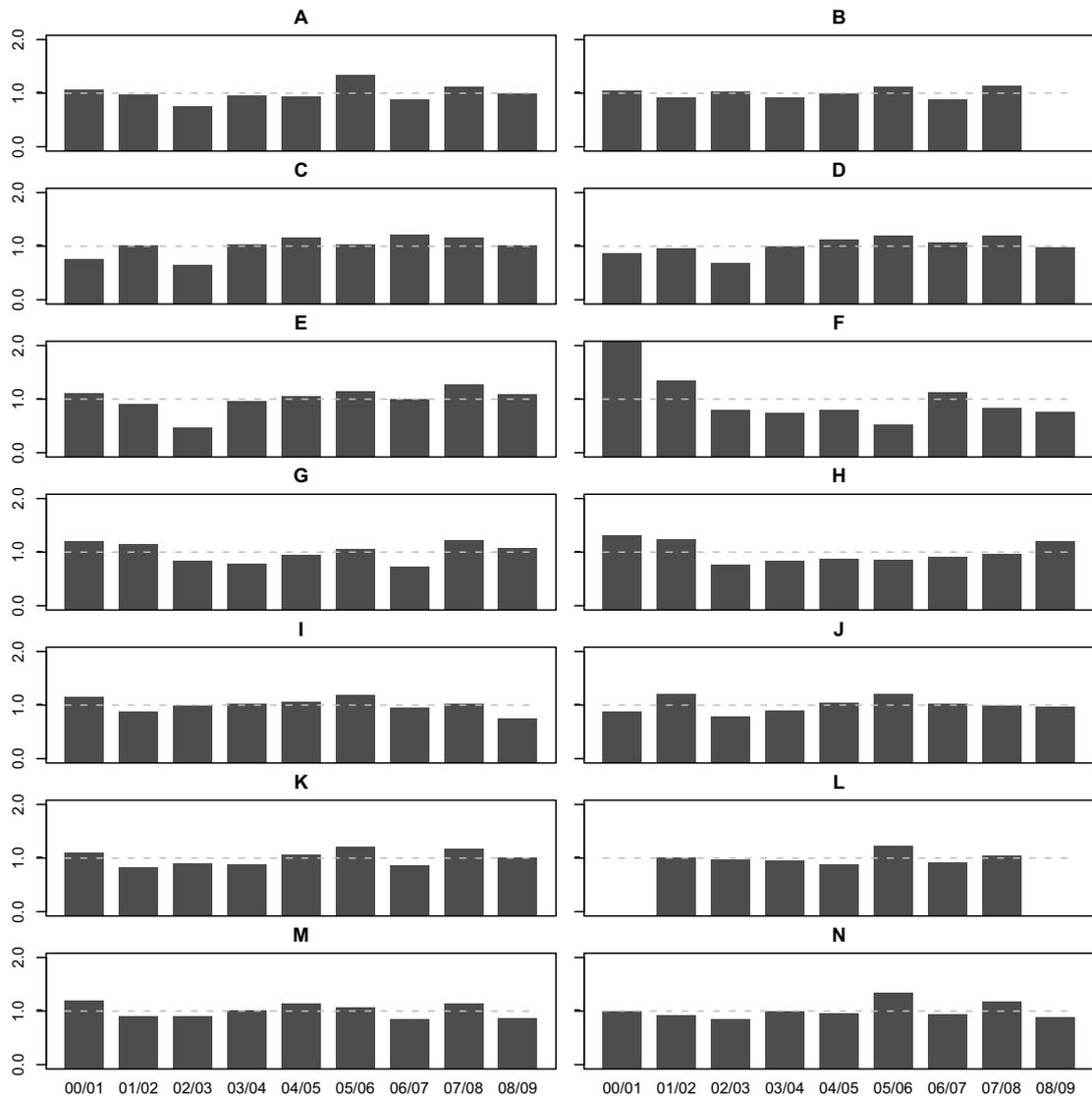
Figure 2.3: Total seasonal inflow relative to historical average for each power station. Note that the value for power station F in season 2000/01 is 2.33, which is higher than the limit of the vertical axis.

between inflow and demand for electricity. During the winter the demand is high, while most of the depreciation comes as snow, causing low inflow. Hence, hydropower stations save water during the summer and fall to meet the upcoming increase in demand. The period where the reservoirs are filled up is referred to as the *filling season*, while the period where the water is used as the *drawdown season*. The timing of these seasons may not be equal for different power stations, depending on geographical placement. However, we use a general definition based on the lowest and highest point of the average reservoir level in our sample; the filling seasons lasting from week 17 to the end of week 39, and the drawdown season from week 40 to end of week 16. Note that reservoir data for season 2000/01 and 2008/09 are missing for power station L.
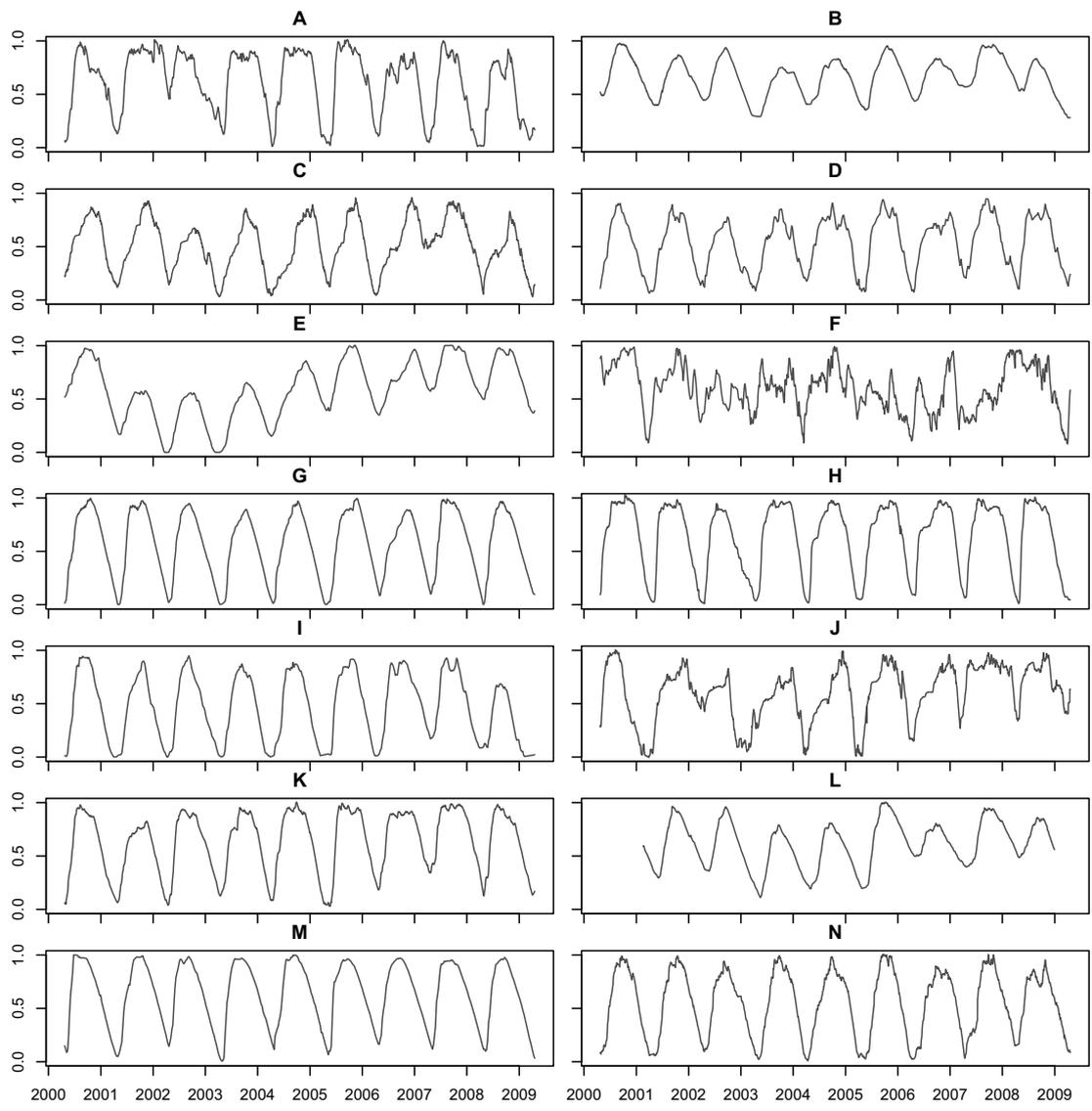
Figure 2.4: Daily reservoir levels relative to reservoir size for each power station.

2.1.3 Production data
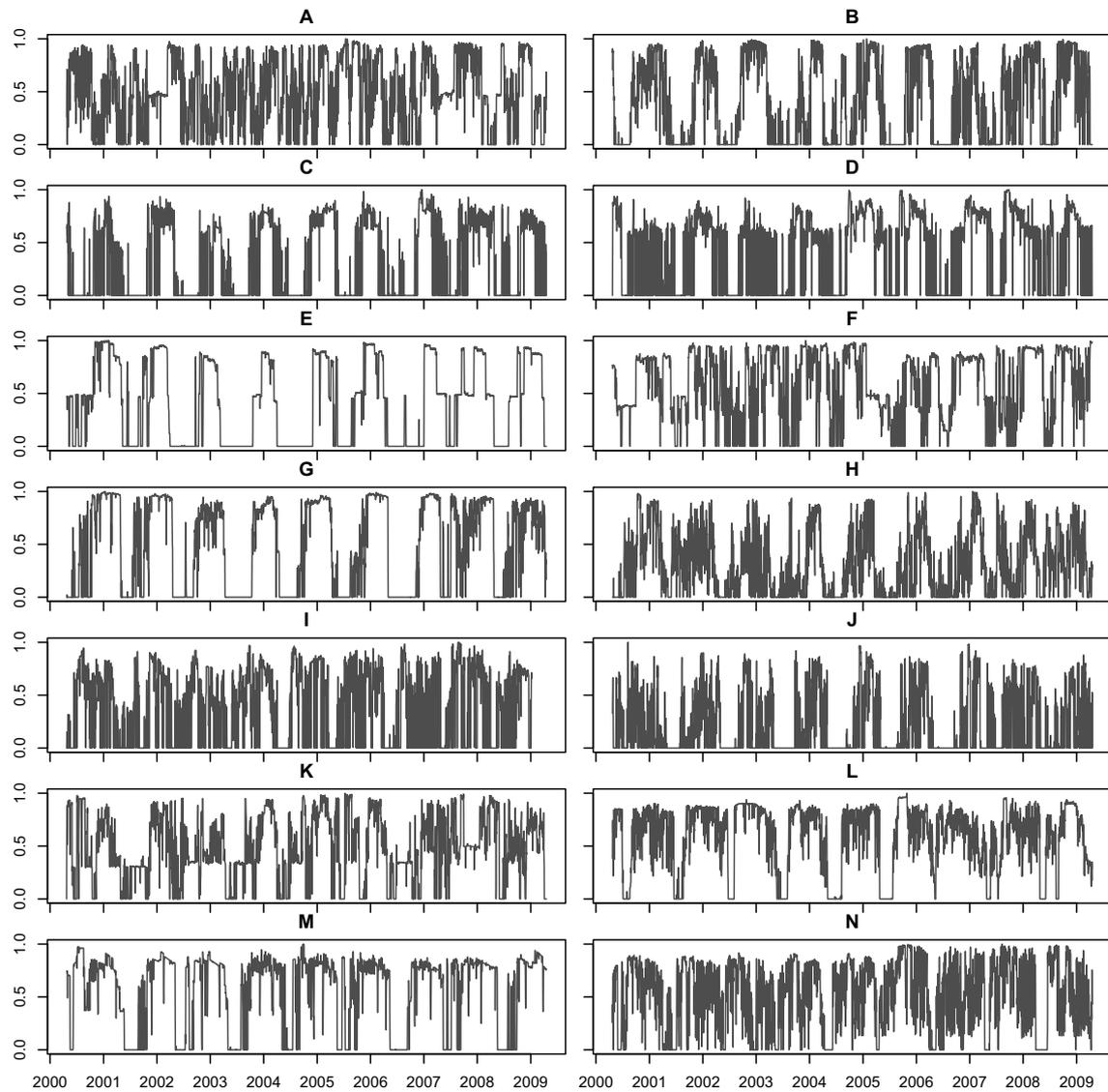
Figure 2.5: Daily production relative to maximum daily production capacity for each power station.

The production data is given on hourly basis for all power stations. Figure 2.5 shows the daily production relative to the maximum production capacity. Several producers shut down operations completely during the summer months, such as C, E, G and J. These power stations have low capacity factors, and hence can concentrate their production in time. We also see that the majority of the power stations choose to produce at a level close to their production capacity or not at all. This is especially seen by the many vertical lines, i most cases representing weekends (with lower prices).
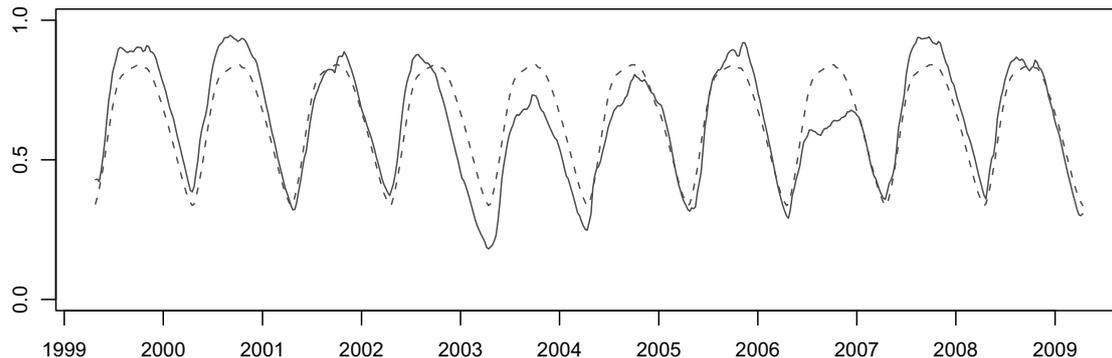
Figure 2.6: Overall reservoir content in Norway relative to total reservoir capacity, from season 2000/01 to 2008/09. The dashed line is the average level for years 1998 to 2009.

2.2 Overall reservoir data

Aggregated weekly reservoir content in Norway is downloaded from The Norwegian Water Resources and Energy Directorate's, see NVE [2010]. Figure 2.6 shows the weekly reservoir content for each season together with the average level. We notice a clear relationship between inflow and overall reservoir content; as pointed out above, most of the power stations experienced relatively low inflow during seasons 2002/03 and 2006/07, coinciding with particularly low reservoir as seen in Figure 2.6. The aggregated reservoir levels for season 2003/04 also where below average, indicating a dry year.

2.3 Electricity price data

Electricity price data are downloaded from the FTP statistical database for Nord Pool [2010]. Nord Pool offers trading of physical and financial power contracts in the Nordic countries. In the physical market, contracts are traded on a day-ahead and intra-day basis. Financial power derivatives such as forward, futures and options with these as underlying are traded in the financial market. We use data from both the physical and the financial market in this study, and the following sections explains the different data sets. Some early prices are listed in Norwegian kroner [NOK], and hence are converted to [EUR] using the historical exchange rate. Note that a future exchange rate should be used to transform the prices of financial products with cash flow occurring at a later point in time, but for products with relatively short time to delivery the use of a spot exchange rate is justified.

2.3.1 Spot prices

Nord Pool spot prices are actually day-ahead future prices for each hour the next day. The price is set based on the balance between supply and demand bids from all market participants in the Nordic countries. This balancing price is called the system price. However, due to grid congestions within Norway and between the Nordic countries, price areas are introduced. Table 2.3 shows statistics for hourly and daily prices for different price areas within Norway as well as the system price. The number of price areas and their borders change over time, however in this paper we fix the areas as of

May 2009[2]. Due to the changes in areas, two or more of the spot price series may be equal in periods when they where merged into one. Prices are given in [EUR/MWh]. We see that prices are somewhat higher in areas NO3 and NO4, which cover the Middle and Northern part of the country. This is mainly due to bottlenecks in the transition lines between the Middle and the South of Norway.

Figure 2.7 shows the daily system spot price and its natural logarithm for season 2000/01 throughout 2008/09. We see that the winter 2002/03, 2006 and winter 2008/09 had especially high prices. This coincides with the abnormal low inflow and low reservoir levels discussed previously. Both hourly and daily prices show well-known qualitative features, as discussed in literature by Geman and Roncoroni [2006]. We will briefly mention some of these features below, while leaving more detailed discussions to other studies.

- **Seasonality:** Spot prices are in general higher during winter than during summer. This follows from the nature of demand for electricity in the Nordic countries; during winter the low temperatures increases the need for heating, which increases demand for electricity. Hence prices tend to follow a seasonal trend, with periods of one season. In addition we see a positive trend during the sample period; the average price level rises throughout the period. It is important to keep this in mind when comparing revenues from different time periods. To account for the change in price level, we propose a method for detrending prices later in this section.

- **Spikes:** Price spikes, or more specifically sudden price changes followed by quick reversion, are fairly normal for prices. An example is beginning of February 2001, where the average daily system price more than tripled in one day, going from 23 to 77 [EUR/MWh]. Two days later it went back to its original level. Such positive spikes can be caused by sudden supply shortages due to unplanned outages of power plants or transmission lines. Sudden increases in electricity demand for heating on abnormally cold winter days can also cause positive spikes. Negative spikes, however, tend to occur more often during summer, when the share of thermal production is high. The presence of start up costs in thermal units increases the threshold for shutting down units during periods of low demand, which can lead to a surplus of supply.

- **Mean reversion:** As pointed out by Geman and Roncoroni [2006] a characteristic of electricity and other commodity prices are mean reversion toward a level that represents the marginal cost. The mean reversion level could be constant or periodic, where the latter can incorporate the seasonality mentioned above. We also see that after the occurrence of price jumps, the prices tend to revert to previous levels after no more than a couple of days. Hence, in addition to a general mean reversion to some constant or seasonal dependent level, spot prices show a faster mean reversion after spikes.

2.3.2 Prices of financial contracts

In this study we use a range of financial contracts, with various length of delivery period; from weekly contracts with delivery next week up to monthly or four week block

---

[2]Price data from Nord Pool where given for five different Norwegian cities: Oslo, Kristiansand, Bergen, Trondheim and Tromsø. We name these price series after the price area they are belonging to as of May 2009.

Table 2.3: Descriptive statistics for daily area spot prices [EUR/MWh] from 01.01.2000 to 19.04.2009. Hourly values are shown in parentheses.

Area	Mean	Min	Max	Stdv
System	29.99	3.89 (2.04)	114.61 (238.01)	14.57 (15.03)
NO1	30.38	2.07 (0.17)	114.61 (238.01)	14.61 (14.88)
NO2	30.38	2.07 (0.17)	114.55 (237.77)	14.62 (14.89)
NO3	32.15	2.99 (2.07)	114.61 (475.75)	15.11 (15.80)
NO4	31.97	2.99 (2.36)	114.61 (475.75)	14.74 (15.36)
NO5	30.38	2.07 (0.17)	114.61 (237.77)	14.62 (14.88)

contracts[3]. Nord Pool uses the term futures for daily and weekly contracts with and the term forwards for contracts with longer delivery periods. In practice these contracts are what financial literature refer to as *swaps*, as they consist of payments settled against the average spot price during a delivery period, see Benth et al. [2008] for more details. The term swap is used for both futures and forwards in the remainder of this study. As the current price of a swap gives an indication of the expected spot price during the delivery period, it provides useful information for the hydropower producers when planning the dispatch. Fleten et al. [2009] show empirically that this information is in fact used by hydropower stations. To obtain realistic results, we incorporate information from the financial market when simulating the dispatch later in this study by using the expected spot price defined below.

Deriving expected spot prices

In the dispatch model described in Chapter 4, we use the average expected spot price for the next six weeks as an input to the production planning. This figure can be derived from market prices for swaps with delivery periods during the upcoming weeks. There exists extensive literature on how to construct smooth forward curves in electricity markets, such as Fleten and Lemming [2003] and Benth et al. [2007]. However, in our case we only look at the average expected spot prices, and thus do not need to implement methods to obtain high-resolution forward price curves. Still, we need to take into account a possible risk premium when deriving expected spot prices from traded swap prices. A widely cited paper by Longstaff and Wang [2004], using empirical data from the PJM electricity market in the US, shows significant variations in these risk premia, depending on the delivery period. Botterud et al. [2009] and Frestad et al. [2010] use data from the Nord Pool market, finding a declining negative risk premium (favoring electricity producers) for short term weekly contracts. Based on the findings in these papers, we assume a constant risk premium of 0.2 % daily for our six weeks horizon[4]. Figure 2.8 shows the constructed average expected spot price, using closing prices from the previous trading day. As expected, the average expected spot for the next six weeks follows the spot price very closely.

---

[3]The range of financial contracts traded on Nord Pool has changed somewhat during our sample period. A three-seasonal contract system has been replaced by quarterly contracts. On more short term, four week blocks have been replaced by monthly contracts.

[4]For simplicity we choose a constant rate. Using the average of the expected spot prices will even out the effects of a declining risk premium

2.3.3 Detrending prices

The positive trend in the spot price series complicates the process of comparing revenues from different time periods. This is solved by *detrending* the price data; converting all prices to the same reference level. In Chapter 4 we model the seasonality in spot prices as a deterministic function containing (among other factors) an exponential trend. This trend estimate is used to detrend the historical prices, and corresponding to 0.03 % per day, or 11.6 % on yearly level[5]. As we expect the forward and future prices to show the same trend as the spot price, these are detrended using the same factor. Figure 2.9 shows the daily spot prices detrended to the level of the last day in our data; 19th of April 2009. The actual historical spot prices are plotted as a reference. For the remainder of this study we use detrended prices.

2.4 Correlation between data series

As pointed out in the earlier sections, we see a correlation between inflow, reservoir content and spot prices for some seasons. For example, the low inflow in season 2002/03, coincides with low reservoir level and high prices. This relationship is not surprising, as lack of water will cause the long term supply of hydropower to decrease, driving up prices. In this section we will investigate the relationship between these series further.

Table 2.4: Pearson's product moment correlation coefficient for deviation from normal for cumulative inflow last 20 weeks against deviation from normal overall reservoir content and spot price, per power station. Given below: Correlation between spot price and deviation from normal of overall reservoir content. Note that the null hypothesis is a correlation coefficient equal to zero.

Power station	$\Delta i_t, \Delta r_t^{overall}$	Significance	$\Delta i_t, S_t$	Significance
A	0.535	***	-0.244	***
B	0.438	***	-0.095	***
C	0.471	***	-0.268	***
D	0.448	***	-0.209	***
E	0.581	***	-0.250	***
F	0.327	***	-0.367	***
G	0.701	***	-0.309	***
H	0.495	***	-0.340	***
I	0.449	***	-0.250	***
J	0.270	***	-0.199	***
K	0.519	***	-0.143	***
L	0.756	***	-0.308	***
M	0.497	***	-0.256	***
N	0.653	***	-0.241	***

Correlation for $\Delta r_t^{overall}$ and S_t: -0.377 ***

Significance codes for Pearson's correlation coefficient: 0 (***), 0.001 (**), 0.01 (*), 0.05 (.), 0.1 ( )

---

[5]Note that growth rate is surprisingly high compared with normal inflation rates, and could be caused by abnormal price variations in our estimation window ranging from April 26th 1999 to April 19th 2009

2.4.1 Correlation between inflow and reservoir level

There is an obvious relationship between inflow and reservoir content of a specific power station. If the inflow has been higher than normal, the reservoir level is also expected to be higher than normal and vice-versa[6]. This means that the *cumulative* inflow during the last weeks, relative to the normal cumulative inflow in the same period, should be highly correlated to the deviation from normal reservoir level. Based on observations in the sections above, we also expect such a relationship between a particular power station's inflow and the aggregated reservoir content in Norway. However, the size of this correlation varies between the power stations, depending mostly on their location. Figure 2.10 shows the deviation from normal for the overall reservoir content and for the cumulative inflow the last 20 weeks[7]. We see a clear positive correlation for all power stations, but as expected the degree of correlation varies. Table 2.4 shows the different correlation coefficients, ranging from 0.27 to 0.76. Power station F and J have the lowest correlation coefficients, both being located in areas with low density of hydropower stations, while power station G and L have the highest correlation. The null hypothesis of no correlation is rejected for all power stations.

2.4.2 Correlation between overall reservoir level and spot prices

As pointed out earlier, low overall reservoir content tends to coincide with high spot prices. This relationship follows from basic microeconomics; in the Nord Pool area, were the share of hydropower is high, the aggregated reservoir content is an indication for the (cost efficient) supply of electricity. Naturally, when reservoir levels are low, prices increase as the result of lower supply. For a more general discussion on the relationship between inventory levels and spot prices for commodities, we refer to Pindyck [2001]. Figure 2.11 shows the daily spot price and the deviation from normal for overall reservoir content. As expected, the correlation coefficient is negative, being equal to -0.38 and significantly different from zero, as found in the bottom of Table 2.4.

2.4.3 Correlation between inflow and spot prices

Based on the two subsections above we expect a negative correlation for the deviation from normal for individual power station's cumulative and the spot price. These two data series are shown in Figure 2.12, while Table 2.4 lists the correlation coefficients for each power station. The values range from -0.09 to -0.37, with power station B having the lowest absolute correlation and F having the highest. Note that the relatively low absolute correlation for some power stations can be caused by positive (negative) price spikes not coinciding with low (high) inflow, reducing the absolute correlation. As discussed earlier, price spikes tend to occur as a result of *sudden* supply/demand shortages/surpluses, and not changes in inflow affecting the medium and long term supply. The null hypothesis of no correlation is also here rejected for all power stations.

---

[6]Especially during the filling period, when production is low

[7]The choice of a window of 20 weeks is based on testing on the sample inflow data, showing a consistent correlation
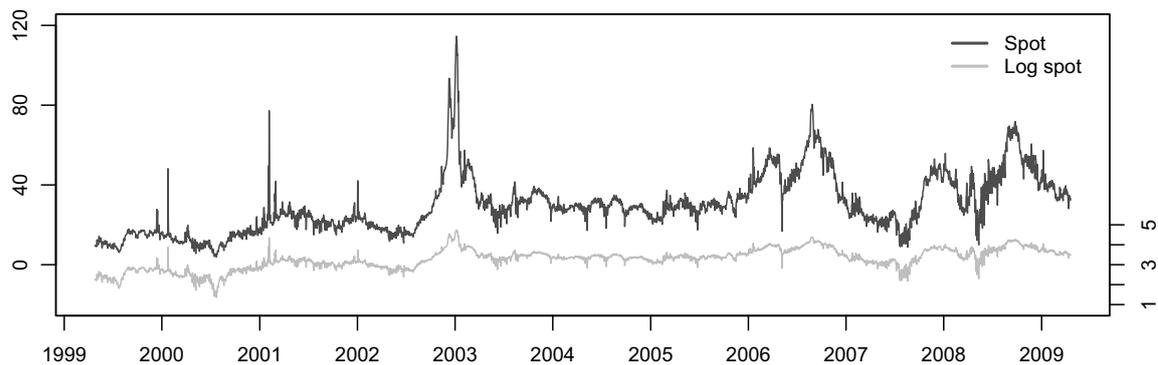
Figure 2.7: Average daily system spot prices (absolute/logarithm) [Eur/MWh].
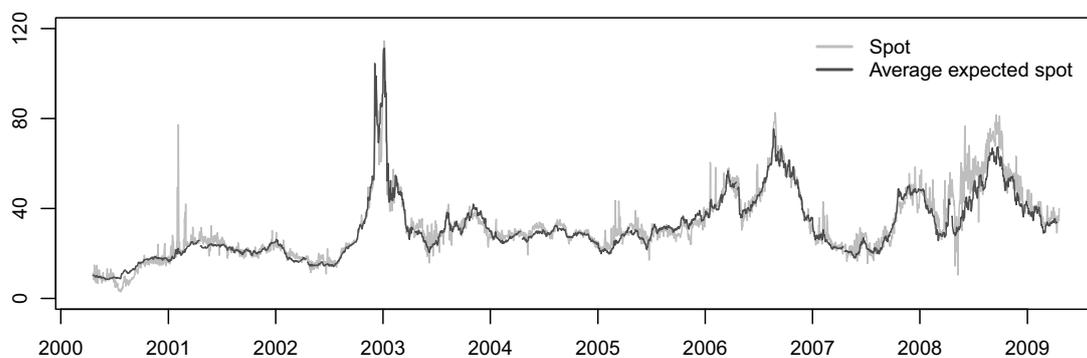
Figure 2.8: Average expected spot price [EUR/MWh] for the next six weeks based on closing prices for swaps from the previous trading day.
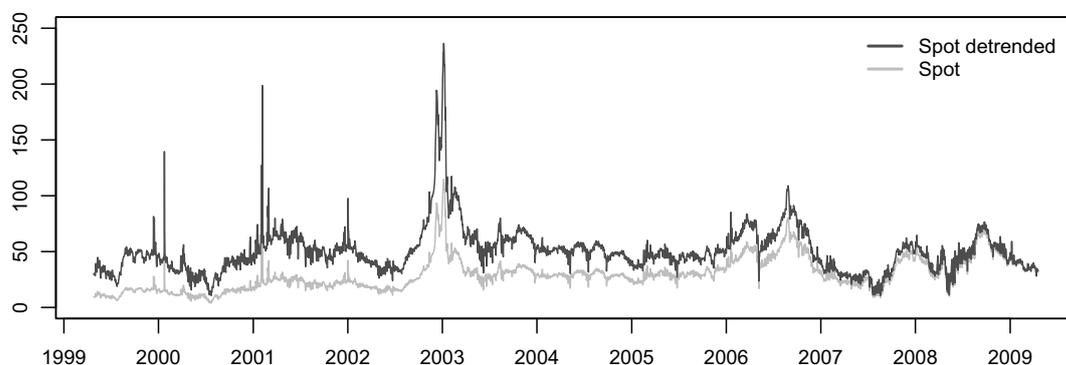
Figure 2.9: Detrended daily system spot prices [Eur/MWh] (blue) together with actual prices (grey). Prices are detrendend to a level corresponding to end of season 2008/09, using a daily trend of 0.03% estimated by regression on the historical data.
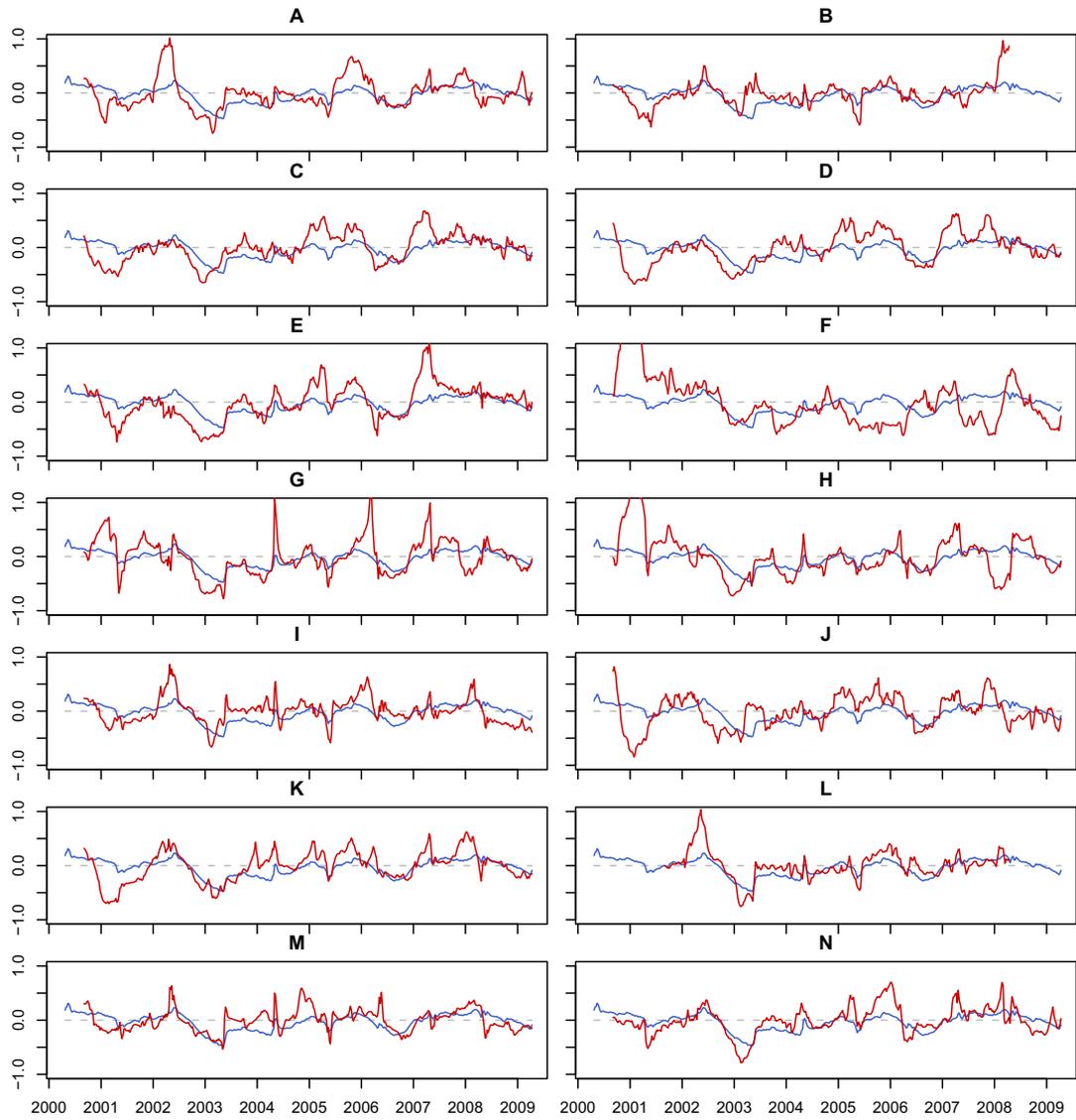
Figure 2.10: Deviation from normal for overall reservoir level (blue) and for each power station's cumulative inflow last 20 weeks (red). The dashed line (zero deviation) corresponds to normal cumulative inflow and overall reservoir levels. Due to the 20 week lag, the cumulative inflow series is 20 weeks shorter than the original inflow data.
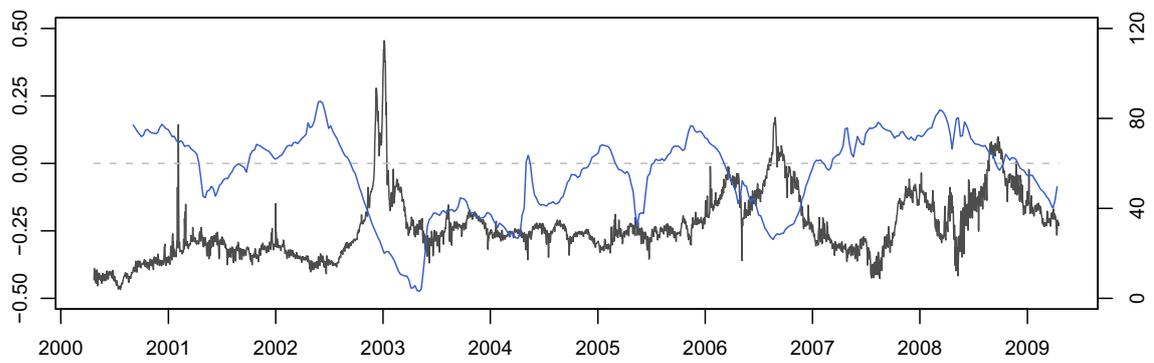
Figure 2.11: Deviation from normal level for the overall reservoir content (blue, left axis) together with daily spot price (dark grey, right axis). The dashed line show the normal level of the reservoir.
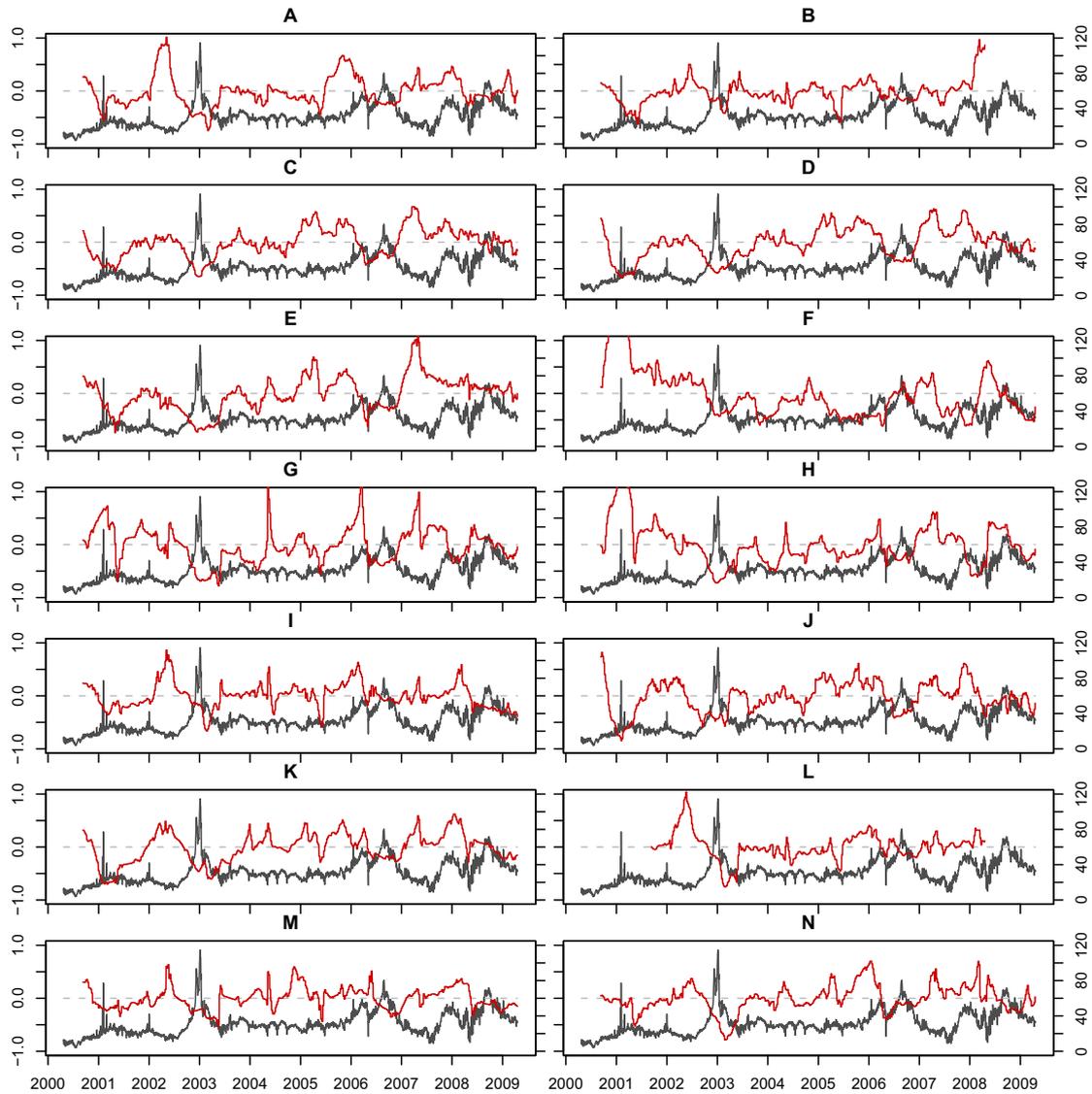
Figure 2.12: Deviation from normal for cumulative inflow last 20 weeks (red, left axis) for each power station, together with the daily spot price (dark gray, right axis). The dashed line corresponds to normal inflow volumes (zero deviation). Due to the 20 week lag, the cumulative inflow series is 20 weeks shorter than the original inflow data.

Chapter 3

Framework for analyzing the impact of flexibility

In the following sections we will develop a framework for analyzing the impact of flexibility on revenues for hydropower producers. By defining different levels of flexibility, and reviewing revenues for each case, we are able to isolate the value of storage flexibility. The results from applying the framework on empirical data will serve as a basis of comparison for the stochastic model developed in Chapter 4.

3.1 Defining flexibility and its value

What exactly does the word *flexibility* mean for a hydropower producer? In general, flexibility is the degree of freedom available when establishing a power station's dispatch plan. Thus, high flexibility gives producers a large possibility space when deciding when and how much to produce, while low flexibility confines their options. For power producers it is necessary to distinguish between short and long term flexibility. Short term flexibility impacts the producers' ability to profit from hourly price variations. A large degree of freedom enables producers to vary their output from hour to hour, and potentially to profit from participating in the regulating power market, a market studied by Skytte [1999]. In the long term, high flexibility enables the producers to obtain higher prices by storing water flowing into their reservoirs throughout the season. As discussed in section 2.3 there are clear seasonalities in spot prices on Nord Pool. By storing the water, producers are able to profit from both these expected price differences and sudden unexpected spikes. In this paper, we choose to focus primarily on long term flexibility.

So how is the *value of storage flexibility* defined? To isolate this value, we must first consider the base value of a hydropower station without storage flexibility. The hypothetical case where a producer must dispatch (by producing and through overflow) all it's water at the exact moment of inflow, could be interpreted as having no flexibility. This is typical for a run-of-river hydropower plant. These type of plants typically have the possibility to vary their output somewhat from hour to hour using a very limited reservoir, but are not able to store the water over longer periods. In other words, run-of-river plants have no long term flexibility. The value of storage flexibility for a specific hydropower station can be defined as the increase in revenues when going from hypothetically having no reservoir to having it's actual reservoir size, everything else being equal. This definition is used in the remainder of this study, and is further specified in Section 3.1.2.

3.1.1 Levels of flexibility and information

We now specify several cases of storage flexibility, with expressions for revenues obtained by producers in each case.

No flexibility

As mentioned above, the case of no flexibility is equivalent to a run-of-river plant. In this case it is assumed that the producer has no ability to store the water, hence all inflow has to be produced at the time of arrival. This case reduces the production scheduling to a deterministic problem, where the production in each time step is equal to the corresponding inflow, with production capacity serving as the upper limit. All inflow above the maximum production capacity is lost as overflow. Thus, the hypothetical revenues during one complete season obtained in this situation are defined by:

$$\Pi^{river} = \sum_{t=1}^{T} \min(I_t, P^{max}) S_t \tag{3.1}$$

where I_t is the historical inflow and S_t the historical spot price (for the specific Nord Pool price area) in period t. T is the number of periods in one season and P^{max} is the production capacity during a single period.

Limited flexibility

In this case it is assumed that the power station has a positive reservoir size, enabling flexibility in terms of when to produce the water that has come as inflow up to the current time. However, with an upper limit of the reservoir, the producer needs to take into account the probability of overflow during periods with high reservoir levels. This restriction limits the producer's flexibility. In other words, this is the real world case for a hydropower station with reservoir.

The production scheduling problem in this case is stochastic, due to the uncertainty of the future inflow and spot prices. The producer must base it's dispatch strategy solely on information available at the current time, taking into account these uncertainties. As shown in Fleten et al. [2009], producers base their decision on several factors, including forward prices traded on Nord Pool's financial market and the expected inflow during the next week. The revenues obtained in this flexibility case are calculated by using the actual historic dispatch plan for each power station:

$$\Pi^{actual} = \sum_{t=1}^{T} P_t S_t \tag{3.2}$$

Where P_t represents the historical production in period t, and the other variables are defined as above. Later, when using this framework in the stochastic model, the actual production P_t is approximated by implementing a model for the dispatch strategy.

In order to obtain a theoretical upper limit for the revenues obtainable under limited storage flexibility, we also calculate the theoretical revenues assuming that producers have *full information*. In this case we assume that producers are clairvoyant, knowing the spot price and the inflow up front. This reduces the production scheduling problem

to a deterministic linear programming (LP) problem. Using the historical reservoir levels at the start and the end of each season as boundary conditions, the revenues for one season are calculated by solving the following LP-problem:

$$\Pi^{clairvoyant} = \underset{p_t, o_t}{\text{maximize}} \sum_{t=1}^{T} p_t S_t \tag{3.3}$$

subject to the restrictions

$$r_t - r_{t-1} = I_t - p_t - o_t \tag{3.4}$$
$$r_t \leq R^{max} \tag{3.5}$$
$$p_t \leq P^{max} \tag{3.6}$$
$$r_t, p_t, o_t \geq 0 \tag{3.7}$$

for $t = 1, 2, \ldots, T$. Here r_t is the reservoir level at the end of period t, while p_t is the production variable and o_t is the overflow during period t. R^{max} is the upper limit for the reservoir level, while the other variables are the same as above. Restriction 3.4 represents the reservoir balance equation, taking into account the historic inflow and the production and overflow variables. The LP-problem is solved for the given boundary conditions

$$r_0 = R^{start} \tag{3.8}$$
$$r_T = R^{end} \tag{3.9}$$

Where R^{start} and R^{end} are the historic reservoir levels at the beginning and end of the season respectively.

Unlimited flexibility

In this case we relax the limits to flexibility be removing the reservoir restriction and assuming that all of the seasonal inflow is available at the beginning of the season. Still assuming full information, the revenues in this case are equivalent to the theoretical upper limit obtainable under what we define as unlimited flexibility. Here, the production scheduling problem is reduced to a simple price duration curve analysis, where the production (being equal to the total seasonal inflow, adjusted for reservoir level changes) is simply allocated to the periods with the highest prices. The revenues are obtained by the following equation:

$$\Pi^{max} = \left[\left(\sum_{j=0}^{\lfloor \tau \rfloor} S_j^{sorted} \right) + (\tau - \lfloor \tau \rfloor) S_{\lfloor \tau \rfloor + 1}^{sorted} \right] P^{max} \tag{3.10}$$

$$\tau = \frac{\sum_{t=1}^{T} I_t + R^{end} - R^{start}}{P^{max}} \tag{3.11}$$

where τ is the load factor[1] and $S_j^{sorted}, j = 1, 2, \ldots, T$ are the historic spot prices in descending order for the given season. See the last plot in Figure 3.1 for an illustrative example.

---

[1]Load factor equals total production in a period divided by production capacity, and are here measured in days.

The four cases described above are summarized in Table 3.1 and the revenues are illustrated in Figure 3.1. From our definition of flexibility, higher flexibility means more possibilities when deciding when and how much water to dispatch from the reservoir. This is influenced by the availability of water (given by the reservoir size and inflow distribution), and the ability to dispatch this water (given by the production capacity and other flow restrictions). Our four cases represent different degrees of *storage* flexibility, by varying reservoir and inflow. Varying production capacity will influence the revenues in all cases, and hence its impact is studied by comparing revenues across producers with different capacity factors. The four cases provide us with both a lower bound for revenues (no flexibility), the real world case, and the theoretical upper bound for limited flexibility. The revenues under unlimited flexibility are included as a reference for studying the effects of limiting storage flexibility.
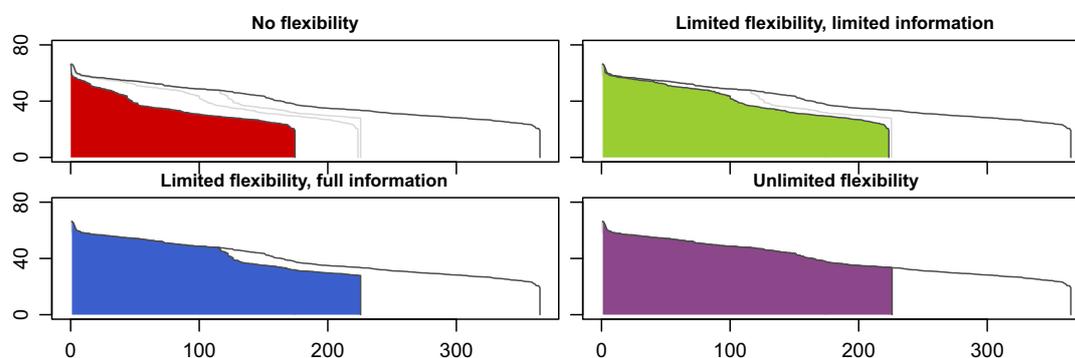
Figure 3.1: Example of revenues in the four different flexibility cases illustrated by price/load duration curves (data for producer A for season 2007/08). The shaded areas represents the revenues relative to production capacity. In all plots the vertical axis represent daily average spot prices [EUR/MWh], while the horizontal axis represent duration [days].

Table 3.1: Flexibility cases with specification of limitations. Note: Limited flexibility with limited information represents the real world case faced by power stations with reservoir.

Flexibility case	Revenue variable	Information	Reservoir	Inflow	Production capacity
No flexibility	Π^{river}	Deterministic	None	Distributed	Limited
Limited flexibility	Π^{actual}	Limited	Limited	Distributed	Limited
	$\Pi^{clairvoyant}$	Full	Limited	Distributed	Limited
Unlimited flexibility	Π^{max}	Full	Unlimited	Concentrated	Limited

3.1.2 Relative flexibility value

To complete the framework, we will now define a measure for relative value of storage flexibility based on the different cases above :

Definition: For a specific hydropower station with reservoir, let Π_i^{actual} be the actual revenues realized with limited flexibility in season i as defined in equation 3.2. Let $\Pi_i^{clairvoyant}$ be the theoretical maximum revenues with limited flexibility in season i as defined in equations 3.3 through 3.9, while Π_i^{river} being the revenues under no flexibility (run-of-river) as in equation 3.1. Then the **relative flexibility value** v for this power station is given by:

$$v^{clairvoyant} = \frac{\sum_{i=1}^{\infty} \Pi_i^{clairvoyant} - \sum_{i=1}^{\infty} \Pi_i^{river}}{\sum_{i=1}^{\infty} \Pi_i^{clairvoyant}} \leq 1 \qquad (3.12)$$

$$v^{actual} = \frac{\sum_{i=1}^{\infty} \Pi_i^{actual} - \sum_{i=1}^{\infty} \Pi_i^{river}}{\sum_{i=1}^{\infty} \Pi_i^{actual}} \leq 1 \qquad (3.13)$$

The clairvoyant flexibility value $v^{clairvoyant}$ represents the share of the revenues in the clairvoyant case that comes from storage flexibility, or more specifically the ability to store water. The actual flexibility value v^{actual} is the same share, using actual revenues for the power station.

When comparing the revenues for the different producers, we use relative figures. Hence we use the total revenues relative to the total water used (both through production and overflow) during all seasons, adjusted for price differences between the Nord Pool price areas[2]. These relative revenues are noted π with the unit [EUR/MWh water]. For the different flexibility cases $a = \{river, actual, clairvoyant, max\}$ the relative revenues are defined as:

$$\pi^a = \frac{\sum_{i=1}^{\infty} \Pi_i^a}{\sum_{i=1}^{\infty} \sum_{t=1}^{T_i} I_{i,t}} \qquad (3.14)$$

To estimate the relative revenues and flexibility values for the 14 power stations in our sample, we will use the historical data described in Chapter 2. We need to take into account the reservoir difference when using a relatively low number of seasons. This

---

[2]Adjustment for price differences between the Nord Pool price areas is done by multiplying each seasonal revenue by the ratio between the average system spot price and average area spot price for that season.

results in the following equation:

$$\hat{\pi}^a = \frac{\displaystyle\sum_{i=1}^{N} \Pi_i^a}{\displaystyle\sum_{i=1}^{N}\sum_{t=1}^{T_i} I_{i,t} + R_1^{start} - R_N^{end}} \tag{3.15}$$

Where N is the number of seasons with available data. Further, the relative flexibility values are estimated as follows:

$$\hat{v}^{clairvoyant} = \frac{\hat{\pi}^{clairvoyant} - \hat{\pi}^{river}}{\hat{\pi}^{clairvoyant}} \tag{3.16}$$

$$\hat{v}^{actual} = \frac{\hat{\pi}^{actual} - \hat{\pi}^{river}}{\hat{\pi}^{actual}} \tag{3.17}$$

3.1.3 Factors influencing revenues in the flexibility cases

There are several factors that are expected to influence the revenues in the different cases above. Obviously, reservoir size is an important factor, or more specifically it's size relative to the expected annual inflow. These and other factors are summarized in the following list:

- **Relative regulation:** The ratio between the reservoir size and the expected seasonal inflow, defines the ability to store the water. A low relative regulation forces the power station to produce more evenly distributed over the season in order to avoid overflow. Hence, a higher relative regulation will increase revenues under limited flexibility, and have a positive effect on the flexibility value.

- **Inflow seasonality:** Strong seasonalities in the inflow increase the need for high relative regulation in order to store the water. As seen in Section 2.1.1, the inflow for a majority of the power stations is high during the spring flood compared to the rest of the season. With a small reservoir, the producer must produce during the spring flood in order to avoid overflow, decreasing the ability to profit from seasonalities in spot prices. However, with a sufficiently large reservoir, strong inflow seasonality could be an advantage rather than problem, as most of the inflow is available early in the season which enables the producer to profit form unexpected price increases. Strong inflow seasonality will impact the revenues in the no flexibility case negatively, due to large amounts of overflow during the spring flood. Hence, for power stations with high (low) relative regulation, strong inflow seasonalities is expected to increase (decrease) the flexibility value.

- **Capacity factor:** The production capacity of a power station sets the amount of time needed to produce the water in the reservoir. High capacity enables the power station to produce more during periods with high prices and thereby increase the revenues under both limited and unlimited flexibility. At the same time, higher production capacity increases the theoretical revenues under no flexibility by decreasing the water lost to overflow in the run-of-river case. Hence, the revenues for all flexibility cases increase with decreasing capacity factor. However, it is not clear how a change in capacity factor influences the value of storage flexibility.

- **Other restrictions:** Several other restrictions limit the flexibility of hydropower stations, such as water flow limits downstream of the generator, time-varying reservoir limits and maintenance periods. These restrictions will clearly impact the

revenues negatively in the limited flexibility cases, and thereby decreasing the flexibility value. However, for simplification we do not account for any of these restrictions when calculating the revenues for the theoretical flexibility cases.

See Table 3.2 for an overview of the expected impact of the different factors on the revenues in each flexibility case and on relative flexibility value.

Table 3.2: Factors influencing flexibility, with expected impact on revenues and relative flexibility value. Increasing the influencing factor is expected to affect the revenues and flexibility according to the arrows (↑ for positive, ↓ for negative). Alternating arrows (↑↓) indicate ambiguous impact.

Influencing factor	π^{river}	π^{actual}	$\pi^{clairvoyant}$	π^{max}	Flx. value υ
Relative regulation	-	↑	↑	-	↑
Capacity factor	↓	↓	↓	↓	↑↓
Inflow seasonality	↓	↑↓	↑↓	-	↑

3.2 Application of framework on empirical data

We now apply the framework defined above on the daily data available for the 14 hydropower stations.[3]. The estimates for the relative revenues for the different flexibility cases and the respective values are shown in Table 3.3. Comparing the relative revenues across the power stations in our sample, provides useful insights. Firstly, we can study how the different factors influence the actual realized revenues (π^{actual}), decomposing this figure into the run-of-river equivalent (π^{river}) and the isolated gains from storage flexibility ($\pi^{actual} - \pi^{river}$). Secondly, this framework quantifies the upper boundary for potential revenues, given each producer's flexibility restrictions ($\pi^{clairvoyant}$). Finally, the additional gain when relaxing the storage flexibility restriction can be found ($\pi^{max} - \pi^{clairvoyant}$).

3.2.1 Relative revenues in each flexibility case

Figure 3.2 graphically shows the revenues for the no flexibility case together with the increase in revenues when going through the other three flexibility cases. The relative revenues in the no flexibility case π^{river} range from 30.9 to 45.6 [EUR/MWh water] for the different producers. Like shown in Table 3.2, the run-of-river equivalent is affected by capacity factor and the degree of inflow seasonality. Power stations C, D and J, all having low inflow seasonality combined with medium to low capacity factors (indicating high production capacity), obtain the highest relative revenues π^{river} of around 44 [EUR/MWh water]. This is expected, as high production capacity combined with low inflow seasonality reduces the chance of overflow in the run-of-river case. Power stations G, L and M all have strong inflow seasonality combined with medium ranged capacity factors, resulting in low relative revenues ranging from 30.9 to 34.8 [EUR/MWh water]. Power station F is a special case, having low seasonalities in inflow combined with the highest capacity factor, obtaining a π^{river} of 32.5 [EUR/MWh water]. Here the high capacity factor (low production capacity) seems to outweigh the low inflow seasonality, resulting in a low relative revenue. In general the empirical results are as expected with

---

[3]Hourly production and price data could also have been used, however the difference in calculated revenues using hourly vs. (average) daily data is small. Figure A.2 in the Appendix shows a deviation from 0 % to 2.5 % for the power stations in our sample.

Table 3.3: Historic relative revenues and flexibility values per producer. Note: The revenues are corrected for trend in prices and differences between the different Nord Pool price areas.

Power station	$\hat{\pi}^{river}$	$\hat{\pi}^{actual}$	$\hat{\pi}^{clairvoyant}$	$\hat{\pi}^{max}$	$\hat{\upsilon}^{clairvoyant}$	$\hat{\upsilon}^{actual}$
	[EUR/MWh water]				[%]	[%]
A	36.6	48.0	58.4	62.4	37.2	23.7
B	41.5	58.1	67.9	67.9	39.0	28.6
C	43.4	51.8	65.9	69.0	34.1	16.2
D	44.3	46.8	59.7	63.3	25.8	5.4
E	41.7	54.9	64.0	64.5	34.8	24.0
F	32.5	32.6	44.0	49.2	26.3	0.5
G	30.9	53.9	63.4	63.7	51.3	42.7
H	37.6	52.7	67.7	71.0	44.5	28.7
I	38.5	53.5	62.7	66.1	38.5	27.9
J	45.6	48.6	68.2	73.2	33.1	6.1
K	38.3	50.0	61.1	63.8	37.3	23.4
L	34.8	55.5	62.2	62.4	44.0	37.3
M	33.8	51.4	60.0	60.7	43.6	34.2
N	40.2	48.3	58.9	62.6	31.8	16.8
Average	38.5	50.4	61.7	64.3	37.2	22.5
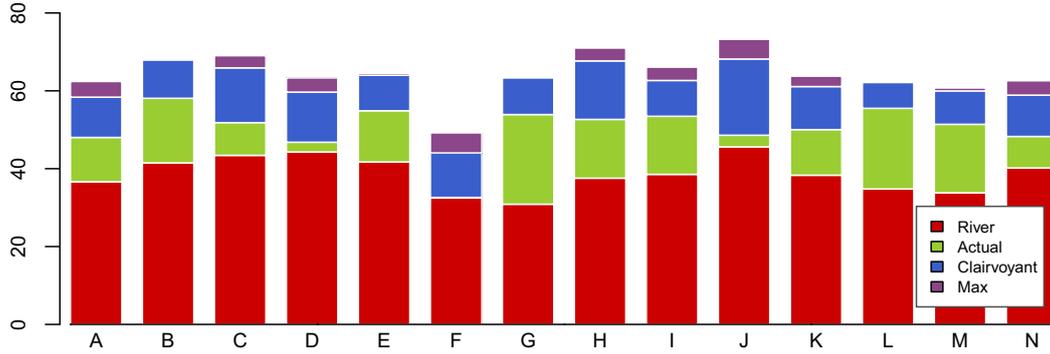
Figure 3.2: Estimated historic relative revenues per producer for each flexibility case in [EUR/MWh water]. Note: The revenues are corrected for trend in prices and differences between the different Nord Pool price areas.

both low seasonality and low capacity factor driving up the relative revenues under no flexibility.

Looking at the actual relative revenues π^{actual}, we see that the different power stations gain from 0.2 to 23.0 [EUR/MWh water] when going from having no storage flexibility to their actual reservoir size. Not surprisingly power station F, which like discussed earlier is fairly close to being a run-of-river plant, achieves the lowest gain. The difference between no flexibility and limited flexibility defines the flexibility value, and will be further discussed below. The total actual relative revenues range from 32.6 to 58.1 [EUR/MWh water]. Power station B, E, G and L achieve the highest realized relative revenues, with G gaining 23.0 [EUR/MWh water] over it's run-of-river equivalent. These four power stations have the highest relative regulation, which combined with medium capacity factors explains the high relative revenues π^{actual}, as predicted from Table 3.2. However, the actual revenues depend highly on the production strategy of each producer and do implicitly take into account the additional restrictions mentioned in Section 3.1.3. Hence, these actual relative revenues should only be used as an indication of performance keeping this in mind.

The clairvoyant case, $\pi^{clairvoyant}$, shows the upper bound for revenues achievable with limited flexibility, and ranges from 44.0 to 68.2 [EUR/MWh water]. Somewhat surprisingly, none of the power stations E, G and L (being among the top four under limited information), are among the top three in this case. Power station B, however, achieve high revenues in both cases under limited flexibility, which with 67.9 [EUR/MWh water] has the second highest revenues in the clairvoyant case. This could indicate that the impact of the factors on relative revenues is somewhat different with full information. Reviewing the characteristics for power station C, H and J, which with relative revenues ranging from 65.9 yo 68.2 [EUR/MWh water] make out rest of the top four in this case, we see that they have the lowest capacity factors in our sample. Comparing with power stations E, G and L, they have significantly lower relative regulation of around 0.5. Thus, the empirical results indicate that relative reservoir size plays a less important role than the capacity factor under full information. This can be explained as follows; under *full information*, the solution of the LP-problem will be situated on the borders

of the feasible region defined by the constraints. With a low relative regulation, it is expected that the reservoir size is a binding constraint in many time periods, thus the reservoir level will be equal to its limit accordingly often. However, under *limited information*, it is risky for a producer to have a reservoir level close to the limit due to the inflow uncertainty. This increases the impact of the relative regulation on the relative revenues under limited information. Hence, it is expected that the difference between limited and full information is larger for power stations with medium or low relative regulation compared to those with high.

Comparing the relative revenues obtained in the two cases of limited flexibility, we see that power station B, E, G, I, L and M are closest to obtaining actual revenues π^{actual}, equal to the upper bound, $\pi^{clairvoyant}$. This result indicates that the operators for these power stations have implemented good dispatch strategies, resulting in the highest actual relative revenues. We should however be careful to deduce that the other operators have poor dispatch strategies, due to other constraints not being taken into account when calculating $\pi^{clairvoyant}$. We will return to possible reasons for large differences between the actual and clairvoyant relative revenues in the next section when discussing the relative flexibility value.

Looking at the revenues under unlimited flexibility, π^{max}, we see that power stations C, H and J also have the highest revenues in this case, ranging from 69.0 to 73.2 [EUR/MWh water]. This is again due to their low capacity factor; they only need to produce around 23 % to 34 % of the time to consume their average seasonal inflow. Given unlimited flexibility (with full information), they will then utilize the 23-34 % highest prices, thus obtaining high relative revenues. On the other end of the scale, with a capacity factor of 76 %, power station F needs to produce at maximum capacity during a majority of the periods and obtain the lowest relative revenue of 49.2 [EUR/MWh water]. The general observation is that the higher capacity factor, the lower is the relative revenue under unlimited flexibility, as predicted in Table 3.2.

The relative difference between π^{max} and $\pi^{clairvoyance}$ indicates how much the power station could gain by relaxing the reservoir and inflow restrictions. The power stations with a low relative difference ($\pi^{clairvoyant}$ close to the theoretical upper bound, π^{max}) have higher storage flexibility than the others. This difference is less than 1.3% for power stations B, E, G, L and M, which also are the power stations with highest relative regulation. Power station A, D, F, J and N show a difference of more than 5.8%, indicating that reservoir size and/or inflow seasonality limits their potential revenues significantly.

3.2.2 Relative flexibility value

The relative flexibility value shows how much of the revenues that are due to the storage flexibility of a power station. Figure 3.3 shows the estimated clairvoyant and actual flexibility values for each power station. The clairvoyant relative flexibility values range from 25.8 % to 51.3 %. These values represent the upper boundaries for the value created through storage flexibility for each producer. Power stations G, H, L and M have the highest clairvoyant flexibility value, all being above 40%. Hence, when utilizing their full potential, over 40 % of the revenues come from storage flexibility. Note that a high flexibility value does not necessary indicate that a power station is more flexible than power stations with lower flexibility value. Power station B and E have a average
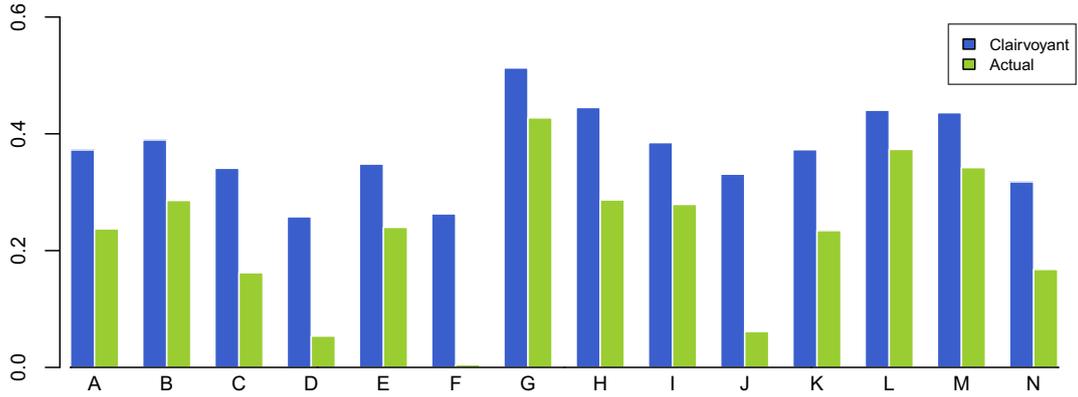
Figure 3.3: Estimated historic relative clairvoyant and actual flexibility value per producer. Relative flexibility value is the share of revenues (actual/clairvoyant) that exceeds the revenues for the run-of-river equivalent

flexibility value, despite being the most flexible in terms of reservoir size as seen by the relative difference between π^{max} and $\pi^{clairvoyance}$. Their high revenues under no flexibility reduces the relative flexibility value, as seen in Table 3.3 and Figure 3.2. Power station D and F have the lowest relative flexibility value. Both have values under 30 %, saying that over 70% of the potential revenues could be achieved without having a reservoir at all.

The estimates for the actual relative flexibility value range from 0.5 % to 42.7 %. On overall there seems to be a strong link between the clairvoyant and actual relative flexibility value, however the range is significantly wider is this case (42 % points compared to 25 % points). The difference between the clairvoyant and actual value is largest for those with low clairvoyant flexibility value, as for example for producer D, F and J. For these power stations only 0.5 % to 6.1 % of the actual revenues are due to storage flexibility. As explained earlier, this large difference is partly caused by the fact that full information gives a greater advantage under strict restrictions. All three power stations have below average relative regulation.

As previously mentioned, a large difference between the clairvoyant and actual relative revenues (or the flexibility values) should *not* be interpreted as a sign of bad operator performance. There are several reasons to expect a varying degree of differences between the actual and clairvoyant case among the power stations in out sample. First of all, the estimates shown here are based on nine seasons with data, leading to uncertainty in the results. Different production strategies in general and/or maintenance periods during high price periods could have large impact on the overall figures with such few seasons. Secondly, we have not taken into account other restrictions such as water flow restrictions up- or downstream of the generator etc., as mentioned earlier. Finally there could be large differences in the ability to forecast inflow for the different producers, affecting the advantage of having full information. For example snow measuring could give a more precise estimate for power stations located higher above sea level than those located in the low lands. By implementing a stochastic model, and using the same dispatch strategy for all producers, we are able to study the difference between the clairvoyant and actual relative flexibility value on more unbiased terms.

3.2.3   Main observations

There are several power stations that distinguish themselves as being more flexible than the rest in our sample. Power station B, E, G, L and M have a low relative difference between the revenues in the unlimited flexibility case and the revenues in the limited flexibility case. These are the plants with the highest relative regulation combined with having strong inflow seasonalities and medium capacity factors, providing them with both high storage flexibility and medium production flexibility. The average actual relative revenues for this group is about 9 % higher than the sample average. Power station C, H, I and J have lower storage flexibility, but all have low capacity factors giving them high production flexibility. Their average actual revenues are about the same as the sample mean. On the lower end of the scale we have power station A, D, F, K and N being the least flexible in terms of both storage and production capacity in our sample. On average they obtain actual revenues about 10 % below average. These results show that both storage and production flexibility have significant impacts on revenues. Note that the variation in revenues is fairly large within each group, and the results should be regarded as preliminary.

The average flexibility values for our sample are 37.2 % and 22.5 % in the clairvoyant and actual case respectively. This means that for the average power station, about one third of the potential revenues come from storage flexibility. For power stations with high clairvoyant flexibility value, there is potentially a lot to gain from investing in good dispatch models to exploit the storage flexibility. The difference between clairvoyant and actual flexibility value is notably large for several of the less flexible stations (D, F and J), indicating particularly challenging production planning conditions under limited information when having strict reservoir constraints.

The analysis above also shows that the defined levels of storage flexibility and the flexibility value provide a good framework for analyzing the revenues for hydropower stations with different degrees of flexibility. However, it is difficult to get an unbiased estimate for the actual revenues, as the actual revenues implicitly take into account restrictions not modeled in the other flexibility cases. To further evaluate the impact of the different factors on the revenues in each flexibility case, we will develop a comprehensive stochastic model for hydropower stations with reservoir. This model will provide us with sensitivities for all factors, and enables the study any type of hydropower station, with arbitrary technical characteristics.

Chapter 4

Stochastic model

In this section we introduce a complete stochastic model for a hydropower station with reservoir. With this model we can study the impact of different factors on the revenues and flexibility value by running simulations for a large number of seasons. The model consists of several components; a producer dependent inflow model, a general spot model for electricity prices and a model for the overall reservoir content in Norway in order to account for the correlation between inflow and spot prices. The data series generated by these models are then used as input in the dispatch model, mimicking production strategy of operators. The model components are shown in Figure 4.1, and explained in detail below. This section is rounded off with a step by step example on how the model is implemented on our empirical data.
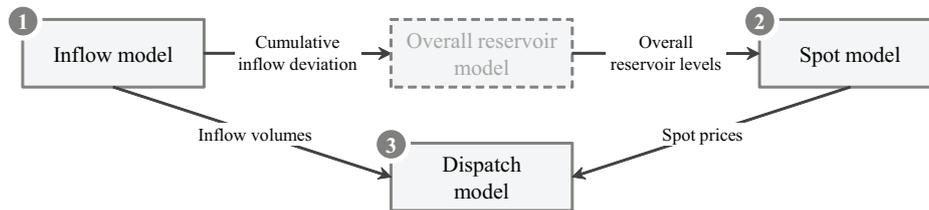
Figure 4.1: Overview of the stochastic model, showing the components and their interactions. The overall reservoir model is included to account for the correlation between each power station's inflow and the spot price.

4.1   Modeling inflow

In order to capture the dynamics of inflow for different hydropower stations in one general model, this model needs to be versatile enough to capture several different inflow dynamics. Recalling the inflow data in Chapter 2, such a model must take into account the varying degree of seasonality and autocorrelation of the inflow series. One way to model inflow for a specific power station is to draw from a pool of historic inflow series, like done in the EMPS model developed by SINTEF Energy Research which is described in for example Doorman [2009]. However, if only a few seasons of inflow data is available, this method results in a relatively small sample space. In such a case it is preferable to use a stochastic model fitted to the historic data available. This enables us to generate large numbers of inflow series using general simulation methods such as Monte Carlo. In the following sections we will formulate a stochastic inflow model, and describe how to estimate its parameters using empirical data.

4.1.1 Model formulation

Although the flow of water is a continuous process, it is only measured over discrete time intervals. Due to the high uncertainty in the measuring, these time intervals are normally quite long (ranging from one day to one week). This motivates the use of discrete models for the inflow series. There are several ways to model the seasonality of the inflow. One way is to add a deterministic seasonality function $f^I(t)$ to an autoregressive (AR) model of order p, X_t, with a long range mean of zero. This model is from now on referred to as *model 1*, with dynamics defined in the following equations:

$$I_t = f^I(t) + X_t \tag{4.1}$$

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \xi_t$$

where I_t is the inflow during time period t, $|\varphi_i| < 1$ are the coefficients for the autoregressive model and ξ_t are assumed to be i.i.d. normal distributed errors with mean zero and standard deviation σ^I. In the model above, the variance of the inflow I_t is assumed to be independent of time t, neglecting a possible seasonal variation. In case of strong heteroskedasticity, the assumption of the error term being equally distributed is violated.

One way to take into account time-varying variance, is to assume that the variance is proportional to the time-varying seasonal mean defined by $f^I(t)$. Here a transformation of the inflow follows a autoregressive process X_t similar to the one above, but with different coefficients φ_i and different errors ξ_t. This model is from now on referred to as *model 2*, with dynamics defined in the following equation:

$$I_t = f^I(t)(X_t + 1) \tag{4.2}$$

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \xi_t$$

Here it is assumed that $f^I(t) > 0$ for all t. In this model, the variance of I_t will depend on the expected amount of inflow at time t. It is also possible to formulate a process σ_t^I for the standard deviation, separately from $f^I(t)$, to achieve time-varying variance. However, to avoid adding more complexity, these options are not further examined.

The seasonality function $f^I(t)$ needs to be versatile enough to fit inflow trends of many different power stations. These trends could consist of one or multiple peaks of different size and width per season. In addition, many power stations experience a strong decay in the general level of inflow during the drawdown season. These dynamics can be captured by the a seasonality function on the following form:

$$f^I(t) = Ae^{-at} + Be^{-bt}\sin(C + ct) \tag{4.3}$$

where A, a, B, b, C and c are constant parameters. Here Ae^{-at} represents a vertical shift of the inflow level, starting at A and with decay in time defined by a. $Be^{-bt}\sin(C + ct)$ represents a sinusoidal peak function with frequency c, horizontal shift C and decay b. B controls the amplitude of the peaks. This function is smooth and continuous in time, and is flexible enough to mimic many different inflow trend forms. See Figure 4.2 for several examples of seasonality functions of this form.
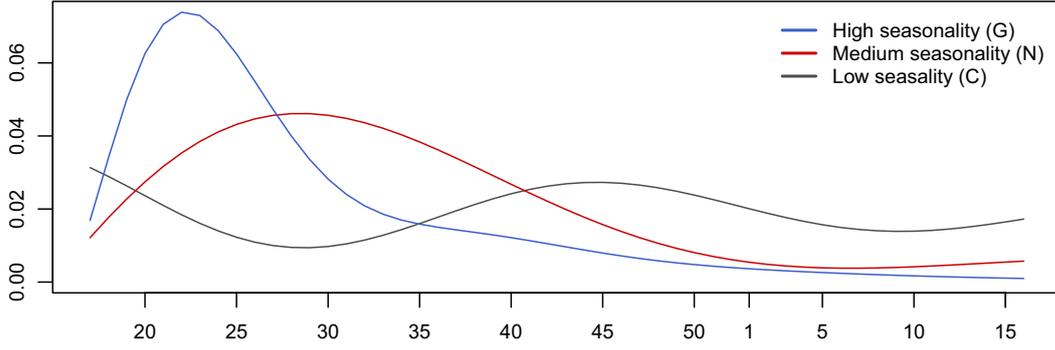
Figure 4.2: Inflow seasonality function examples with high, medium and low degree of seasonality, using estimated parameters for power stations G, N and C respectively.

4.1.2 Model estimation

In order to estimate model parameters based on historic inflow series, we perform a regression on the following equation for model 1:

$$I_t = Ae^{-at} + Be^{-bt}\sin(C + ct) + X_t \tag{4.4}$$

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \xi_t$$

and for model 2:

$$I_t = \left(Ae^{-at} + Be^{-bt}\sin(C + ct) \right)(X_t + 1) \tag{4.5}$$

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \xi_t$$

Due to the non-linearity of the function $f^I(t)$, it is not possible to use linear regression to estimate all of the coefficients. Hence we use the solution of first finding an estimate for the seasonal trend function, denoted $\hat{f}^I(t)$ through non-linear optimization, before doing a regression on the residuals (in the case of model 1) and on a transformation of the inflow dependent variable $(I_t/\hat{f}^I(t) - 1)$ (is the case of model 2). This is summarized in the following equations:

$$\hat{f}^I(t) = f(\hat{A}, \hat{a}, \hat{B}, \hat{b}, \hat{C}, \hat{c}, t) = f^I(\hat{\Theta}, t) \tag{4.6}$$

where

$$\hat{\Theta} = \operatorname*{arg\,min}_{A,a,B,b,C,c} \sum_{\forall t} \left(I_t - f^I(A, a, B, b, C, c, t) \right)^2 \tag{4.7}$$

Since the problem above is non-convex (due to the sinusoidal function), the (local) solution found from standard non-linear optimization depends on the starting values. Hence, we manually adjust the function to fit the historic inflow through visual observation, and define ranges for the different parameters, before running the optimization. This way we ensure that our solution is close to the global minimum. To estimate the coefficients φ_i in each case, we use the following linear equation:

$$\hat{X}_t = \sum_{i=1}^{p} \varphi_i \hat{X}_{t-i} + \xi_t \tag{4.8}$$

40

where $\hat{X}_t = I_t - \hat{f}(t)$ for model 1 and $\hat{X}_t = (I_t/\hat{f}(t)-1)$ for model 2. The equation above can be solved using ordinary least squares to obtain estimates for both the coefficients φ_i and the standard deviation σ^I. The order of the AR process is determined depending on the significance of the coefficients φ_i.

4.2 Modeling spot prices

In this section we present a model for *daily* spot prices. As discussed in Chapter 2, daily electricity spot prices (and their logarithm) have certain characteristics such as mean reversion, seasonality and spikes. To account for these special characteristics, several models are proposed in literature. We will briefly mention some different types of stochastic models, before presenting the model of choice. The modeling of seasonality is done separately from the stochastic component, and hence the following models should be regarded as models for *deseasonalized* prices.

- **One-factor mean reversion models:** As a starting point we consider a Brownian motion driven Ornstein-Uhlenbeck process with possibly time-varying mean:

$$dS_t = \beta(\mu_t - S_t)dt + \sigma dB_t$$

  This one-factor model reverts to a mean μ_t with speed of reversion β. The standard deviation of the diffusion part B_t is defined by σ. A geometric version of the process above is used to model commodity prices by Schwartz [1997]. Although this model captures the mean reversion of electricity spot prices, it does not allow for jumps which are obviously apparent in the data. However, this model could suit well to price series with few or no jumps.

- **Jump diffusion models:** Several models extend Schwartz' model (or other similar one-factor models) to include a jump component. In the general case this can be expressed in the following way:

$$dS_t = \beta(\mu_t - S_t)dt + \sigma dB_t + dZ_t$$

  where $Z_t = \sum_{i=1}^{N(t)} D_i$ is the jump component, with $\{N(t), t > 0\}$ being a counting process and $\{D_i, i \geq 1\}$ the random jump sizes. The other factors are as defined above. One example of such a process is the one used by Geman and Roncoroni [2006]. In their model the sign of the jump is determined by the current price level; if the price is below (above) a certain threshold, jumps are expected to be positive (negative). This ensures that prices return to normal levels after the occurrence of spikes. However, this model does not account for negative jumps when prices are at "normal" levels. A similar jump diffusion model is proposed by Weron et al. [2004a], where positive jumps are immediately followed by a negative jump to capture the fast mean reversion. However, this does not allow the mean reversion to last longer than a predefined number of time steps.

- **Regime switching models:** To allow for spikes to last longer than one time step, Weron et al. [2004b] propose a regime switching model. Two dynamics are defined in this case; one for the normal regime and one for the spike regime.

$$S_t = \begin{cases} Y_{1,t} & \text{if in the base regime} \\ Y_{2,t} & \text{if in the spike regime} \end{cases}$$
$$dY_{1,t} = (c_1 - \beta_1 Y_{1,t})dt + \sigma dB_t$$
$$Y_{2,t} \sim \text{logN}(c_2, \sigma)$$

The variable that determines the current state follows a Markov chain with two possible states, and transition between the regimes are determined by the transition probabilities. Several other authors propose different types of regime switching models with two or more regimes, such as in Haldrup and Nielsen [2006]. Although regime switching models capture the changing behavior of the spot price, they have limitations. The price can never be in more than one regime at the time, implying that in the spike regime it will only follow the dynamics of this regime, independent of variations in the base signal.

- **Multi-factor models:** Benth et al. [2008] and Lucia and Schwartz [2002] model spot prices with multi-factor models, assuming the price be a sum of several Levy-driven (geometric) Ornstein-Uhlenbeck processes:

$$S_t = \sum_{i=0}^{N} Y_{i,t} \tag{4.9}$$

$$dY_{i,t} = \theta_i(\mu_{i,t} - Y_{i,t}) + dL_{i,t} \tag{4.10}$$

Where N is the number of factors. For the processes $Y_{i,t}$ for $i = 1, \ldots, N$, $\theta_i > 0$ is the mean reverting rate and $\mu_{i,t}$ is a possibly time dependent mean reverting level. This allows for the spike component to be modeled as one factor with possibly high and rare (absolute) jumps with fast mean reversion, and the "base" component as a Brownian motion driven process with slower mean reversion. Being independent processes, the mean reversion of the Brownian motion driven process is taken into account throughout the series, also during spikes. It is also relatively easy to add complexity and refine the model by including more factors.

4.2.1 Model formulation

In this study we apply a multi-factor model, due to both it's simplicity and good ability to capture the dynamics of electricity prices. Specifically, daily *deseasonalized log prices* are modeled as a sum of two factors, using the logarithm to avoid negative spot prices. The following model formulation is to a large degree based on the approach of Meyer-Brandis and Tankov [2008], who apply a two-factor model for EEX[1] spot prices. Note that we use capital S for the spot price and small s for its logarithm; $\ln S_t = s_t$ in the remainder of this study.

The seasonality of spot prices is modeled with the deterministic function $f^S(t)$ is estimated and subtracted from the original log price data. This function consists of a constant, a linear trend and a linear combination of sine and cosine functions with periods of 6 and 12 months:

$$f^S(t) = \beta_0 + \beta_1 t + \beta_2 \sin(\frac{2\pi t}{365}) + \beta_3 \cos(\frac{4\pi t}{365}) + \beta_4 \sin(\frac{2\pi t}{365}) + \beta_5 \cos(\frac{4\pi t}{365}) \tag{4.11}$$

where t is time in days. The deseasonalized spot price, s_t^{des}, is modeled as the sum of two independent Levy driven Ornstein-Uhlenbeck processes:

$$s_t^{des} = s_t - f^S(t) \tag{4.12}$$

$$s_t^{des} = Y_{1,t} + Y_{2,t} \tag{4.13}$$

---

[1]European Energy Exchange.

where $Y_{1,t}$ and $Y_{2,t}$ are as defined in Equation (4.10). In our model we chose $Y_{2,t}$ to be a spike process to model the high jumps and their fast mean reversion, while $Y_{1,t}$ models the remaining base variations in log prices. According to our observations in Chapter 2, both this processes are mean reverting, but with possibly different mean levels and reversion speeds. The process $Y_{1,t}$ is referred to as the base process and is assumed to be an Ornstein-Uhlenbeck process driven by a Brownian motion, B_t with a possibly time dependent mean level, $\mu_{1,t}$. This leads to the following dynamics for the base signal:

$$dY_{1,t} = \theta_1(\mu_{1,t} - Y_{1,t})dt + \sigma^S dB_t \tag{4.14}$$

The spike process is assumed to revert to zero, $\mu_{2,t} = 0$, and is driven by a compound Poisson process $Z_t = \sum_{i=1}^{N(t)} D_i$ where $\{N(t); t > 0\}$ is a Poisson process with intensity λ, and the jump sizes $\{D_i; i \geq 1\}$ are i.i.d random variables:

$$dY_{2,t} = -\theta_2 Y_{2,t} + dZ_t \tag{4.15}$$

The absolute spot price is given as the exponential of the sum of the components defined above:

$$S_t = e^{f^S(t) + Y_{1,t} + Y_{2,t}} \tag{4.16}$$

Before going into more detail on how to separate and estimate the different factors of this price model, we will look at how to calculate expected spot prices assuming the dynamics defined above.

Expected spot price

A view of the price development in the future is essential for hydropower station's when doing the production planning. Having a model for the spot price, we can derive the expected spot at a future time t. Assuming that the components of the spot price are uncorrelated, the expectation of each component can be computed separately. Hence, being at time s, the expected spot price at time $t > s$ is defined as:

$$\begin{aligned}
\mathbb{E}[S_t|\mathcal{F}_s] &= \mathbb{E}[e^{f^S(t) + Y_{1,t} + Y_{2,t}}|\mathcal{F}_s] \\
&= e^{f^S(t)}\mathbb{E}[e^{Y_{1,t}}|\mathcal{F}_s]\mathbb{E}[e^{Y_{2,t}}|\mathcal{F}_s]
\end{aligned} \tag{4.17}$$

where $\mathcal{F}_s$ is the filtration at the current time s, including knowledge of $Y_{1,s}$ and $Y_{2,s}$ In this case we assume the mean $\mu_{1,t}$ to be deterministic. In cases where the mean is stochastic, as later defined in Section 4.2.3, this is a unrealistic assumption when $t - s$ is large. In that case the expression for expected spot price should only be used for relatively short horizons.

The expectation of the base component is calculated using the closed form of the geometric Brownian motion driven Ornstein-Uhlenbeck process (also referred to as the Dixit and Pindyck model in financial literature):

$$\ln \mathbb{E}[e^{Y_{1,t}}|\mathcal{F}_s] = e^{-\theta_1 t}Y_{1,s} + \int_s^t e^{-\theta_1(t-u)}\theta_1\mu_{1,u}du + \ln \mathbb{E}[e^{\int_s^t \sigma e^{-\theta(t-u)}dB_u}|\mathcal{F}_s] \tag{4.18}$$

where the last expectation is given by

$$\ln \mathbb{E}[e^{\int_s^t \sigma^S e^{-\theta_1(t-u)}dB_u}|\mathcal{F}_s] = \frac{1}{2}\sigma^2 \int_s^t e^{-2\theta(t-u)}du = \frac{\sigma^2}{4\theta_1}(1 - e^{-2\theta_1(t-s)}) \tag{4.19}$$

Note that for notational convenience, σ stands σ^S in the equation above. For the spike component, being driven by a compound Poisson process, the expectation is computed as follows:

$$\ln \mathbb{E}[e^{Y_{2,t}}|\mathcal{F}_s] = e^{-\theta_2(t-s)}Y_{2,s} + \ln \mathbb{E}[e^{\int_s^t e^{-\theta_2(t-u)}dZ_u}|\mathcal{F}_s] \tag{4.20}$$

As the last expectation depends on the distribution of Z_u, and hence on the chosen distribution for the jump sizes, D, it is in general complex to solve analytically. However, it is possible to estimate this expectation by simulating large samples for Z for all wanted intervals of $(t-s)^2$. For a proof of the general solution of the expectation for geometric multi-factor spot models, see Benth et al. [2008]. Note that the expected spot price is not necessarily equal to the forward price, which needs to be calculated using risk-neutral probabilities taking into account a possible risk premium.

4.2.2 Model estimation

In this section we explain how to estimate the parameters of the spot model using discrete historical data. The estimation of the seasonality function defined in Equation (4.11) is relatively straight forward: the function $f^S(t)$ is fitted to historical log prices using ordinary least squares on the following formula:

$$s_t = f^S(t) + u_t \tag{4.21}$$

Where the residuals u_t from this regression form the deseasonalized log prices. These residuals contain spikes and hence are obviously not Gaussian, violating the normality assumption. However, note that the purpose of this regression is to estimate the seasonal component, not to derive confidence intervals or test the estimated parameters. Hence the normality assumption is not necessary, as it has no impact on the parameters.

Separating the spike and base process

To estimate the parameters of the base signal, $Y_{1,t}$ and the spike process, $Y_{2,t}$, we first separate these two components for the historic deseasonalized log price series. The challenge is to determine which price variations that are caused by jumps, and which that are caused by variations in the underlying base signal. Relying on Meyer-Brandis and Tankov [2008], we choose a procedure referred to as *hard thresholding*, to filter out spikes.

The basic idea is to filter out the spikes one by one, by optimally placing one custom sized spike at the time, taking into account the mean reversion of both components. We assume that the spike series is a deterministic function $g(t)$;

$$g(t) = \sum_{i=1}^{M} D_i \mathbf{1}_{t \geq \tau_i} e^{-\theta_2(t-\tau_i)} \tag{4.22}$$

where M is the number of spikes, D_i are the jump sizes and τ_i are the jump times. The part $\mathbf{1}_{t \geq \tau_i} e^{-\theta_2(t-\tau_i)}$ represents the mean reversion of the placed spikes. The speed of mean reversion, θ_2, is at this point assumed known, and we will come back to how to determine this later. Initially this spike series is set equal to zero, $M = 0$. The first spike is placed at the time τ_1 and with size D_1, which reduces the variance of the price series

---

[2]The expectation for such a jump process with low intensity is expected to converge very slowly, due to large variation in the sample. Hence a large number simulations, in our case 10 000 000, are needed to achieve convergence

the most. After finding the first spike, denoted $\{D_1, \tau_1\}$, the spike function is updated and subtracted from the initial deseasonalized log prices. Then the process starts over again, placing spike number two etc. A target variance of the base signal determines when to stop the spike filtering. When this target level is achieved, the two separate price components are obtained by setting:

$$Y_{2,t} = g(t) = \sum_{i=1}^{M} D_i \mathbf{1}_{t \geq \tau_i} e^{-\theta_2(t-\tau_i)} \tag{4.23}$$

$$Y_{1,t} = s_t^{des} - g(t) \tag{4.24}$$

The spike intensity λ is estimated by the ratio of the total number of spikes found, M, and the total number of time steps in our data, N; $\hat{\lambda} = \frac{M}{N}$. Further, the distribution for the spike sizes is found using the series of placed spike sizes D_i. For the mathematical background, and a more detailed description of spike detection through hard thresholding, we refer to Section A.1 in the Appendix. and Meyer-Brandis and Tankov [2008]. The spike filtering procedure is illustrated in Figure 4.3, showing an example of daily deseasonalized log prices separated into the base and the spike process after placing two spikes.
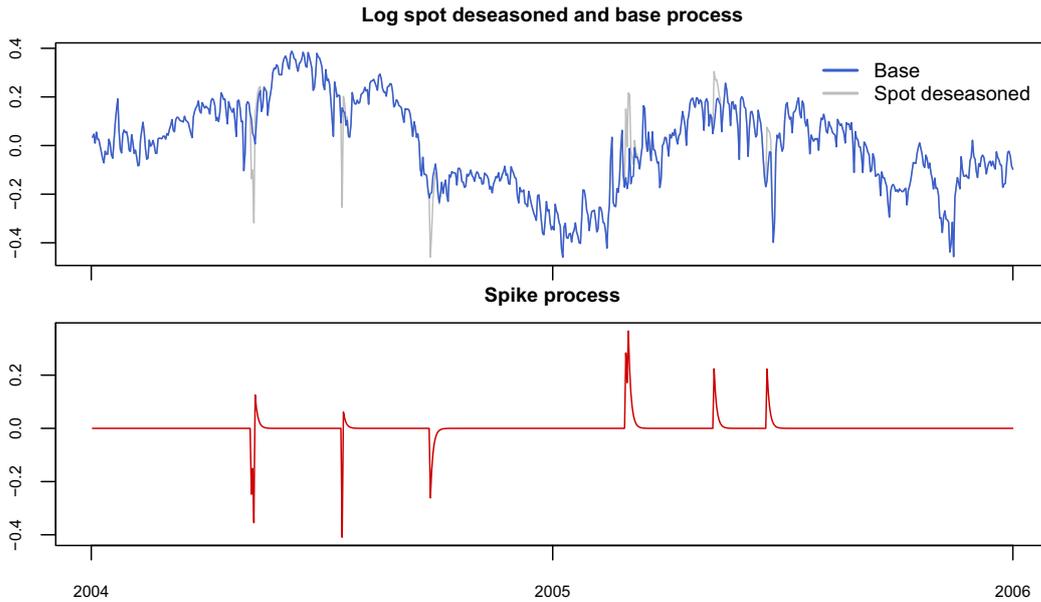
Figure 4.3: Illustration of spike detection procedure for spot prices. The plot show the daily deseasonalized log price (gray) separated into the base process (blue) and the spike process (red) after placing two spikes.

Base signal

After subtracting the estimated spike series from the log price series, the parameters of the base signal can be estimated. In our model we assume that the mean reverting level of the base process is known, and we can subtract this from the original base process to

obtain the following dynamics:

$$X_t = Y_{1,t} - \mu_{1,t}$$
$$dX_t = d(Y_{1,t} - \mu_{1,t}) = dY_{1,t} \tag{4.25}$$

where the last equality only holds when the mean reverting level is constant. X_t will then be a mean reverting process with mean reverting level 0^3. The discrete dynamics when following Dixit and Pindyck [1994] are:

$$dX_t = -\theta_1 X_t dt + \sigma^S dB_t \tag{4.26}$$
$$X_{t+1} - X_t = (e^{-\theta_1} - 1)X_t + u_t \tag{4.27}$$
$$X_{t+1} = e^{-\theta_1} X_t + u_t \tag{4.28}$$
$$X_{t+1} = \beta X_t + u_t \tag{4.29}$$

where u_t are assumed to be i.i.d normal distributed errors with mean zero standard deviation σ^u. To estimate the parameters θ_1 and the variance of the Brownian motion σ^S, we perform a OLS regression on the log price series remaining after we have filtered out the spikes and subtracted the mean reverting level. The variance σ^S is estimated using the standard error of the regression, σ^u, and the two parameters are determined by the following equations

$$(\sigma^S)^2 = \frac{2\theta_1(\sigma^u)^2}{1 - e^{-2\theta_1}}$$
$$\theta_1 = -ln\beta$$

as in Dixit and Pindyck [1994].The mean reverting level is preliminary set equal to zero, $\mu_{1,t} = 0$. This level is reasonable since the deseasonalized log prices are distributed around zero (follows from Equation (4.12)), and the spike series around zero. However, to incorporate the correlation with inflow, we will include a time dependent mean reversion level below.

4.2.3 Spot and inflow correlation model

As shown in Chapter 2, there is a correlation between a power station's inflow and the spot price. The inflow and spot processes presented above are assumed to be independent, and hence do not take this into account. Since in reality, high inflow often coincides with low spot, this would result in unrealistic high (low) revenues in cases with high (low) inflow. To make our model more realistic, we incorporate the correlation with inflow by extending the spot model to take into account the *overall reservoir level*.

Spikes are often a result of supply shortage (or surplus) do to plant outages rather than sudden changes in inflow. Hence the correlation is incorporated in the base process, $Y_{1,t}$, or more specifically in the mean reverting level of this process, $\mu_{1,t}$. It is possible to incorporate the correlation by letting each power station's inflow influence the mean reverting level of the spot directly. However, since the degree of correlation varies between power stations, this would require us to estimate separate spot models for each power station. This is not desirable, and hence we propose to link inflow and spot through an model for the overall reservoir level instead, as explained below.

---

[3]We will also use this approximation when $\mu_{1,t}$ is time varying, but with a derivative close to zero.

Inflow and overall reservoir dependency

The empirical data shows a positive correlation between deviation from normal cumulative inflow the last 20 weeks and the deviation from normal overall reservoir level.

$$\Delta i_t^{cum} = \frac{I_t^{cum} - \bar{I}_t^{cum}}{\bar{I}_t^{cum}} \tag{4.30}$$

$$\Delta r_t^{overall} = \frac{R_t^{overall} - \bar{R}_t^{overall}}{R_t^{overall,max}} \tag{4.31}$$

where I_t^{cum} is the actual cumulative inflow the last weeks and $\bar{I}_t^{cum}$ is the normal (historic average) cumulative inflow during the same period. $R_t^{overall}$ is aggregated reservoir level at the end of period t, and $\bar{R}_t^{overall}$ is the normal (historic average) aggregated reservoir level. Their relationship is here modeled in the following way:

$$\Delta r_{i,t}^{overall} = \beta_{i,1}\Delta r_{i,t-1}^{overall} + \beta_{i,2}\Delta i_{i,t}^{cum} + \epsilon_{i,t} \tag{4.32}$$

where t is measured in weeks, $\epsilon_{i,t}$ are assumed to be i.i.d errors with mean zero and variance σ_i^R, and i indicates the power station. After estimating $\beta_{i,1}$ and $\beta_{i,1}$, we can simulate the corresponding overall reservoir level $r_{i,t}^{overall}$ when the inflow series $I_{i,t}$ is given. Daily overall reservoir deviations are obtained by linear interpolation.

Spot and overall reservoir dependency

We let the mean reverting level of the base process depend on each power station's inflow through the overall reservoir level. This relationship is formulated as follows:

$$\mu_{1,t} = \eta\Delta r_{t-1}^{overall} \tag{4.33}$$

where $\Delta r_t^{overall}$ is the deviation from normal for the overall reservoir level in Norway at the end of period t and η is a parameter that determines to which degree the correlation between the base process $Y_{1,t}$ and $\Delta r_t^{overall}$ should be incorporated. η is estimated in such a way that the correlation between the simulated base process of the price and the deviation from normal reservoir level converge to the historic value. Note that $r_t^{overall}$ (and hence $\mu_{1,t}$) is autocorrelated and with relatively small variance in the error terms from day to day. Hence we estimate the base process by subtracting $\mu_{1,t}$ from $Y_{1,t}$ and assume that the resulting process X_t is mean reverting to zero, as previously explained.

4.3   Modeling dispatch

In order to simulate the actual revenues obtained by each power station, a model for the dispatch strategy is needed. When planning the dispatch under limited information, power stations need to take into account the uncertainty of inflow and spot prices. In practice, producers try to maximize their expected revenues using advanced models to estimate the marginal value of the water in their reservoir. These models are normally divided into long, medium and short term models, which are re-ran on a regularly basis in order to do long term planning and to determine the production strategy for the next day or week, as explained in Doorman [2009]. These models are highly complex and computer-intensive, resulting in long simulation times, and thus they are not applicable for this study. Instead we try to replicate the operators production planning by using two different (simplified) dispatch strategies. The first is a simple and intuitive decision rule based on normal reservoir level, while the second dispatches water according to a threshold function which approximates the marginal water value. Both these dispatch strategies are simple to implement and can easily be used in practice.

4.3.1 Simple decision rule

The simple decision rule is based on the assumption that producers want to stay as close to the normal reservoir level as possible. Hence, the amount of water to dispatch/produce P_t is determined with the aim of minimizing the difference between the current reservoir level and the normal target level $\bar{R}_t$:

$$P_t = min\{P_{max}, max(0, R_{t-1} + I_t - \bar{R}_t)\} \tag{4.34}$$

where P_{max} is the maximum production capacity during one time step t, R_t the reservoir level at end of time step t and I_t the inflow occurring in the current period (assumed to be known to the producer). The target level $\bar{R}_t$ can be determined by using the historic average reservoir, or through simulations assuming full information. The term within the max clause ensures that if the reservoir is below the target level, the production should be zero. If on the other hand the reservoir is higher, the power station should produce the amount needed to reach the target level. The production is limited by the installed capacity, as accounted for by the min clause. This simple decision rule is naive in the way that it ignores inflow forecasts and information about future spot prices. Still, it should at least provide a lower bound for the actual revenues obtained by the power stations.

4.3.2 Threshold function

Several sources justify the use of a threshold function in the production planning. In Chapter 2 we saw that power stations often choose to produce close to maximum capacity or not at all, which suggest that the producers use a bang-bang strategy for the dispatch planning. This is also in line with the study of Kolsrud and Prokosch [2009], who use a real option approach to estimate the value of the price threshold. This threshold function estimates the marginal water value depending on several factors, and the producer should produce if the current spot price is higher than this value, and wait if else. Näsäkkälä and Keppo [2008] argue for the use of such a threshold function, and provide a parameterization of this function based on empirical data for one Norwegian power station. We will use a customization of this parameterization, taking into account the findings of Kolsrud and Prokosch [2009] on how this function should depend on the current reservoir level. This leads to the following parameterization of the price threshold Λ:

$$\Lambda(t, \tilde{S}_t, \tilde{I}_t, R_t) = \alpha_s \tilde{S}_t e^{-\frac{\alpha_t}{T-t} - \alpha_i \Delta \tilde{i}_t} + \alpha_r \ln(\frac{1}{r_t} - 1) \tag{4.35}$$

$$\alpha_s, \alpha_t, \alpha_i, \alpha_r \geq 0$$

where T is the last time period in the current season and r_t is the relative reservoir level at time t. $\Delta \tilde{S}_t$ is the average of the expected spot price for a specific time interval ahead, calculated using Equation (4.17) for each day in the interval. $\tilde{I}_t$ is the average daily inflow the next seven days, relative to normal inflow in the same period. Hence we assume that the inflow the upcoming seven days is known to the producers when planning the dispatch. In practice good weather forecasts are available in the short run, which justifies this assumption.

While Näsäkkälä and Keppo [2008] suggest that the water value should be exponential with respect to the current reservoir, we choose instead to include a s-shaped function $\ln((r_t)^{-1} - 1)$ based on the findings in Kolsrud and Prokosch [2009]. When the reservoir

is close to its minimum (maximum), the threshold function goes to infinity (minus infinity). This reflects the producer's desire to avoid extremely low or high reservoir levels, due to the risk of running dry or spilling water by overflow. The rest of the variables are incorporated as in Näsäkkälä and Keppo [2008]; the expected spot price has a linear positive effect on the threshold, the inflow a negative exponential, and when the time approaches end of the season the first term in the threshold will decrease exponentially to zero. In Section 4.4.1 Figure 4.9 represents an example of the estimated threshold function, showing dynamics that are close to the theoretical threshold function found in Kolsrud and Prokosch [2009].

Estimation of threshold parameters

To find the optimal parameters α_s, α_t, α_i and α_r, we try a large number of different combinations within a set of realistic ranges for each, as done by Näsäkkälä and Keppo [2008]. The dispatch for each parameter combination is simulated using historical inflow and price data. The expected spot price is calculated using forward and future prices on Nord Pool, as discussed in Chapter 2. For each power station the parameter set which gives the highest revenues is chosen, and used for the threshold function in the stochastic model.

4.4   Implementing the model

It's now time to implement the different components described above and integrate them into one complete model. The implementation mainly consist of three steps; first we estimate the producer *independent* parameters, or more specifically the parameters of the spot model. Second we estimate the producer *dependent* parameters related to the inflow, overall reservoir and the threshold function used in the dispatch model. Finally we implement simulation methods for the different components and integrate the modules into a single procedure. All of the calculations and operations described below are performed using the statistical computing environment R on standard desktop computers. We want to point out that even though some of the model assumptions do not seem to hold for our empirical data, as described in more detail below, the verdict on the model's overall performance is left for the Chapter 5.

4.4.1   Estimated producer *independent* parameters

The parameters for the spot model are estimated based on historic daily prices, using the historic overall reservoir content to account for the correlation with inflow. Figure 4.4 shows the estimated components of the historic log prices. The upper plot shows the estimated seasonal function, while the lower plot shows the deseasonalized log prices separated into the spike and base process. The estimated parameters of the seasonal function and the base process are shown in Table 4.1 and 4.2 respectively. Note that the estimated linear trend has a relatively high coefficient of 0.0003, which corresponds to a annual trend in absolute spot prices of 11.6 %. This is surprisingly high, and may be due to the relatively low number of seasons in our sample in combination with particular low (high) prices in the beginning (end) of the sample period.

Descriptive statistics for the residuals from the OLS regression on the base process are also given in Table 4.2. The high excess kurtosis indicates fat tails, and the positive skew of 0.38 shows asymmetry in the distribution of the residuals. In addition, the

Table 4.1: Regression results for the seasonal component in the spot model. The significance of the coefficients is calculated under the assumption of normal distributed residuals, which is clearly violated as seen by the p-value of the Jarque-Bera normality test (null-hypothesis is that the residuals are normal distributed). The p-value of the Breusch-Pagan test with the null-hypothesis of constant variance in the residuals is also shown.

Coefficient	Estimate	Std. error	t-value	p-value	Significance
β_0	2.7080	0.0116	233.12	0.0000	***
β_1	0.0003	0.0000	57.16	0.0000	***
β_2	-0.1125	0.0082	-13.71	0.0000	***
β_3	-0.1158	0.0082	-14.14	0.0000	***
β_4	-0.0127	0.0082	-1.55	0.1217	
β_5	0.0268	0.0082	3.27	0.0011	**

Adjusted R-squared: 51.2 %, Jarque-Bera: 0.0, Breusch-Pagan: 0.0

Significance codes: 0 (***), 0.001 (**), 0.01 (*), 0.05 (.), 0.1 ( )

Table 4.2: Estimated parameters of the base process, $Y_{1,t}$, and p-values from Jarque-Bera normality test (under the null hypothesis of normal distributed residuals) and Breusch-Pagan heteroskedasticity test (under the null hypothesis of homoskedastic residuals). The kurtosis shown is actually *excess* kurtosis.

Coefficient	Estimate	Std. error	t-value	p-value	Significance
X_t	-0.0210	0.0034	-6.23	0.0000	***
	Mean	Variance	Skewness	Kurtosis	
u_t	-0.0001	0.0048	0.3804	1.9606	

Adjusted R-squared: 1.0 %, Jarque-Bera: 0.0 , Breusch-Pagan: 0.0

Significance codes: 0 (***), 0.001 (**), 0.01 (*), 0.05 (.), 0.1 ( )

Jarque-Bera test (see Jarque and Bera [1987]) does reject the null hypothesis of normal distributed residuals. This suggests that optimally, a distribution different from normal should be chosen to provide a better fit for the historic data. However, this would complicate the derivation of the expected price, and hence is not desirable. For simplicity, we assume the base process to be Brownian motion driven.

The descriptive statistics for the spike sizes are given in Table 4.3. The distribution of

Table 4.3: Estimated parameters of the spike process, $Y_{2,t}$ and descriptive statistics and fitted distribution parameters of the *absolute* spike sizes. Both negative and positive absolute spikes sizes are log-normal distributed, with μ and σ as shown. θ_2 are estimated found on the parameters used by Meyer-Brandis and Tankov [2008] and using visual inspection on both the autocorrelation function and how the spike detection algorithm performs under different values. The kurtosis shown is actually *excess* kurtosis.

	Mean	*Variance*	*Skewness*	*Kurtosis*	μ	σ
$D_i \geq 0$	0.3761	0.0433	2.63	7.5717	-1.0740	0.3982
$D_i < 0$	0.3288	0.0105	1.53	1.8455	-1.1513	0.2672
$\hat{\lambda}$ =0.029, θ_2 =0.5, $P(D_i \geq 0)$ =0.606						

spike sizes are determined by optimally fitting a distribution to the histogram of spike sizes. In our data, with both positive and negative spikes, we fit a different distribution for positive and negative jumps. Both spike sizes were fitted to log-normal distributions, but with different mean and standard deviation. The detected spikes shown in Figure 4.4 show indication of clustering, however for simplicity we still assume a constant spike intensity $\hat{\lambda}$ in our model.

Figure 4.5 shows one simulation of the spot price and its components together with the historical prices. On overall, the stochastic spot model seems to show the same characteristics as the historical prices. The red line shows the mean reversion level driven by the deviation from normal for the overall reservoir level. By incorporating the correlation with overall reservoir content, the model manages to recreate periods of historical high prices not caused by spikes or seasonal variations.

### 4.4.2 Estimated producer *dependent* parameters

In the example below we estimate the needed parameters for the power station denoted by N in our sample. The procedure is the same for all power stations in our sample, see Table A.1 in the Appendix for estimated producer dependent parameters for the other stations.

Estimating inflow model

Figure 4.6 shows the historic relative weekly inflow for power station N, with data from each season representing one gray line. The seasonality function, fitted to the historical data using non-linear least squares, seems to give a good representation of the mean inflow. The fitted parameters are given in Table 4.4. An AR process of order one fits best, as p-values turn out high when using higher orders. The p-value for the Breusch-Pagan test (see Breusch and Pagan [1979]) on residuals for model 1 suggests the rejection of the hypothesis of the residuals being homoscedastic. As seen in the empirical data, the
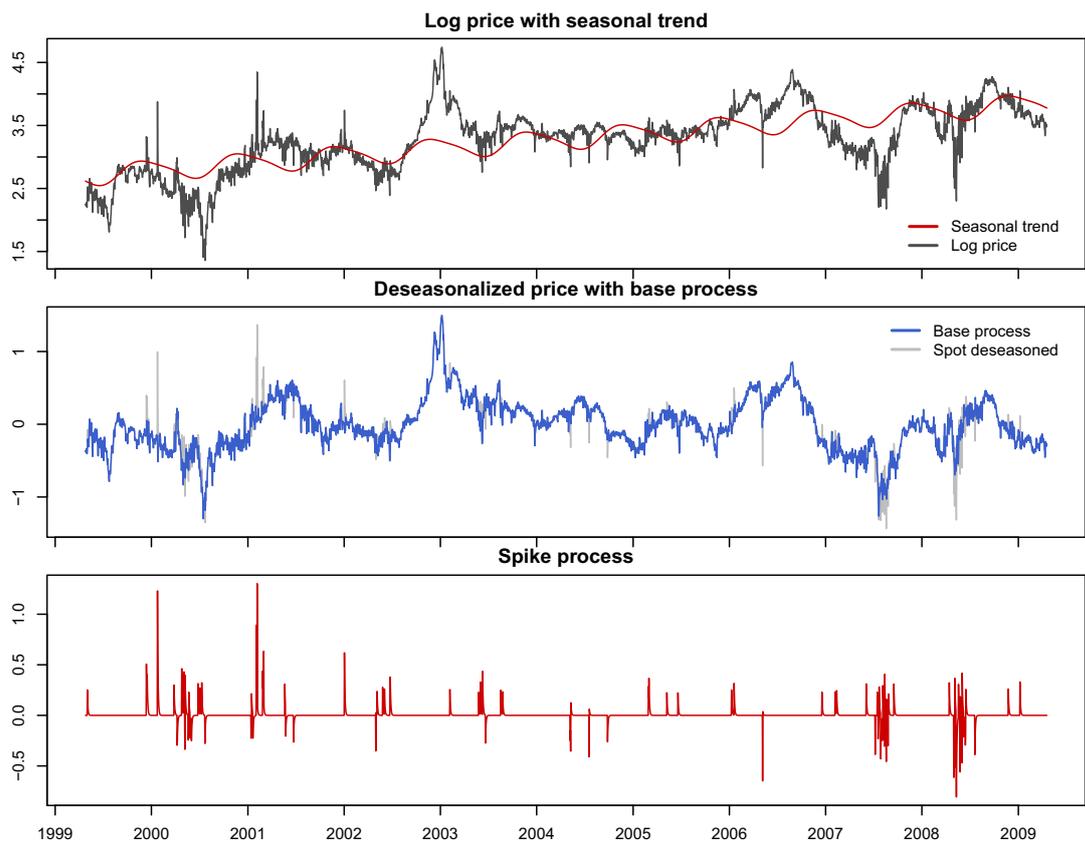
Figure 4.4: Separation of the components of the log spot price; the upper plot shows the daily log prices together with the estimated seasonal trend. The plot in the middle shows the deseasonalized log price with the estimated base process, while the lower plot shows the spike process.
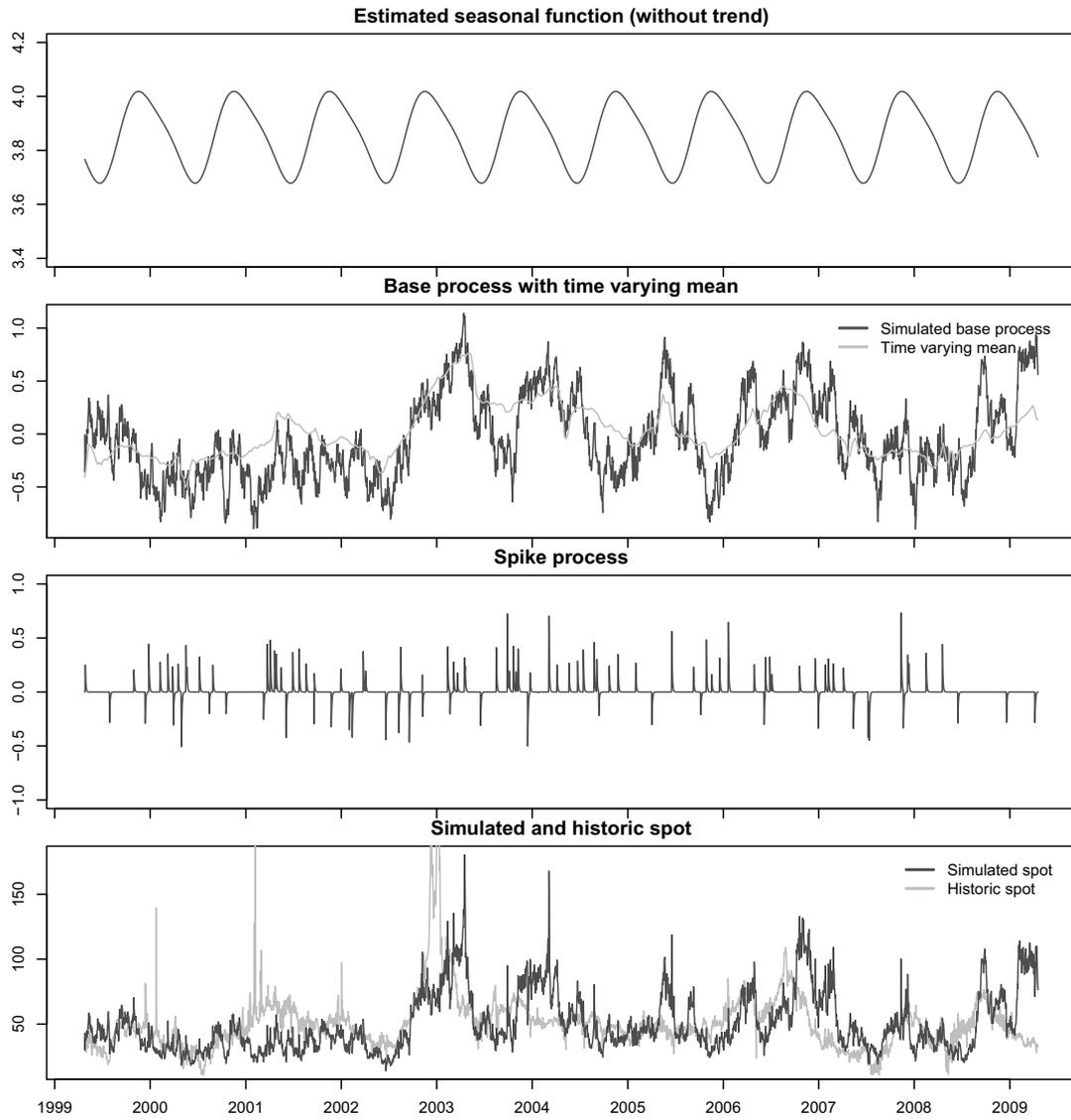
Figure 4.5: Simulated spot price and each of its components. From top to bottom: the seasonal function, the base process with the time varying mean reverting level, the spike process and last the resulting spot price. The historical spot price is shown as a reference. Both the simulated and the historical price are detrended to end of season 2009 level.

variance is clearly dependent on the general level of inflow. This is somewhat accounted for in model 2, as seen from the higher p-value of 0.052 from the same test. None of the two models pass the Jarque-Bera test of normality, but implementing alternative distributions for the error terms is left for further work. Hence model 2 is chosen as the best representation of the inflow our in model.

Table 4.4: Estimated parameters for the inflow dynamics using weekly data. P-values from Jarque-Bera normality test (under the null hypothesis of normal distributed residuals) and Breusch-Pagan heteroskedasticity test (under the null hypothesis of homoskedastic residuals).

	Coefficient	Estimate	Std. error	t-value	p-value	Significance
Model 1	y_{t-1}	0.3183	0.0439	7.25	0.0000	***
	Adjusted R-squared: 9.9 %, Jarque-Bera: 0.00 , Breusch-Pagan: 0.000					
Model 2	y_{t-1}	0.2896	0.0444	6.53	0.0000	***
	Adjusted R-squared: 8.2 %, Jarque-Bera: 0.00 , Breusch-Pagan: 0.052					

Significance codes: 0 (***), 0.001 (**), 0.01 (*), 0.05 (.), 0.1 ( )
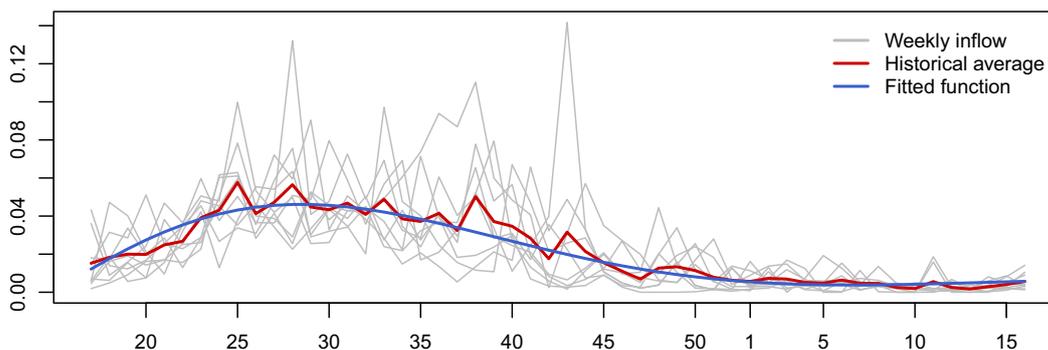
Figure 4.6: Inflow seasonality function fit. The gray lines show historic weekly inflow (one per season). The average and the fitted seasonality function is given by the red and blue line, respectively.

Estimating overall reservoir model

The overall reservoir model is estimated using the inflow data for power station N, together with historic data for overall reservoir level. Figure 4.7 shows the historic deviation from normal inflow together with the deviation from normal overall reservoir level. The results of the regression on Equation (4.32) are given in Table 4.5. Again the tests for heteroskedasticity and normality suggest us to reject the respective null hypotheses. However, from looking at simulations, like the one shown as a blue line in Figure 4.7, the model is still deemed acceptable.

Implementing dispatch models

Implementing the dispatch strategies, or more specifically the threshold function method, is the most time consuming part of the model estimation. While the simple decision rule is implemented using the average reservoir level from 100 simulation under full information (clairvoyance), the threshold function is estimated by using historic inflow and price

Table 4.5: Estimated parameters of overall reservoir dynamics, and p-values from Jarque-Bera normality test (under the null hypothesis of normal distributed residuals) and Breusch-Pagan heteroskedasticity test (under the null hypothesis of homoskedastic residuals).

| Coefficient	Estimate	Std. error	t-value	p-value	Significance
$\Delta r_{t-1}^{overall}$	0.9647	0.0085	113.95	0.0000	***
Δi_t^{cum}	0.0234	0.0050	4.63	0.0000	***

Adjusted R-squared: 98.0 %, Jarque-Bera: 0.0 , Breusch-Pagan: 0.0

Significance codes: 0 (***), 0.001 (**), 0.01 (*), 0.05 (.), 0.1 ( )

data. A large number of possible combinations for the four parameters in the function are tried out (as seen in Table 4.6), resulting in long calculation times. The historic production and reservoir level when using the parameters that achieve the highest revenues are shown in Figure 4.8. The same profiles for the simple decision rule are given as a reference. The results show that the reservoir level tends to be lower when using threshold function as production strategy, compared to actual historic levels (dotted line). A probable reason for this is that our revenues calculations do not take into account the effect of changing turbine head on the amount of energy produced per volume of water. For our sample, the average difference in head between upper and lower reservoir boundary is 9 % of maximum head, which means that the obtained revenues per MWh is approximately 9 % lower when the reservoir is close to empty compared to when the reservoir is full. The differences in head between upper and lower reservoir boundary for all power stations are shown in Figure A.4 in the Appendix. By assuming a constant energy equivalent, this model obtains the same amount of energy at all reservoir levels. This dispatch model could be improved by implementing a varying energy equivalent, depending on the reservoir level, but this is here left for further work. Figure 4.9 shows the threshold function using the optimal parameters given in Table 4.6.

Table 4.6: Ranges for threshold function parameters, and estimated parameters for power station N. Note due to computational limitations, larger step size were used in the first iterations, and reduced iteratively until the desired step size was obtained. Estimated parameters for all power stations are shown in the Appendix.

	α_s	α_t	α_i	α_r
Range	[0.8, 1.3]	[0, 35]	[0, 0.4]	[0, 8]
Step size	0.05	1	0.05	0.5
N	1.2	25	0.20	3.0
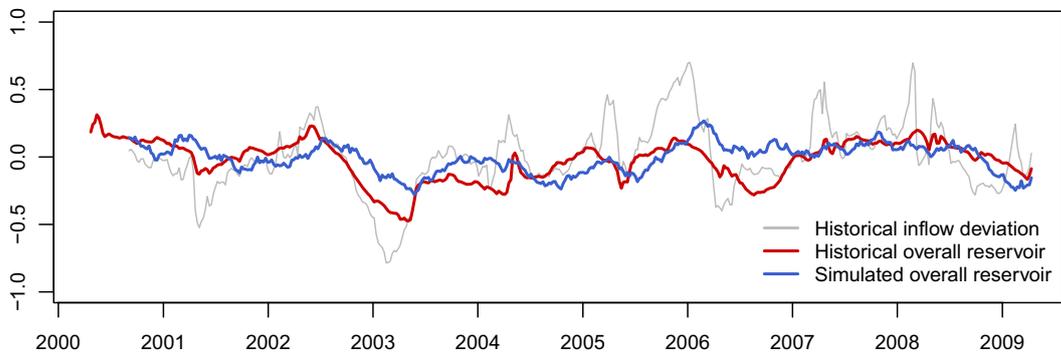
Figure 4.7: Overall reservoir deviation model estimation. The blue line represents one simulation based on the estimated parameters for power station N.
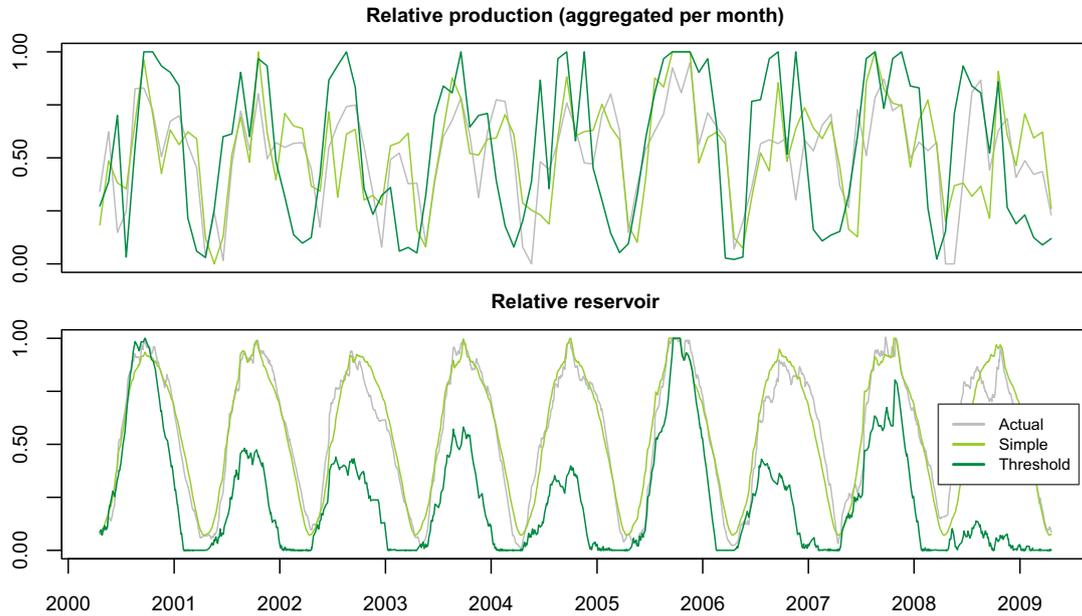
Figure 4.8: Historic production and reservoir profiles for the two dispatch strategies; the simple decision rule and the threshold function method. The actual reservoir is plotted in a gray line.
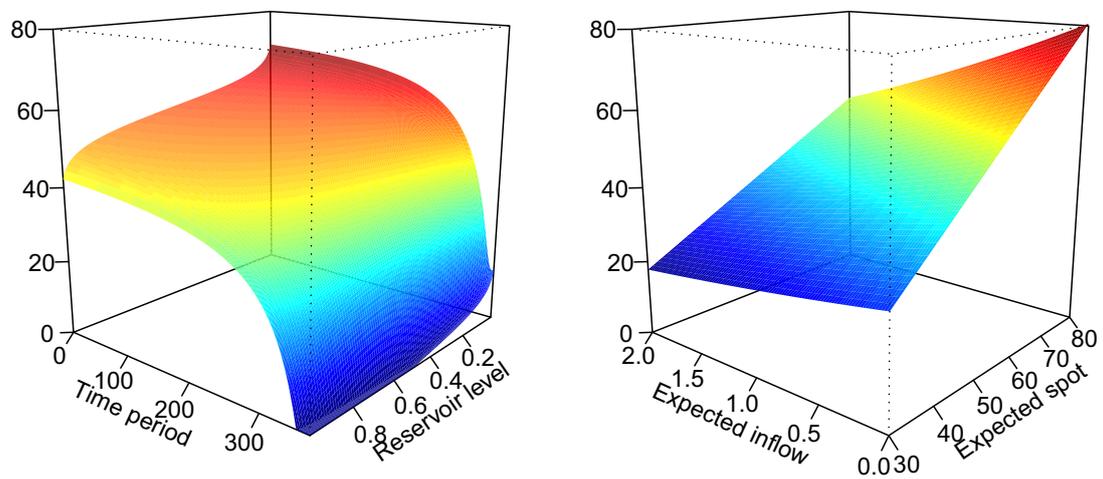
Figure 4.9: Fitted threshold function for power station N with varying time left of the season, current relative reservoir level, deviation from expected inflow for the next week and the average expected spot for the next 60 days.

4.4.3 Performing simulations

Figure 4.10 shows the simulation procedure step by step. In the following walk-through M is referred to as the number of simulations, and N the number of time steps (here 365 days).

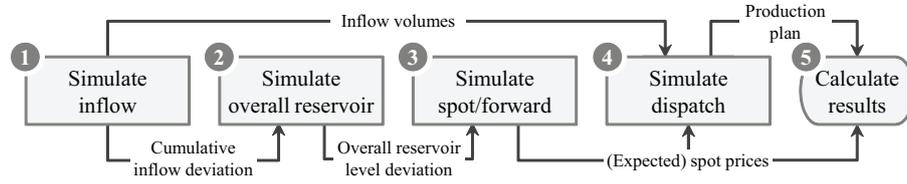**Figure 4.10:** Overview of simulation process

1. **Simulate inflow:** The model generates M time series, with 52 weeks in each, of inflow relative to the seasonal average. The relative inflow is bounded below by zero in each time step. The volumes for each weeks inflow are distributed evenly over seven days[4], giving a step-function with $N = 365$ time steps. The time series are multiplied with the seasonal average inflow to obtain absolute figures. In addition, M simulations for the 20 last weeks in the previous season are simulated in order to calculate cumulative inflow at the beginning of the season.

2. **Simulate overall reservoir deviation:** Using the seasonality function and the $M \times (N + 20)$ inflow matrix generated above, the deviation from normal for cumulative inflow is calculated based on the last 20 weeks. A number of M series for deviation from normal for the overall reservoir are simulated, using estimated parameters for the specific power station.

3. **Simulate spot prices:** Using the $M \times N$ matrix with deviation for overall reservoir, M spot price series with N time steps are simulated. This function also returns a $M \times N$ matrix for each component; base, spike and seasonality. The latter are used in the calculation of the average expected spot for the next 60 days.

4. **Simulate dispatch:** Using the absolute inflow series, spot price series and input for technical characteristics of the specific power station, the production profile for each flexibility case is calculated[5]. The simple decision rule uses an average reservoir level obtained by separate simulations under full information as reference level. The threshold function uses the expected spot and forward looking inflow in each time step to estimate the water value. The production profile and reservoir level for the run-of-river equivalent, the simple dispatch rule and the threshold function, are generated by iterating through time step 1 to N. For the case with limited flexibility and full information, the solution of the LP-problem returns the corresponding profiles[6].

5. **Calculate revenues:** Each production profile is multiplied with the corresponding spot price series to obtain the revenues. For the case of unlimited flexibility,

---

[4]The last week in the season is distributed over 8 days

[5]There is no production profile for the case with unlimited flexibility, as this is a pure duration curve analysis

[6]For the reservoir level boundary conditions for LP-problem we use the average historic reservoir level at the beginning of each season for the specific power station

the total amount of inflow is used to determine the revenues through price duration curves. The estimates for the relative revenues in each case are calculated by dividing the total revenues (for all M simulations) with the total amount of inflow. This corresponds to the weighted average with respect to water available.

Figure 4.11 shows an example of the simulation of one season for power station N. The simulated inflow is shown in the top plot, showing seasonality in both level and variance. The next plot shows the simulated deviation from normal levels for the overall reservoir content in Norway, influencing the simulated spot price below. The expected spot (red line) is then calculated. The last three plots show the results of the dispatch simulation for the two dispatch models and two other flexibility cases; no flexibility and limited flexibility with full information. We see that the production under no flexibility (red line) is linked directly to the inflow pattern, as storage of water is not possible. Not surprisingly, this flexibility case also results in the highest overflow. Note that the production strategy under full information deviates from the dispatch models in the way it "foresees" the price spike during the end of the season and thus keeps a higher reservoir level. Still, all three reservoir curves show the typical seasonal shape seen in empirical data.

We now perform $M = 1000$ simulations like the one above, and review the convergence of the results. Figure 4.12 shows the same plots as for the single simulation, but now with aggregated results. The 95 % quantiles are given as dotted lines where relevant. The simulations are performed per season, neglecting all interdependencies between seasons. This also implies that the spot price will start at the same value for each season. Figure 4.13 shows the convergence of the revenues calculated for each flexibility case (including the two dispatch strategies). We see that for even a relatively small sample of 1000 simulation, there is a clear convergence in both the revenues and their seasonal standard deviation. This is a welcoming results, as the simulation algorithm is rather calculation intensive[7].

---

[7]The most important bottleneck is the linear optimization problem to be solved for the revenues under limited flexibility with full information.
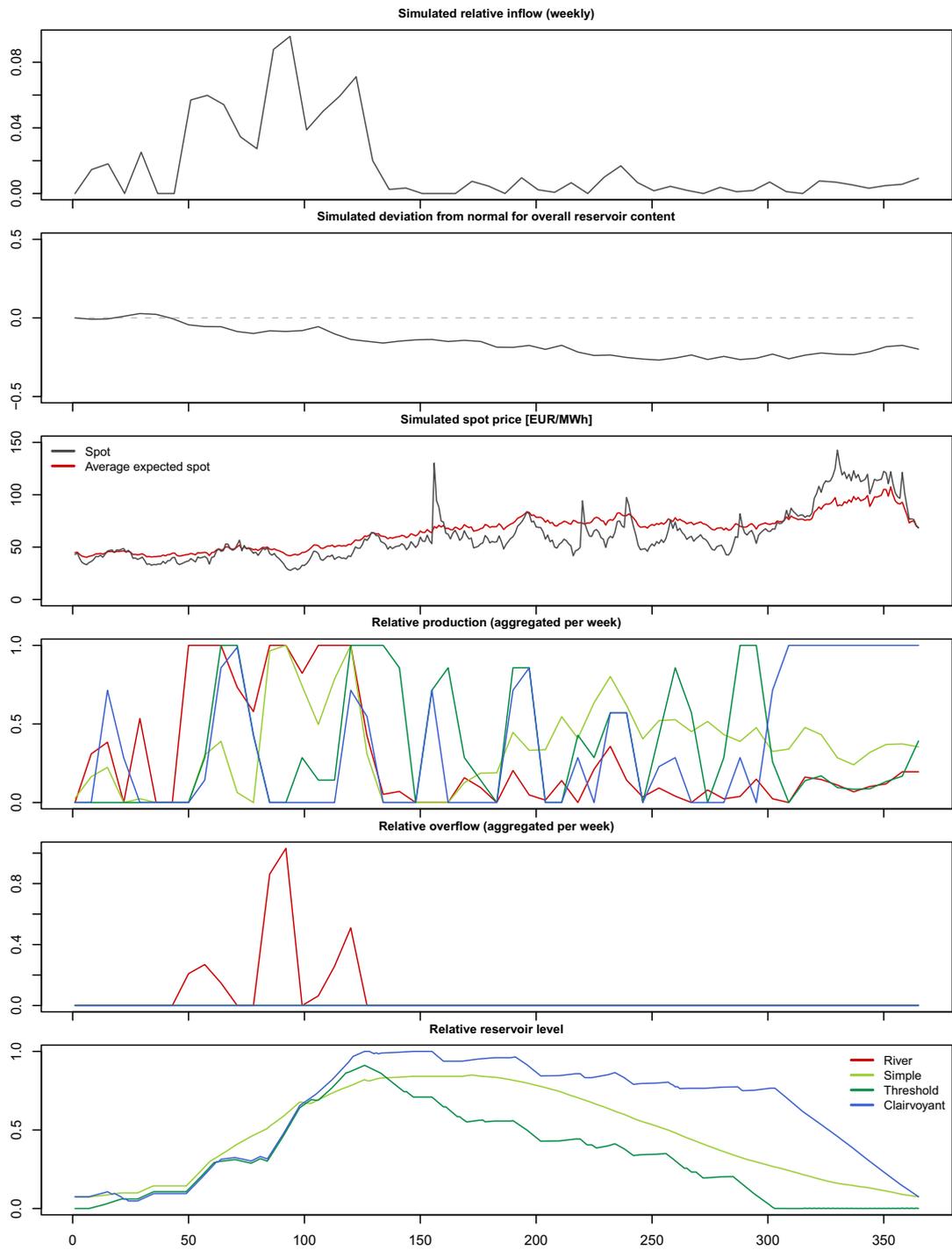
Figure 4.11: Single simulation example. Note that overflow is here given relative to production capacity. The legend in the bottom figure relates to the last three plots.
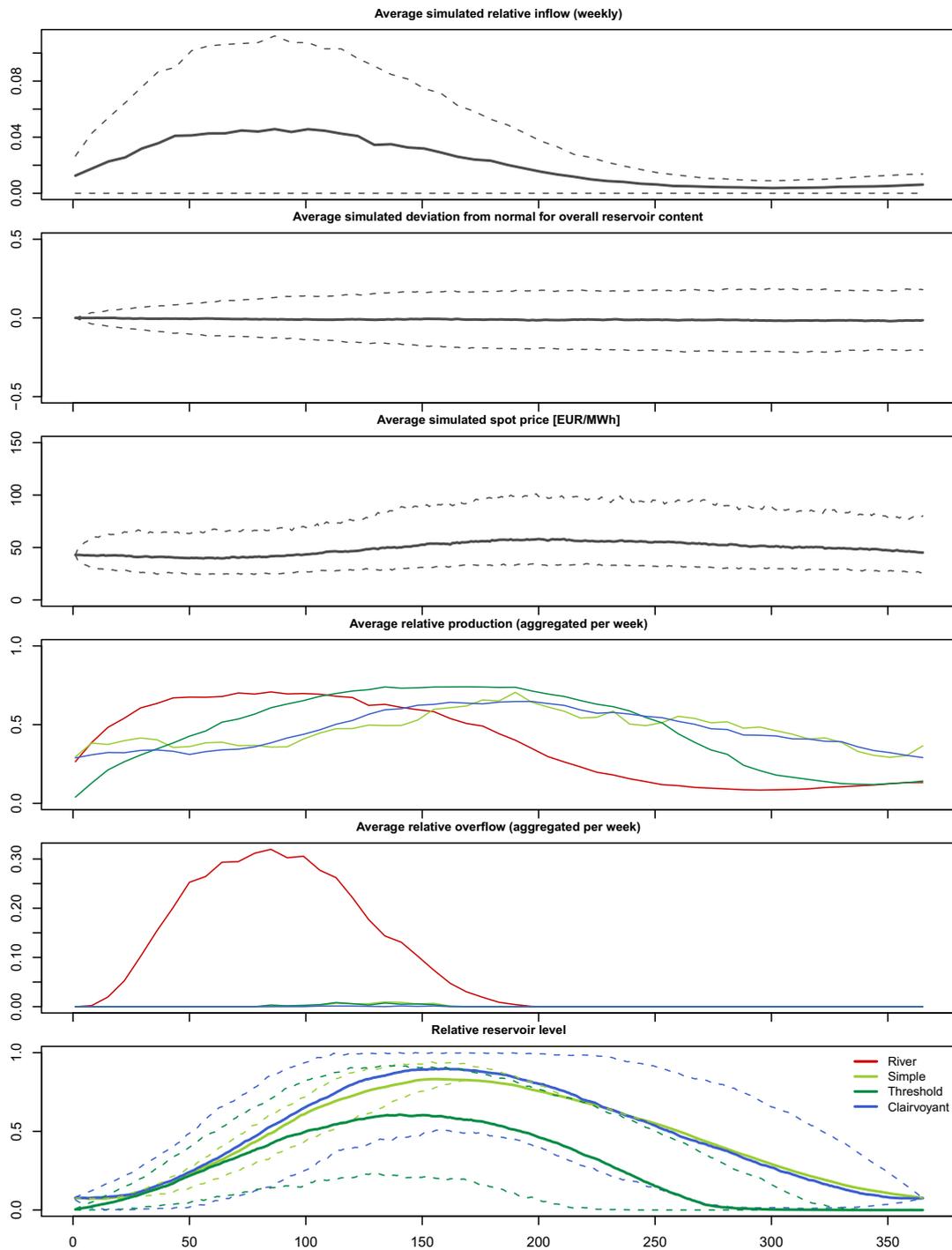
Figure 4.12: Simulation example with M=1000 simulated seasons. Dotted lines represent 95 % quantiles. Note that overflow is here given relative to production capacity. The legend in the bottom figure relates to the last three plots.
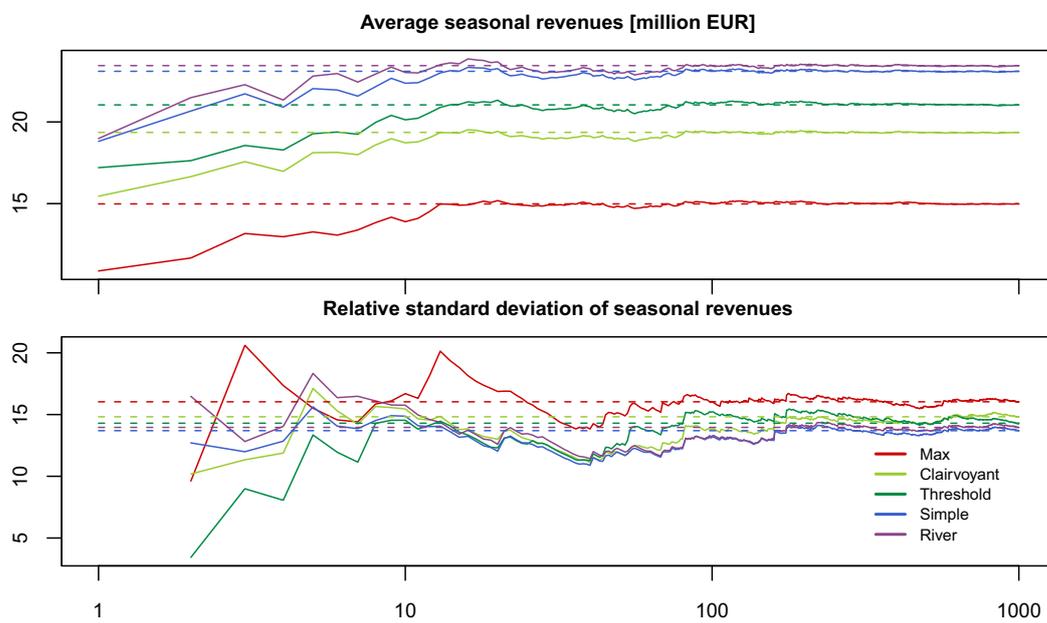
Figure 4.13: Simulation convergence for seasonal revenues and their standard deviation for each flexibility case

Chapter 5

Results

5.1 Model performance

The model described in the previous chapter is used to simulate $M = 1000$ seasons for each of the power stations in our sample. In order to evaluate the overall performance of the stochastic model, the results are compared with those obtained from the empirical data. This is done by calculating the relative revenues in each flexibility case and the corresponding flexibility values using the flexibility framework presented in Chapter 3. Note that with a perfect model, the results obtained from the simulations should converge to the revenues obtained in practice. However, due to simplifications in the inflow and spot dynamics, and with imperfect models for the production strategies, our model is not a perfect imitation of the dynamics in the real world. Hence we do expect the simulated results to deviate from the empirical results. The simulation results are given in Table 5.1, with deviations from empirical results provided in parentheses. Figure 5.1 shows the relative revenues together with the empirical results (shaded bars).
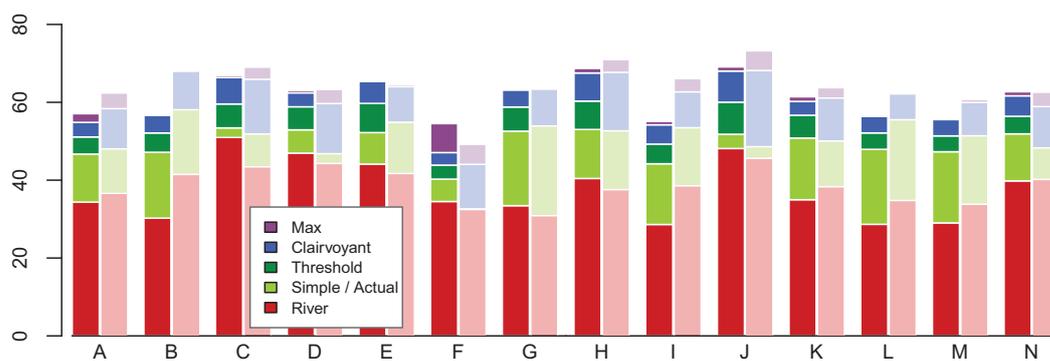
Figure 5.1: Simulated relative revenues for the different flexibility cases for each power station compared to empirical results (shaded bars).

Table 5.1: Simulation results for the different flexibility cases for each power station. Deviation from empirical results [%] are given in parentheses. Revenues obtained with the simple decision rule and by using the threshold function are both compared to the *actual* revenues from the empirical analysis.

Power station	π^{river}	π^{simple}	$\pi^{threshold}$	$\pi^{clairvoyant}$	π^{max}	$\upsilon^{clairvoyant}$	υ^{simple}	$\upsilon^{threshold}$
	[EUR/MWh water]					[%]	[%]	[%]
A	34 (-6)	47 (-3)	51 (6)	55 (-6)	57 (-9)	37 (0)	26	33
B	30 (-27)	47 (-19)	52 (-10)	57 (-16)	57 (-16)	47 (20)	36	42
C	51 (17)	53 (3)	60 (15)	66 (1)	67 (-3)	23 (-32)	5	14
D	47 (6)	53 (13)	59 (26)	62 (5)	63 (-1)	25 (-4)	11	20
E	44 (6)	52 (-5)	60 (9)	65 (2)	65 (1)	33 (-7)	16	26
F	35 (6)	40 (23)	44 (34)	47 (7)	55 (11)	27 (2)	14	21
G	33 (8)	53 (-2)	59 (9)	63 (0)	63 (-1)	47 (-8)	36	43
H	40 (8)	53 (1)	60 (15)	68 (0)	69 (-3)	40 (-10)	24	33
I	29 (-26)	44 (-17)	49 (-8)	54 (-13)	55 (-17)	47 (23)	35	42
J	48 (6)	52 (7)	60 (24)	68 (0)	69 (-6)	29 (-12)	7	20
K	35 (-9)	51 (2)	57 (13)	60 (-1)	61 (-4)	42 (12)	31	38
L	29 (-18)	48 (-14)	52 (-6)	56 (-9)	57 (-9)	49 (12)	40	45
M	29 (-14)	47 (-8)	51 (0)	56 (-7)	56 (-8)	48 (10)	39	44
N	40 (-1)	52 (8)	56 (17)	62 (5)	63 (0)	35 (12)	23	29
Average	37 (-3)	49 (-2)	55 (9)	60 (-3)	61 (-5)	38 (2)	25 (9)	32 (43)
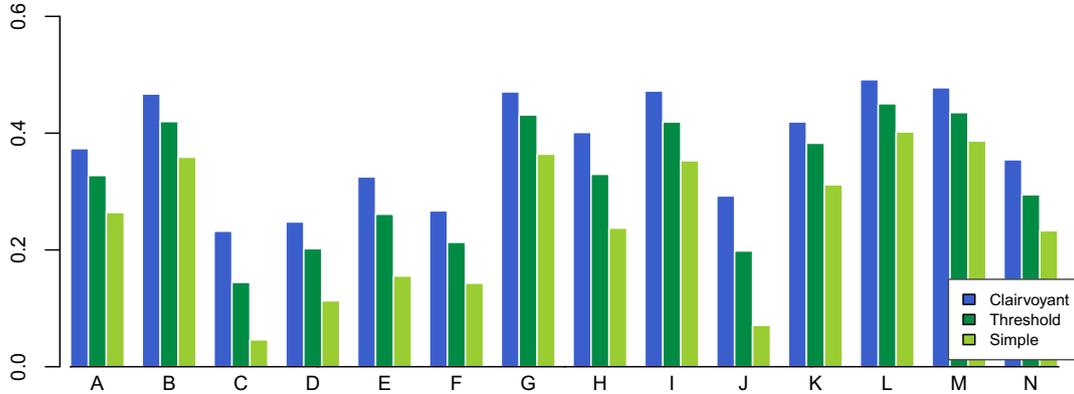
Figure 5.2: Simulation results for the flexibility value of each power station.

Note that when we now analyze the performance of the stochastic model, we will not discuss deviations from the actual revenues for π^{simple} and $\pi^{threshold}$, as these are highly dependent on the production strategy implemented by the specific power station. This is left for a separate discussion about the dispatch strategies at the end of this section.

Looking at the relative revenues obtained under no flexibility, π^{river}, power stations B, C, I and L stand out with the largest deviations from empirical results. The model overestimates π^{river} for power station C by 17 %, while B, I and L obtain significantly lower relative revenues in the model (between -18 % and - 27 %). From the fact that the revenues for the run-of-river equivalent are mainly linked to the inflow series, we deduce that these deviations are most likely caused by unrealistic inflow model dynamics for the

particular power stations. Further, a bad fit for the inflow model affects the simulated spot prices through their correlation, enhancing the deviation from empirical results. However, despite high deviations for several producers, the relative revenues under no flexibility do not show a bias in any specific direction; seven power stations obtain higher revenues from the model than in practice, while the opposite is the case for the other seven. The average deviation from empirical results for π^{river} is -3 %.

The revenues obtained under limited flexibility with full information ($\pi^{clairvoyant}$) and unlimited flexibility (π^{max}), also show large deviations from the empirical results for some power stations. Like before, power station B, I and L stand out with large deviations, and again this is most likely linked to the inflow model not being able to fully represent the actual dynamics. However, on average the model mimics the real world dynamics satisfactorily, showing an average deviation of -3 % and -5% for $\pi^{clairvoyant}$ and π^{max} respectively. Further, if we exclude the five power stations with the largest deviations (B, C, F, I and L), the average deviation of each flexibility case, π^{river}, $\pi^{clairvoyant}$ and π^{max} is 0 %, 0 % and -3 % respectively.

Figure 5.2 shows the simulated flexibility values for the different power stations. Due to the already mentioned deviations, the flexibility value of several power stations, especially B, C, and I, deviates from the empirical results. However, if we disregard these three, the relationship between the power stations' flexibility value in the clairvoyant case is in line with the findings from the empirical analysis shown in Figure 3.3 in Chapter 3.

Despite of some significant deviations, most likely caused by a poor fit for the inflow model, the stochastic model shows good overall performance. The results for the 14 power stations show mainly the same patterns as in the empirical results, indicating that the impact of factors such as reservoir size, inflow seasonality and production capacity is satisfactorily captured in the stochastic model.

Before moving on to analyzing the impact of specific factors on revenues, we discuss the performance of the two different dispatch models implemented to simulate the actual revenues obtained by the power stations; the simple decision rule with relative revenues π^{simple}, and the threshold function with relative revenues $\pi^{threshold}$. When comparing with actual revenues, we see that the threshold function model overestimates the revenues for nearly all producers, with power station B, I and L being the exceptions[1]. This indicates that the assumptions behind this dispatch strategy are not satisfied in practice. Specifically, the knowledge of the deviation from normal for overall reservoir levels 60 days ahead, which is used to calculate the expected spot price, may be too unrealistic. In addition, the fact that a power station's energy coefficient increases with reservoir level (head) is neglected. This way the energy generated per amount of water (directly influencing revenues) is overestimated for low reservoir levels[2].

While the threshold function method overestimates the actual revenues, the revenues obtained with the simple decision are similar to those obtained in practice for the majority of the power stations. Hence this method seems to provide a good approximation

---

[1]The model underestimates the relative revenues in *all* flexibility cases for these power stations. As mentioned above, this is most likely linked to poor inflow model performance

[2]The average head difference between maximum and minimum reservoir level is 9% of maximum head for our sample. See Figure A.4 in the Appendix for more details

of the actual dispatch strategy. This also indicates that producers do in fact use historical reservoir levels as a reference level when planning future dispatch. The simple decision rule also has clear implementation advantages over the threshold function; it only requires one estimated factor, the normal reservoir level. The threshold function needs to be estimated for each power station in a rather calculation intensive process. When changing underlying characteristics of power stations, such as reservoir size or production capacity, new sets of parameters need to be estimated, resulting in long computation times. In the next section, we want to be able to change these factors to see how they impact the relative revenues in the different flexibility cases. Hence, only the simple dispatch rule is applicable as an approximation of the actual production strategy in further analysis.

5.2 The impact of flexibility

The different factors impacting the revenues and flexibility of a hydropower station were discussed in Chapter 3. Having a complete stochastic model for hydropower stations, we can now quantify their impact. This is done by reviewing the changes in relative revenues and flexibility value in each of the flexibility cases when varying the following impacting factors; relative regulation, degree of inflow seasonality and capacity factor. Note that relative regulation and capacity factor are varied by changing the reservoir size and the production capacity respectively, while holding average annual inflow constant. When running these sensitivities, we need a set of power station parameters to represent the base case. This power station should be representative for the sample, with close to average characteristics in terms of reservoir size, inflow volumes and seasonality, and production capacity. Based on these criteria, power station N is chosen as the base case. The deviation between empirical and simulated results for this power station are relatively small, indicating that the model provides realistic results.

Table 5.2 shows the base values and sensitivity ranges for the parameters. In the next sections, the impact of the different factors is analyzed in isolation (keeping the other factors constant). Further, we show two-dimensional sensitivity plots and analyze the impact on revenues and relative flexibility value when changing two factors simultaneously. The results below are obtained by simulating 1000 seasons for each scenario.

Table 5.2: Ranges for one-dimensional sensitivities. The degree of seasonality in the inflow distribution decreases within the range; inflow distribution 1 (power station G) corresponds to a strong seasonal distributed inflow, while inflow distribution 5 is approximately flat (power station C).

	Base case	Range	Steps
Relative regulation	0.465	[0.047,0.931]	20
Capacity factor	0.471	[0.09,1.00]	14
Inflow distribution	3	{1,2,3,4,5}	5

5.2.1 Impact of relative regulation

In order to quantify the impact of reservoir size on relative revenues, we run simulations for 16 different scenarios, for which the results are shown in Figure 5.3. As expected,
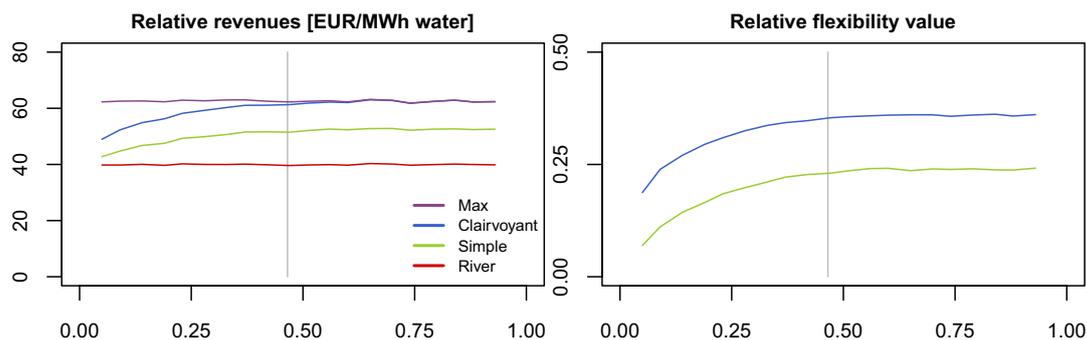
Figure 5.3: The impact of relative regulation on relative revenues (left plot) and flexibility value (clairvoyant and simple, right plot). The horizontal axis represents relative regulation, while the gray line marks the base case.

relative regulation has a positive impact on the relative revenues[3]. However, the degree of impact clearly decreases; with a relative regulation of approximately 0.5, a further increase in reservoir size has little to no impact on the upper boundary for revenues, assuming everything else being equal. For power stations with low regulation, on the other hand, an increase in storage flexibility could increase potential revenues significantly; increasing the relative regulation from 0.2 to 0.5 results in an increase in revenues of around 10 %, the other factors being equal to the base case.

The value of storage flexibility clearly converges as the relative regulation increases; in the case of full information, the relative value of storage flexibility converges to 36%, while converging to 24 % under limited information. Hence, for a power station with similar production capacity and inflow characteristics as N, having sufficient storage flexibility could account for an increase in revenues of about 50 % beyond the run-of-river equivalent[4].

5.2.2 Impact of inflow seasonality

The impact of inflow seasonality is analyzed by implementing five different seasonality functions for the inflow model, referred to as inflow case 1 through 5, while keeping the average total inflow constant. The first inflow function has a strong degree of seasonality, with around 90 % of the total inflow occurring in the filling season. Moving through the functions, the degree of seasonality decreases, the last function having approximately evenly distributed inflow throughout the season. The inflow distribution cases 1 through 5 corresponds to the inflow seasonality functions for power station G, B, N, D and C respectively.

Figure 5.4 shows that the impact of inflow seasonality on relative revenues decreases as the inflow becomes more distributed throughout the year; going from inflow function 1 (strong seasonality) to inflow function 2 (medium seasonality), increases potential revenues with 8 %, while there is no gain in revenues by further reducing the seasonality of the inflow function. Hence, for power stations with similar characteristics as N, with medium relative regulation and medium capacity factor, the degree of inflow seasonal-

---

[3]Note that in the case of no or unlimited flexibility the revenues are independent on relative regulation.
[4]50 % increase beyond the run-of-river revenues corresponds to a flexibility value of 33 %

ity does not affect the potential revenues significantly. We do however expect different results for power stations with less storage flexibility, and this is explored further by conduction two-dimensional sensitivity analyses later in this chapter.

Looking at the relative flexibility values, we see that it decreases as inflow seasonality decreases. This is as expected; the higher the degree of seasonality, the more valuable is the ability to store water (mainly driven by the reduction of overflow). With strong degree of inflow seasonality, the flexibility value is as high as 46 % in the clairvoyant case. When going from strong seasonality to no seasonality, the flexibility value decreases with 23 % points and 29 % points for the clairvoyant case and simple decision rule respectively. Hence, the degree of inflow seasonality has a large impact on the value of storage flexibility.
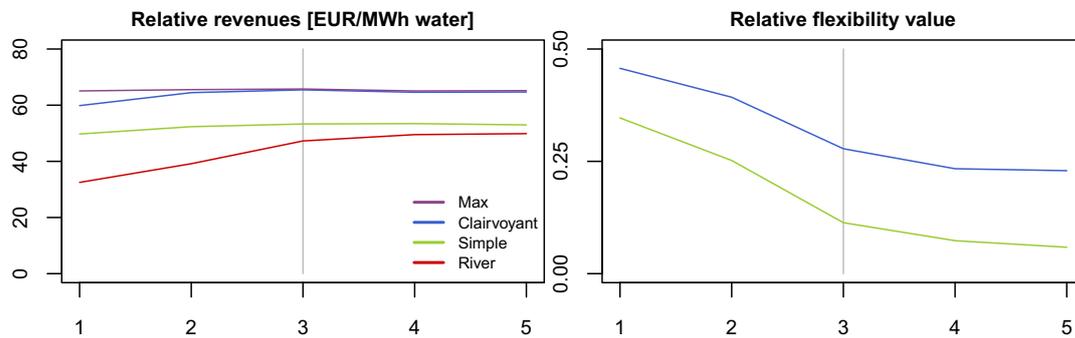
Figure 5.4: The impact of inflow seasonality on relative revenues (left plot) and flexibility value (clairvoyant and simple, right plot). The horizontal axis represents the degree of seasonality, with 1 representing a strong seasonality in the inflow with a large share of the total inflow occurring at the beginning of the season, and 5 representing an approximately evenly distributed inflow throughout the season. The gray line marks the base case.
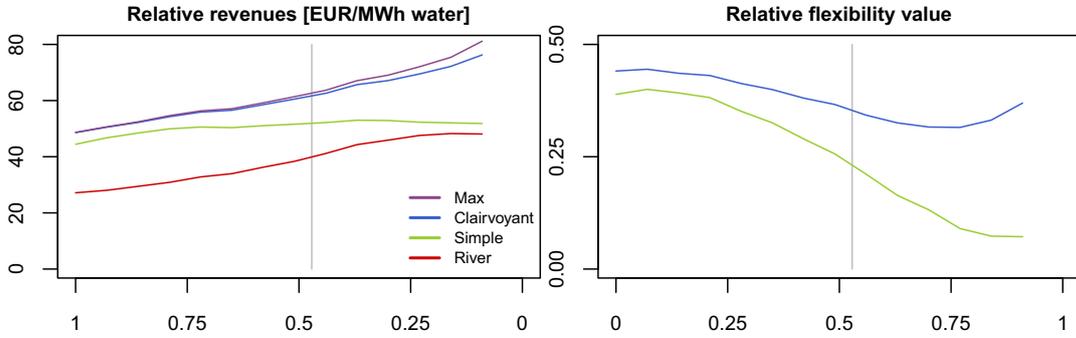
5.2.3   Impact of capacity factor

The impact of the capacity factor on revenues is analyzed using a range of capacity factors from 0 to 1, equivalent to a production capacity of around 50 % of power station N's capacity and up to about 500 %. Note that the relationship between the capacity factor and the production capacity is non-linear; doubling the production capacity only reduces the capacity factor with 50 %. Figure 5.5 shows that relative revenues in most flexibility cases increase linearly with decreasing capacity factor. The only exception is the simple decision rule case, where the relative revenues, π^{simple}, converge when the power station has sufficient capacity to reach the reservoir target level each day[5]. The revenues under full information, $\pi^{clairvoyant}$, which represent the upper boundary for the actual revenues obtained with a more comprehensive dispatch strategy, continue to increases with decreasing capacity factor. This indicates that power stations could increase their revenues significantly by increasing the production capacity, but only if the implemented dispatch strategy is complex enough to exploit the gain in production flexibility. As production capacity increases, the share of time a power station needs to produce, assuming equal total inflow, is reduced. This increases the importance of choosing the right periods for when to dispatch the water, and hence implementing a

---

[5]The target level will only change slightly when decreasing the capacity factor

good production scheduling strategy becomes crucial.

The benefit from decreasing the capacity factor does not depend on the current level; decreasing the capacity factor by 0.1 increases the potential revenues ($\pi^{clairvoyant}$) with approximately 4 %. Hence, all power stations could achieve revenue gains by increasing their current production capacity. However, due to the inverse relationship between capacity factor and production capacity, the marginal impact on revenues decreases when the production capacity increases. Reducing the capacity factor from 0.5 to 0.4, is equivalent to a 25 % increase in production capacity, while reducing the capacity factor twice as much, from 0.5 to 0.3, corresponds to a 67 % production capacity increase. For the average power station in our sample, with a capacity factor of 0.46, an investment in doubling production capacity could increase revenues with 13 %. This is seen in Figure 5.6, which shows the same results as above, but with production capacity instead of capacity factor.

The storage flexibility value $v^{clairvoyant}$, decreases with decreasing capacity factor. In this case the higher production flexibility (as a result of lower capacity factor), reduces the importance of storage flexibility. This can be explained as follows: the value of storage flexibility is driven by two main components; the ability to store water for dispatch at a later point in time (during periods of high prices) and the ability to avoid overflow at the time of inflow. Increasing the production capacity, will reduce the overflow volumes, and hence reduce the need for (value of) storage flexibility. For the simple decision rule, the relative flexibility value decreases even more when decreasing the capacity factor, as this strategy fails to benefit from the increase of production flexibility, while the run-of-river equivalent obviously does.

Figure 5.5: The impact of the capacity factor on relative revenues (left plot) and flexibility value (clairvoyant and simple, right plot). The horizontal axis represents capacity factor, while the gray line marks the base case. Note that a low capacity factor indicates high production capacity.

5.2.4 Cross-sensitivities

In order to evaluate the impact on relative revenues when changing more factors simulatenuously, we perform cross-sensitivity analysis. To avoid long calculation times, the ranges and step size are adjusted as shown in Table 5.3. Figure 5.7 shows the simulated relative revenues obtained under limited flexibility, π^{simple} and $\pi^{clairvoyant}$, which are the lower and upper bound for the revenues that can be obtained in the real world,
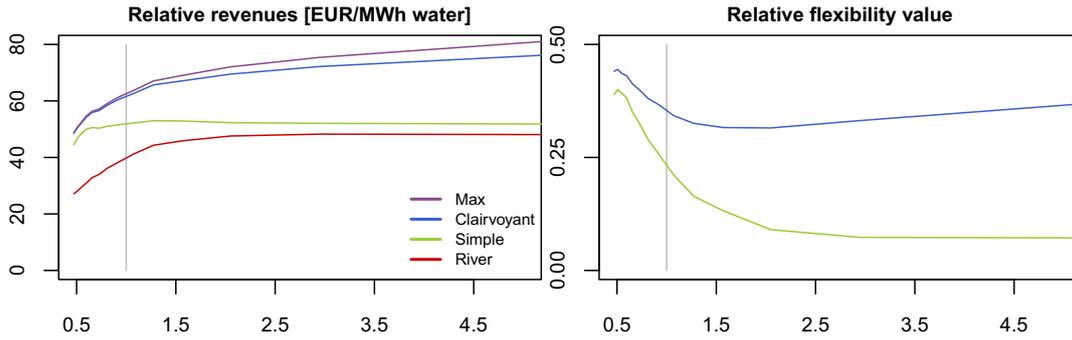
Figure 5.6: The impact of production capacity on relative revenues (left plot) and flexibility value (clairvoyant and simple, right plot). The horizontal axis represents production capacity relative to the base case (gray line). Note that the distance between data points increases for increasing production capacity

and the corresponding relative flexibility values. The relative revenues for the two other flexibility cases are shown in Figure A.3 in the Appendix.

Table 5.3: Ranges for cross-sensitivities. The degree of seasonality for the inflow decreases within the range; inflow distribution 1 (power station G) corresponds to a strong seasonal distributed inflow, while inflow distribution 5 is approximately flat (power station C). The base case for relative regulation is adjusted to 0.4 (from 0.465), while the base case for the capacity factor is rounded to 0.5 (from 0.471) to better fit with the defined ranges and step sizes.

|	*Base case*	*Range*	*Steps*
Relative regulation	0.4	[0.1,1.0]	4
Capacity factor	0.5	[0.25,1.0]	4
Inflow distribution	3	{1,2,3,4,5}	5
Total number of scenarios: 80			

The relative revenues for varying relative regulation and degree of inflow seasonality are shown in the left perspective plots in Figure 5.7a and 5.7b. We identify a clear dependence; relative regulation has a significantly higher impact on revenues under strong inflow seasonality (1). This impact declines with decreasing degree of inflow seasonality. With an approximately flat inflow distribution (5), the revenues are independent of relative regulation. If a large share of the inflow occurs in the beginning of the season, sufficient storage flexibility in terms of relative regulation is important to be able to avoid overflow and achieve higher revenues by being able to chose when to produce. In this case the increase in relative revenues is as high as 35 % under both limited and full information (π^{simple} and $\pi^{clairvoyant}$) when increasing the relative regulation from 0.1 to 1.0. For the same reasons, revenues depend highly on the distribution of inflow when the relative regulation is low, while being independent of this in the opposite case. From looking at the value of storage flexibility in the leftmost plots in Figures 5.7c and 5.7d, we see that the combination giving the highest flexibility value, is a high degree of inflow seasonality with high relative regulation. For this combination the relative flexibility value is between 37 % to 48 %. Hence, having sufficient storage flexibility

70
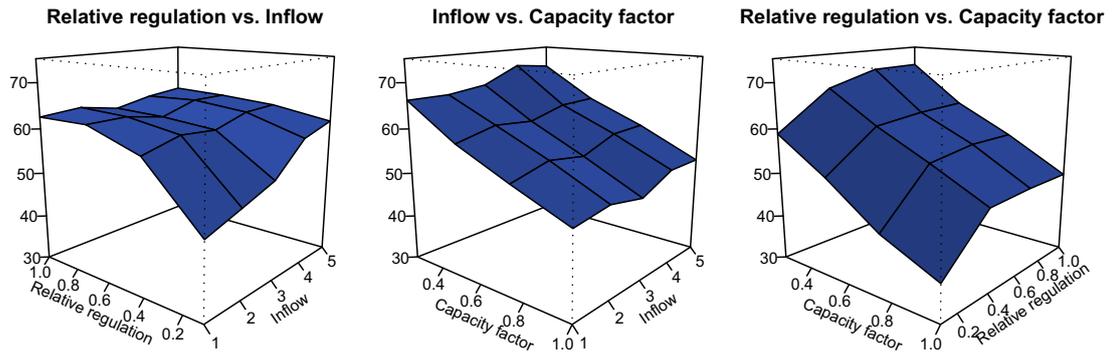
(a) Relative revenues $\pi^{clairvoyant}$
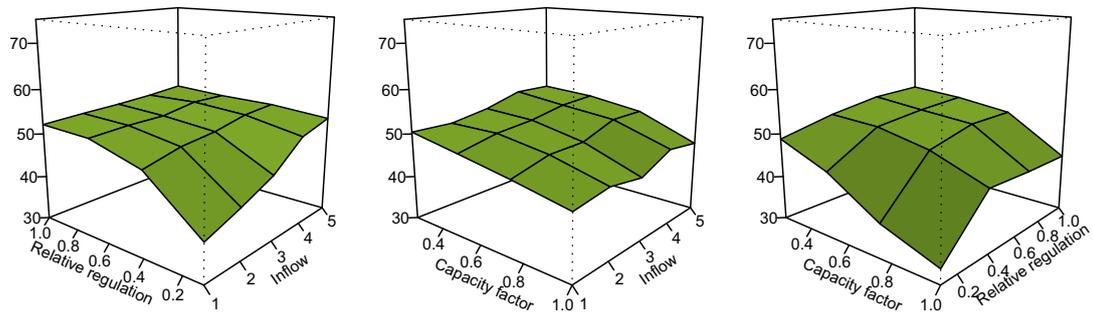
(b) Relative revenues π^{simple}
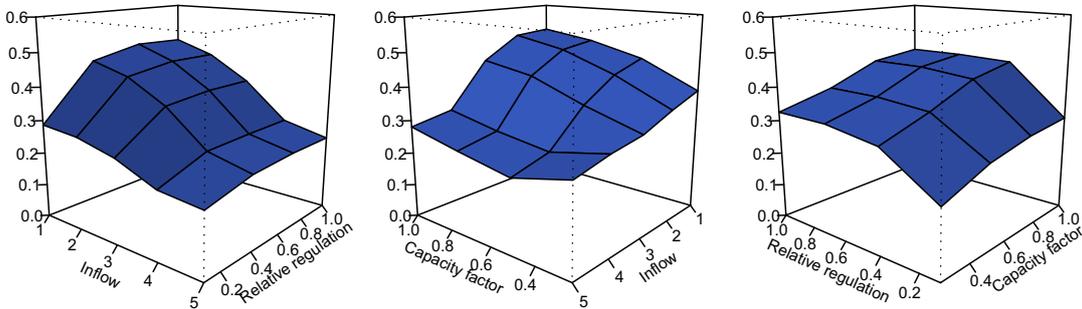
(c) Relative flexibility value $\upsilon^{clairvoyant}$
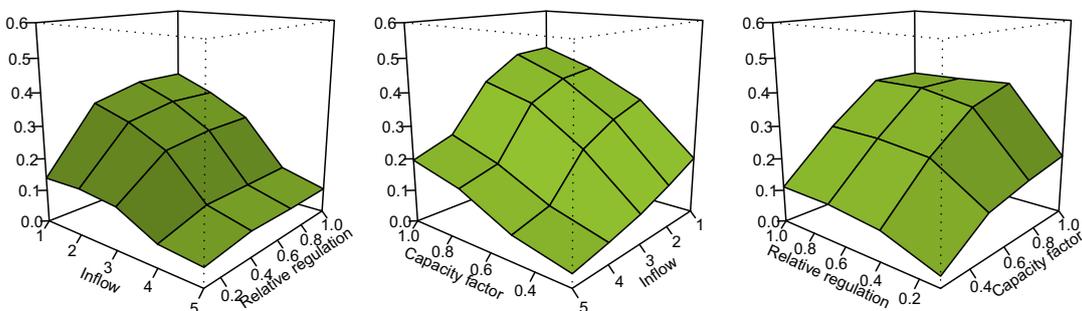
(d) Relative flexibility value υ^{simple}

Figure 5.7: Relative revenues and flexibility values (clairvoyant and simple) for different cases of relative regulation, inflow seasonality, and capacity factor. Note that the scale on the inflow axis corresponds to the degree of seasonality, with 1 being strong and 5 being flat.

could account for around 40-50 % of revenues, given the current capacity factor (0.5).

While the impact of relative regulation is highly dependent on the degree of inflow seasonality, the effect of changing the capacity factor is not. As shown in the middle plots of Figure 5.7a and 5.7a, the potential relative revenues increase linearly with decreasing capacity factor, independent of the degree of seasonality in inflow. However, this in not the case for the corresponding flexibility values seen in the middle plots of Figures 5.7c and 5.7c. The relative value of storage flexibility is more dependent on the degree of inflow seasonality when the capacity factor is high, than when it is low. This can be explained by the higher importance of storage flexibility in order to avoid overflow when the production capacity is low. We also see that with a low seasonality in inflow (5), the capacity factor have no significant impact on the relative value of storage flexibility, as the need for high capacity to reduce overflow is reduced.

The right plots in Figures 5.7a and 5.7b show the impact of capacity together with relative regulation, and also here we see a similar dependency. Relative regulation has a lower impact on relative revenues $\pi^{clairvoyant}$ when the capacity factor is low. In this case the high production capacity reduces the importance of having a large reservoir. Like mentioned previously, we see that revenues obtained under limited information, π^{simple}, are constant when the capacity increases and relative regulation increase above certain values, as this dispatch strategy fails to exploit the increase in flexibility. We also see that a low capacity factor (high production capacity) is crucial when having low relative regulation[6]; the increase in revenues are as high as 47 % and 65 % when reducing the capacity factor from 1.0 to 0.25 in the limited information and clairvoyant case.

To summarize the interdependencies, we see that the impact on revenues and flexibility values due to changes in inflow distribution and relative regulation is closely linked. Having a high relative regulation or low inflow seasonality reduces the impact of the other factor to a minimum. However, with a high degree of inflow seasonality, a high relative regulation is critical to be able to obtain high revenues. The impact of capacity factor and relative regulation are also interlinked; when flexibility is high, either in terms of storage (high relative regulation) or production (low capacity factor), the other factor is of less importance to the revenues.

5.3   Spot model variations

The underlying spot model is fairly comprehensive, in terms of both incorporating jumps and inflow correlation. But how much impact does the complexity of the spot model actually have on the results? In order to answer this question, we rerun the model above with three different spot models. The multi-factor model defined and discussed previously, with a seasonal component, a spike process and a mean reverting base process depending on inflow, forms the base model. In addition we propose two simplified models. In the first, the correlation between inflow and spot is removed by setting the mean reverting level of the base process equal to zero, $\mu_{1,t} = 0$. The second (and most simple) model is a version where also the separate spikes process is removed, leaving the spot model to only consist of a seasonal component and a geometric mean reverting

---

[6]And given the degree of inflow seasonality of power station N. However, similar results are obtained for the other inflow scenarios.

base process. The parameters for the two simplified spot models are estimated using the same procedure as before, and are show in Table A.3 in the Appendix.

Figure 5.8 shows the average relative revenues for for our sample, with variations of the spot model. Comparing the relative revenues obtained using the multi-factor model with all three components (seasonal function, spike process and base process) with and without correlation with inflow, we see that removing the correlation increases the relative revenues in all flexibility cases significantly. For the clairvoyant case, this increase equals 6 %. The increase is in line with the expectations as low (or zero) correlation increase the possibility for high prices to coincide with high inflow, which increases the revenues. Using the simple spot model without spikes does not seem to influence revenues significantly, indicating that the modeling of spikes is not important given the current precision level of our model. However, to rule out potential bias caused by unrealistic spot dynamics after refining the precision level of the other model components, we still suggest to include the modeling of spikes in similar models and analyses.

We have previously pointed out some weaknesses in the way we model the inflow of each power station. As the inflow and spot model are coupled through the correlation with overall reservoir level, unrealistic inflow dynamics will in turn also influence the spot model. The importance of a good spot model should therefore be further investigated, possibly with a different inflow model, as the results above may be influenced by the poor performance of the inflow model for some power stations. However, as previously shown, the model does not purely over- or underestimate revenues. Hence, we can assume that by removing the correlation between inflow and spot on, the potential revenues are on average overestimated with 6 %.
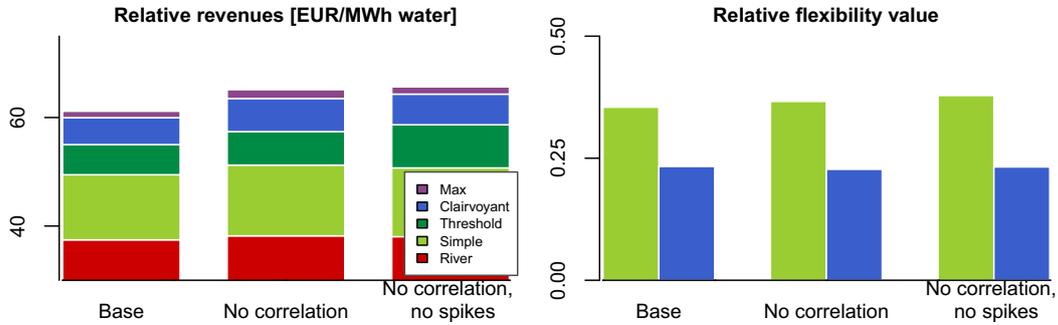
Figure 5.8: Sample average relative revenues and relative flexibility value for the three different spot models.

5.4 Application: Valuation of hydropower assets

The stochastic model provides estimates for expected annual revenues given todays electricity price dynamics. It gives approximations for both the actual (π^{simple}, $\pi^{threshold}$) and the theoretical upper boundary ($\pi^{clairvoyant}$) of the relative expected revenues under different degrees of flexibility. These multiples can serve as a basis for the valuation of existing hydropower stations with reservoirs. Below we show a very simple example of the valuation of operating results for power station N.

The following formula gives an estimate of the net present value (NPV) of the operating results (before taxes of any kind) for a power station, assuming a long (infinite) lifetime. The future annual operating results A are discounted using the discount factor r and the annual revenue growth rate g:

$$A = \pi \bar{I}^{tot} - C \times P^{max} \tag{5.1}$$

$$V = \frac{A}{r - g} \tag{5.2}$$

Where π is the relative revenue multiple, $\bar{I}$ is average annual inflow, P^{max} is the installed capacity and C are the operating costs relative to installed capacity. $P^{max} = 90$ [MW] and $cf = 47\%$ for power station N. According to NVE [2007], the typical operating costs for a hydropower station with reservoir are around 1 % of the capital expenditures, or operating costs C relative to installed capacity of around 12.5 [EUR/kW installed][7]. For power station N, this figure is equivalent to relative costs of around 3 [EUR/MWh water], which is around 6 % of the estimated actual relative revenues (from the simple decision rule). Further, NVE [2007] suggests the use of a discount factor r of 6.5 % when evaluating investments in hydropower in Norway. This figure is used in the example, without going into the debate on how to accurately determine the required rate of return for this type of assets. Note that when using the revenues from the stochastic model for valuation, the annual trend of spot prices needs to be taken into account. As previously mentioned, the estimated annual nominal trend of 11.5 % is likely to be too high in the long run. For simplicity we set the real growth g rate equal to zero in this example. Using these figures, and the multiples for relative revenues for power station N from our model, the infinite lifetime value of the operating results ranges from 279 [EUR million] (simple decision rule) to the upper limit of 336 [EUR million] (clairvoyant). This suggest an absolute value of storage flexibility ranging from 68 [EUR million] to 125 [EUR million], when subtracting the value for the run-of-river equivalent.

Note that a valuation based on the expected cash flows from the stochastic model does not consider potential increases or decreases in prices due to for example new transition lines to other markets or changes in the cost structure of electricity production in the future. In other words, the expected revenue flows presented here only represent today's marked conditions. This motivates the use of a real option approach when considering investments in generation assets. For more information, see rland [2007] who uses the model of Dixit and Pindyck [1994] for an assessment of investments in hydropower in Norway.

5.5 Main shortcomings of the model

The modeling approach has several shortcomings, which leave room for further improvements. First of all, the implemented inflow model shows clear weaknesses; the residuals fail to pass the normality tests, which indicates that the proposed model does not fit sufficiently well to real world dynamics. More importantly, we identify several deviations between simulated and empirical results that to a large degree are explained by an imperfect inflow model. For further work, we suggests a more detailed study of the inflow dynamics, and how they can be modeled. As possible improvements, we propose to consider other distributions for the residuals, and/or to expand the model with a separate process for time-varying variance. Using a larger dataset could also improve

---

[7]We use a exchange rate of 8 [NOK/EUR].

the model performance. In addition, a larger dataset would allow for an alternative modeling approach; like done in the EMPS model, the inflow series could be randomly selected from a large pool of historic series.

The model lacks a sufficient realistic dispatch strategy to be used under limited flexibility and information, which is capable of exploiting high degree of storage and production flexibility. The threshold function which is implemented, clearly overestimates the actual revenues due to two main shortcomings; firstly we assume that the deviation from normal aggregated reservoir level is known 60 days ahead. Although a slowly varying reservoir level to a certain degree can justify this assumption, this is an obvious simplification of the real world. Secondly, neglecting the energy coefficient's dependency on the turbine head, favors low reservoir levels. The realism of this dispatch strategy could be improved by making the energy coefficient (and hence the revenues) dependent on the current reservoir level by incorporating reservoir level curves (giving reservoir level as a function volume).

In practice the flexibility of a hydropower station is often restricted by time-varying constraints on reservoir level and/or dispatch volumes due to environmental reasons. These restrictions are not accounted for in our model, which leads to an overestimation of the actual revenues. Hence, we encourage future studies on the value of flexibility of hydropower stations to take these restriction into account.

Chapter 6

Conclusion

In this study we quantify the impact of storage flexibility on revenues for hydropower plants with reservoir. We develop a framework for analysing flexibility, and isolate the value of storage flexibility by decomposing the revenues into different flexibility cases.

By applying the framework on data for 14 Norwegian hydropower stations, we gain preliminary insights about the value of storage flexibility. This accounts for on average 22 % of the actual revenues achieved during the sample period (ranging from 0 to 43 %). The empirical analysis also shows that the power stations with the highest storage flexibility achieve on average 9 % higher relative revenues than the sample average.

Further, this study shows that it is possible to recreate the dynamics of hydropower stations with a fairly simple stochastic model. This model enables us to quantify the isolated impact of the three main drivers of revenues; relative regulation, inflow seasonality and capacity factor. By conducting sensitivity analyses, we show that an increase in relative regulation increases revenues, but that the effect is dependent on the degree of inflow seasonality. Given high degree of inflow seasonality, increasing relative regulation from 0.1 to 1.0 increases the revenue potential with as much as 35 %, while such an increase has no effect on revenues under low inflow seasonality. These findings emphasize the importance of storage flexibility for power stations with high degree of inflow seasonality. Our analysis also illustrates that the marginal gain of increasing the relative regulation decreases with its the level, which is aligned with the intuition that increasing storage flexibility only ads value up to a certain point.

The potential revenues increase linearly with decreasing capacity factor. However, due to the non-linear relationship between capacity factor and production capacity, the impact on revenues decrease with higher production capacity. For the average power station in our sample, reducing the capacity factor with 50 % (corresponding to a doubling in production capacity) increases potential revenues with 13 %. With the increase of production capacity, having good dispatch strategies becomes more important. Our simple decision rule, using normal reservoir levels as a reference, fails to exploit increases in production flexibility above a fairly low threshold.

We gain some relevant insights from the modeling approach itself. Applying the threshold function strategy of Näsäkkälä and Keppo [2008] will result in high realized revenues for all power stations in our sample. However, it is possible to question these result; this model neglects the effect of turbine head on the energy equivalent, and hence fails to penalize low reservoir levels. Further, we learn that the modeling of the inflow dynamics

is particularly challenging, mostly due to the time-varying variance.

This study has several practical implications. As reservoir size and inflow seasonality are generally given by nature, the dynamics of their impact are less important when considering investments in new infrastructure. However, the importance of reservoir size and inflow seasonality clearly has to be taken into account when doing a valuation of existing power stations. By providing revenue estimates depending on these two factors, the stochastic model can provide a basis for this type of valuation. The impact of production capacity on revenues is different, as this can be increased by replacing existing or adding new generators. The stochastic model quantifies the potential increase in revenues when increasing capacity, and hence is useful when considering investments in new or upgraded generators. Providing both an upper and lower boundary for revenues, this model can also be used to benchmark different dispatch strategies. Further, by quantifying the standard deviation of yearly revenues, it can be used in risk-assessments and to review the need for different hedging strategies. However, to better model the real world dynamics of hydro power stations, this model should be expanded to take into account additional restrictions faced by operators.

The flexibility framework and stochastic model developed in this study embraces a broad research field; valuation of hydropower plants, production scheduling under uncertainty and stochastic modeling of electricity prices and inflow, and the incorporation of the correlation between these. As this broad perspective limits the possibility of doing in-depth analysis within each of these fields, we encourage further studies within one or more of these fields to further develop the presented framework and stochastic model from this analysis. In addition we hope that the broadness and creative approaches of this study inspires for more analysis on the storage flexibility of hydropower producers.

Bibliography

F. E. Benth, S. Koekebakker, and F. Ollmar. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *Journal of Derivatives*, 15(1):52–66, 2007.

F. E. Benth, J. S. Benth, and S. Koekebakker. *Stochastic modelling of electricity and related markets*, volume 11 of *Advanced Series on Statisitical Science Applied Probability*. World Scientific Publishing Co. Pte. Ltd., 2008.

P. Bjerksund, G. Stensland, and F. Vagstad. Gas storage valuation price modelling v. optimization methods. Discussion paper, ISSN: 1500-4066, October 2008, Norwegian School of Economics and Business Administration, Norway, 2008.

A. Boogert and C. de Jong. Gas storage valuation using a Monte Carlo method. *The Journal of Derivatives*, 15(3):81–98, 2008.

A. Botterud, T. Kristiansen, and M. D. Ilic. The relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics*, In Press, Corrected Proof, 2009.

T. S. Breusch and A. R. Pagan. A simple test for heteroskedasticty and random coefficient variation. *Econometrica*, 5:1287–1294, 1979.

A. S. Cahn. The warehouse problem. *Bulletin of the American Mathematical Society*, 54:1073, 1948.

A. K. Dixit and R. S. Pindyck. *Investment under uncertainty*. Princeton University Press, 1994.

G. L. Doorman. Hydro power scheduling. Course material for ELK15, Department of Electric Power Engineering, Norwegian University of Science and Technology, Norway, 2009.

S.-E. Fleten and J. Lemming. Constructing forward price curves in electricity markets. *Energy Economics*, 25(5):409–424, 2003.

S.-E. Fleten, J. Keppo, J. Sollie, H. Lumb, and V. Weiss. Derivative price information use in hydroelectric scheduling. Working paper, Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Norway, 2009.

F. R. Foersund. *Hydropower economics*. Springer Verlag, 2007.

O. B. Fosso, A. Gjelsvik, A. Haugstad, B. Mo, and I. Wangensteen. Generation scheduling in a deregulated system: The Norwegian case. *IEEE Transactions on Power Systems*, 14(1):75–81, 1999.

D. Frestad, F. E. Benth, and S. Koekebakker. Modelling the term structure dynamics in the Nordic electricity swap market. *The Energy Journal*, 31(2):53–86, 2010.

H. Geman and A. Roncoroni. Understanding the fine structue of electricity prices. *The Journal of Business*, 79(3):1225–1262, 2006.

N. Haldrup and M. O. Nielsen. A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1-2):349–376, 2006.

IHA. The International Hydropower Association. http://www.hydropower.org. Last visited 2010.01.06, 2010.

C. M. Jarque and A. K. Bera. A test for normality of observations and regression residuals. *International Statistical Review*, 55(2):163–172, 1987.

C. Kolsrud and M. Prokosch. Dynamics of hydropower scheduling: Producers' response to forward prices and other uncertainties. Project thesis, Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Norway, 2009.

F. A. Longstaff and A. W. Wang. Electricity forward prices: A high-frequency empirical analysis. *The Journal of Finance*, 59(4):1877–1900, 2004.

J. J. Lucia and E. S. Schwartz. Electricity prices and power derivatives: Evidence from the Nordic power exchange. *Review of Derivatives Research*, 5:5–50, 2002.

M. Ludkovski and R. Carmona. Valuation of energy storage: An optimal switching approach. *Quantitative Finance*, 10(4):359–374, 2009.

T. Meyer-Brandis and P. Tankov. Multi-factor jump-diffusion models of electricity prices. *International Journal of Theoretical and Applied Finance*, 11(5):503–528, 2008.

E. Näsäkkälä and J. Keppo. Hydropower with financial information. *Applied Mathematical Finance*, 15(5):1–27, 2008.

Nord Pool. http://www.nordpool.com. Last visited 2010.03.06. 2010.

NVE. *Kostnader ved produksjon og varme*. The Norwegian Water Resources and Energy Directorate, 2007.

NVE. The Norwegian Water Resources and Energy Directorate. http://www.nve.no. Last visited 2010.01.06, 2010.

R. S. Pindyck. The dynamics of commodity spot and futures markets: A primer. *The Energy Journal*, 22(3):1–30, 2001.

REN21. Renewables global status report 2009. Technical report, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH., 2009. http://www.ren21.net/.

F. Kjærland. A real option analysis of investments in hydropower: The case of Norway. *Energy policy*, 35:5901–5908, 2007.

E. S. Schwartz. The stochastic behavior of commodity prices: Implications for valuations and hedging. *The Journal of Finance*, 52(3):923–973, 1997.

N. Secomandi. Optimal commodity trading with a capacitated storage asset. *Management Science*, 56(3):449–467, 2010.

K. Skytte. The regulating power market on the Nordic power exchange Nord Pool: an econometric analysis. *Energy Economics*, 21(1):295–308, 1999.

M. Thompson, M. Davison, and H. Rasmussen. Valuation and optimal operation of electric power plants in competitive markets. *Operations Research*, 52(4):546–562, 2004.

M. Thompson, M. Davison, and H. Rasmussen. Natural gas storage valuation and optimization: A real options application. *Naval Research Logistics*, 56(3):226–238, 2009.

R. Weron, M. Bierbrauer, and S. Truck. Modeling electricity prices: Jump diffusion and regime switching. *Physica A: Statistical Mechanics and its Applications*, page 182, 2004a.

R. Weron, I. Simonsen, and P.Wilman. Modeling highly volatile and seasonal markets: evidence from the nord pool electricity market. In H. Takayasu, editor, *Toward Control of Economic Change - Applications of Econophysics*, page 182. Springer-Verlag, 2004b.

O. Q. Wu, D. D. Wang, and Z. Qin. Seasonal energy storage operations with limited flexibility. Working paper, Stephen M. Ross School of Business, University of Michigan, Michigan, 2010.

Appendix A

Appendix

A.1   Hard thresholding method

The spike filtering algorithm is based on a method used by Meyer-Brandis and Tankov [2008]. We will here explain this procedure. First of all, we assume that the spike path is a deterministic function, $g(t)$

$$g(t) = \sum_{i=1}^{M} D_i \mathbf{1}_{t \geq \tau_i} e^{-\theta_2(t-\tau_i)} \tag{A.1}$$

where M is the number of spikes, D_i are the jump sizes and τ_i are the jump times. When subtracting this spike series, $g(t)$ from the deseasonalized log price series, we should achieve the base signal.

$$s_t^{des} = Y_1(t) + g(t) \tag{A.2}$$

$$s^{des}(j) = Y_1(j) + g(j), j = 1, \ldots, N \tag{A.3}$$

where N is the number of time steps in our daily price series.

Now consider the case where we want to place only one spike. Let $X(j)$ be the series from which we want to subtract *one* spike, that is that for this case $X(j)$ is equal to the initial deseasonalized price path, $X(j) = s^{des}(j)$. This one spike, with size D^* and initial jump time τ^* is placed optimally according to the following function

$$(D^*, \tau^*) = \arg \inf_{D, \tau} \sum_{j=1}^{N} (\Delta X(j) - \Delta g(j))^2 \tag{A.4}$$

where for a function $X(j), \Delta X(j)$ is defined as

$$\Delta X(j) = X(j) - e^{-\theta_1} X(j-1) \tag{A.5}$$

$$= (Y_1(j) + g(j)) - e^{-\theta_1}(Y_1(j-1) + g(j-1)) \tag{A.6}$$

$$= \Delta Y_1(j) + \Delta g(j) \tag{A.7}$$

$\Delta X(j)$ has a high absolute value where a sudden upward or downward change in the price path occurs and the spike are placed where the absolute difference between $\Delta X(j)$ and $\Delta g(j)$ can be reduced the most. To account for price variations caused by the base signal, the decay factor of the base signal $e^{-\theta_1}$ is included.

We have now separated the base signal, $Y_1(t)$ from the spike series $g(t)$ containing one spike, by setting

$$g(t) = D^* \mathbf{1}_{t \geq \tau^*} e^{-\theta_2(t-\tau^*)} \tag{A.8}$$

$$Y_1(t) = s_t^{des} - g(t) \tag{A.9}$$

If we now want to place more spikes, we do the same one more time, but now using
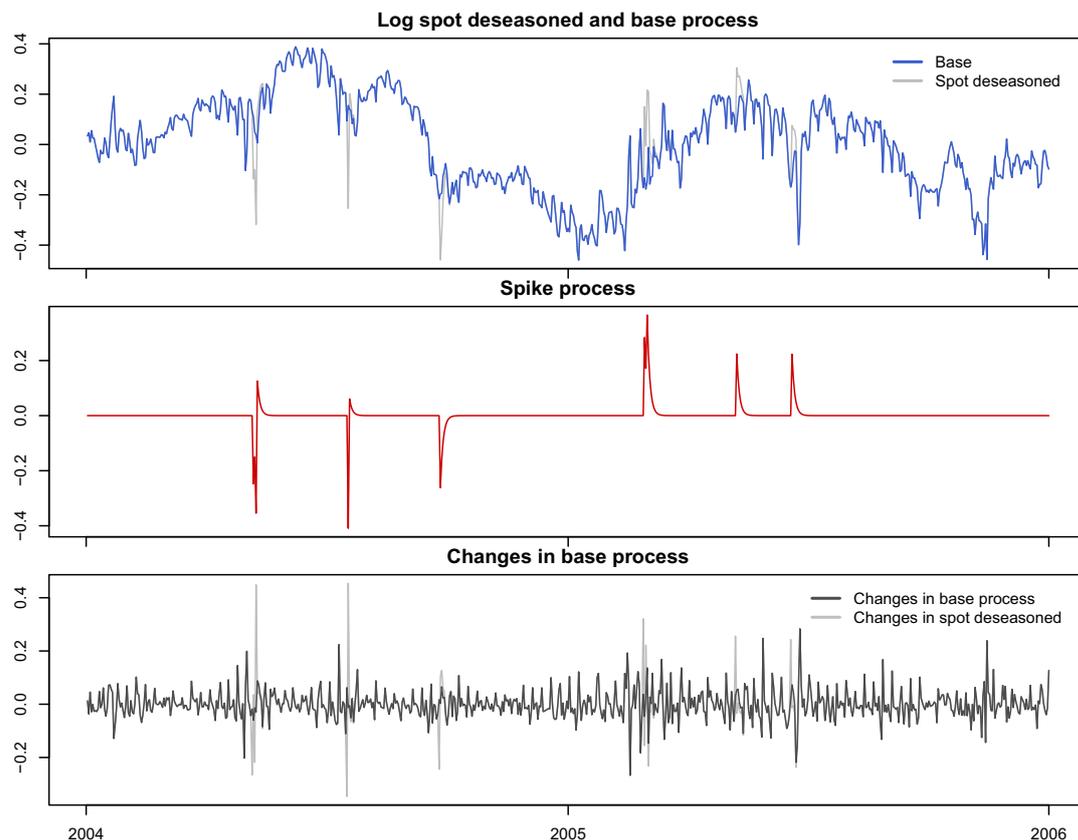
Figure A.1: Results from the spike detection algorithm. The upper plot show the daily deseaonalized log price (gray) separated into the base process (blue) and the spike process (red). The lower plot show (j) before (gray) and after (blue) the spikes are placed. Note that the spike series plus the base signal equals the original deseasonalized log price (gray).

the new base signal $Y_1(t)$ as the initial price path, $X(j) = Y_1(j)$. The spikes found are removed from the initial base signal, and added to the spike series one by one. Figure A.1 illustrate this procedure. The upper plot show the daily deseaonalized log price and the base process, while the middle plot shows the spike process. The lower plot show $\Delta X(j)$ before and after the spikes are placed. We summarize the procedure in the algorithm below. The spike detection ends when we have reached a predefined desired variance of the returns in the base signal.

1. Calculate the desired variance of returns in the base signal; σ_{target}^2, as the variance of the returns in the deseasonalized log prices after removing ϵ of the highest absolute returns.

2. Set $m = 1$, $X(j) = s^{des}(j)$ and $g(j) = 0$ for $j = 1, \dots, N$

3. Find the spike parameters $(_m, \tau_m)$ using formula A.4.

4. Set

$$g(t) = \sum_{i=1}^{m} D_i \mathbf{1}_{t \geq \tau_i} e^{-\theta_2(t - \tau_i)} \qquad (A.10)$$

5. Set

$$X(j) = s^{des}(j) - g(t) \qquad (A.11)$$

6. Calculate the variance of the returns of the base signal after m spikes

$$\sigma_m^2 = var(X(j) - X(j-1)), \ j = 2, \ldots, N \qquad (A.12)$$

If $\sigma_m^2 > \sigma_{target}^2$, set $m = m + 1$ and go to 3. Else; go to next step.

7. Set $M = m$ and $Y_1(t) = s_t^{des} - g(t)$. Calculate spike intensity $\lambda = M/N$.

As seen we find the spike intensity by dividing number of spikes, M by the total number of data points. This calculation assumes a constant spike intensity throughout the season, which we will use in our model. However, it is also possible to calculate a time dependent spike intensity, for example one during summer and one during winter. Next, the distribution of spike sizes is determined. We will not go into detail about this, but rather refer to the general literature about fitting distributions to data.

It follows from equations A.1 and A.5 that we need to know the mean reversion coefficient of both the base signal and the spike series, θ_1 and θ_2 respectively, to perform this spike filtering. To estimate these, we investigate the autocorrelation function, $h(t)$, of the deseasonalized log prices. The autocorrelation function are often representable as a sum of exponentials, and according to our model with two different mean reversion coefficients a sum of *two* exponentials:

$$h(t) = corr(S(t + \tau), S(\tau)), \ \tau = 1, \ldots, N \qquad (A.13)$$
$$= \omega_1 e^{-\theta_1 t} + \omega_2 e^{-\theta_2 t} \qquad (A.14)$$

where the first exponential corresponds to the mean reverting base signal, and the second to the faster reverting spike series. Based on the parameters of Meyer-Brandis and Tankov [2008] for other electricity price series, and using visual inspection on both the autocorrelation function and how the spike detection algorithm performed under different values for θ_1 and θ_2, we found that $\theta_1 = \frac{1}{150}$ and $\theta_2 = \frac{1}{2}$ fits well to our data. As Meyer-Brandis and Tankov [2008] also point out; the performance of the algorithm is rather robust to certain values of these parameters. For $\theta_1 \in [1/20, 1/200]$ and $\theta_1 \in [1/8, 1]$, there is almost no difference in the number of spikes found and the resulting base signal. Note that the value estimated for θ_2 are the actual value that are used in the spot model, while θ_1 are estimated again once the separation of the components is finished. With $\theta_2 = \frac{1}{2}$, 60% of the spike remains after one day, while only 5% after six days.

A.2 Additional tables and figures

Table A.1: Estimated producer dependent parameters for the inflow, overall reservoir and threshold dispatch models for all power stations in the sample.

Power station	Inflow model								Overall reservoir model			Threshold function			
	A	a	B	b	C	c	ϕ_i	σ^I	$\gamma_{1,i}$	$\gamma_{2,i}$	σ^R	α_s	α_t	α_i	α_r
A	0.048	0.046	0.058	0.133	-1.027	0.363	0.358	0.044	0.967	0.024	0.021	1.15	15	0.10	0.5
B	0.066	0.068	0.099	0.110	-0.899	0.161	0.126	0.140	0.979	0.018	0.021	1.15	27	0.05	1.0
C	0.021	0.002	0.015	0.023	2.161	0.193	0.272	0.027	0.972	0.022	0.021	1.20	27	0.05	0.5
D	0.025	0.011	0.016	0.035	1.184	0.235	0.435	0.035	0.975	0.017	0.021	1.20	28	0.05	2.0
E	0.050	0.044	0.024	0.087	-2.072	0.409	0.284	0.083	0.965	0.020	0.021	1.15	25	0.05	1.0
F	0.021	0.000	0.024	0.030	2.431	0.189	0.521	0.131	0.986	0.003	0.022	1.05	15	0.15	4.0
G	0.100	0.089	0.097	0.190	-1.517	0.339	0.574	0.015	0.957	0.022	0.021	1.15	26	0.05	0.5
H	0.055	0.058	0.089	0.239	-0.453	0.593	0.579	0.028	0.978	0.010	0.022	1.15	20	0.05	0.5
I	0.008	0.019	0.166	0.094	-0.138	0.105	0.373	0.216	0.973	0.025	0.021	1.15	22	0.05	0.5
J	0.018	0.000	0.051	0.101	1.795	0.198	0.315	0.004	0.981	0.017	0.021	1.20	23	0.05	6.0
K	0.065	0.066	0.071	0.194	-0.865	0.290	0.553	0.049	0.970	0.020	0.021	1.20	21	0.10	0.5
L	0.078	0.075	0.083	0.098	-1.446	0.176	0.411	0.178	0.969	0.021	0.022	1.00	33	0.00	3.0
M	0.104	0.086	0.106	0.222	-2.421	0.507	0.076	0.226	0.969	0.033	0.021	1.10	25	0.00	0.5
N	0.034	0.030	0.052	0.052	-0.562	0.123	0.290	0.039	0.965	0.023	0.021	1.20	25	0.20	3.0

Table A.2: Top 10 power producers in Norway, 2006, Source: Norwegian Water Resources and Energy Directorate (NVE). *indicates owned by Statkraft.

Power producer	Mean annual production	Domestic market share	Nordic market share
	[TWh]	[%]	[%]
Statkraft	35.9	30.0	9.4
BKK Produksjon AS	6.9	5.8	1.8
Norsk Hydro ASA	6.9	5.8	1.8
E-CO Vannkraft AS	6.8	5.7	1.8
Lyse Produksjon AS	5.9	4.9	1.5
Agder Energi Produksjon AS	5.6	4.7	1.5
Skagerak Kraft AS	4.0	3.3	1.0
Nord-Trøndelag Elektrisitetsverk FKF	3.3	2.8	0.9
Trondheim Energiverk Kraft AS*	3.2	2.7	0.8
Otra Kraft AS	2.6	2.2	0.7
Total top 10	81.1	67.7	21.1
Total	119.9	100.0	31.2

Table A.3: Parameters of the alternative spot models. Alternative model 1 is a multi-factor model without correlation with inflow. Alternative model 2 consists of a seasonal function and a mean reverting base signal driven by a Brownian motion.

	θ_1	σ^S	S_0
Model 1	0.0206	0.0699	3.14
Model 2	0.0438	0.1034	3.16
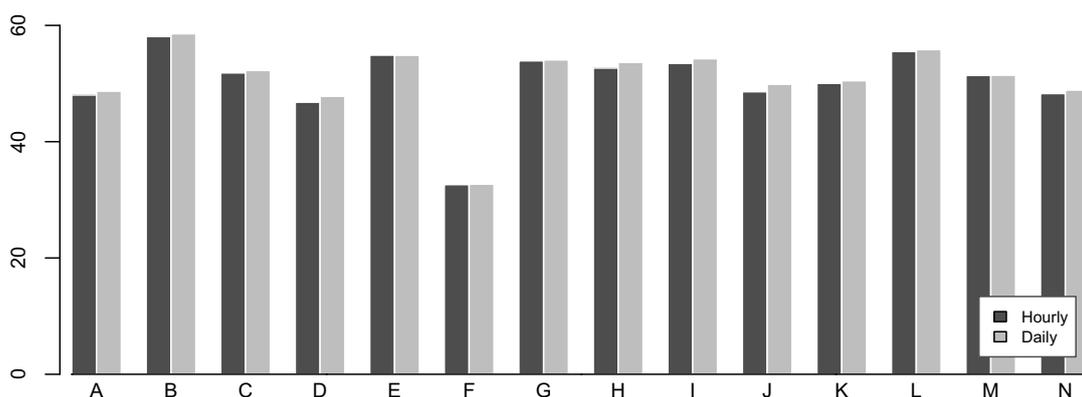
Figure A.2: Differences in actual historic relative revenues [EUR/MWh]] between using hourly and spot prices (dark gray)) and using daily total daily production and daily average spot prices (light gray)

(a) Relative revenues (river)

(b) Relative revenues (max)
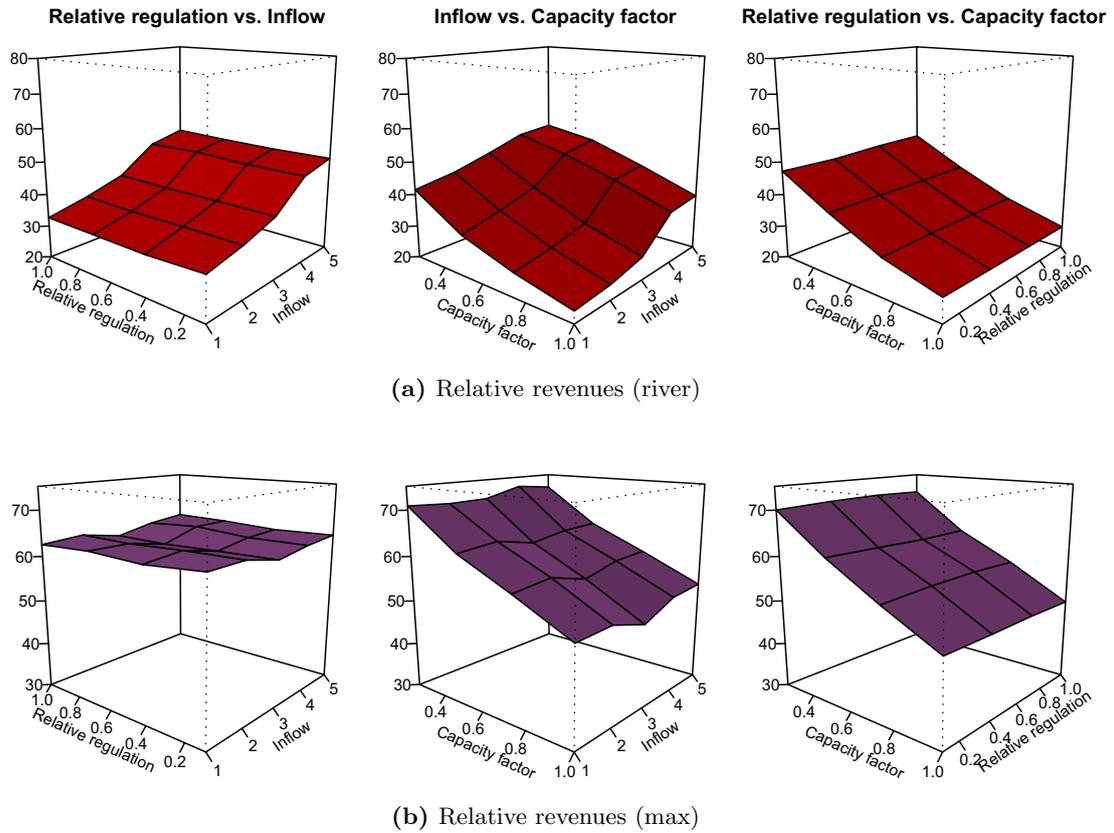
Figure A.3: Relative revenues (river and max) for different relative regulations, inflow seasonalities, and capacity factors. Note that the scale on the inflow axis corresponds to the degree of seasonality, with 1 being strong and 5 being flat.
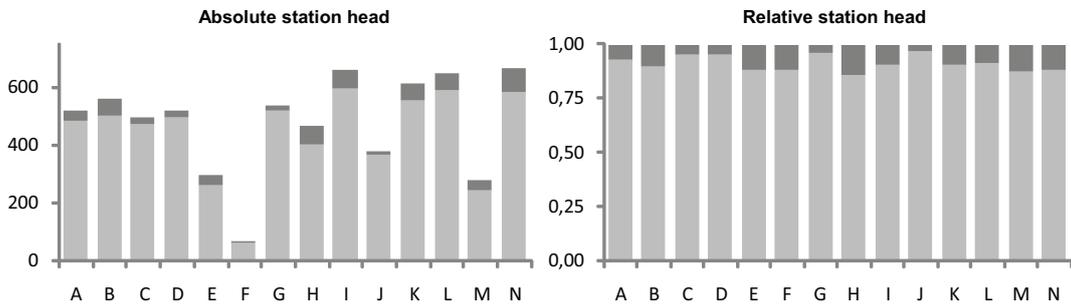
Figure A.4: Left: Absolute power station head with upper and lower reservoir boundary (dark gray area). Right: Upper and lower reservoir boundaries relative to maximum station head. Average difference in head between upper and lower reservoir boundary is 9 % of maximum head.