NTNU
Norwegian University of
Science and Technology

Empirical Analysis of Hydropower Scheduling

Sebastian Brelin
Morten Adrian Lien

Industrial Economics and Technology Management
Submission date: June 2017
Supervisor: Stein-Erik Fleten, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Preface

This master thesis was written within the field of Financial Engineering at the Norwegian University of Science and Technology (NTNU), Department of Industrial Economics and Technology Management.

We would like to thank our supervisor, Professor Stein-Erik Fleten for his time and guidance. His insight and expertise in the field has been especially valuable.

Abstract

A reservoir manager at a hydropower plant has to decide whether or not to release water in order to produce electricity, and the level at which to produce. These production levels have different related efficiencies as well as other related technical aspects. Often, the plant will produce at the most efficient, i.e. release water at a rate that produces the highest amount of electricity per unit of water. In this thesis, a structural estimation model was applied to an undisclosed hydropower plant in the Norwegian electricity price zone NO5, in order to discover the managers' preferences related to the different production levels. This model is based on time series models in order to replicate the managers' expectations of future conditions. The results show a greater willingness of the manager to produce at levels below than above the best efficiency point, which we argue is mainly due to the increased level of cavitation. They also imply that the reservoir managers' preferences have changed over time, showing an increased willingness to produce at production levels both above and below the most efficient level.

Sammendrag

En produksjonsplanlegger ved et vannkraftverk må bestemme når det skal slippes ut vann for å produsere elektrisitet og på hvilket nivå det skal produseres. Produksjonsnivåene har forskjellige effektivitetsnivå, i tillegg til andre tekniske aspekter som kavitasjon. Vannkraftverket vil ofte produsere på det mest effektive nivået, det vil si at det slippes ut vann ved en hastighet som produserer den største mengden elektrisitet per enhet vann. Det kan allikevel observeres at dette ikke alltid stemmer. I denne oppgaven blir en strukturell estimeringsmodell brukt på et vannkraftverk i det norske prisområdet NO5 for å avdekke produksjonsplanleggerens preferanser relatert til de forskjellige produksjonsnivåene. Denne modellen er basert på tidsseriemodeller for å gjenskape planleggerens forventninger om framtidige forhold. Resultatene viser en større tilbøyelighet til å produsere på nivåer under enn over nivå et med høyest effektivitet, hvilket vi argumenterer for at hovedsakelig er på grunn av økte kavitasjonsskader på disse nivåene. Resultatene antyder også at planleggerens preferanser har endret seg over tid ved en økt tilbøyelighet til å produsere på både høyere og lavere nivåer enn det mest effektive.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

A manager of a hydropower plant has the choice of releasing water and generating electricity or saving this water for later and, hopefully, making a greater return. His choices depend on his current conditions and on his expectations for future conditions, such as inflow, the water collected in the reservoir from nearby areas, and the price of electricity. In hydropower production, water can be considered the fuel used for generating electricity. This fuel, however, is not bought and filled into the reservoir at a time of convenience. Rather, it is determined by the precipitation in nearby areas and the melting of snow.

Hydropower plants use different forecasting methods to predict the future expected conditions and determine a production schedule which will provide them with the greatest returns. The forecasted conditions are essential to the choices being made by the reservoir manager. By making our own expectations of the future conditions and applying a structural estimation model, we attempt to analyze these choices. By comparing the choices made, the current surrounding conditions, and the expected future conditions, parameters related to the manager's preferences can be estimated. The main goal of this model is to describe and analyze the choices the producer makes and the power plant's production policy, rather than improve returns.

The Nordic countries have a deregulated market for trading of electrical energy. In addition to a deregulated market, a hydropower producer also faces technical difficulties as to which production level to choose in order to make the end return as high as possible. A turbine has a specific efficiency curve that incentivizes production at the best efficiency point (BEP). However, producing at a higher level when prices are high might result in a higher total return despite the loss of efficiency. Likewise, producing at a lower level in order to store water when anticipating a rising price level, might also result in higher returns. Further, the production efficiency curve changes over time due to damages to the turbine. In addition, some levels of production entail certain problems, such as increased maintenance costs, cavitation etc. The production policy that the operator uses might hide economic preferences which are not explained by the mechanical loss of efficiency itself. Discovering these preferences can be valuable in the analysis of a hydropower plant, both

for outsiders and for the reservoir managers themselves. The results can be used in internal discussions of whether they reflect the intended operational policy of the plant.

In this thesis we apply structural estimation theory of Markov decision processes to a hydropower planning problem in order to discover and describe the behaviour and preferences of the reservoir manager at a hydropower plant. The model is based on previous work by Su and Judd (2012), Fleten et al. (2015), and Boger and Vestbøstad (2016). The model is applied on two different hydropower plants located in the Norwegian electricity price zones NO4 and NO5. The results show how the producer prioritizes the production levels.

Structural estimation was first applied to a switching problem by Rust (1987). The problem considered the optimal switching of bus engines for the maintenance department. He analyzed the action made by the decision maker, which were either to replace the engine, incurring the connected cost, or to wait. Waiting resulted in an increased risk of engine failure, which in itself incurred a cost in addition to damages to customer goodwill. Rust approached the decision-making problem inversely. Rather than using specific assumptions about the objective function and the stochastic process, he tried to infer these from the data. This way it was possible to uncover unobservable preferences and expectations of the decision-maker. According to Rust (1994), structural estimation models are able to predict more accurately the impacts of policy changes than reduced form models.

In order to estimate the structural parameters in a stochastic dynamic programming problem, Rust used an algorithm called the Nested Fixed Point (NFXP) algorithm. This algorithm is separated in two loops. The inner loop solves the stochastic dynamic programming model and the outer loop searches for structural parameters that maximizes the likelihood of the data. Su and Judd (2015) found the NFXP-algorithm computationally demanding due to iterating over all structural parameter values and then solving the underlying stochastic dynamic programming problem for each structural parameter value. They proposed a constrained optimization model referred to as the Mathematical Program with Equilibrium Constraint (MPEC) approach. This approach does not have an inner loop that runs for each iteration of the outer loop, but uses instead the stochastic dynamic programming model as a constraint in the maximization of the likelihood. Using this approach, the only stochastic dynamic programming model to be solved accurately is the one corresponding to the final estimate of the structural parameter. MPEC's computational advantage over NFXP is the reason why it is utilized when applying the structural estimation model in this thesis.

The planning problem of a hydropower producer can be treated as a dynamic problem, since a certain decision changes the level of the reservoir, thereby affecting future potential production. The problem is stochastic because of the uncertainty in the variables influencing the decisions, such as inflow and electricity price. The result is a stochastic dynamic problem (SDP), suitable for the MPEC approach by Su and Judd (2012). Modelling the hydropower planning problem as an SDP has earlier been done by Fosso et al. (1999), Mo et al. (2001), Fleten et al. (2002) and Wolfgang et al. (2009).

A parametric approach is used for the transition probabilities between the states in the structural estimation model, as suggested by Fleten et al. (2016). A state is characterized by the current electricity price, the inflow, the reservoir level, the deviation from cumulative inflow, and the deviation from the aggregate reservoir level. The parametric approach involves time series modeling of the state variables as Markovian processes, where the next state is only dependent on the previous one. The goal is to capture the dynamics of these state variables in order to successfully apply the structural estimation model. The state variables in this thesis are modeled in a similar way as done by Kolsrud and Prokosch (2010) and Boger and Vestbøstad (2016).

Structural estimation has earlier been utilized on a wide number of different problems. Rust and Rothwell (1995) and Rothwell and Rust (1997) applied it to electricity related problems, like regulatory shifts and optimal lifetime for a nuclear power plant. Rapson (2014) studied the timing of appliance investment, Kellog (2014) applied structural estimation to well drilling and Lin and Thome (2013) used it for corn-ethanol plant investment. In 2016, Fleten et al. applied it to maintenance and switching costs of peak power plants. Structural estimation was applied to hydropower planning by Boger and Vestbøstad (2016).

The next chapter presents the Nordic energy market and the technical basics of hydropower production. Further, structural estimation, Markov Decision processes and the state space transitions are explained. We then elaborate on the underlying time series models for the state variables and how they are connected. The results of the model in terms of marginal water value and production preferences are then presented and analyzed.

Chapter 2

Hydropower and the Nordic energy market

2.1 The Nordic energy market

The Nordic countries decided in the 1990's to deregulate the market for trading electrical energy, leaving behind the landscape of a traditional monopolistic and government-controlled power sector (Weron and Misiorek, 2008). As of today, Nord Pool is Europe's leading power market. Nord Pool offers trading, clearing, and settlement in both day-ahead markets and intraday markets across nine European countries. The Elspot market, also referred to as the day-ahead market, is the primary market for power trading. Agreements are made between sellers and buyers for the delivery of power, hour-by-hour the following day. The intraday market, also called the Elbas market, supplements the Elspot market and helps secure the necessary balance between supply and demand. In this market, hourly contracts are continuously traded in the period between clearance in the day-ahead market and up to one hour before the hour of operation (IBP, 2015).

A power market needs a connected and well-developed grid with access for all players to be effective. The operation and management related to the grid is monopolistic, and there is no competition within these sectors (Weron and Misiorek, 2008). To prevent grid companies to exploit this position, authorities have put in place strict regulations on their operations as monopolies. As of today, Statnett is the main grid operator and owner in Norway.

The power market distinguishes between wholesale and end-users. In the wholesale market, large volumes are bought and sold between power producers, power suppliers, brokers, energy companies and large scale consumers. In the Nordic countries, these players trade on Nord Pool Spot. In the end-user market, individual consumers enter agreements to purchase power from a supplier of their own choice. The end-user market in Norway consists of one third household customers, one third industry and one third medium sized

consumers. By medium sized consumers, we consider hotels and chain stores (NordReg, 2014).

2.1.1 Price Formation

Electricity differs from other commodities, because it cannot be easily stored. This implies that a constant balance between supply and demand is required, in order to have power system stability (Kaminski, 2012). The market price of power is determined each day on the Nord Pool Spot exchange. Players who wish to buy or sell power on the Elspot power market send their orders to Nord Pool Spot before noon the day before the power is to be delivered to, or withdrawn from, the power grid. At Nord Pool spot, purchase orders are aggregated to a demand curve and sale offers are aggregated to a supply curve. The intersection of these two curves indicate the market price of power for a specific hour, as shown in figure 2.1. This way of calculating the price is referred to as *double action*, because both buyers and sellers submit orders. At most other auctions, only the bidder places orders.

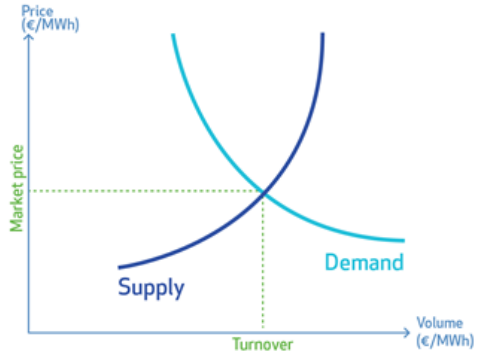

Figure 2.1: Supply and demand Curve (NordPool, 2017)

The rising supply curve indicates the amount of power that producers are willing to produce at different prices, thereby reflecting the marginal production cost of power in each type of plant. Hydropower, wind power and nuclear power have the lowest marginal costs in the Nordic region and are offered at low prices. Gas, biopower, and coal-based power have higher marginal costs and are therefore located further to the right on the supply curve. The electricity spot price for each period is set by the most expensive generator required to satisfy demand (Hagfors et al., 2016).

2.1.2 Area Prices

Nord Pool Spot operates with one system price and several area prices. The system price is determined by the balance between total supply and demand in the Nordic market. The

price does not consider the physical capacity constraints in the transmission networks, which is why Nord Pool Spot also operates with area prices. As of today, Norway is divided into five areas where the prices are determined by area-specific market characteristics. Some areas have a surplus of power, while others have a deficit. Power system stability is determined by the balance between import and export across the different bidding areas. Insufficient capacity in the grid system for import or export results in different prices between price zones, and are referred to as bottlenecks. If there were no restrictions in power flow, all bidding areas in the Nordic region would have the same price.

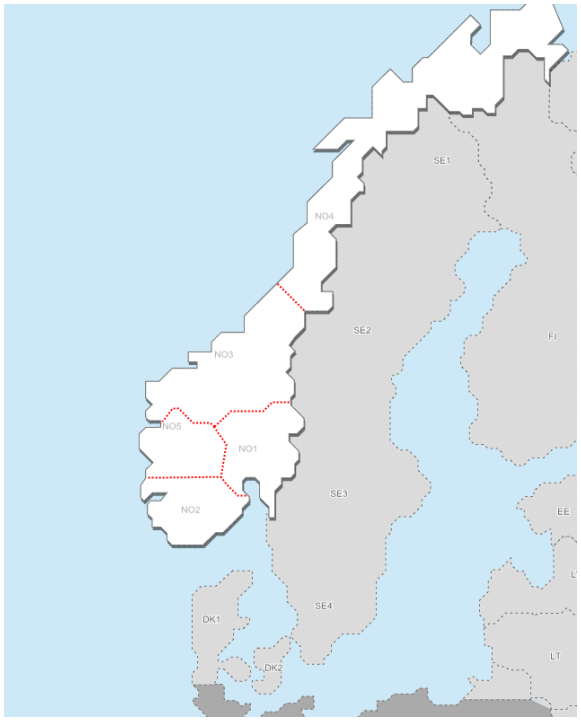

Figure 2.2: The price areas for Nord Pool (Nord Pool, 2017)

2.1.3 Power production in Nord Pool

Table 2.1 shows the average production split from 2002-2014. Hydropower is the single largest contributor to the power distributed through Nord Pool. Norway is the largest provider of electricity generated by hydropower with 32.16% of the total production. Hydropower contributes with 52.9% of the total power distributed through Nord Pool (NordPool, 2012).

Country Energy Source	Denmark	Finland	Norway	Sweden	Sum	Share of total generation
Hydropower	0.0	12.3	121.4	65.8	199.4	52.9
Nuclear power	0.0	22.3	0.0	58.0	80.3	21.3
Fossil fuels	21.8	24.2	4.8	5.4	56.1	14.9
Wind power	8.9	0.5	1.3	6.1	16.7	4.4
Other renewables	2.4	10.5	0.0	11.2	24.1	6.4
Non-identifiable	0.0	0.7	0.0	0.0	0.7	0.2
Total production	33.1	70.4	127.4	146.4	377.4	100.0

Table 2.1: Production Split 2004-2012 (Nord Pool, 2017)

2.2 Hydropower planning

As previously mentioned, hydropower is the main source of electricity distributed through Nord Pool. The value of the water stored in the reservoirs is determined by calculating the opportunity cost of using water immediately as opposed to storing it for later use. The uncertainties are approached with complex stochastic dynamic optimization tools. These tools are used to calculate water values, optimize power production, manage the reservoir, and to forecast prices. The objective is to maximize the profits in the spot market and the value of water at the end of the planning period. The optimization problem is often solved with stochastic dynamic programming (SDP) and stochastic dual dynamic programming (SDDP), as in Fleten and Wallace (2003). Often, the planning period for these models is a couple of years with a resolution of one week. The most important values are marginal water values and production schedule (Botterud et al., 2002).

The optimization models used for hydropower planning depend on different factors due to the size, complexity, and revenue of the power company. Models may depend on weather forecasts for precipitation, wind and temperature, reservoir levels, inflow forecasts, the value of future and current snow pack, hydro balance, fuel and emission prices, power plant transmission outages, market prices, in-house and external price forecasts and import/export expectations. The technical specifications and conditions for each power plant also affect the production schedule. The efficiency of the turbines changes through the life of the power plant. Production levels and their related efficiencies, that is, the electricity produced per m^3 of water released, is also an important aspect of the production schedule to maximize profits (Botterud et al., 2002).

2.3 Basics of hydro power production

A Hydropower turbine converts water pressure into mechanical shaft power, which can be used to drive an electricity generator. The power available is proportional to the product of volume flow rate and pressure head. The general formula for a hydro system‘s power output is

$$P = \eta \rho g Q H$$

Where P is the mechanical power produced at the turbine shaft, η is the efficiency of the turbine, ρ is the density of water (kg/m^3), g is the acceleration due to gravity (m/s^2), Q is the volume flow rate passing through the turbine (m^3/s), and H is the effective pressure head of water across the turbine (m). Some of the best turbines can have hydraulic efficiencies of 80 to 90% (Paish, 2002).

When comparing turbine types, an important aspect is their efficiency at different flow rates. The operator of the power plant will consider the combined efficiency of the turbine and the generator when planning the production schedule. The turbines show declined performance after few years of operation as they get severely damaged. One main reason is erosive wear of the turbines due to high content of abrasive material during floods and cavitation (Kumar and Saini, 2010).

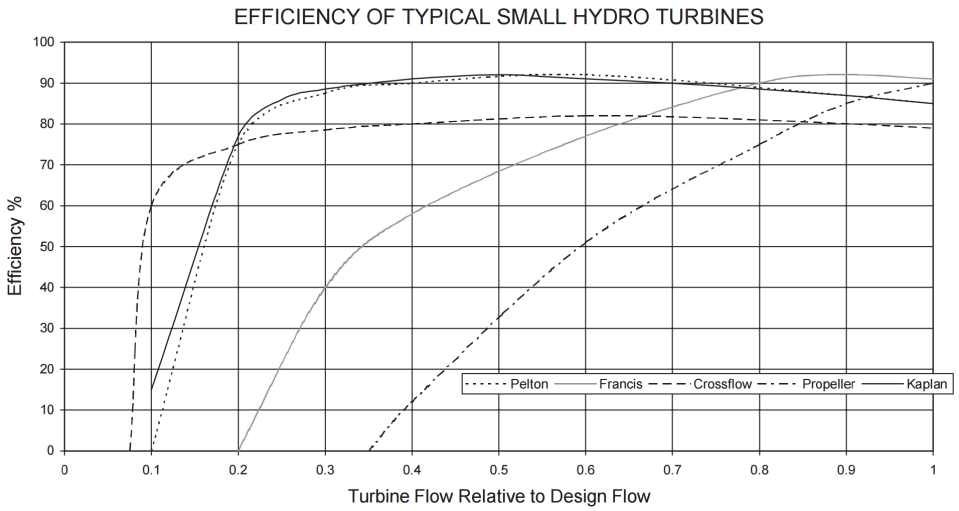

Figure 2.3: Efficiency curves for different turbines (Okot, 2012)

Impulse and reaction turbines are the two main categories of turbines in hydropower production (U.S. Department of Energy, 2017). According to Kjøll (2003), the selection of turbine for a hydropower project is determined on the head, flow of water, how deep the turbine must be set, efficiency, and cost. Damages concerning water turbines are caused mainly by cavitation problems, material defects, sand erosion, and fatigue. Damages occur primarily in turbines with a head of 250 m and above. These damages are consequences of high pressure, pressure variations, and high water velocities. Impulse turbines are hard to operate at production plants where the head is relatively low and discharge is medium or high because of the necessary speed of the flow in order to drive the generator. Reaction turbines are preferred for medium and low head plants. However, these turbines are more prone to cavitation. A zone in the operating range is heavily affected by cavitation, and

is therefore considered a ”forbidden production zone” (Kumar and Saini, 2010). This is especially true for Francis turbines.

According to Kumar and Saini (2010), small bubbles of vapors are formed when the pressure in a part of the turbine drops below the evaporation pressure. Vortexes that contain voids or bubbles may appear as a result of streams of water cutting their paths short. These bubbles, formed by low pressure, are carried by the stream to parts with higher pressure and the bubbles suddenly collapse, since the vapor are condensed to liquid again. This results in formation of cavity and the surrounding liquid rush in to fill the gap. The stream of liquid arriving from different location to fill this gap collides at the center point of cavitation and gives rise to a high local pressure, this might be as high as 7000 atm. This formation of cavity and high pressure is repeated many thousand times per second. This causes pitting on the metallic surfaces of the runner blades and the material fails by fatigue (Kumar and Saini, 2010).

By applying our structural estimation model to the time series for the two power plants, we are able to investigate their production preferences. These preferences are influenced by the economic incentives and downsides related to the mechanical specifications of the plant. The time series for the power plants range approximately 20 years each, so when comparing production preferences for the beginning of the time series and the end of the times series, we are able to identify how these preferences change over time.

Chapter 3

Structural Estimation of Markov Decision Processes

3.1 Introduction

This thesis is largely based on the work on structural estimation of Discrete Choice Dynamic Programming (DCDP) models by John Rust. Rust (1987) estimates structural parameters in a problem of optimal replacement of bus engines. He analyzes the actions made by decision-maker Harold Zurcher, which are either to replace an engine and incurring the connected costs or to wait. Waiting results in an increased risk of engine failure, which in itself incurs costs in addition to damages to customer goodwill.

A key insight of Rust (1987) is the inverse way of approaching the decision-making problem. Rather than using specific assumptions about the objective function and the stochastic process in order to derive an optimal strategy, we try to infer the objective function and the stochastic process from the data. This way, we can uncover the unobservable preferences and expectations of the decision-maker (Rust, 1988). Rust (1987) also introduces the analysis of counterfactuals in structural estimation models: *...we obtain a rich behavioural model that can be used to answer a wide range of what if? policy questions.* According to Rust (1994), structural models are able to predict more accurately the impacts of policy changes than reduced-form models.

Reduced-form models can be used to uncover an agents historical decision rule. When predicting future behaviour, the environment needs to be stationary in order for this decision rule to be applicable. Therefore, the reduced-form models do not account for the changes in an agents preferences when subject to policy changes in the same way as structural models:

The hope is that the changes in θ induced by policy changes will occur slowly, and that conditional forecasting based on tracking models will therefore be roughly accurate for a

few periods. This hope is both false and misleading. (Lucas, 1976)

3.2 Markov Decision Processes

We call d_t the choice of the decision-maker at time t, and x_t the observed state variables of the decision maker at time t. We model $\{d_t, x_t\}$ as a stochastic process and use an assumption of expected discounted profit maximization. That is, we assume that the decision-maker is profit maximizing and that future reward is discounted according to a discount factor β. The process $\{d_t, x_t\}$ is then generated from a solution to a dynamic programming problem. This solution can only be found recursively, using Bellmans principle of optimality (Rust, 1987). The process is *Markovian* because the decision rule δ only depends on the current state and not on previous decisions. However, this does not mean that the decisions that are made do not influence the expectation of the next state. The next state x_{t+1} depends both on the current state x_t and on the decision $d \subseteq D$ that is made.

Our problem is a single-agent MDP. The utility function $g(s_t, d_t)$ represents the agent's preferences at time t, the Markov transition probability $p(s_{t+1}|s_t, d_t)$ represents her beliefs about future states, and β the rate at which she discounts future vs. present value. This way, the agent can be represented by her primitives g, p and β. The goal is to estimate hidden, structural parameters that characterize these primitives.

Using the notation of Rust (1994), the discrete-time MDP consists of:

Table 3.1: Components of the Markov Decision Process

$t \in \{0, 1, 2, ..., T\}, T \leqslant \infty$	Time index
S	State space
D	Decision space
$\{D_t(s_t) \subseteq D\}$	Constraint sets
$\{p_{t+1}(\cdot\|s_t, d_t) : \mathscr{B}(S) \Rightarrow [0, 1]\}$	Transition probabilities
$\{g_t(s_t, d_t)\}$	Utility function
β	Discount rate

3.2.1 Observable vs. Unobservable state variables

We make a distinction in the state variables between those which are observed by both the agent and the researcher, x_t, and those which are observed only by the agent ε_t. This way, the complete state $s_t = (x_t, \varepsilon_t)$. The ε_t may also be referred to as the error term, and represents the variation which we cannot explain. If we had complete information about s_t, there would be no need for the error term ε_t, and the behaviour of our decision-maker

would be completely deterministic. Since this is an unobtainable goal, the error term has to be included. It explains why the agent, who we assume is a rational decision-maker, may decide differently when in two states which seem identical to us.

3.2.2 Stationary MDPs and infinite horizon

Notice how the time horizon may be finite or infinite. When the utility functions and transition probabilities are not dependent on the time t, and the discount rate β is set to a constant $\in [0, 1)$, the MDP is said to be *stationary*. When we have a stationary MDP with infinite horizon, simplifications can be made because the future looks the same given that the state is the same. The current state is therefore the only variable that affects the optimal decision rule and the corresponding value function.

In our case the data contains seasonal effects, which implies that the time of year will actually play a role in the decision-making process. This is solved by using the week as an additional state variable and adding the seasonal effect to the other state variables based on which week it is. This is explained in chapter 4.

3.2.3 Framework and assumptions

There are 6 important assumptions that are necessary to make in order to use the framwork developed by Rust (1987). This model has also been called the "dynamic programming conditional logit model" (Aguirregabiria and Mira, 2009). These are:

AS	Additive separability: The utility function can be separated and the components can be added together. That is, $G(d_t, x_t, \varepsilon_t) = g(d_t, x_t) + \varepsilon_t$.
IID	Independently, identically distributed unobserved state variables ε_t
CI-X	Conditional independence of x: The next observable state variables x_t are not dependent on the current unobserved state variables ε_t.
CI-g	Conditional independence of g: The payoff variable g is independent of the unobserved state variables.
CLOGIT	Conditional logit: The unobserved state variables ε_t are independent of the different possible decisions d_t and have an extreme value type 1 distribution.
DIS	The observed state variables x_t must be discretely supported and there must be a finite number of possible states.

3.2.4 Bellman equation

As pointed out in section 3.2.2, the optimal decision rule is only affected by the current state. Bellman equations are used to find the optimal policy of Markov processes. These equations are based on suboptimality and solved recursively.
The optimal policy is then:

$$\delta(s) = \underset{d \in D}{\operatorname{argmax}} \left[g(s,d) + \beta \int V(s') p(ds'|s,d) \right] \tag{3.1}$$

And the value achieved is:

$$V(s) = \max_{d \in D} \left[g(s,d) + \beta \int V(s') p(ds'|s,d) \right] \tag{3.2}$$

Equation (3.2) is the Bellman equation. It shows that we are selecting decision d in every state such that the total value of the immediate profit and the continuation value is maximized. We are integrating over the value function for the next state, $V(S')$, multiplied by the transition probability of landing in that state given the current state and the decision made, $p(ds'|s,d)$. In other words, when making a decision we are taking into account:

1. The immediate profit achieved by making decision d in the current state, s.
2. The payoffs that we can achieve in future states, s'.
3. The probability of reaching these future states $p(ds'|s,d)$.
4. How much we value future vs. immediate earnings, β.

We can write the Bellman equation as done by Fleten (2016):

$$V(x) = \max_{d \in D} \Big(g\left(x; d\right) + \beta \mathbb{E}_d \big(V\left(X_1\right) | X_0 = x \big) \Big) \tag{3.3}$$

We know from section (3.2.1) that we need to include an error term because of the unobservable state variables. The additive separability assumption allows us to express the immediate profit as a combination of the observable and unobservable states:

$$g(x, \varepsilon; d) = g(x; d) + \varepsilon(d) \tag{3.4}$$

where ε is an idiosyncratic shock. Continuing with Rust's framework (p. 3103, 1994), we use assumption CI-X and factor the state transition probability:

$$\pi(dx_{t+1}, d\varepsilon_{t+1} | x_t, \varepsilon_t, d_t) = \mathcal{E}(d\varepsilon_{t+1} | x_{t+1}) p(dx_{t+1} | x_t, d_t) \tag{3.5}$$

Where $\pi(\cdot)$ is the total state transition probability, $\mathcal{E}(\cdot)$ is the transition probability of ε, and $p(\cdot)$ is the transition probability of the observed part, as above in (3.2). We now have to include the error term in the value function (3.3). Using assumption AS, the first term in the maximization is straightforward: $g(x; d) + \varepsilon(d)$ as in (3.4). The β is, of course, unchanged and the term inside the expectation is factored using assumption CI-X. We are left with the value function:

$$V(x, \varepsilon) = \max_{d \in D} \{ g(x; d) + \varepsilon(d) + \beta \mathbb{E}_d \left(\int V(X_1, \varepsilon_1) \mathcal{E}(\mathrm{d}\varepsilon_1 | X_1) \middle| X_0 = x \right) \} \tag{3.6}$$

Following Fleten (2016), we define the expected value function,

$$v(x,d) := \mathbb{E}_d \left(\int V(X_1, \varepsilon_1) \mathcal{E}(\mathrm{d}\varepsilon_1 | X_1) \middle| X_0 = x \right), \tag{3.7}$$

so that the value function can be written:

$$V(x,\varepsilon) = \max_{d \in D} g(x;d) + \varepsilon(d) + \beta \cdot v(x,d) \tag{3.8}$$

We need a fixed point equation for v. This can be obtained by inserting $V(x,\varepsilon)$ into the expected value function in (3.7):

$$v(x,d) = \mathbb{E}_d \left(\int \max_{d \in D} \{ g(X_1;d) + \varepsilon_1(d) + \beta \cdot v(X_1) \} \mathcal{E}(\mathrm{d}\varepsilon_1 | X_1) \middle| X_0 = x \right) \tag{3.9}$$

3.2.5 Gumbel distribution

Rust (1994) uses the restriction that ε is an IID extreme value process. This is reasonable because of the maximization in the expected value function $v(x,d)$. Using this restriction, our choice probability formula is the according to the logit model:

$$P(d|x) = \frac{exp\left\{ \frac{g(x,d)+\beta EV(x,d)}{b} \right\}}{\sum_{d' \in D(x)} exp\left\{ \frac{g(x,d')+\beta EV(x,d')}{b} \right\}} \tag{3.10}$$

As Fleten (Working paper, NTNU, 2016) remarks, the Gumbel distribution is closed under maximization. The closed form formula is given by:

$$\int \max_{d \in D} (\varepsilon(d) + c_d) \mathcal{E}(\mathrm{d}\varepsilon | x) = b \cdot \log \left(\sum_{d \in D} \exp\left(\frac{c_d}{b} \right) \right) \tag{3.11}$$

Where b is the scale parameter i.e. a number which can be interpreted as the degree of uncertainty. We observe that we can substitute ε with ε_1, c_d with $g(X_1;d) + \beta \cdot v(X_1)$ and $\mathcal{E}(\mathrm{d}\varepsilon_1 | X_1)$ with $\mathcal{E}(\mathrm{d}\varepsilon | x)$. Using the closed form formula, we can then write the expected value function in (3.9) as:

$$v(x,d) = \mathbb{E}_d \left(b \cdot \log \left(\sum_{d' \in D} \exp\left(\frac{g(X_1;d') + \beta \cdot v(X_1)}{b} \right) \right) \middle| X_0 = x \right) \tag{3.12}$$

For notational convenience we define the operator

$$t_\theta(v)(x,d) := \mathbb{E}_d \left(b \cdot \log \left(\sum_{d' \in D} \exp\left(\frac{g(X_1;d') + \beta \cdot v(X_1)}{b} \right) \right) \middle| X_0 = x \right) \tag{3.13}$$

This way we can write the fixed point Bellman equation as

$$v = t_\theta(v) \tag{3.14}$$

3.2.6 Estimation of the conditional expectation

Simplifying the conditional expectation
What remains is now to evaluate the conditional expectation. $v(x, d)$ is currently an expectation conditional on $X_0 = x$. In order to use $v(x, d)$ as a constraint in the structural estimation formulation, we need to have a closed form formula of $v(x, d)$. We are using the parametric approach suggested by Fleten et al. (2016). This involves using an autoregressive scheme with one lag (AR1-processes) in a time series model. We can only have one lag in order for the model to maintain the Markovian property. The autoregressive scheme

$$X_{k+1}^{'} = \mu + \rho X_k^{'} + \sigma \mathcal{N}_t \tag{3.15}$$

has independent, random variables $\mathcal{N}_t$ as the error term. To simplify the conditional expectation to a simple expectation, we can write:

$$\mathbb{E}(f(X_{k+1}^{'})|X_k^{'} = x) = \mathbb{E}f(\mu + \rho x + \sigma \mathcal{N}_t) \tag{3.16}$$

State space transition
We can extend this to our case by denoting the state transition:

$$x_{t+1} = f(x_t, d_t) = Ax_t + Be + c \tag{3.17}$$

That is, when in state x_t, choosing action d_t will result in state x_{t+1} through the function f. This function is specified by matrices A, B, e, and c. Matrix A contains the parameters for the state vector, e contains the error terms of independent random variables, B contains the parameters for the error terms, and c contains the seasonal components.
In order to simplify notation, we denote:

$$h(x) := b \cdot \log\left(\sum_{d \in D} \exp\left(\frac{g(x; d) + \beta \cdot v(x)}{b}\right)\right) \tag{3.18}$$

Since the state variables follow an AR scheme, we can extend the simplification in (3.13) and (3.14) by Fleten et. al. (2016) to the value function in (3.12):

$$t_\theta(v) = \mathbb{E}_d(h(X_1)|X_0 = x) = \mathbb{E}h(Ax + c + BLe) \tag{3.19}$$

Further simplification through discretization of error terms
In order to avoid the expectation in the expression altogether, we need to discretize the error terms in e, and give appropriate probability weights to the different discrete levels. Using N state variables, the value function is then:

$$t_\theta(v) \sim \sum_{i_1=1}^{M_1} \cdots \sum_{i_N=1}^{M_N} w_{i_1} \cdots w_{i_N} h(Ax + c + BL\tilde{e})), \tag{3.20}$$

M_1 represents the number of discretization levels for $\tilde{e}_{i_1}$, M_2 represents the number of discretization levels for $\tilde{e}_{i_2}$ etc., while w_{i_1} is the probability weight on $\tilde{e}_{i_1}$, w_{i_1} is the probability weight on $\tilde{e}_{i_2}$ etc. There is one error term for each state variable, i.e.:

$$\tilde{e} = \begin{bmatrix} \tilde{e}_{i_1} \\ \cdot \\ \cdot \\ \cdot \\ \tilde{e}_{i_N} \end{bmatrix} \tag{3.21}$$

The expected value function now has a closed form solution, so that it can be used as a constraint in the constrained optimization.

3.3 Structural estimation

The structural parameters are estimated using the approach by Su and Judd (2012). Instead of guessing the structural parameters and then solving the corresponding endogenous variables with high accuracy, Su and Judd formulated the maximum-likelihood estimation problem as a constrained optimization problem. This involves using the Bellman equation as a constraint.

The likelihood-function l is based on the structural parameters, the value function and the state for every observation $n \in N$:

$$l(\theta, v_\theta, (X_n, d_n)_{n=1}^N) = \prod_{n=1}^{N} P_v(d_n|X_n) \tag{3.22}$$

We use the choice probability formula $P_v(d_n|X_n)$ from equation (3.5). The likelihood function should reflect the likelihood of all of the observations combined. The probabilities are therefore multiplied with each other. Using the log-transform on the expression will yield:

$$L(\theta, v_\theta, (X_n, d_n)_{n=0}^N) = \sum_{n=1}^{N} log(P_v(d_n|X_n)), \tag{3.23}$$

Where L is the log-likelihood function, which is computationally much easier to solve. The log-transform is monotonically increasing, and the log-likelihood function will therefore have a maximum value in the same point as the likelihood function. As a result, L will yield the same values for θ as l.

The constrained optimization problem can be summarized by:

$$\begin{aligned} \underset{\theta}{maximize} \quad & L(\theta, v_\theta, (X_n, d_n)_{n=1}^N) \\ s.t. \quad & v_\theta = t_\theta(v_\theta) \end{aligned} \tag{3.24}$$

Structural parameters to be estimated The vector θ collects the structural parameters that we are interested in estimating, in our case θ_1 and θ_2, which correspond to the left and right side penalties for deviating from the most efficient production level, as explained in section 4.8. The parameters in θ are set in such a way that the our data set has the highest possible likelihood.

State space transition While we introduce the state variables used in this section, the derivations of their underlying time series models are found in chapter 4. An overview of the state transitions can be found in (A.3). The state variables are:

I_t	Inflow
C_t	Deviation from normal cumulative local inflow
R_t	Deviation from normal overall reservoir level
P_t	Price level
S_t	Local reservoir level

These are collected in the state vector x_t, so that the state transition when going to state x_{t+1} can be denoted:

$$x_{t+1} = f(x_t, d_t) = (I_{t+1}, C_{t+1}, R_{t+1}, P_{t+1}, S_{t+1}) \tag{3.25}$$

We then use the closed form formula in (3.20) for these state variables in order to get our specific expected value function:

$$t_\theta(v)(x, d, t) \sim \sum_{i=1}^{M_i} \sum_{r=1}^{M_r} \sum_{p=1}^{M_p} w_i w_r w_p h(A_t x + c_t + BL\tilde{e})), \tag{3.26}$$

Expanding h, we get:

$$\begin{aligned} t_\theta(v)(x, d, t) \sim & \sum_{i=1}^{M_i} \sum_{r=1}^{M_r} \sum_{p=1}^{M_p} w_i w_r w_p \\ & b \cdot \log\left(\sum_{d' \in D} \exp\left(\frac{g(A_t x + c_t + BL\tilde{e}; d) + \beta \cdot v(A_t x + c_t + BL\tilde{e})}{b} \right) \right) \end{aligned} \tag{3.27}$$

The only remaining part of the expected value function is then the immediate profit function, $g(x; d)$.

Profit function, $g(\cdot)$ By multiplying the current system price with the amount of electricity produced by the power plant, we can formulate a simple profit function:

$$g(X_t; d_t) = P_t u(X_t, d_t) \tag{3.28}$$

This formulation relies on the assumption that there are no marginal production cost, no start-up cost, that the head is constant, and that the power producer is a price taker.P_t is the system price in EUR, and $u(X_t, d_t)$ is the release of water as a function of the current state and decision.

Release, $u(\cdot)$ The release of water from the reservoir is a function of the choice d_t, scaled with a production factor Q. In addition, we use an efficiency function that takes the structural parameters θ_1 and θ_2 that we are trying to estimate, as inputs.

$$u(X_t, d_t) = min\{d_t Q E(\theta_1, \theta_2, \xi), S_t - S^{min} + I_t\} \tag{3.29}$$

If there is not enough water in the reservoir for the desired production level, the amount released is the available water that is above the minimum reservoir level, as seen on the right hand side in the min-operator.

Efficiency function, $E(\cdot)$ The efficiency function is included in order to analyze the power producer's resistance to deviating from the best efficiency point (BEP). The efficiency function is dependent on three factors: the BEP, ξ, the efficiency for production levels beneath the BEP, θ_1, and the efficiency for production levels above the BEP, θ_2.
For production levels below the BEP, $d_t < \xi$ the following equation applies:

$$E(\theta_1, \xi) = (1 - (\xi - d_t)\theta_1), \tag{3.30}$$

and for production levels above the BEP, the efficiency function is:

$$E(\theta_2, \xi) = (1 - (d_t - \xi)\theta_2) \tag{3.31}$$

To test different prioritization, the production efficiency is implemented in 3 different ways: Linear, square root, and squared.

Linear production efficiency: $E(\theta_1, \theta_2, \xi)$

Square root production efficiency: $\sqrt{E(\theta_1, \theta_2), \xi}$

Squared production efficiency: $(E(\theta_1, \theta_2), \xi))^2$

These equations are later included in the profit function so the lost efficiency is relative to the chosen production level.

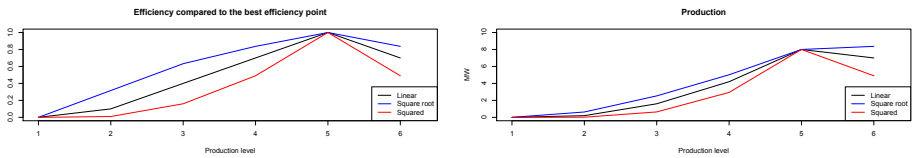

Figure 3.1: Example of efficiency and production curves with θ_1 and $\theta_2 = 0.3$, $\xi = 5$

Figure 3.1 shows an example with $\xi = 5$ and $\theta_1 = \theta_2 = 0.3$. The left plot in the figure shows the efficiency for the three different efficiency functions. The plot to the right shows the production as a result of the efficiency. The production is discretisized to a number between 1 and 6. Where 1 is no production and 6 is maximum production. The actual efficiency and production curves are shown in the results and analysis section.

Final expected value function Expressing the expected value function in terms of the state variables results in:

$$t_\theta(v)(x,d,t) \sim \sum_{i=1}^{5}\sum_{r=1}^{5}\sum_{p=1}^{5} w_i w_r w_p \\ b \cdot \log\left(\sum_{d' \in D} \exp\left(\frac{(A_t^{(4)}x + c_t^{(4)} + (BL)^{(4)}\tilde{e}) \cdot u(A_t x + c_t + BL\tilde{e}, d')}{b} \right.\right. \\ \left.\left. + \frac{\beta \cdot v(A_t x + c_t + BL\tilde{e})}{b}\right)\right), \quad (3.32)$$

Where P_t is expressed as $A_t^{(4)}x + c_t^{(4)} + (BL)^{(4)}$, since the price is represented by the fourth row in the matrices in the appendix, A.3.

Chapter 4

Data description and state variables

The empirical estimation and analysis in this chapter is based on information from two disclosed hydropower plants. Power plant A is located in price zone NO5. The power plant has a reservoir of 35 Mm^3 and a production capacity of 10 MW. Egre and Milewski (Egre and Milewski, 2002) categorize hydropower plants as small when their production is less than 10 MW. The second power plant, B, has two turbines and two generators, a reservoir capacity of 180 Mm^3 and a production capacity of 128 MW. This plant is located in price zone NO4. Since both hydropower plants are approached with the same methods, there will only be a detailed elaboration for power plant A in this chapter. Some details for power plant B are elaborated due to the differences between the two.

The state variables presented in this chapter are based on a model developed by Kolsrud and Prokosch (2010) and later modified by Boger and Vestbøstad (2016).

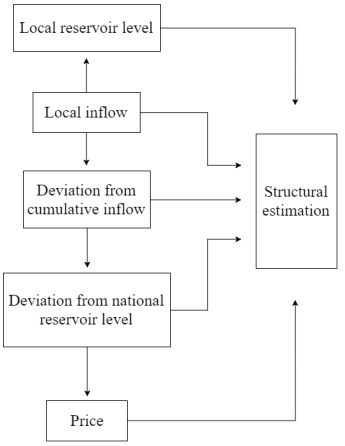

Figure 4.1: Model overview

	Energy Coefficient $[KWh/m^3]$	Production Capacity $[MW]$	Reservoir Capacity $[Mm^3]$	Average Inflow $[m^3/s]$
A	0.94268	10	35	3
B	1.2385	128	180	15

Table 4.1: Power plant coefficients for power plant X and Y

4.1 Inflow, I_t

The time series for inflow was given with an hourly resolution and later aggregated to a weekly resolution. The data has a high grade of uncertainty as it contains negative values for inflow, still after aggregating to a weekly resolution. To eliminate the negative values, these are set to zero in the time series. This leads to a higher average than was originally observed. By lowering the seasonal component we can counteract the risen mean. The time series for inflow is from 1.1.1993 to 31.12.2014 and is shown in figure 4.2.

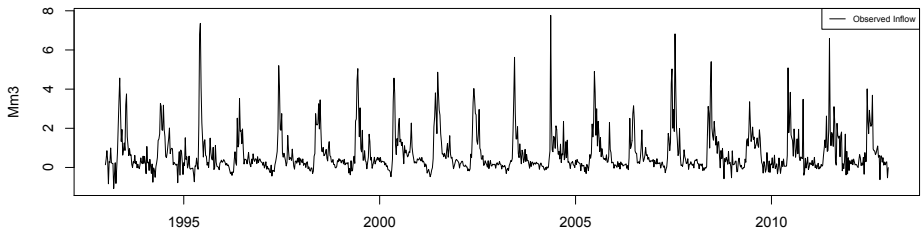

Figure 4.2: Observed weekly inflow for power plant A

4.1.1 Seasonality

The figure shows clear tendencies for seasonal variations as well as high peaks during spring floods. When the temperature is below 0°C, the precipitation is in the form of snow, which does not move through downhill terrain like water, resulting in local snow packs. The snow that is located at higher altitude will affect the reservoirs inflows when they melt. The melting of snow will often result in a spring flood, which explains the high peaks. This often happens between easter and summer, but is not deterministic in time, it is dependent on temperature.

To include and model the anticipated seasonal fluctuations in the observed inflow data, a seasonal function has been applied. This function is a second order Fourier transform, that has been fitted to best capture the seasonality. The data has a weekly resolution, hence a frequency of 52. The coefficients are estimated using the nonlinear least squares method in R and can be found in the appendix. Equation 4.1 shows the function for the seasonal component.

$$f^I(t) = A_1^I \cos(\frac{2\pi}{52}t + \phi_1^I) + A_2^2 \cos(\frac{4\pi}{52}t + \phi_2^I) + D^1 \tag{4.1}$$

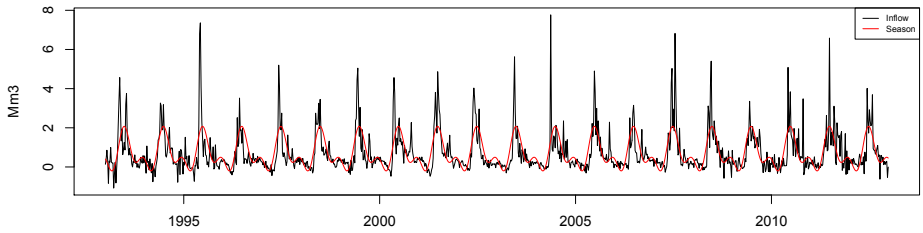

Figure 4.3: Observed weekly inflow and seasonal component for power plant A

As shown in figure 4.3, the seasonal function seems to capture some of the inflow. The peaks are not captured by the model and are underestimated or smoothed. This can potentially be a result of the unpredictable timing of the spring flood.

4.1.2 Stationarity and Independence

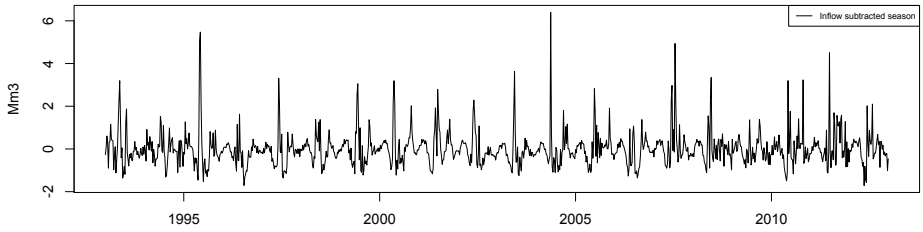

Figure 4.4: Deseasonalized weekly inflow for power plant A

Figure 5.3 shows the remains of the inflow after the seasonal component is subtracted. As shown by the figure, remains of the original seasonal component is still visible. This is because the seasonal component suggested in the previous section did not capture the extreme peaks of the inflow.

The model requires the Markov property, where the next term is only dependent on the previous. An autoregressive process with one lag, an AR(1)-process, satisfies this property. These processes require stationary time series.

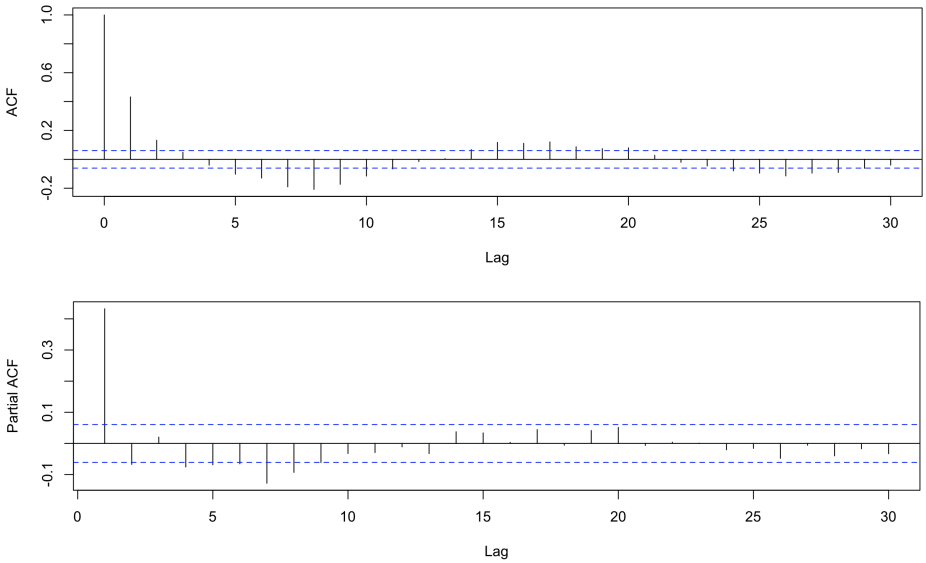

Figure 4.5: ACF and PACF for deseasonalized inflow for power plant A, 95% siginificance level

Test	Value	Critical Value	Assumption satisfied
ADF-test	-18.355	-1.95	+

Table 4.2: Test values for stationarity for Inflow

The data series was subjected to an Augmented Dickey-Fuller test. Table 4.2, shows stationarity for a significance level of 95%. The stationarity requirement is satisfied and details of this test is further elaborated in the appendix.

Order of autoregressive terms

Figure 4.5 shows the auto-correlation function (ACF) and the partial auto-correlation function (PACF) for the deseasonalized inflow. The autocorrelation function shows correlation between the time series and lags of itself. The partial autocorrelation function shows partial correlation coefficients between the series and lags of itself.
Robert Nau (2017), explains partial autocorrelation as:

"... the amount of correlation between a variable and a lag of itself that is not explained by correlations at all lower-order-lags."

The PACF-plot in figure 4.10 shows a significant spike only at lag 1, meaning that all the higher-order autocorrelations are effectively explained by the lag-1 autocorrelation, indicating that an AR(1)-model should be used.

4.1.3 Autoregressive process, X_t^I

The seasonal component is subtracted from the observed inflow and the rest is submitted to an autoregressive process with one lag.

$$X_{t+1}^I = \gamma X_t^I + \epsilon_{t+1}^I \tag{4.2}$$

γ is the AR-coefficient, and ϵ_t^I is the random error term. The random error has a mean of 0 and a standard deviation of σ^I. The parameters are estimated by using the Arima-function in R. The negative observations from Inflow is set to zero as there can be no negative observations in the model. As a consequence, the mean will rise slightly. To counteract the risen mean, we have lowered the mean for the seasonal function slightly.

4.1.4 The model

$$I_{t+1} = f^I(t+1) + X_{t+1}^I \tag{4.3}$$

By adding the autoregressive process and the seasonal component, a simple model of the inflow is obtained. Equation 4.3 shows the model for inflow. Figure 4.6 shows the seasonal component together with a simulation of the AR(1)-process in red and the observed inflow in black. The model is unable to reach the isolated tops due to very large peaks in inflow and the inconsistency in the timing of these. By increasing σ^I, the variance for ϵ^I, we would be able to reach the peaks, but this would negatively affect the overall fit of the model. The descriptive statistics in A.2 shows the mean, standard deviation, minimum value, maximum value and the median for both the observed inflow time series and the simulated time series.

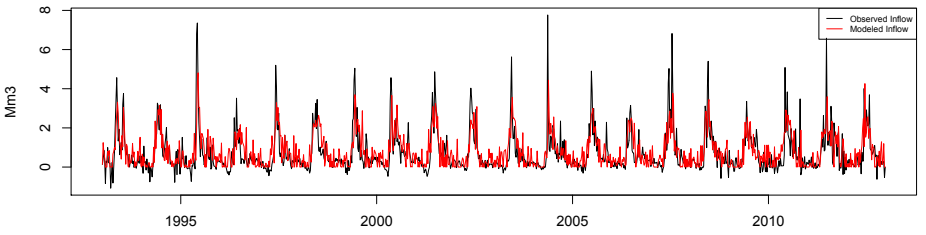

Figure 4.6: Observed and modeled weekly inflow for power plant A

4.1.5 State variable

Using the alternative form of (4.3):

$$I_t = f^I(t) + X_t^I \tag{4.4}$$

Rearranging and expressing by X_t^t:

$$X_t^I = I_t - f^I(t) \tag{4.5}$$

Inserting (4.5) into (4.1), the state variable transition expressed only in terms of itself, is obtained:

$$I_{t+1} = \gamma I_t + \epsilon_{t+1}^I + f^I(t+1) - \gamma f^I(t) \tag{4.6}$$

4.2 Deviation from cumulative inflow, C_t

There is a high correlation between the Nord Pool spot price, the local inflow and the overall reservoir level in Norway. Kolsrud and Prokosch (2010) present a model which includes this relationship. In their model, they use the deviation from cumulative inflow. This tells the reservoir manager if he has received more or less inflow the last f weeks than what is to be expected of this period.

Boger and Vestbøstad (2016) defined the cumulative inflow as:

$$i_t := \sum_{j=1}^{\infty} \rho^{j-1} I_{t-j} \tag{4.7}$$

In equation 4.7, i_t is the cumulative local inflow. $I_t - j$ is the inflow at time $t - j$ and ρ is a weighting factor deciding how much influence the cumulative inflow of last week should affect this weeks level. Equation 4.7 can be written as:

$$i_t = I_{t-1} + \rho i_{t-1} \tag{4.8}$$

i_t should optimally reflect a period that yields a high correlation with the national reservoir deviation, denoted R. The length of this period is adjusted by f, which we estimated to be 20.

$$i_t^* = \sum_{k=1}^{f} I_{t-k} \tag{4.9}$$

ρ in equation 4.8 is estimated by using equation 4.9. f is set previously and for each f, one ρ is calculated based on correlation between the data sets given by equation 4.8 and equation 4.9. These data sets can be compared to the deviation from the aggregate national reservoir level by making different data sets for f. RStudio was used to estimate f equal to 20 and ρ to 0.936. i and i^* with $f = 20$ is plotted in figure 4.7.

$\bar{i}_t$ is the historical cumulative average inflow for each week. The deviations from historical cumulative average is showed in equation 4.10.

$$C_{t+1} := \frac{i_{t+1} - \bar{i}_{t+1}}{\bar{i}_{t+1}} \tag{4.10}$$

Using equation 4.10, inserting equation 4.8 and use $i_t = C_t * \bar{i}_t - \bar{i}_t$, one obtains:

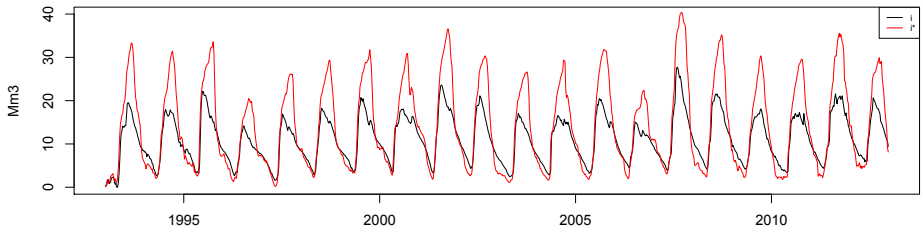

Figure 4.7: Weekly cumulative inflow with f=20

$$C_{t+1} = \frac{I_t + \rho C_t \bar{i}_t + \rho \bar{i}_t}{\bar{i}_{t+1}} - 1 \tag{4.11}$$

Equation 4.11 gives the state variable transition for C_t, expressed only in terms of itself.

Figure 4.8 shows the simulated and observed deviation from cumulative inflow.

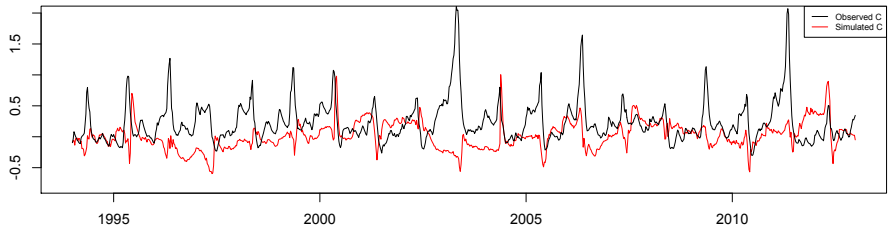

Figure 4.8: Simulated and observed deviation from cumulative weekly inflow for power plant A, C

4.3 Deviation from aggregate Reservoir level, R_t

The reservoir level used to calculate R_t is Norway's national reservoir level, which is an aggregation for all reservoir levels in Norway. The data series is provided by The Norwegian Water Resource and Energy Directorate and given in a weekly resolution. In section 2.1.3 the power production in Nord Pool is shown. Norwegian hydropower contributes to 32.16% of total electricity distributed through Nord Pool. Deviation from the historical national aggregated reservoir level will, therefore, have a correlation with the Elspot price, as the aggregated reservoir level is a large driver for the supply side of the market. In the

previous section, a relationship between the deviation from average reservoir level and the deviation from cumulative inflow was presented. Equation 4.12 shows R_t as the deviation from the average reservoir level.

$$R_t = \frac{r_t - \bar{r}_t}{\bar{r}_t} \tag{4.12}$$

Figure 4.9: Deviation from aggregate reservoir level, R_t

The aggregate national reservoir level is r_t, while $\bar{r}_t$ is the average national reservoir level for that particular t. An important note is that r_t is the aggregated reservoir level for all of Norway combined and not just price zone 5. Figure 4.9 shows deviation from aggregated average national reservoir level, R_t

4.3.1 Stationarity and independence

As with inflow, a requirement for using an AR(1)-process is that the data series is stationary. The deviation from the aggregated reservoir level is subjected to an ADF test for stationarity. The value is below the critical value for a significance level of 95%, indicating stationarity. This seems reasonable as there is no observable trend in the data series shown in figure 4.9.

Test	Value	Critical Value	Assumption satisfied
ADF-test	-5.1118	-1.95	+

Table 4.3: Test values for stationarity for deviation from aggregated reservoir level

Order of AR terms

Figure 4.10 shows the autocorrelation function and the partial autocorrelation function for the deviation from aggregate reservoir level. The autocorrelation function shows significant correlations for a large number of lags. To test the amount of lags necessary in the autoregressive process, the time series is subjected to a PACF-test. When testing the partial correlation, high correlations with the first two lags are found. This indicates that an

AR(2) model would be better, but it would violate the Markov Property. In order to model the deviation from aggregated reservoir level, an AR(1)-process is used.

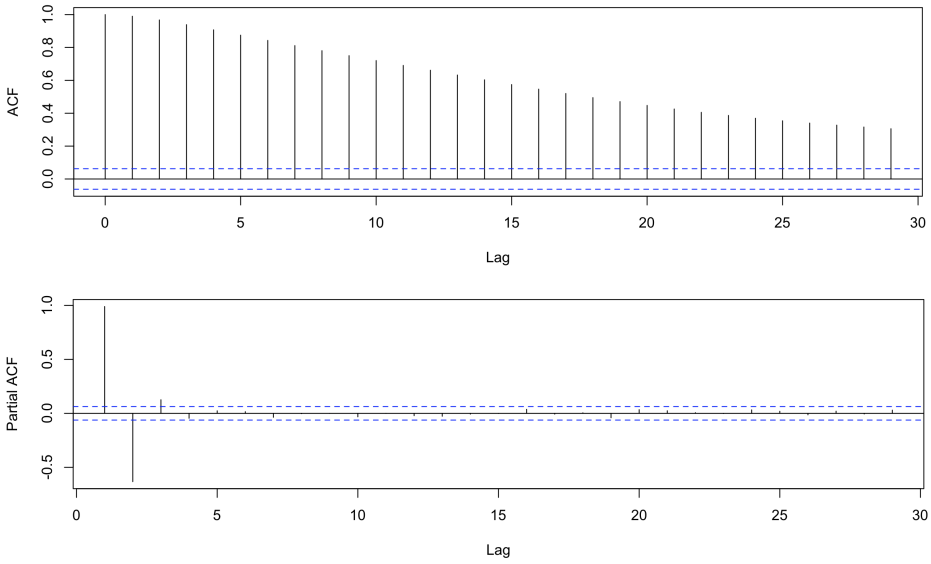

Figure 4.10: ACF and PACF for deviation from aggregated reservoir level

4.3.2 Autoregressive process with an exogenous component

Figure 4.11 shows both R_t and C_t. The correlation between the data series is 0.602. The high correlation enables the incorporation of C_t in the modelling of R_t.

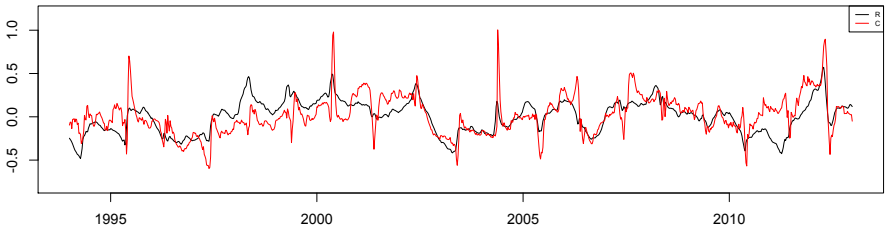

Figure 4.11: R and C with a correlation of 0.602.

Using an autoregressive process, while adding an exogenous term for the cumulative inflow, yields a process with an autoregressive component for itself and the exogenous term C_t, equation 4.13. An autoregressive process with a one-lagged exogenous component, an

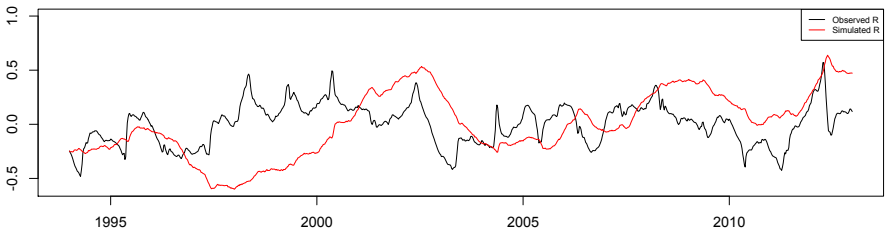

Figure 4.12: Simulated and observed deviation from aggregated reservoir level.

ARX(1) model, depends on today's deviation from the aggregate reservoir level, R_t and today's deviation from the cumulative inflow, C_t. The C_t works as an additional predictor. β_1 is the autoregressive coefficient and β_2 is the coefficient for the predictor C_t. The error, ϵ^R_{t+1}, is a normal random variable with mean 0 and standard deviation σ^R, estimated to be 0.005.

$$R_{t+1} = \beta_1 R_t + \beta_2 C_t + \varepsilon^R_{t+1} \tag{4.13}$$

Using the Arima-function in R, the estimates for β_1 and β_2 were found to be 0.99 and 0.04, respectively. This is shown in figure 4.12. The simulation of the ARX(1)-process seems to capture some of the dynamics. Descriptive statistics for the simulation and the observed R_t can be found in the appendix.

4.4 Price, P_t

The electricity price that directly influences the hydropower plant's revenues is the local area price. As mentioned, hydropower plant A is located in price zone NO5 and plant B is located in zone NO4. Therefore, these respective prices are used in the structural estimation model.

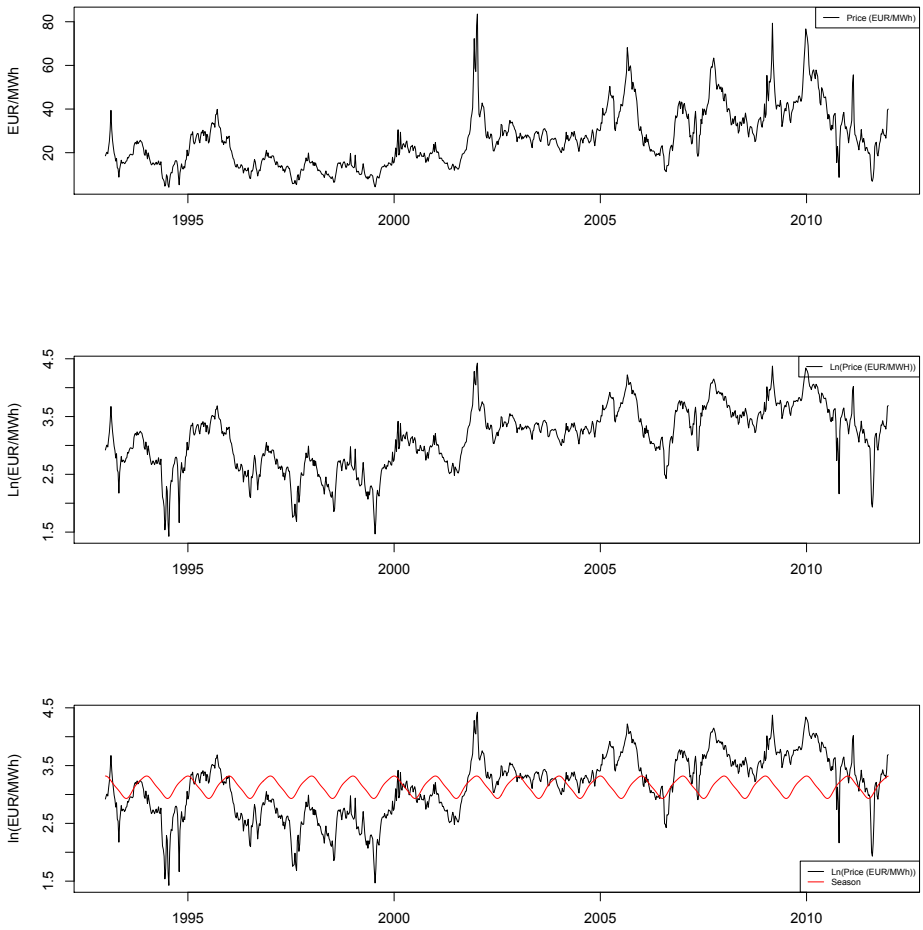

Figure 4.13: Elspot price, Ln(Elspot price) and Ln(Elspot price) with a seasonal component

4.4.1 Seasonality

The logarithm of the Elspot price is subjected to a seasonal component. The seasonal component for the price is captured by a third order Fourier series, presented in equation 4.14.

$$f^P(t) = A_1^P \cos(\frac{2\pi}{52}t + \phi_1^P) + A_2^P \cos(\frac{4\pi}{52}t + \phi_2^P) + A_3^P \cos(\frac{6\pi}{52}t + \phi_3^P) + D^P \quad (4.14)$$

The parameters are estimated using the non linear least squares method in R. The resulting

coefficients are listed in A.2. The results shown in the bottom of figure 4.13 show the seasonal component obtained, which is clearly not as prominent as with inflow.

4.4.2 Stationarity and independence

The logarithm of the price was first subtracted the seasonal component, $f^P(t)$, before tested for stationarity with an Augmented Dickey-Fuller test. Based on the results from the test shown in table 4.4, the requirement of stationarity is satisfied.

Order of AR terms

The logarithm of the price was subjected to an ACF and a PACF-test with a significance level of 95%, shown in figure 4.14. The autocorrelation test shows significant correlations with the first 30 lags. The results from the PACF-test indicates that an AR(1) should be used.

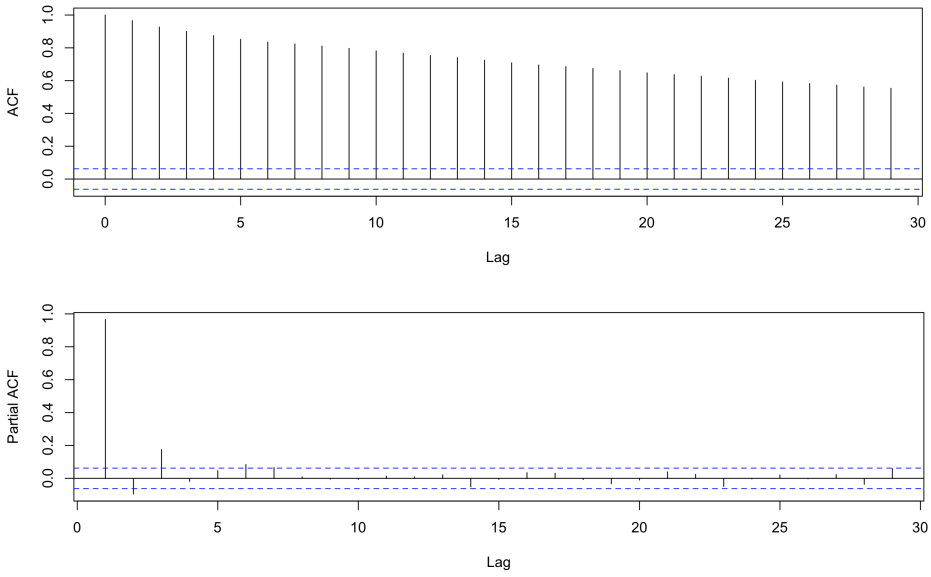

Figure 4.14: ACF and PACF for the price

Test	Value	Critical Value	Assumption satisfied
ADF-test	-4.4765	-1.95	+

Table 4.4: Test values for stationarity for P_t

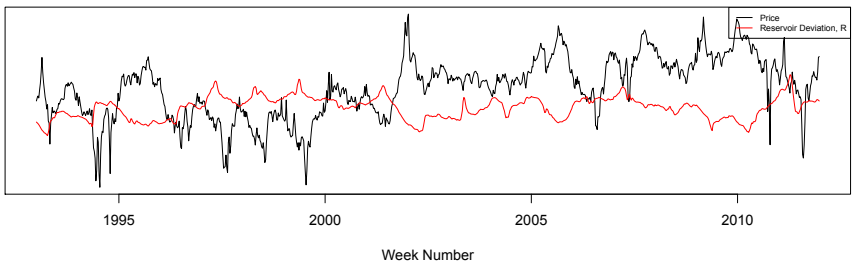

Figure 4.15: Log Elspot price and R

4.4.3 Autoregressive process

Aggregated reservoir level is a fundamental driver for power prices in Norway, as stated by Kaaresen and Husby (Kaaresen and Husby, 2000), Haldrup and Nielsen (Haldrup and Nielsen, 2006) and Povhet al. (Povh et al., 2010). The logarithm of the price and the deviation from aggregate reservoir level have a correlation of -0.4217 and are plotted in figure 4.15.

The correlation between the price and the deviation from aggregate reservoir level is incorporated in the model by using an ARX(1)-process, like the one used for R_t. The exogenous component is R_t and the process is autoregressive with one lag. The process is shown in equation 4.15:

$$X^P_{t+1} = \alpha_1 X^P_t + \alpha_2 R_t + \epsilon^P_{t+1} \tag{4.15}$$

α_1 and α_2 are estimated to be 0.95025 and -0.09516, respectively. The negative parameter α_2 makes sense due to the negative correlation between the price and the deviation from aggregate reservoir level. The error, ϵ^P, is a normal random error with mean 0 and standard deviation σ^P, estimated to be 0.05.

4.4.4 The model

By adding the seasonal component and the autoregressive process, the model for the price, shown in equation 4.16 is achieved. The model is shown in figure 4.16 together with the observed price.

$$P_{t+1} = X^P_{t+1} + f^P(t+1) \tag{4.16}$$

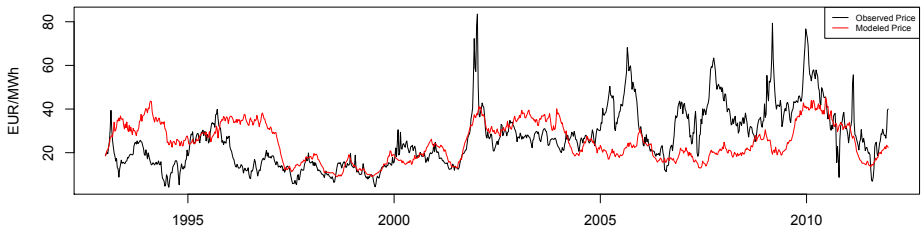

Figure 4.16: Observed and simulated Elspot price

4.4.5 State variable transition

Starting with the equation for the model:

$$P_{t+1} = X^P_{t+1} + f^P(t+1)$$

Using:

$$X^P_t = P_t - f^P(t)$$

The latter can be inserted into equation 4.15, which yields:

$$X^P_{t+1} = \alpha_1(P_t - f^P(t)) + \alpha_2 R_t + \epsilon^P_{t+1}$$

Inserting this into the first equation results in the final model:

$$P_{t+1} = \alpha_1(P_t - f^P(t)) + \alpha_2 R_t + f^P(t+1) + \epsilon^P_{t+1} \tag{4.17}$$

Equation 4.17 shows the price, P_t, as a Markovian process which only depends on the previous state of itself and of the exogenous component.

4.5 Local Reservoir level, S_t

The local reservoir level is calculated endogenously and is dependent on the production and inflow. When the plant manager chooses to produce, water is released from the local reservoir and goes through a power producing turbine. Water is added to the reservoir by the local inflow. Equation 4.18 explains this relationship where S_t has a unit of Mm^3. Since the reservoir level can not exceed the maximum level, S^{max}, S_t will be the smallest value of the level after inflow and production, $S_t + I_t - u(d_t)$, and the maximum reservoir level, S^{max}.

$$S_{t+1} = min\{S_t + I_t - u(X_t, d_t), S^{max}\} \tag{4.18}$$

$$S^{min} \leq S_t \leq S^{max} \tag{4.19}$$

Equation 4.19 describes the relationship between the minimum and maximum level of the reservoir.

4.6 Production, d_t

The production data was given in MWh with an hourly resolution, and is shown in figure 4.17. As shown by the figure, production levels seem to concentrate around 8 MWh or no production at all. There seems to be a tendency of increased distribution at around 8 MWh over time. That is, the production seems more binary at the beginning than at the end of the period.

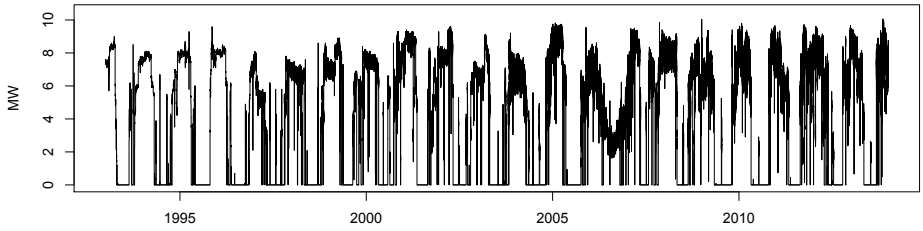

Figure 4.17: Production data for power plant A

4.6.1 Aggregating and discretizing the production levels

The production level is considered the decision variable in our MDP. The model is, as mentioned, a DCDP, and the production data must, therefore, be discretized. When trying to discover the structural parameters linked to the loss of efficiency when deviating from the efficient production level, it is important to consider how the aggregation and discretization of the production levels will affect the results. When aggregating from an hourly to a weekly resolution, a lot of information may be lost. If the power plant is starting up and shutting down many times during one week, it may appear as if it is producing at a lower level than the efficient one.

For example, if the efficient production level is at 8 MW and the decision-maker is switching every hour between releasing water at the efficient level and releasing no water, it will seem like the power plant was producing at 4 MW the entire week after aggregation. In our case, the discretized, aggregated levels should have similar occurrences to the corresponding production levels from before aggregation. We discretize to 6 levels to be able to assign one of these levels as the efficient one. Most likely, the efficient level will be around 8 MW, which will approximately correspond to the discretized production level 5. After aggregating to a weekly resolution and discretizing the production into 6 levels, we obtain the following histogram of the occurrences:

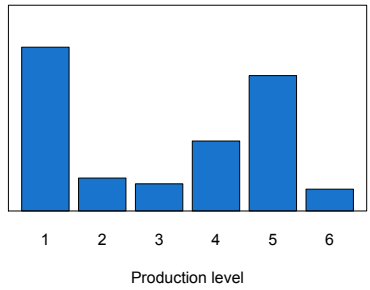

Figure 4.18: Discretized production levels used for power plant A

We want to investigate if this is really reflecting the levels at which the power plant operates on an hourly basis. Using bin sizes of 0.5 MWh, a histogram of the actual production with an hourly resolution is given:

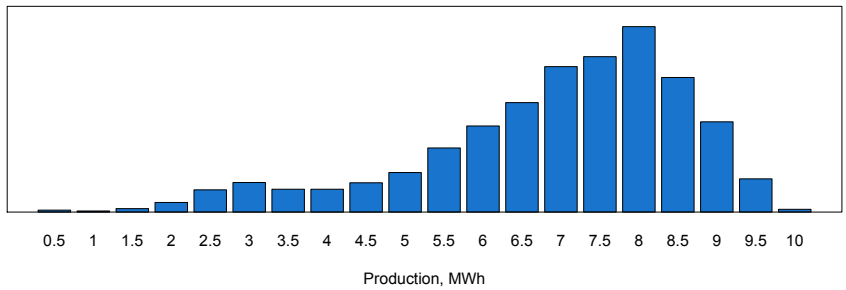

Figure 4.19: Histogram of production levels with hourly resolution, power plant A

The occurrences when the production is zero have been excluded since this number is very large and would make the other levels difficult to observe. We note that the preferred level of production is around 8 MWh, as expected. To see if we capture the actual production levels after aggregation, we also discretize the production on an hourly level and compare the two histograms:

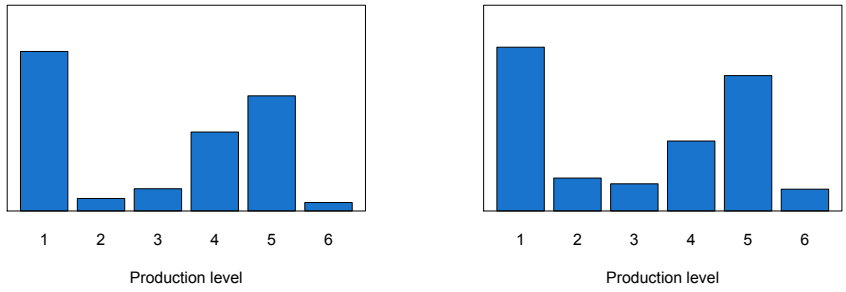

Figure 4.20: Comparison between hourly (left) and weekly (right) discretized production levels, power plant A

We observe that the production levels are quite similar. This is because the production level is generally quite stable, and does not have fast-moving fluctuations. Had this been the case, it would have been difficult to uncover the decision maker's preferences, since our data set would not be representative of the underlying, hourly production. In order to emphasize the importance of having a simple hydropower plant, we perform the same aggregation and discretization for power plant B, displayed in figure 4.21

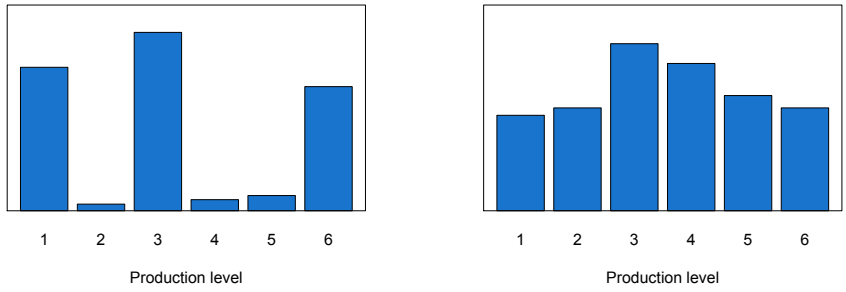

Figure 4.21: Comparison between hourly (left) and weekly (right) discretized production levels, power plant B

This demonstrates one of the main advantages in using a smaller hydropower plant with only one turbine. Production level data from larger hydropower plants with several turbines and generators will get distorted when aggregating to a higher level. Power plant B has two turbines. Therefore, after discretizing to six levels, both level 3 and 6 can be considered best efficiency points. This is because the power plant can produce at the efficient level in one turbine while the other is not producing. The other efficient point is when both turbines are producing at their efficient levels. The efficient points, however,

are only visible with an hourly discretization. After discretizing to a weekly resolution, the efficient production levels completely disappear, as can be seen in figure 4.21.

Chapter 5

Results and analysis

The model was implemented in AMPL (A Mathematical Programming Language), (Fourer et al., 1985). The environment that was used was the AMPL IDE, and the solver used was the Artelys Knitro solver. The Knitro solver is used in large-scale nonlinear optimization, and it was not limited in memory usage. The computers had 32GB of memory installed, which allowed us to discretize the production sufficiently so that we could capture the differences between efficient and non-efficient production levels. The model had a total of 25442 variables and 25440 constraints. The two extra variables are θ_1 and θ_2. The CPU was an Intel Core i7-6700 with a maximum speed of 3.40GHz. Solving the model for a manually set value of θ_1 and θ_2 takes approximately 40 seconds. Alternatively, solving the model with variable values for θ_1 and θ_2 will take between 5 and 12 hours.

We present the three main results obtained from the model. These are related to the implied efficiency loss, the changes in θ-values over time, and the power plant's implied water values. We present a discussion for each of the results, what they imply and whether or not they make intuitive sense.

5.1 Efficient production level - ξ

When discretizing the production levels for power plant A in section 4.6.1, the level with the most occurrences was production level 5, corresponding to 83 percent of maximum production. It is natural to assume that the production level with the most occurrences is the most efficient one. In order to test if the model also finds production level 5 as the most likely BEP, it was tested for the highest obtainable likelihoods for different BEPs. By setting the BEP to a certain level, iterating over different θ-values and then selecting the highest log-likelihood value for each, we obtain the results summarized in table 5.1. We observe that production level 5 has the highest likelihood value for power plant A.

The same process was applied to the second power plant, B. As explained in section 4.6.1, the efficient production levels completely disappear as a result of discretizing to weekly

resolution. We therefore provide no further analysis of power plant B.

	Log-likelihood value	
ξ	Power plant A	Power plant B
2	-1774.73	-1675.18
3	-1755.23	-1670.92
4	-1766.13	-1868.87
5	-1625.95	-1788.13
6	-1784.50	-1692.87

Table 5.1: Efficient production levels and corresponding likelihoods

5.2 Choice of efficiency function

In the previous section the most preferred production level was established as production level 5 for power plant A. Before we analyze the θs that return the highest likelihoods, we have to identify the best fitting efficiency function for our data set. We apply a linear, a square root, and a squared efficiency function. These are elaborated in section 4.8, and yield the following log-likelihood values:

	ξ	Square root	Linear	Squared
Plant A	5	-1633.25	-1625.95	-1641.37

Table 5.2: Log-likelihood values for the three different efficiency functions

The results tell us how the producer values the different production levels relative to the distance from the best efficiency point. That is, a squared efficiency function would punish small deviations from the BEP more than a square root function, with a linear function being in between. We observe that a linear relationship yields the highest likelihood and base the remaining analysis on the assumption that this relationship is correct. The linear efficiency function is therefore used when estimating θ_1 and $\theta 2$.

5.3 Implied efficiency loss - θ_1 and θ_2

The main goal of our research is to investigate the loss of efficiency that is implied from the data. The results let us analyze how the producer values the different production levels, or conversely, to what degree the producer is penalized for deviating from the BEP. The reader should keep in mind that θ_1 and θ_2 are coefficients of efficiency loss for production levels below and above the BEP, respectively.

By comparing the implied efficiency levels to the anticipated mechanical efficiency of the plant, we are able to discuss and discover hidden economic parameters related to the

producer-specific preferences. First, we show how the values were obtained. We then present a discussion of possible explanations for the estimated variables and how they deviate from a purely mechanical efficiency loss.

5.3.1 Manual case

In order to avoid being stuck in a local optimum, we want to examine the likelihood values for many different combinations of values for θ_1 and θ_2. We manually set these values in the range [0, 0.85] and with a step size of 0.05. Setting the θ-values manually implies solving the model 324 times and results in the likelihood values given by the plots in figure 5.1.

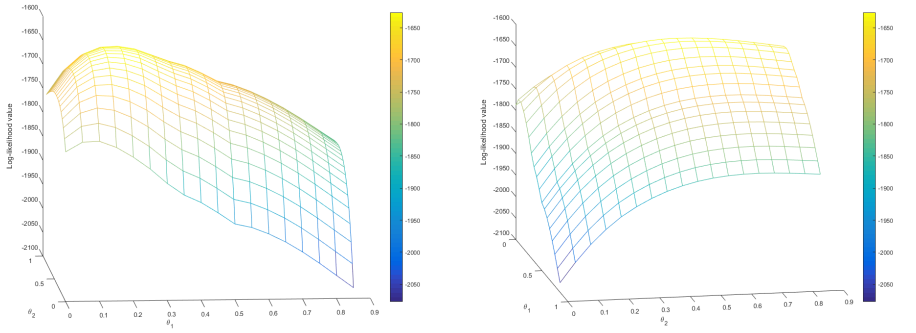

Figure 5.1: Log-likelihood values for θ_1 and θ_2

As indicated by the figure, the likelihood has a maximum around $\theta_1 = 0.15$ and $\theta_2 = 0.5$. The log-likelihood value at this point is -1625.95. As the model has to be solved 324 times, each taking 40 seconds, the total solve time was approximately 3.5 hours.

5.3.2 Variable case

The model should, hopefully, return similar values for θ_1 and θ_2 when setting these as variables in order to discover more precise values. Therefore, we solve the model with θ_1 and θ_2 as variables, the efficient production level, $\xi = 5$, and we use the linear efficiency function. The variable solution yields:

$$\theta_1 = 0.1702$$

$$\theta_2 = 0.505242$$

The results seem to confirm that we are not getting stuck in a local optimum and that these are, in fact, the most likely values for θ_1 and θ_2 when in their logical range. θs outside of this range would not make intuitive sense. For example, a negative θ value would imply an extra reward for deviating from the most efficient production. $\theta > 1$

would imply that when deviating from the point of efficient production, we are producing negative electricity, which also would not make sense.

The likelihood value of -1625,100426 seems to indicate that we have found values for the $\theta's$ which are somewhat improved from the manual case, as expected. The estimation requires 1287 iterations and has a run-time of 6.5 hours, which is almost twice as long as solving the model manually 324 times. In return, it yields more accurate values for θ_1 and θ_2.

5.3.3 Discussion

The θ-values that were found for power plant A indicate that the reservoir managers require a 51% higher reward for deviating one production level above the BEP, i.e. producing at 100% instead of 83% of maximum production. They require a 17% higher reward for deviating one production level below the BEP, i.e. producing at 67% of maximum production. Further, since the relationship is assumed to be linear, they require a $2 \cdot 17\% = 34\%$ higher reward for deviating two production levels below the BEP, i.e. producing at 50% of maximum production, and so on.

The implied efficiency loss of producing above the BEP is clearly larger than the loss when producing below the BEP. Since the efficiency of the turbine-generator scheme probably does not lose as much as 51% when producing at maximum, there has to be some other economic parameter behind the producer's choice. A plot of the implied penalty is shown in figure 5.2, together with the efficiency curve of a Francis turbine. The efficiency curve of a Francis turbine is included in the plot, as this is the most used turbine for hydropower plants that share the same technical specifications as power plant A.

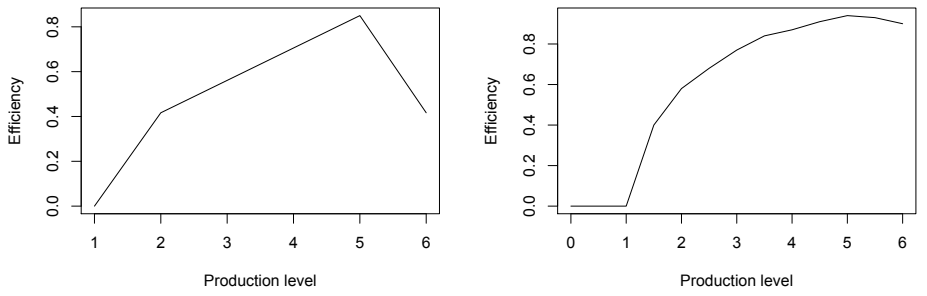

Figure 5.2: Comparison between the implied efficiency and the efficency of a Francis turbine

It is important to keep in mind that the θs discovered represent the behavior of the reservoir managers, given that the model is correct. The implied efficiency does not have to match the mechanical efficiency of the turbine-generator scheme but can include other hidden economic incentives that explain the dissimilarities between the two curves. There

can be several reasons why a power producer does not want to produce at other production levels. Examples of these are increased maintenance cost, lowered durability, and cavitation problems.

The results of the implied efficiency seen in figure 5.2 can be the basis for internal discussions at the hydropower plant as to whether or not the extra required reward reflects their operating policy and if their current policy is reasonable. The implied efficiency should be the result of potential mechanical failure and related costs, the production efficiency and the mechanical fatigue this production implies. Ultimately, the power producer wants to produce power at levels and at times which generate the highest total profits.

5.3.4 Explaining the implied efficiency

Establishing the turbine used by the plant
We earlier assumed that the turbine used in the power plant is a Francis turbine, but will in this section elaborate on this assumption.

By using the general formula for a hydro system's power output from section 2.2.1, we are able to find the effective head of water across the turbine:

$$\begin{aligned} P &= \eta \rho g Q H \\ 8.8MWh &= 1000 \cdot 0.8 \cdot 9.81 \cdot 3 \cdot H \\ H &= 373.77m \end{aligned}$$

A Francis turbine is typically used for hydropower plants with heads of water between 30 m and 600 m. The rough calculations give an effective head of 374 m, which matches these specifications. Further, the Francis turbine is the most commonly used turbine in Norwegian hydropower plants. A third argument for why we believe our plant is using this turbine can be drawn from their electricity capacity and their related prevalence in Norwegian hydropower plants. Keeping in mind that our hydropower plant has a 10 MW capacity, table 5.3 shows that Francis turbines are used by 73% of small (1-10 MW) hydropower producers in Norway.

Turbine type	Total number	Quantity above 100 MW	Quantity between 1 - 10 MW
Francis turbine	685	45	345
Pelton turbine	190	20	70
Kaplan turbine	95	2	45
Bulb turbine	20	-	8

Table 5.3: Turbine types and their quantities in Norwegian hydropower plants

Cavitation as a potential explanation

Cavitation was introduced earlier in section 2.3. If we assume that the turbine used by the hydropower plant is, in fact, a Francis turbine, cavitation could explain the implied efficiency related to θ_2, considering the large deviation from the producer's implied loss of efficiency and the actual mechanical loss of efficiency in such a turbine.

According to Kumar and Saini (2009), the point of critical cavitation is above the BEP. The implied efficiency suggests a BEP of production at 83%. As seen in figure 5.3, the critical cavitation point is denoted σ_c, efficiency is along the y-axis, and Thoma's cavitation factor σ is along the x-axis. The important implication of this is that producing at a higher level than the BEP leads to a much higher cavitation than producing at or below the BEP. As such, it is obviously not just the pure efficiency loss that is being taken into account by the decision-makers at the plant. Damage to the turbine becomes an important factor when deciding whether or not to produce above the BEP. This may explain why the reservoir managers require such a high reward for doing so. Damaging the turbine can be costly, and as a result, the compensation must be high.

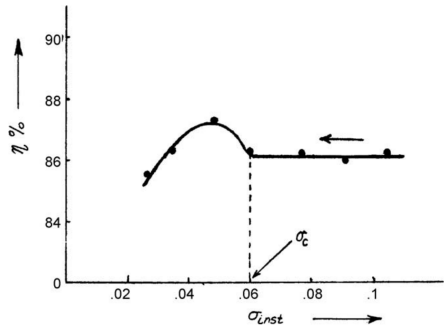

Figure 5.3: Variation of efficiency with respect to cavitation factor, σ

5.4 Examining changes in θ-values over time

The data series used to estimate the θs for the power producer contains over 20 years of production data. To examine if the plant's preferences have changed over time, the data series was split in half and the model was solved for the first and second halves with the results displayed in table 5.4. The goal was to identify if the willingness to deviate from the efficient production level has changed during this time.

	First half	Second half
θ_1:	0.5456	0.0794
θ_2:	1	0.398
Log-likelihood:	-579.69	-676.54

Table 5.4: Changing values of θ_1 and θ_2 for power plant A over time

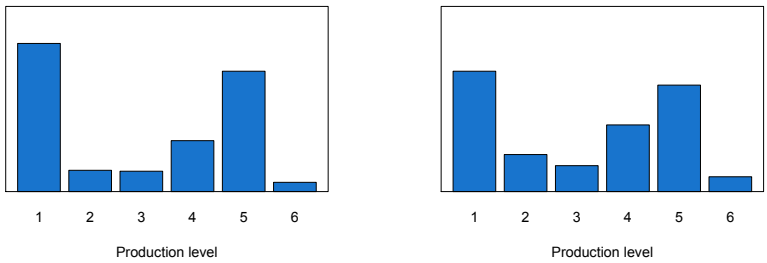

Figure 5.4: Discretized production levels for the first and second half of observation, power plant A.

The results displayed in Table 5.4 show that the willingness to deviate from the BEP has increased over the years, as both θs are lower for the second half than the first half. A potential explanation is the change in monitoring technology. It is possible that the power plant had a much stricter policy for production at levels above the BEP due to uncertainty regarding the potential economic downside. A θ_2 value of 1 signifies that the producer should not want to produce above the BEP, as this would mean that the producer receives no power in return for the released water. A potential scenario is that the precision of monitoring technology related to turbine maintenance and wear was increased. The results of such monitoring may have changed their willingness from being completely unwilling to produce at levels above the BEP in fear of, among other things, cavitation, to producing at levels above the BEP if the economic relative gain was above 39,8%.

The histograms displayed in Figure 5.4 show the discretized production levels for the first and the second halves of our time series. They confirm that the amount of production regarding production level 6 has doubled and that production on level 4 also has increased. Production on the best efficient point has as a consequence decreased.

5.5 The marginal value of water

The marginal value of water for the producer tells us the price that the reservoir manager requires for one unit of water. If he values his water lower than the current price of power, he will release and gain an immediate profit. Otherwise, he will continue to store water and release it at a more profitable time. To see how the reservoir level and seasonality influences the marginal water values, the 3D plot displayed in Figure 5.5 and Figure 5.6 was made. The water value is the expected marginal value of the energy stored in the reservoir. This value is found by dividing the change in expected value by the change in reservoir level.

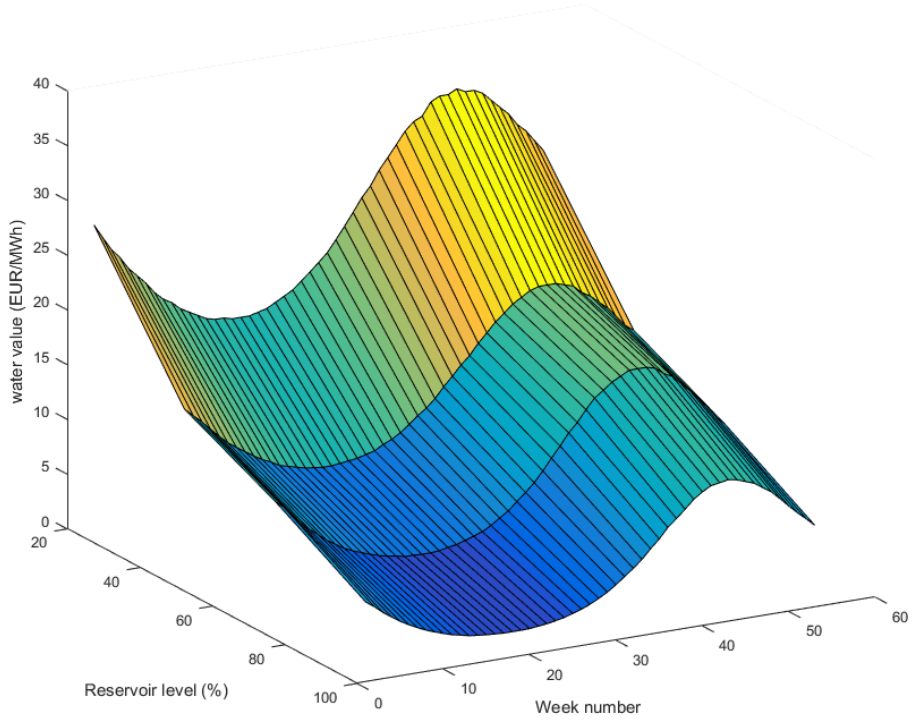

Figure 5.5: Marginal water values

The figure shows the reservoir level on one axis. This is an important validation of the model because the producer should value the water relative to the amount of available water in the reservoir. The reservoir level dynamics are best seen in figure 5.6. The relationship between the reservoir level and the water value is not linear.

The weeks are displayed on the other axis, in order to show the marginal water value in comparison with seasonal changes. The season influences the expectation for both power prices and inflow, which directly affects the marginal value of water. The figure is in accordance with knowledge of the Norwegian hydropower market, as the time around week 18 is when we would usually expect the spring flood. This is the time when the snow packs in the mountains at higher altitudes melt and fill up the reservoirs for hydropower producers. One unit of water at this time has very little value, since the expectation of large amounts of inflow gives little incentives for having water stored in the reservoirs. According to Olsson et al. (2015), long-term hydrological forecasts are used by the hydropower industry to make sure there is sufficient remaining capacity to handle sudden inflows, where the spring-flood forecast is considered the most important one.

In the autumn temperatures drop and the precipitation is in the form of snow instead of

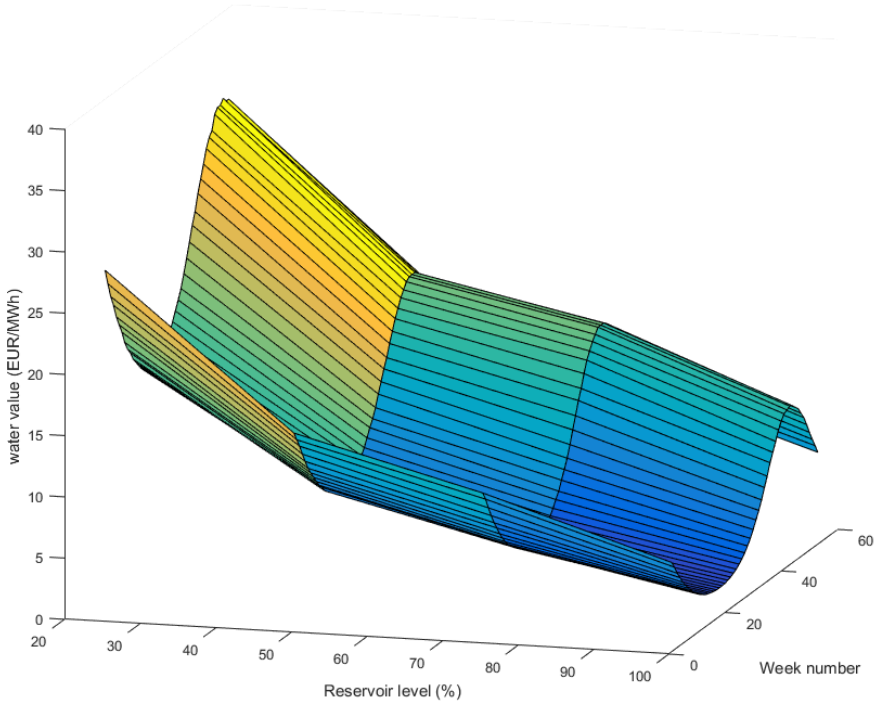

Figure 5.6: Marginal water values

rain. Since snow does not move through terrain in the same way as water, the reservoir receives less water. This makes the marginal value of water for the producer increase since the expectations for future inflow in this period is low. As a result, the value of the remaining water increases drastically as seen in the figure. These seasonal changes vary each year, but will often begin around weeks 40-50. This is in accordance with the results displayed in figure 5.5, which shows that the highest water values are found in week 45.

The water value decreases with higher reservoir levels, which is in accordance with intuition. Further, the curve of the water values along the reservoir level is dependent on the season. The water value has a steeper slope closer to both edges of the reservoir level around week 45. Around week 18, however, the slope of the water value gets gradually less steep as the reservoir level increases. The decrease in the water value when increasing an already high reservoir level is higher for weeks when the inflow is low than for weeks when the inflow is high.

.

5.6 Further development

Model validation When the model estimates θ-values it uses 989 observations for each of the state variables. To see how large the variance for these θ-values actually are, 20 data series were made. These 20 data series each consisted of 50 randomly picked observations. The model was then solved for each of the data series. The model had an average run-time of 18 hours with these data sets, which is noticeably longer than the 5-12 hours it takes with the original data set. This could be due to inconsistencies as a result of the power plant's changing policy over time. Due to the increased run-time, the model was not run with the 20 data sets, but only 5. The results are displayed in table 5.5.

	Mean	St.dev	Min	Max	Median
θ_1	0.128	0.091	0.006	0.282	0.114
θ_2	0.421	0.257	0.014	0.664	0.504

Table 5.5: Descriptive statistics for model validation

As further development we suggest that the model validation is run with all of these 20 data sets or more in order to validate the θ's.

Increasing the resolution of the state variables With a daily or even hourly resolution for the state variables, it would be possible to estimate θ-values with even greater precision. As mentioned earlier, the hydropower plant can start and stop production several times within one week. Even though the histogram of weekly occurrences of different production levels is similar to the histogram of the hourly occurrences, as shown in section 4.6.1, we assume that some precision has been lost in the aggregation. We imagine that it would be possible to keep the weekly resolution of the seasonal component in order to avoid an increase in dimensionality.

Including more observed state variables Hydropower producers often make expectations for the spot price in-house or buy forecasts from analytic companies. To better capture the producers' expectations for the spot price, adding more price drivers could be beneficial. Some important drivers for the Nord Pool area price are: local temperature, the amount of production from fossil fuels, nuclear power production, and the EU ETS carbon price.

Chapter 6

Conclusion

In this thesis, a structural estimation model has been applied to data series from two disclosed hydropower plants in Norway. This has been done in order to discover certain aspects of the reservoir managers' preferences related to the production levels at which they operate. Time series models have been developed for the data series in order to replicate the managers' expectations of future conditions.

Specifically, the discovered preferences were related to the producer's willingness to produce at other levels of than the production level with the highest efficiency. The results yield an implied efficiency curve that considers the producer's preferences instead of the mechanical efficiency. The implied efficiency is a valuable finding and can be used as a basis for internal discussion within the hydropower plant's management.

The results show a greater willingness of the manager to produce at levels below than above the best efficiency point. Our hypothesis is that this is mainly linked to cavitation issues when producing at higher levels. Changes in these preferences over time was also investigated, which showed an increased willingness to produce both above and below the best efficiency point at the end of the time period.

Bibliography

Boger, M., Vestbøstad, E., 2016. Structural estimation analysis in hydropower scheduling. Master thesis, Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management.

Botterud, A. K., Bhattacharyya, A., Llic, M., 2002. Futures and spot prices - an analysis of the scandinavian electricity market. Proceedings of the 34th Annual North American power symposium.

Egre, D., Milewski, J., 2002. Diversity of hydro power projects. Energy Policy.

Fleten, S., Haugom, E., Pichler, A., Ullrich, C., 2015. Nonparametric structural estimation of operational switching options. Article submitted to Operations Research.

Fleten, S., Wallace, S., Ziemba, W., 2002. Hedging electricity portfolios via stochastic programming. Decision making under uncertainty, 71–91.

Fosso, O. B., Gjeldsvik, A., Haugstad, A., Mo, B., Wangsteen, I., 1999. Generating scheduling in deregulated system. the norwegian case. IEEE Transactions on Power Systems 14 (1), 75–81.

Hagfors, L., Kamperud, H., Paraschiv, F., Prokopczuk, M., Sator, A., Westgaard, S., 2016. Prediction of extreme price occurrences in the german day-ahead electricity market. Quantitative Finance 16 (12), 1929–1948.

Haldrup, N., Nielsen, M., 2006. Directional congestion and regime switching in a long memory model for electricity prices. Nonlinear analysis of Electricity Prices.

IBP, I., 2015. Norway Energy Policy, Laws and Regulations Handbook Volume 1 Strategic Information and Basic Laws. International Business publications, USA.

Kaaresen, F., Husby, E., 2000. A joint state space model for electricity spot price and futures prices. Norsk Regnesentral.

Kaminski, J., 2012. The development of market power in the polish power generation sector: A 10-year perspective. Energy Policy 42, 136–147.

Kellog, R., 2014. The effect of uncertainty on investment: Evidence from texas oil drilling. American Economic Review 104 (1), 1698–1734.

Kjølle, A., 2001. Hydropower in norway, mechanical equipment. http://www.ntnu.no/documents/381182060/641036380/Mechanical+Equipment+Kjolle+ny.pdf/83c2e69f-de23-4579-b5d5-4f3ed1ca4660, [Online; accessed 20-May-2017].

Kolsrud, C. W., Prokosch, M., 2010. Reservoir hydropower: The value of flexibility. Master thesis, Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management.

Kumar, P., Saini, R., 2010. Study of caviation in hydro turbines - a review. Renewable and Sustainable Energy Reviews.

Lin, C., Thome, K., 2013. Investment in corn-ethanol plants in the midwestern united states: An analysis using reduced-form and structural models. Working Paper, University of California at Davis.

Mo, B., Gjeldsvik, A., Grundt, A., Kresen, K., 2001. Optimisation of hydropower operation in a liberalised market with focus on price modelling. IEEE Porto Power Tech Conference.

NordPool, 2012. The Nordic Production Split 2004-2012. Nord Pool Spot.

NordPool, 2017. Price calculation. http://www.nordpoolspot.com/TAS/Day-ahead-market-Elspot/Price-calculation/, [Online; accessed 25-May-2017].

NordReg, 2014. Nordic Market Report. Nordic Energy Regulators.

of Energy, U. D., 2017. Types of hydropower turbines. https://energy.gov/eere/water/types-hydropower-turbines/, [Online; accessed 29-May-2017].

Olsson, J., Uvo, C., Foster, K., Yang, W., 2015. Technical note: Initial assessment of a multi-method approach to spring-flood forecasting in sweden. Journal of Hydrology.

Paish, O., 2002. Small hydry power: technology and current status. Renewable and Sutainable Energy Reviews.

Povh, M., Golob, R., Fleten, S., 2010. Modelling the structure of long-term electricity forward prices at Nord Pool. Springer Berlin Heidelberg.

Rapson, D., 2014. Durable goods and long-run electricity demand: Evidence from air conditioner purchase behavior. Journal of Environmental Economics and Management 68 (1), 141–160.

Rothwell, G., Rust, J., 1997. On the optimal lifetime of nuclear powe plants. Journal of Business Economic Statistics 15 (10), 195–208.

Rust, J., Rothwell, G., 1995. Optimal response to a shift in regulatory regime: The case of the us nuclear power industry. Journal of Applied Econometrics 10, 75–118.

Weron, R., Misiorek, A., 2008. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. International Journal of Forecasting 24, 744–763.

Wolfgang, O., Haugstad, A., Mo, B., Gjeldsvik, A., Wangsteen, I., Doorman, G., 2009. Hydro reservoir handling in norway before and after deregulation. Energy 34 (10), 1642–1651.

Appendix A

Appendix

A.1 Unit root test A

Statistical tests of the null hypothesis that a time series is non-stationary against the alternative hypothesis that it is stationary, are called *unit root tests:*

$$H_0\text{: } X_t \sim I(1)$$
$$H_1\text{: } X_t \sim I(0).$$

The interpretation of this is that an autoregressive process is stationary if and only if the roots of its characteristic polynomial lie strictly inside the unit circle (Olsson et al., 2015). The *Dicky-Fuller (DF) test* is the most basic unit root test. The test is based on the *Dicky-Fuller regression*, which is a regression of the form

$$\Delta X_t = \alpha + \beta X_{t-1} + \varepsilon_t.$$

The test statistic is the *t* ratio on $\hat{\beta}$, and it is a one-sided test for

$$H_0\text{: } \beta = 0$$
$$H_1\text{: } \beta =< 0.$$

If a test statistic falls into the critical region, we conclude that the process is stationary at the confidence level prescribed by the critical region. However, a problem with the Dickey-Fuller tests is that their critical values are biased if there is autocorrelation in the residuals of the Dickey-Fuller regression. The *augmented Dicky-Fuller (ADF) test* solves this issue by including as many lagged dependent variables as necessary to remove any autocorrelation in the residuals. The ADF-test of order *q* is based on the regression

$$\Delta X_t = \alpha + \beta X_{t-1} + \gamma_1 \Delta X_{t-1} + ... + \gamma_q \Delta X_{t-1} + \varepsilon_t.$$

The test proceeds as in the ordinary DF-test above, except that the critical values depend on the number of lags, q, that has been included. In this paper, we run all ADF-test with one lag,i.e. q = 1.

A.2 Descriptive statistics

Power plant A		Mean	St.dev.	Min.	Max.	Median
I	Observed	0,7151472	1,064891	-1,07798	7,766352	0,342648
	Simulated	0,751687	0,843074	0	4,373952	0,5568258
C	Observed	0,01462579	0,2187709	-0,5969365	1,003681	-0,00486093
	Simulated	0,227499	0,3268673	-0,3029207	2,104676	0,1476972
R	Observed	0,00E+00	0,1887571	-0,4811746	0,5708791	0,0308818
	Simulated	0,0156275	0,2160682	-0,5664131	0,7145599	-0,000471952
P	Observed	3,138499	0,5256082	1,427228	4.425137	3.204928
	Simulated	3,082369	0,3152591	2,346197	3,752814	3.064955

Table A.1: Descriptive statistics for the state variables processes for power plant A

Power plant B		Mean	St.dev.	Min.	Max.	Median
I	Observed	8,842896	10,62155	-13,3225	60,46186	4,9464
	Simulated	8,936684	6,151064	0	23,24549	7,273236
C	Observed	0,01480243	0,3172844	-1,02264	1,0478	-0,01281955
	Simulated	0,01855873	0,3069211	-1,097067	1,10942	-0,00044
R	Observed	-0,004508971	0,164374	-0,4700683	0,4615402	0,01634631
	Simulated	0,02265908	0,2654173	-0,8235879	0,6411824	0,01990878
P	Observed	3,232482	0,5020268	1,468325	4,425138	3,312216
	Simulated	3,168794	0,4115622	2,593862	3,79042	3,164708

Table A.2: Descriptive statistics for the state variables processes for power plant B

Table **??** ant table **??** shows the descriptive statistics for the state variable processes. These processes are elaborated in chapter 4. Table **??** shows the different coefficients for the state variables, for both power plant A and B.

	Power plant A	Power plant B
γ	0,432	0,5532132
$\overset{I}{1}$	-0,8281	-6,944
ϕ_1^I	6,0815	6,123
$\overset{I}{2}$	0,5579	-5,309
ϕ_2^I	0,1638	3,694
D_1	0,7151	8,856
σ^I	0,5158	0,532
f	20	20
ρ	0,936	0,932
β_1	0,99	0,93
β_2	0,04	0,07
σ_R	-1,53E-02	0,005
A_1^P	0,18059	0,17325
ϕ_1^P	-0,03969	-0,06570
A_2^P	-0,01348	0,01800
ϕ_2^P	-13,04764	1,55786
A_3^P	-0,01620	0,01322
ϕ_3^P	2,30658	5,33779
D^P	5,33572	5,39782
α_1	0,95025	0,9692
α_2	-0,0916	-0,067
σ^P	0,01	0,01459
Q	1,69	5

Table A.3: Coefficients for the state variables

A.3 State transition matrices

$$x_{t+1} = \begin{bmatrix} I_{t+1} \\ C_{t+1} \\ R_{t+1} \\ P_{t+1} \end{bmatrix}, \qquad x_t = \begin{bmatrix} I_t \\ C_t \\ R_t \\ P_t \end{bmatrix}, \qquad A_t = \begin{bmatrix} \gamma & 0 & 0 & 0 \\ \frac{1}{\bar{i}_{t+1}} & \frac{\rho \bar{i}_t}{\bar{i}_{t+1}} & 0 & 0 \\ 0 & \beta_2 & \beta_1 & 0 \\ 0 & 0 & \alpha_2 & \alpha_1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad e' = \begin{bmatrix} \varepsilon^I \\ 0 \\ \varepsilon^R \\ \varepsilon^P \end{bmatrix}, \qquad c_t = \begin{bmatrix} f^I(t+1) - \gamma f^I(t) \\ \frac{\rho \bar{i}_t}{\bar{i}_{t+1}} - 1 \\ 0 \\ f^p(t+1) - \alpha_1 f^P(t) \end{bmatrix}$$

$$L = \begin{bmatrix} cov(\varepsilon_t^I, \varepsilon_t^I) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ cov(\varepsilon_t^I, \varepsilon_t^R) & 0 & cov(\varepsilon_t^R, \varepsilon_t^R) & 0 \\ cov(\varepsilon_t^I, \varepsilon_t^P) & 0 & cov(\varepsilon_t^R, \varepsilon_t^P) & cov(\varepsilon_t^P, \varepsilon_t^P) \end{bmatrix} \tag{A.1}$$

The state variable transition for the reservoir level, (4.22), is not included in (4.30), but it is still part of the state transition on vector form in (4.29). The reason for not including it in the matrix form, is only because of the difficulty surrounding the min-operator.