

Hydropower scheduling: Correlated price and inflow as stochastic variables in policy generation

Maren Boger
Einar Midttun Vestbøstad

Project assignment TIØ4550
Norwegian University of Science and Technology
Supervisor: Stein-Erik Fleten, IØT

December 18, 2015

Abstract

This study explore the effects of including correlated and stochastic price and inflow in a stochastic dynamic programming model for decision making in reservoir management for hydropower. The performance of the release policies, when changing the correlation assumed during policy generation, is analyzed. We find that a policy with correlation between price and inflow embedded performs better than a policy that ignores it. An emphasis is placed on the importance of obtaining the correct correlation. Assuming a correlation that is too high leads to a worse performance than estimating it correctly.

Table of Contents

1	Introduction	2
2	Water resources management	4
2.1	Choosing power plant	4
2.2	Data description	7
3	Model formulations	9
3.1	The stochastic dynamic programming problem	10
3.2	SDP algorithm	12
3.3	Dynamic path analysis	17
3.4	Modeling inflow	18
3.5	Modeling spot prices	19
3.6	Finding the correlation between inflow and price	20
3.7	Generating a lattice for inflow and spot prices	21
4	Empirical analysis	23
4.1	Model applicability for our case hydropower plant	23
4.2	The effect of correlation between inflow and price	26
4.3	Shortcomings and improvements for the future	31
5	Conclusion	33
	Bibliography	33
	Appendix	36
5.1	Power plant descriptive data	36
5.2	Power plant capacity calculations, \bar{U}	36
5.3	Regression analysis	37
5.4	Lattice analysis	40
5.5	Effect of running until steady state	42

Introduction

The main focus of this study is to gain an understanding of how incorporating stochastic and correlated price and inflow in a water power planning problem affects revenues and optimal release policy of a hydropower plant. Due to the stochastic nature of price and inflow, we use a stochastic dynamic programming approach. The scheduling problem yields a release policy through time, as a function of the current price and reservoir levels used to describe the system. We have explored the performance of an optimal policy, on simulations where price is correlated with inflow. Our main interest is in comparing a release policy where the correlation is included, to a policy where it is ignored. This is to see if hydropower producers should take the correlation into account when developing the release policy, to get a better performance from the policy.

The correlated inflow and price simulations, used as a basis for our scheduling problem, is largely based on an empirical model by Kolsrud and Prokosch (2010). They incorporate a correlation between the combination of simulated inflow and overall reservoir levels with modeled spot prices. Their model is meant to simplify scheduling procedures, as the common approaches is based on complicated bottom-up models (Kolsrud and Prokosch, 2010). An example of such an approach is the EMPS-model.

The EMPS-model consists of two parts, the strategy part and the simulation part (Doorman, 2015). During the strategy part the release strategy is determined by finding the water values, which is the expected marginal value of the energy stored in the reservoirs. Different elements affect the computation of the water values, including expected prices. The EMPS-model does not take the correlation between inflow and prices into consideration in the strategy part, however. This correlation is first embedded during the simulation part, where the water values found in the strategy part are used to optimize a larger area, and how the water should be distributed between the reservoirs. Here the EMPS takes the correlation between inflow and prices into account, by simulating a large system where demand and transmission constraints, among other things, are inputs. We want to look into how including the correlating already when developing the release policy might be an improvement on how it is done now.

To our knowledge, no other study on hydropower have previously explored the effect

of including correlated and stochastic inflow and price explicitly when generating release policies for a hydropower plant.

The literature on stochastic dynamic scheduling for hydropower commonly utilize a cost minimization approach where only inflow are treated as stochastic and demand is included as an exogenously given deterministic variable. The commonality of deterministic treatment of demand and a cost minimization perspective can be explained by the fact that the energy market in most countries historically has been regulated. A regulated market entails that power companies are obliged to supply electricity to the geographical area they are responsible for. In such a situation, a cost minimisation perspective is applicable.

Norway was one of the first countries to deregulate their energy market, when the Energy act of June 1990 was introduced (Wolfgang and Doorman, 2009). Deregulation shifted the focus from a cost minimizing approach to a profit maximization one, with stochastic price inputs. Generating companies can now be assumed to only have the objective to produce electricity and sell at maximum profit. The power market is comprised of the spot market, the regulating market and the future markets. However, the degree of freedom to trade electricity imposed by deregulated markets make it reasonable to assume that all electricity is sold to the spot market in our problem formulation (Fosso, 1999).

Fosso (1999) present a model similar to one presented here in terms of using stochastic dynamic programming with a stochastic price model included as an additional state. To forecast prices they use a simulation model based on the EMPS model (EFIs Multiarea Power Scheduling Model), developed by SINTEF (Flatabø, 1988). EMPS is the commonly used as a tool for price-forecasting in the Scandinavian power market (Mo and Kåresen, 2001). The model can be used to analyze the whole Nordic market, taking into account demand, transmission constraints and thermal power production, and as a result outputs, among other things, price forecasts (Doorman, 2015). According to Mo and Kåresen (2001) the EMPS model has a link between spot price and local inflow through linking historical inflow years directly to a price scenario. However, the approach demand a historical inflow scenario for each new price scenario generated, limiting the amount of new scenarios possible to generate (Mo and Kåresen, 2001).

We find that a release policy where the correlation between inflow and prices are accounted for when generating the policy, performs better than a policy where it is not. The observed effect increases linearly for higher levels of actual correlation. We also highlight the importance of using a correlation as close to the observed correlation as possible, when generating a policy. This is because an overestimation of the correlation in the policy also shows a worse performance than if the actual correlation between inflow and prices is used.

This paper is organized as follows: In section 2 we explain the mechanisms in water resources management and how available data relates to our analytical purposes. In section 3 we elaborate on the stochastic dynamic model, simulation model and lattice formulation applied in our analysis. In section 4 the resulting policies is analysed and the implication of correlation between inflow and prices explored.

Water resources management

2.1 Choosing power plant

Our models utilize inflow data from a Norwegian hydropower plant and system spot prices for electricity and overall Norwegian reservoir levels. In order to exemplify our problem formulation data from an appropriate hydropower plant is selected. We have available data for a number of different power stations throughout Norway, and choose the one that best fit the following needs:

Suitability of chosen inflow data

First of all we need available inflow data to fit with model assumptions and provide a sufficient base for data analysis. To capitalize on both long- and short term flexibility we look at a planning period of one year, with a daily resolution. Daily resolution entails that it is important for the producer to have recorded inflow series with a daily, or more frequent, resolution. Furthermore it is desirable to have data for as many years back as possible to obtain adequate foundation for estimating our model parameters. Our power plant have daily inflow recordings, and available data from 2001 to 2008.

Suitability of reservoir composition

Another important factor for the choice of power plant is number of reservoirs connected to the power plant. Solving the Stochastic dynamic programming problem becomes much more complex with the inclusion of more than one reservoir, usually referred to in literature as the curse of dimensionality (Nandalal and Bogardi, 2007). Hence it is advantageous to choose a plant that has one single reservoir, or at least one where we can make the assumption that it has only one reservoir, within reasonable accuracy. As shown in figure 2.1 our power plant has three main reservoirs, in addition to several smaller lakes and rivers. The combined maximum capacity of the two smaller reservoirs is only 13.1% relative to the largest. Since one reservoir is so much larger than the others, we assume it is a suitable approximation to only look at them as one reservoir.

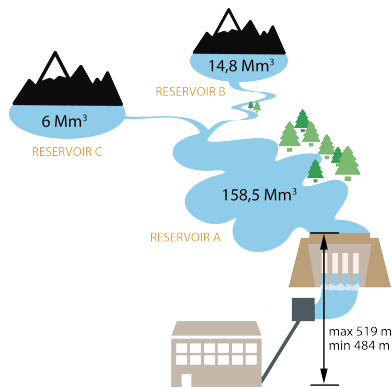


Figure 2.1: Illustration of the power station, used as case example, with reservoirs

Constant head assumption

We are assuming a constant energy equivalent and thereby also a constant head in our model. The energy coefficient determine how much energy is produced from water released, and is dependent on the head. See appendix for more detail on energy equivalent calculations. Head is defined as the difference in height between the reservoir and the turbine, thereby it changes as reservoir level changes. In order to be able to assume a constant head, our data should have negligible change in head from max to min levels in reservoir filling. The assumption is more reasonable the higher the head is compared to reservoir level difference. In our case company the change in head is less than 10% of the average head, which we deem negligible.

Binary production assumption

We assume the production to be binary, meaning the hydropower plant are assumed to produce at maximum capacity or not produce at all. Näsäkkälä and Keppo (2008) refer to this as a bang-bang strategy and state that such a policy, when being based on threshold functions of the spot price, is suitable. Such a strategy is relevant for hydropower production planning as reservoir are highly flexible also for short term changes in production. Therefore an almost immediately and costless change in production from 0 to max capacity is reasonable to assume.

A binary release regime is not necessarily applicable for our power plant. When analysing our power plant production, seen in figure 2.2a, we can see that the production is largely centered around minimum, maximum and half of maximum production. The heavy weight around half of maximum production is due to the power plant having two turbines, running independent of each other. This means that when one runs at max and the other is shut down, the total production is half of the max. Because of this, the assumption that we either produce max or nothing seems to hold for one turbine. For the daily resolution the power plant production is much more scattered, 2.2b. We still see that there is a small concentration around max and min, but not at all the same as for hourly resolution. Nevertheless we continue assuming binary production, to simplify. A next step would be to go from daily resolution to hourly in our problem, to better match assumed

production with the actual one.

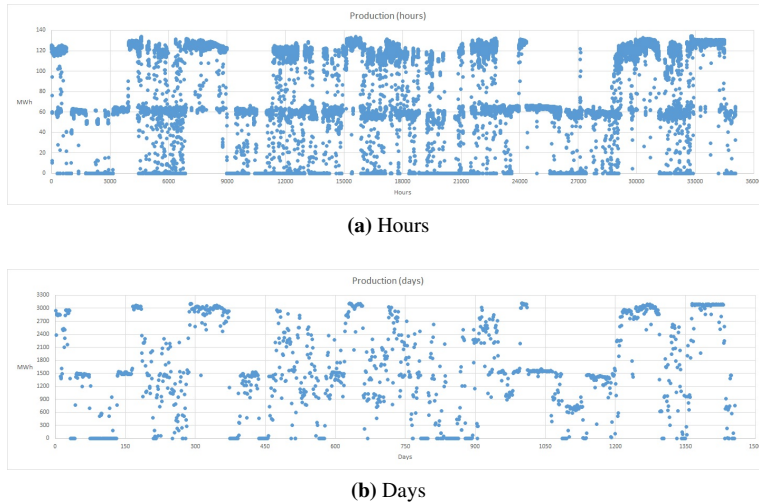


Figure 2.2: Production

Degree of regulation and capacity factor

Water reservoirs affects flexibility by functioning as a tool for storing power and moving production capabilities to periods with a high need for power, and thereby high price levels. The flexibility of hydropower plants depend on the size of the reservoir, the amount and distribution of inflow and the production capacity. If the size of the reservoir are small compared to the the accumulated inflow the production will be loose its ability to postpone production for periods with a higher price.

The degree of regulation of a reservoir provides an indication on how much water the plant are able to store. If the aggregated inflow minus production exceeds the reservoir volume the inflow will be non-storable and have to be used continuously. Our power plant have a degree of regulation of about 37%, and 68% during flooding season, which means that a third of the yearly inflow can be stored in the reservoir.

The yearly capacity factor tells us how much of the incoming inflow the turbines are able to produce, and is measured in average yearly inflow divided by production capacity. The capacity factor for our case reservoir is 53% and just over 200% during flooding season. If the accumulated inflow over a period raise the capacity factor above 100% at the same time as the reservoir levels are high, the reservoir will face a risk of spillage and thereby wasted water. We can see from figure 2.3 that our case power plant do not face any major limitations concerning production capacity, as long as they make sure to have room in reservoirs during the spring flood.

Both the degree of regulation and the capacity factor of our case study hydropower reservoir indicate that the plant have to handle their water release strategically for not overflowing their reservoir during the spring flood.

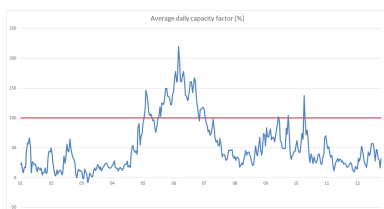


Figure 2.3: Daily capacity factor, averaged over 7 years.

2.2 Data description

Description of inflow data

The inflow from 2001 to 2008 for the relevant catchment area is shown in figure 2.4. We have data for the surrounding area all the way back to 1931, but that is not the exact same catchment areas, so we choose to only use the data from 2001, and onwards. We observe that some days there is negative inflow. That doesn't make much physical sense, but can be a result of the method commonly used to measure inflow. The data shows clear seasonal tendencies. The low periods can be explained by precipitation coming down as snow in during winter, resulting in low inflows. The high periods can be explained by snow melting during spring and summer, resulting in high inflows and spring floods.

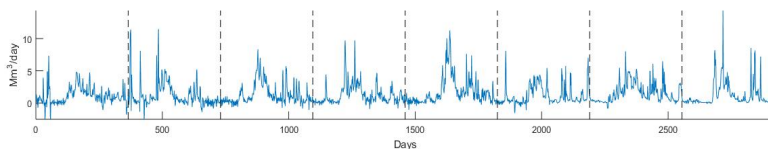


Figure 2.4: Inflow data from 2001 to 2008

Description of reservoir data

Overall reservoir level for 2001-2008, is shown in figure 2.5. This is the total amount of water in all of Norwegian reservoirs at a given week, in Gwh. When comparing price with the reservoir data we observe a correlation in term of high prices in periods of lowering reservoir levels. High inflow in our region corresponds with an increase in overall reservoir level. The data is downloaded from the NVE (2015) website.

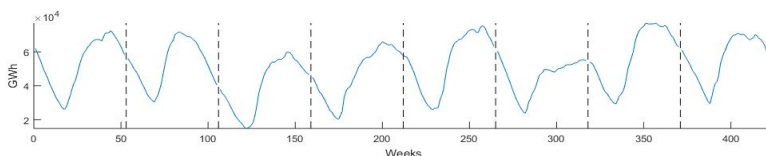


Figure 2.5: Overall reservoir level from 2001 to 2008

Description of spot price data

Figure 2.6 shows the spot log price in Norway in €/MWh, over the period 1993-2004. From the plot we can see that there is seasonal dependencies in the price. Hydropower producers experience seasonal and short term variations in demand and supply and hence price variations. The price is usually lower in the summer, and higher in the winter.

The high electricity prices during winters in Norway are mainly driven by the need for heating during cold winters. As 99% of the total electricity in Norway is generated by hydropower, and accumulated national inflow levels during winters are small, electricity becomes a scarce resource at this time (Kolsrud and Prokosch, 2010). The combined effect of low national inflow levels and electricity high demand during the winter is causes high prices during winter. This effect is partly counteracted by other hydropower plants saving production capacity for high demand periods, increasing the supply during this periods.

The low prices during the spring are influenced by the restriction imposed by the degree of regulation of Norwegian power plants. For power plants with low degree of regulation, very high inflow levels will force them to produce, driving prices down. During this period, demand for heating is low, further driving prices down.

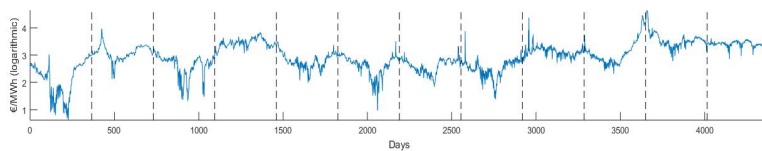


Figure 2.6: Log price data from 1993 to 2004, in €/MWh

Description price inflow correlation

In figure 2.7 we can observe slight indications of a negative correlation between inflows and price. In periods of high inflow the spot price are likely to be low because of high availability of water. Oppositely, during low inflow periods, electricity prices are likely to be driven upwards as water becomes a scarce resource.

Modelling the price without including this effect, would result in a policy that overestimate prices during high periods of inflow, thereby proposing to produce more electricity than optimal during the respective period. The opposite is the case for low inflow periods, driving prices up. In order to correct generation policies, so that they are not overproducing during high inflow periods, we build our models taking the correlation into consideration.

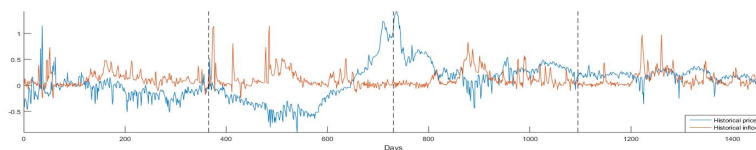


Figure 2.7: The blue line is spot log price data from 2001 to 2004, in €/MWh, and the red line is the corresponding inflow data for our case power plant

Model formulations

The aim of this chapter is to provide an overview of the elements included in our model and to explain the background for specific model detail decisions. Figure 3.1 provide an overview of the model structure used for our analysis. The overall model use a combination of sub-models as a basis for making decisions regarding release policies for a hydropower plant with a single reservoir.

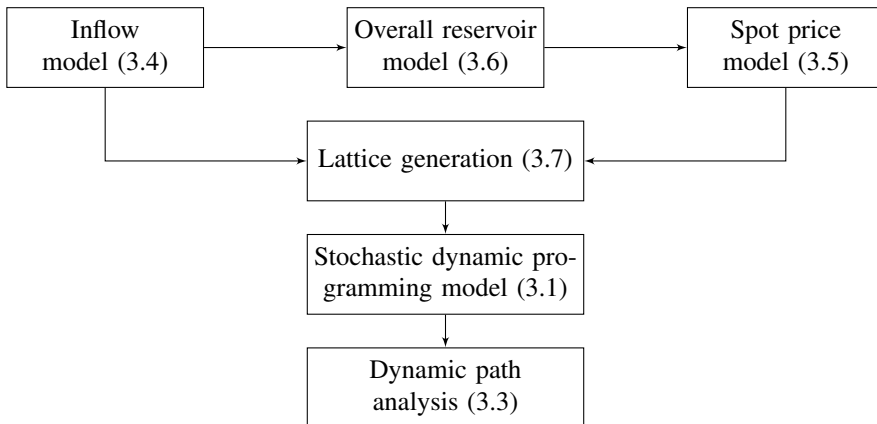


Figure 3.1: Overview of the model structure

As a first step in analysing the hydropower environment, simulation models for price and inflow are explained in chapter 3.4 and 3.5. Further on an option to include correlation between price and inflow are embedded through correlating the price model in chapter 3.5 with the results from the overall reservoir model and the local inflow model described in chapter 3.6 and 3.4. The results from the simulation models of price and inflow are embedded in a three dimensional lattice, described in chapter 3.7. The lattice are mainly used as a tool for allocation transition probabilities to a discredited set of states in order to be able to use the simulations in a stochastic dynamic programming (SDP) model described

in chapter 3.1. The output of the SDP model consists of two functions, a value function $V_t(R, P)$, containing resulting values and an optimal policy function $x_t(R, P)$, containing what actions should be taken. These functions work as decision making tools given different input scenarios of reservoir and price state levels. Lastly, dynamic path analysis, described in chapter 3.3, is used for simulation actions taken in simulated scenarios. The simulated actions provide a basis for discussing the quality of the decision making tool.

3.1 The stochastic dynamic programming problem

In this chapter we describe the formulation of the stochastic dynamic programming problem for our hydropower planning problem. The choice of dynamic programming for this analysis is due to its suitability for time-sequential decision problems, as it breaks down multistage decisions to a sequence of sub-problems. Thereby the method is able to derive the value of saving water for future periods. Since uncertainty is an inherent characteristic of water resources systems a stochastic approach to dynamic programming, stochastic dynamic programming (SDP) is best suited. For our study it is suitable to treat both price and inflow as stochastic. This is due to our planning period of one year, something we consider a long term planning horizon. In long term scheduling (Fosso, 1999) arguments both price and inflow should be considered as stochastic variables.

Indexes, parameters and variables

Set and indexes:

- t : time interval in days
- \mathcal{T} : stage space of discrete time intervals, $t = 1, \dots, T$
- h : inflow/price scenario at time t
- \mathcal{H} : state space of inflow/price scenarios at time t , $h = 1, \dots, H$

Parameters and variables:

- R_t : discretized reservoir levels in Mm^3 for each t
- I_t : inflow scenarios in Mm^3/day at time t
- P_t : price scenarios in $euro/kWh$ at time t
- R^{max} : max allowable storage volume in reservoir
- R^{min} : min allowable volume in reservoir
- Q : discharge capacity in Mm^3/day
- e : average energy equivalent in kWh/m^3
- β : discount factor
- S_t : hydro spillage in period t

Decision variables:

- x_t : binary release, decision variable, $x_t \in \{0, 1\}$
-

Objective function

Profit function

$$g_t(x, P) = P_t \cdot x_t \cdot Q \cdot e \quad \forall t \in \mathcal{T} \quad (3.1)$$

Bellman equation

$$V_t(R, P) = \max_{x_t} \{g_t(x, P) + \beta \sum_{I'} V_{t+1}(R + I' - x_t, P') \cdot P(I_{t+1} = I', P_{t+1} = P' | P_t = P)\} \quad (3.2)$$

The profit function is the product of the price at time t , the binary decision at time t , the discharge capacity and the energy equivalent (3.1). It is in the value €/day.

Since this is a discrete stochastic problem different scenarios for the price, P_t , and inflow, I_t , at each stage are taken into account. This means we get a new value function for each price/inflow scenario. The different inflows at stage t , together with the different decisions the producer can take, result in a range of different reservoir levels possible at each stage and price/inflow scenario, which restriction 3.3 shows. This means that we not only get a value function for each price/inflow scenario, but also for all different reservoir levels.

For a given stage, price and reservoir level, the objective function is the maximized sum of the profit function and the expected value function over all prices and reservoir levels in the next stage, with respect to the decision x_t (3.2).

Hydro reservoir balance

Restriction 3.3 is the reservoir balance together with the maximum and minimum level for the reservoir. This ensures that the reservoir level at stage $t + 1$ is equal to the sum of the reservoir level at stage t , the inflow into the reservoir in the next stage, $t + 1$, and the amount of water released for production at stage t (Doorman, 2015). The hydro reservoir balance could also reflect evaporation loss. In large reservoirs evaporation effects are relatively small and they do not effect the nature of our problem. Of that reason they are not included (Tejada-Guibert and Stedinger, 1993).

$$R_{t+1} = R_t + I_{t+1} - Q \cdot x_t - S_t \quad \forall t \in \mathcal{T} \quad (3.3)$$

Reservoir storage constraint

Restriction 3.4 defines the allowed values for the reservoir level at stage t , between the minimum reservoir capacity and the maximum capacity. Fosso (1999) and Doorman (2015) also use a formulation like this for the reservoir constraint.

$$R^{min} \leq R_t \leq R^{max} \quad \forall t \in \mathcal{T} \quad (3.4)$$

Release constraint

We assume that the producer can decide to either release water to the turbines at maximum of discharge capacity, Q , or chose to do nothing, i.e. 0 release. The decision variable x_t is therefore binary, either be 0 or 1, which restriction 3.5 shows. This is then multiplied

with the discharge capacity, Q , to get the actual amount of water released. Näsäkkälä and Keppo (2008) have been using such a policy, and arguments it is suitable for the hydropower planning problem.

$$x_t(R, P) \in \{0, 1\} \quad \forall t \in \mathcal{T} \quad (3.5)$$

3.2 SDP algorithm

Figure 3.2 provide an overview of the main steps in the SDP algorithm used for our analysis. The numbered elements refer to sub chapters containing discussions regarding case specific detail of the algorithm. Summed up the algorithm will iterate, solving argmax of the policy function for so to calculate the corresponding values, for all states from the terminal stage T to the starting point t_0 . By doing this the value of future decision will be taken into account when deciding on an optimal policy.

In sub-chapter 1 the discretization grid of the algorithm and the discretized transition probabilities are discussed. Sub-chapter 2 briefly elaborate on the benefits of releasing water at a current stage. Sub-chapter 3 explains the value function for our problem and sub-chapter 4 elaborates on the properties of the transition probabilities from one stage to the next. Sub-chapter 5 puts the transition probabilities back into context with the value function. In sub-chapter 6 our state transition calculations are explained and put into context of what have been done before. Sub-chapter 8 explain the form of the the optimal results, who consists of a value function, $V_t^*(R, P)$, containing resulting values and a policy function $x_t^*(R, P)$, containing what actions should be taken.

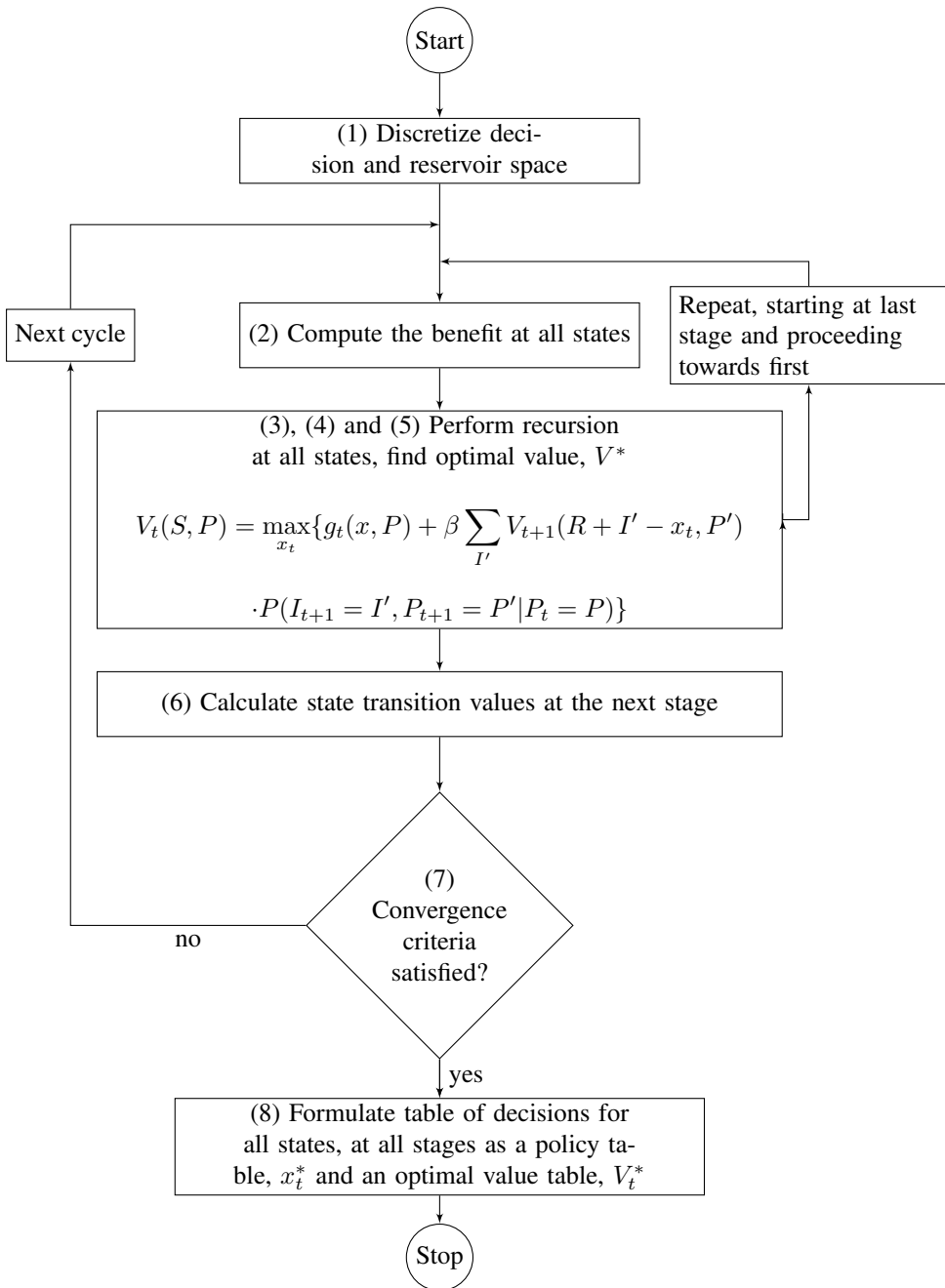


Figure 3.2: SDP algorithm. See corresponding subsections for more details.

1. Discretization of state space

The price and inflow simulation is discretized at each stage, t , in the lattice. The number of discrete states can change from one stage to the next, depending on user preferences. State values also change for each stage, with random increments, making infinite possible states. We limit the state space for inflow and price to all positive numbers, \mathbb{R}^+ . The state space for the reservoir levels have fixed state values, so the state space, \mathcal{R} , is a finite set. The discretized values are denoted by:

$$R \in \mathcal{R} \equiv \begin{bmatrix} R^{min} \\ R_2 \\ \vdots \\ R^{max} \end{bmatrix}, \quad I_t \in \mathbb{R}^+ \equiv \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_H \end{bmatrix} \quad \text{and} \quad P_t \in \mathbb{R}^+ \equiv \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_H \end{bmatrix}$$

2. Profit function

We assume our case company to be a price taker, something that is common when modeling Norwegian power plants. As the Norwegian electricity market has about 70 generating companies this is a reasonable approach as a single producer will not have large enough share of the market to have a significant impact on prices. The price taking assumption is a necessary condition for a free market to be economically efficient. Such an assumption also avoid considerable modeling complexity as a result of oligopoly (Fosso, 1999). A price taking assumption let us formulate the profits from production as a simple linear relationship between price and produced electricity .

Fosso (1999) state that, in a deregulated market, a generation company has in principle no other objective then to produce and sell electricity with a maximum profit. This implies that we can assume no production requirements or start-up costs for our scheduling problem. Formulation the profit function as strictly profit maximizing function is therefore reasonable.

At a given stage, t , the profit function matrix is:

$$g_t(x, P) = P_t \cdot x_t \cdot Q \cdot e$$

where x_t is the binary release, Q the discharge capacity and e the average energy equivalent.

3. Value function formulation

We want to find the optimal value of the reservoir for each reservoir level and each price/inflow level at the starting point, t_0 . This can be written as the maximum, given optimal choice of x_t , of the conditional expected value of of all profits from now until the terminal stage, T , given you have price P_0 and reservoir level R_0 now and discounted with β .

$$V_0(R, P) = \mathbb{E}\left(\sum_{t=0}^T \beta^t g_t | R_0 = R, P_0 = P\right)$$

We can then write optimal value at any given stage t , as

$$V_t(R, P) = \max_{x_t} \{ \mathbb{E} \left(\sum_{t'=t}^T \beta^{t'} g_{t'} | R_t = R, P_t = P \right) \}$$

The profit at stage t is known, so we can take it outside of the expectation and write

$$V_t(R, P) = \max_{x_t} \{ g_t(x, P) + \mathbb{E} \left(\mathbb{E} \left(\sum_{t'=t+1}^T \beta^{t'} g_{t'} | R_{t+1} = R, P_{t+1} = P \right) | R_t = R, P_t = P \right) \}$$

Then we can write the inner expected value as the value function in the next stage, $t+1$, in a recursive manner according to Bellman's principle of optimality. This states that the optimality of the remaining decisions is incorporated in the continuation value (Dixit and Pindyck, 1994). By choosing the optimal decision at t , you will have an optimal policy from that stage until the termination stage.

$$V_t(R, P) = \max_{x_t} \{ g_t(x, P) + \beta \mathbb{E} (V_{t+1}(R_{t+1}, P_{t+1}) | R_t = R, P_t = P) \}$$

Now the conditional expectation is the sum over all value functions in the next stage, times the conditional probability of ending up in the next reservoir level and price, given the price and reservoir level you are in today

$$V_t(R, P) = \max_{x_t} \{ g_t(x, P) + \beta \sum V_{t+1}(R', P') \cdot P(R_{t+1} = R', P_{t+1} = P' | R_t = R, P_t = P) \}$$

4. Transition probabilities

We can rewrite the conditional probability, using our reservoir balance restriction

$$R_{t+1} = R_t + I_{t+1} - x_t$$

Then

$$P(R_t + I_{t+1} - x_t = R', P_{t+1} = P' | R_t = R, P_t = P)$$

Rewriting the first part of the probability we can express it as the inflow instead of the reservoir level. This means the probability no longer is conditional on what the reservoir level is today, but rather what the inflow is. However, since each price is linked to one inflow the probability doesn't have to be conditional on both price and inflow, but just one of them. We choose the price, which means the conditional probability is

$$P(I_{t+1} = R' + x_t - R, P_{t+1} = P' | P_t = P) \tag{3.6}$$

This is the probability of ending up at inflow I_{t+1} and price P_{t+1} in the next stage, given you are at price P_t (and inflow I_t) now. Hence this is the transition probability from going from one stage to the next, inheriting the Markov-I property of serial correlation (Nandalal and Bogardi, 2007).

5. Optimal value and policy

We can now put equation 3.6 back into the value function. Using that $R' = R + I' - x$ we write

$$V_t(R, P) = \max_{x_t} \{g_t(x, P) + \beta \sum_{I'} V_{t+1}(R + I' - x_t, P') \cdot P(I_{t+1} = I', P_{t+1} = P' | P_t = P)\}$$

Here we sum over all possible inflows, I' , at the next stage. This equation can now be used to find optimal value at the starting point, t_0 , by working backwards from the terminal stage, T . This also yields the optimal policy x , for each stage t .

6. State transitions

If employing multiple state variables, dense discretizations or multiple reservoirs the stochastic dynamic programming problem will quickly reach extremely high dimensionalities, again leading to large computational requirements. This is commonly referred to as “the curse of dimensionality” (Nandalal and Bogardi, 2007). Our problem has a dense discretization of reservoir level states, combined with discretizes price/inflow state variables and a high number of time steps.

The inflow comes in discrete grid points from the Markov chain, generated by the lattice, but the grid point positions vary through time. When we then use 3.3 to calculate the reservoir level in the next stage, R_{t+1} , values in next time step likely fall outside the grid points already established for the state space R , creating a new possible reservoir level. This causes our state dimensionality to grows rapidly as we move from one time step to the next.

In order to limit the growth in number of state variables, we round R_{t+1} to the nearest predefined gridpoint. Maidment and Chow (1981) utilize a strategy that has a similar effect on the dimensionality of the problem as the rounding approach we have applied when going from one stage to the next. In their approximation they limit the growth of reservoir level states by mapping the expected values of inflows to the reservoir levels in next state. Their approach illustrate the need for reducing the growth in number of states when using stochastic state variables.

Rounding R_{t+1} to the nearest grid point entails rounding errors, but reduces the growth in dimensionality drastically so that we can increase the density of the reservoir level state grid in order to compensate for the error. The impact on operational performance by increasing the number of storage and inflow classes of SDP's have been studied by Bogardi and Nandalal (1988). Their results showed a positive effect up to a certain limit. After this limit the system performance would have diminishing returns when increasing the number of classes. If further improvement in system performance is required, they suggest rather synchronizing the number and sizes of the storage and inflow classes as a future study. For our model, a synchronization approach could have been applied in order to avoid rounding errors. However, Tejada-Guibert and Stedinger (1993) indicate that such an approach is seldomly attractive. The reason is that the feasible system dynamics gets distorted and a very fine state space grid becomes necessary in order to match to the resolution of the release and inflow variables.

7. Convergence criteria check

If we set the termination value V_T equal to zero, the optimal policy will want to empty out all the water before reaching this period, since then it will be worthless. The water needs a continuation value to avoid this problem (Wolfgang and Doorman, 2009). A way to obtain such a value is to find a convergence value, through backwards recursion.

It is shown by Su and Deininger (1972) that since the profit flow is discounted the recursive computation always converges to a steady state, which yields the maximum total expected profit, in the long run. For hydropower plants a 50-100 year life time is reasonable to assume (IEA, 2015). Therefore we find it reasonable to running the one year backward recursive computation over multiple cycles, k , until convergence.

Initially we guess a values in the terminal value function, $V_T(R, P)$, and run a backwards recursion. In the lattice-generation we set the number of inflow/price states to be the same in the first and last time step. This results in equally many resulting value functions after backwards recursion as terminal value states. The resulting value functions for state 1 to H can thereby be set as new terminal values for state 1 to H .

The backwards recursion is then repeated by resetting the optimal value after each cycle to $V^*(R_0, P_0)$. The change, δ_k , is calculated as the percentage change of average value for each backwards recursion iteration, k . Tejada-Guibert and Stedinger (1993) are, like us, using the average benefits over all state variables at $t = 0$ from one cycle to the next when estimating convergence. The convergence criteria (Nandalal and Bogardi, 2007) is set to:

$$\delta_k = \frac{|V_k^* - V_{k-1}^*|}{V_1^* - V_0^*} \leq \varepsilon \quad \forall k = 1, \dots, K \quad (3.7)$$

This is calculated for every k th iteration of the cycle and is terminated if the convergence criterion is not satisfied in K iterations. If any of the changes in value goes below the convergence criteria ε the cycle will be terminated, and the last given value will be considered as the optimal solution. For this study ε is set to 0.00125.

8. Optimal policy matrix

The optimal values at each stage is stored in a policy matrix $x^*(R, P)$ and a corresponding value matrix, $V^*(R, P)$. These matrices can be used as look up tables. If you have a given reservoir level, R_t , and price, P_t at a time t the optimal decision and its corresponding value can be found from the policy matrices.

3.3 Dynamic path analysis

When an optimal policy is found it has the form of a collection of data points representing the optimal policy given values for each price/inflow and reservoir level, R , in current time step. This provides a static picture of the controlled process. In order to learn more about its dynamic behavior, dynamic path calculations is used for finding the optimal policy given a input scenario (Miranda and Fackler, 2002).

We start by going forward in time from an initial value to the terminal point, first finding the optimal production based on what prices and reservoir level we are at in the

current time step. The optimal decision are then used to find the profit, which we sum up to get the total value at the end point, $V(R_{iT}, P_{iT})$.

In order to know the optimal production in this time step, we need to find price and inflow levels that fit to the grid points created by the lattice. Thereby, when using simulated values, we have to find the point that best matching our simulation. In order to do this we find the point with the lowest distance weighted on both price and inflow.

$$l_{closest} = \left(\operatorname{argmin} \left\{ \left| \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} - \begin{pmatrix} I_{t,1}^{Lattice} / I_t^{Sim} \\ \vdots \\ I_{t,l}^{Lattice} / I_t^{Sim} \end{pmatrix} \right| + \left| \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} - \begin{pmatrix} P_{t,1}^{Lattice} / P_t^{Sim} \\ \vdots \\ P_{t,l}^{Lattice} / P_t^{Sim} \end{pmatrix} \right| \right\} \quad \forall i \in \mathcal{I}, m \in \mathcal{M}, k \in \mathcal{K} \right)$$

Here the first part is the percentage difference between our simulated inflow and all the different inflow levels in the lattice. The same is done with prices for the second part. The sum of these are then the total percentage difference for both price and inflow. The lowest of these differences then gives us the closes point in the lattice, with index $l_{closest}$, combined for both price and inflow. The index is used to find the optimal production, $x^*(R_{it}, P_{l_{closest}})$. Then the profit at this stage can be calculated and the total value up to this point in time updated, $V(R_{it}, P_{l_{closest}})$.

3.4 Modeling inflow

A large sample size, or inflow and price paths, is needed to generate a Markov chain, in order for the model to have sufficient statistical power. With a sample size too small, there might be properties in the total population that is not captured in the sample. Our planning period is one year. The number of paths available, is then equal to the number of years our power plant has been in operation. In order to obtain a sufficient number of scenarios of inflows, we need to introduce a stochastic model based on our available historical data. This allows us to run simulations for a large number of seasons, or scenarios.

Model structure

Annual and seasonal inflow series are seldom highly correlated, but monthly, weekly and especially daily or hourly flows generally exhibit high serial correlations (Nandalal and Bogardi, 2007). As we are working with a daily resolution we find it reasonable to incorporate the serial correlation attribute. From figure 2.4 we observe seasonal variations. To include both the serial correlation and the seasonal variations we build a model consisting of two parts. In the same way as Kolsrud and Prokosch (2010) we have an auto-regressive base process, and a seasonal function.

$$I_t = f_{season}(t) + X_t^I$$

where $f_{season}(t)$ is the function describing the seasonality of the inflow, and X_t^I is the underlying base function.

Unlike Kolsrud we use weekly dummy variables to create the seasonal function.

$$f_{season}(t) = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \dots + \beta_{51} D_{51} \quad (3.8)$$

Here D_1, \dots, D_{51} are the dummy variables of week 1 to 51. $D_1 = 1$ and 0 elsewhere, $D_2 = 1$ if we are in week 2 and 0 elsewhere, and so on. Week 52 doesn't need a dummy variable, because at that time all other dummies are 0. We can then estimate the coefficients by doing a linear regression on the historical data, I_t with this equation:

$$\hat{f}_{season}(t) = \hat{\beta}_0 + \hat{\beta}_1 D_1 + \hat{\beta}_2 D_2 + \dots + \hat{\beta}_{51} D_{51}$$

To find the base function we use a simple auto-regressive model, like Kolsrud and Prokosch (2010):

$$X_t^I = \sum_{i=1}^p \varphi_i X_{t-i}^I + \xi_t \quad (3.9)$$

where c is a constant, p is decided from significance of the coefficients, φ_i , and ξ_t is the error with mean 0 and standard deviation σ_ξ .

The coefficients, φ_i , can be estimated by first subtracting the seasonalities from the historical data, $\hat{X}_t^I = I_t - \hat{f}_{season}(t)$. We can then do a new linear regression with:

$$\hat{X}_t^I = \sum_{i=1}^p \hat{\varphi}_i X_{t-i}^I + \xi_t$$

From the inflow data we register that the daily inflow sometimes drops below zero, as mentioned in section 2.2. This can create problems in our SDP algorithm, since negative inflow would mean there is a possibility to reach a negative reservoir level. Hence we want to avoid this problem in our model. To do this we set all negative value to zero. Now the total inflow over a year is too high, however. This is solved by decreasing all other values by the average amount removed below zero.

3.5 Modeling spot prices

When modeling spot prices we want to use the logarithm of the price, to avoid the possibility of negative prices in our model. We assume price is serial correlated between days, and that there are seasonal variations, as we can observe from figure 2.6. Price is then modelled in the same way as with inflow, using an auto-regressive process as a base process, and adding a seasonal function with weekly dummy variables, see equation 3.8.

We are especially interested in the correlation between inflow and price. To include this effect into the price model, we follow the same approach as Kolsrud and Prokosch (2010) by including a mean reverting level in the base process of the price, Y_t . First we assume a known mean reverting level, and then remove it from the base process

$$X_t^P = Y_t - \mu_t \quad (3.10)$$

The mean revering level, μ_t , is initially set to zero. This is reasonable since removing the seasonal function from the actual spot price would leave the base process around zero Kolsrud and Prokosch (2010). In the next section a time dependent mean reverting level is included, to account for correlation with local inflow. X_t^P can be modeled as an autoregressive process, as we did with the base process of the inflow, see equation 3.9, and using linear regression to find the coefficients. The final model for the price will then look like

$$P_t = f^{season}(t) + Y_t \quad (3.11)$$

3.6 Finding the correlation between inflow and price

In the previous sections, price and inflow have been modeled separately. In reality, this is not the case as high inflow levels usually coincides with low spot prices. In the EMPS model this correlation is not included before generating the release policy, but after in the simulation phase (Doorman, 2015). In order to account for the correlation between inflow and price already in the policy generation, we incorporate the model proposed by Kolsrud and Prokosch (2010) for overall reservoir, into the spot price model. The overall reservoir level is here dependent on the local inflow. The reason for using overall reservoir levels is that we model the system prices, and the overall reservoir level in Norway affects the system price, as mentioned earlier.

They calculate the deviations from cumulative inflows, Δi_t^{cum} , and the national deviations from normal reservoir level, $\Delta r_t^{overall}$.

$$\Delta i_t^{cum} = \frac{I_t^{cum} - \bar{I}_t^{cum}}{\bar{I}_t^{cum}} \quad (3.12)$$

$$\Delta r_t^{overall} = \frac{R_t^{overall} - \bar{R}_t^{overall}}{\bar{R}_t^{overall}} \quad (3.13)$$

Where I_t^{cum} is the cumulative inflow of the respective day for the last year and \bar{I}_t^{cum} is the historical average cumulative inflow for the same period. $R_t^{overall}$ is the aggregated reservoir level at the end of period t , and $\bar{R}_t^{overall}$ is the historical average of the aggregated reservoir level at time t . This deviation calculations is used to model the relationship between local inflow variations and aggregated reservoir levels.

$$\Delta r_t^{overall} = \beta_1 \Delta r_{t-1}^{overall} + \beta_2 \Delta i_t^{cum} + \epsilon_t \quad (3.14)$$

t is measured in weeks, for so to be interpolated into daily reservoir deviations. ϵ_t are independent and identically distributed random variables with mean zero and variance σ^R for our power station. β_1 and β_2 are estimated using linear regression. After estimating β_1 and β_2 the overall reservoir level deviations, $\Delta r_t^{overall}$, can be simulated with a given inflow series, I_t .

The mean reverting level of the base process in the spot price simulations can now be modeled to be dependent on inflow level at our hydropower plant by using a scaled version of the simulated overall reservoir level.

$$\mu_t = \eta \Delta r_{t-1}^{overall} \quad (3.15)$$

Where η is a parameter determining how much the national deviation from normal reservoir levels should influence the base process (Kolsrud and Prokosch, 2010).

3.7 Generating a lattice for inflow and spot prices

The set of data series, simulated using historical spot prices and inflow data, is used as scenario inputs for modeling a lattice. The lattice is used as a tool for describing the probabilities of going from a given price/inflow state, or node, to a given node at the next stage. The information generated in the lattice is a discrete Markov chain, hence it can be used as input for price and inflow development in the dynamic programming problem. Both the value of the nodes and the transition probabilities vary with time when using this kind of lattice, resulting in different state grids for inflow and price for each time step in our model.

We use a model developed by Alois Pichler (2015) to generate a lattice for both price and inflow, interlinking them in pairs with the same respective transition probabilities. This model starts by defining the possible nodes directly from what happens in the first scenario. The node values are not statically determined beforehand, but dynamically generated and adapted using a distance criterion built on transportation theory. By applying this technique, if the nodes on the next scenario is very far from the first, new nodes are generated. Else, if a scenario is close enough to a node, that node is moved a little bit closer to the scenario. The probability of ending up on that node, given the node the scenario was closest to in the previous stage, is then updated by one. The percentage transition probabilities are then found by dividing the probability of ending up on each node in a given stage, by the total number of scenarios. For every new scenario input, the possible node positions will be adjusted together with their transition probabilities. As more scenarios are used, the lattice will gradually improve (Pflug and Pichler, 2015).

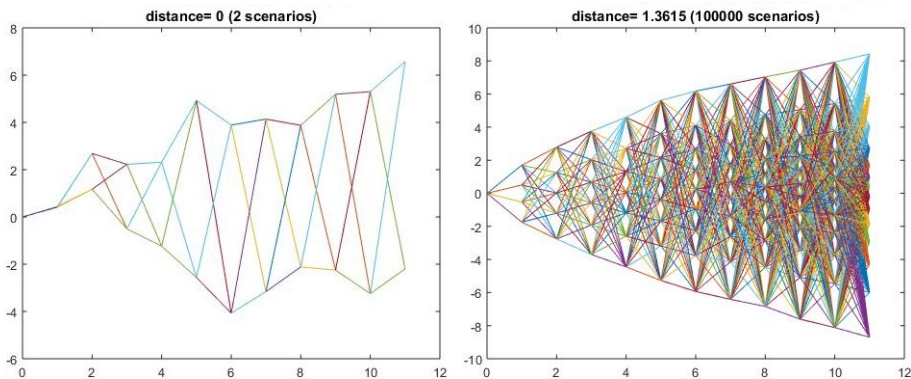


Figure 3.3: Lattice plot examples

Empirical analysis

The aim of this chapter is to explore the difference in performance of using a policy that does take into account the correlation of inflow on price. We compare the performance to the performance of a policy that does not take the correlation into account, when both policies are used on simulations that has the correlation embedded. In addition we analyze the stochastic dynamic programming model in general, to test whether it is applicable for our case power plant.

4.1 Model applicability for our case hydropower plant

As a tool for analysing the models applicability for our case reservoir we compare average simulated policy results from our SDP with their realized historical counterparts, shown in figure 4.1. The simulated reservoir levels are obtained by using the dynamic path analysis described in chapter 3.3. By simulated reservoir level, we refer to the reservoir level we obtain through the year by running simulations through the optimal policy given by the SDP. We here start at the same initial level as the historical average.

Physical description of reservoir path

We see from figure 4.1 that the policy recommends to produce electricity during winter in order to capture on high prices. Before flooding season our police almost empties out the content of the reservoir in order to be prepared for the expected incoming water. During flooding season the reservoir level rises again, since the inflow is larger than the maximum discharge, and prices are low. Towards the end of the year the precipitation comes down as snow in the mountains and reservoir levels decrease as demand for electricity increases. The observed simulation behavior is close to what is expected in a real life situation. As we can observe a strong similarity between the simulated value and historic values we conclude our models applicability for our case reservoir is within reason.

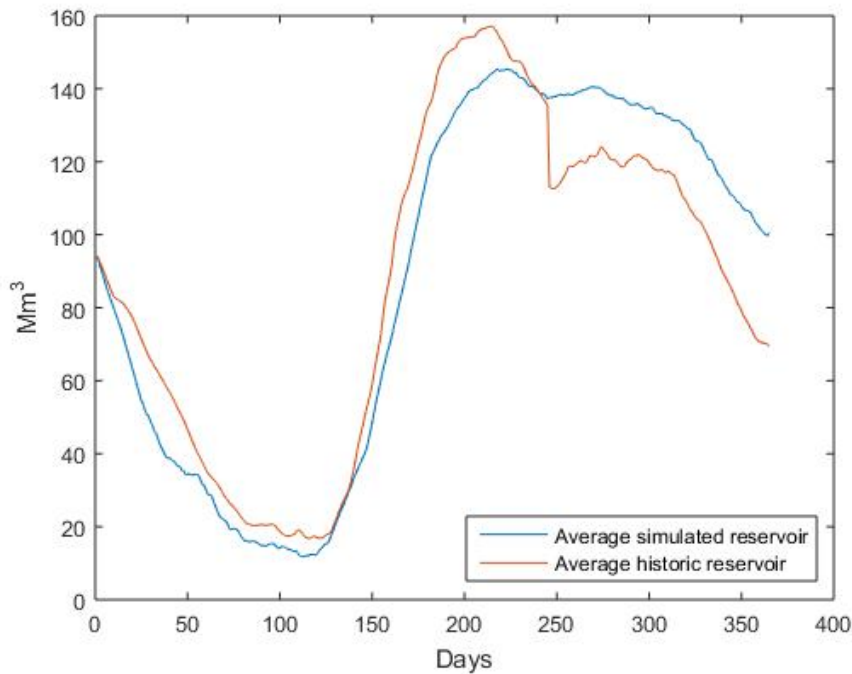


Figure 4.1: Average reservoir path given optimal policy plotted against historical average

Effect of assuming constant head

By assuming a constant head we expect our model to overestimate the value of water slightly when reservoir level are low and underestimate the value for high reservoir levels. This can partly be explained by our assumption of constant head. In reality the energy equivalent is dependent on the head, a value that is lowered as the reservoir level goes down. Our energy equivalent account for a head of about 70% of max. This lead to an overestimation of the energy equivalent when reservoir levels are lower than this level, and an underestimation for higher levels. See appendix 5.2 for energy equivalent calculations. Looking at figure 4.1 we see that the simulated reservoir path lies below the historical average in the first half of the year, as expected. For the second part the comparison between realised results have to be explained by something else.

The effect of inflow variance on risk aversion when reservoir levels are high

We assume our policy to be more risk averse at high reservoir levels than the previously observed policy. This expectation is due to a higher variance in our simulated inflow levels during dry periods then the observed inflow variance levels in this period. A high variance in inflow during winter will increase the risk of spillage when reservoir levels are high. The modeled risks of spillage are illustrated as percentiles in reservoir level occurrence in figure 4.2. In our model 90-percent of the cases avoid spillage. To make the model less

risk averse during winter, heteroscedasticity could be employed in the inflow simulation. For the hydropower case a lower variance during the winter season is applicable. In fig. 4.1 we can observe average peak reservoir levels during winter to reach $150Mm^3$ in our model. Historically realized peak reservoir levels are $160Mm^3$, where $180Mm^3$ is the maximum capacity before spillage. As assumed, average simulated SDP peak reservoir levels are observed to be slightly low, compared to historically realized peak reservoir levels. Secondly we assume the observed risk aversion to be slightly higher in the average

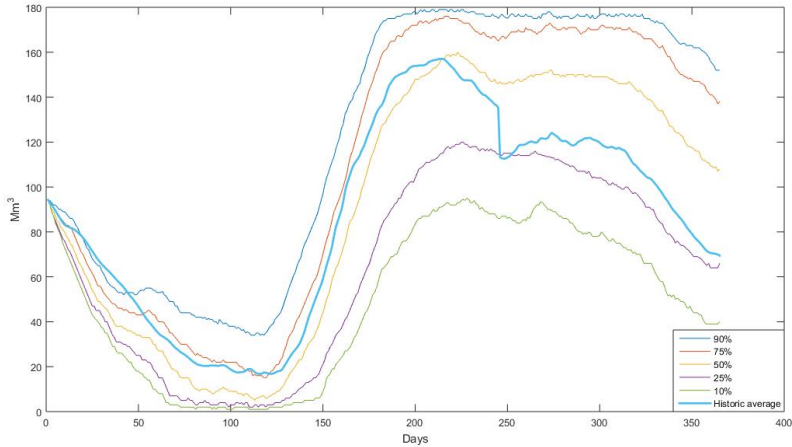


Figure 4.2: Percentiles for simulated reservoir paths

observed case, as opposed to our model, during the autumn. We assume this to be a consequence of a risk of very high cases of precipitation, observed in historical data in the form of high variance levels during autumn. If such is the case, and our model underestimates the variance during the autumn, we expect our policy to be less risk averse than previously observed policy realizations during autumn. We can observe a drop in average previously observed reservoir levels in September-October in figure 4.1. We envisage the difference between our simulations and average historical reservoir levels to be partly explained by a difference in risk aversion, as a consequence of the observed increase in variance in the data during the autumn. Nevertheless, this is only part of the explanation. We also believe our historical sample size to have an effect on the observed release policy, as is explained in the next section.

Effect of sample size in reservoir data

When comparing the observations with average simulated reservoir levels, in figure 4.1, we expect there to be some deviations as a consequence of differences in sample size. The average previously observed reservoir levels are comprised of a data set with 6 years of data. Such a data set will most likely not be completely generalizing and to some extent illustrate time specific incidents. This can partly explain the sudden drop seen in the average historic reservoir plot. Also the variance of a small sample size should be

higher than if using a larger one. We can observe such a tendency in 4.1 and conclude from this that these differences can not disprove our model. If we compare our modeled reservoir levels with modeled reservoir levels by Mo and Kåresen (2001), we can observe that both models have excluded the sudden drop. Also we can see from figure 3.6 of overall reservoir levels that the sudden drop is not represented. From this we conclude that our model deviations from observed sudden drops is reasonable.

4.2 The effect of correlation between inflow and price

A key part of this analysis is to explore the effect of incorporation a correlation between inflow and prices. The correlation of inflow on price is embedded through a mean reverting level for the price, see section 3.6. The mean reverting level is the deviation from normal for the overall reservoir level in Norway, where accumulated local inflow is affecting the overall reservoir level. The inclusion of the mean reverting level is weighted with a factor η , (equation 3.15). η can be adjusted to see the effect of the correlation on the result of the SDP model. We are mainly interested in comparing a release policy where the correlation is included, to a policy where it is ignored. This is to see if hydropower producers should take the correlation into account when developing the release policy, to get a better performance from the policy.

In order to analyze effects of correlation, we make a series of generation policies based on different assumptions on the correlation. We generate different lattices, each based on a price simulation with a different η . The SDP is then run on the lattices to make an optimal policy for each assumption on the correlation. By using these policies on different price and inflow paths, we can analyse the effect of the correlation. A useful measure for these comparisons is the achieved price. That is the total profit for a year divided by the total production, and tells us how well we took advantage of high prices, while avoiding low ones.

Comparing value of differently correlated policies on historical data

We expect historical price and inflow to be negatively correlated. In order to explore this assumption we would like to see what happens if we apply our policies on historical data. If we use a policy where we have modelled this relationship, and assume we have done it perfectly, the value we get from the water should be the optimal. If we on the other hand have a policy where this correlation is not taken into account, we would produce more than we should for periods with low inflow. That is because the prices here are actually higher than what we expected, which gives us an incentive to produce. During periods of high inflow, the prices are lower than we expected, so we decide to save the water for later, producing less than we should have done, if we had accounted for correlation. The combined effect is that the achieved price is lower if we don't take correlation into account while developing the policy.

Figure 4.3 shows how the average achieved price varies relative to the average price using historical data, for different values of η . The figure shows a lower achieved price for some of the η 's than when it is 0. This is inconsistent with our assumption that achieved price is higher for higher values of η . Even though we might be using the wrong number for η we expect to at least get a higher value than for no correlation, since we assume the

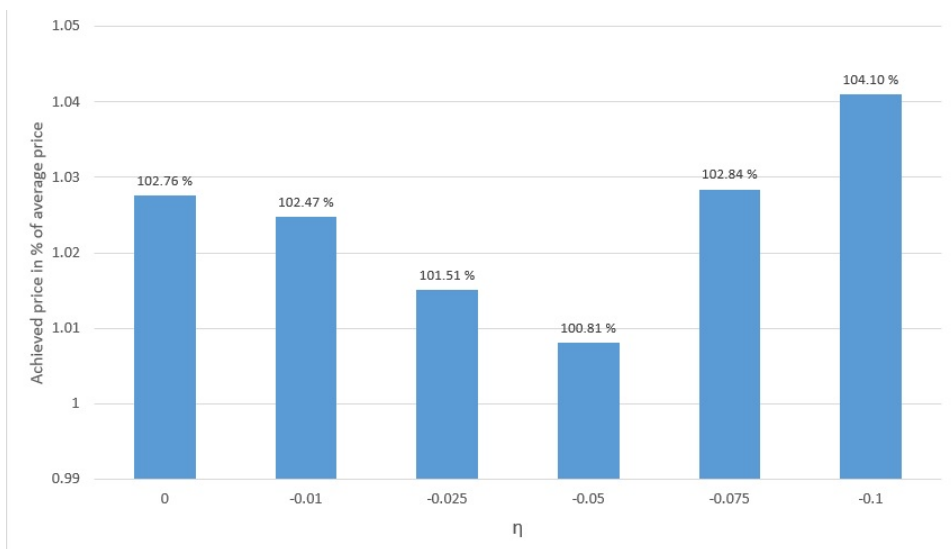


Figure 4.3: Using different policies on historical data for price and inflow, from 2001 to 2004

historical values has a correlation. The reason for this fault is, most likely, due to sampling errors. The data used is only from 2001 to 2004, four sample paths, because these are the only years with overlapping data for price and inflow. Only four years of data might not represent all the information needed. Given we have sampling errors, the conclusion of this assumption is left inconclusive. To get a more accurate measure we need more years of data.

Comparing value of differently correlated policies on the same simulation

Instead of historical data we here use a simulation on all policies, and compare the results in the same way as we did for historical data. We expect achieved price to vary for different policies, and that the policy developed with the same correlation as the simulation will have the highest value.

Figure 4.4 shows the achieved price as a percentage of average price through a year, for different policies on a simulation with η equal to -0.05 . We notice that the highest value is for the policy with an assumption on η equal to -0.05 , the same as for the simulation used on the policies. That is not unexpected since assuming the correct correlation, with respect to the simulation used, when developing the policy. The value of policies with lower assumption on correlation is as expected. The policies generate less than they should during high inflow, because the price is lower than they expected, and vice versa. For policies with a higher correlation than the simulation, the opposite is true. Now the producers generate more than they should during periods of high inflow, because the actual price is higher than what they expected. During low periods of inflow they expected a higher price than what is reality, and decide not to produce. The figure indicates that for high assumed correlation there is a severe drop in value.

From the results illustrated in figure 4.4 we conclude that finding the correct η is impor-

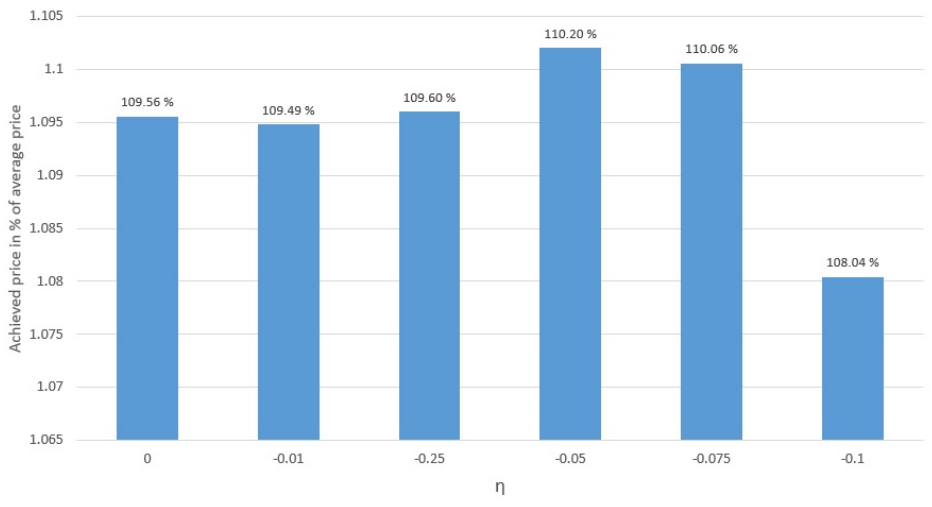


Figure 4.4: Using the same simulation on different policies, with $\eta = -0.05$

tant when making a policy. Assuming a lower correlation when generating policies then the correlation of the data you are using the policy on, lead to a drop in average achieved price. The same tendency results from assuming a correlation that is higher than for the data you are using your policy on.

The value of a correlated and an uncorrelated policy on different simulations

We now compared the difference between using a policy that accounts for correlation and one that don't, for a number of different assumptions of correct correlation. Each time running a simulation using the corresponding η on both policies, respectively. We expect the policy without correlation to have a lower production during periods of high inflow, and vice versa, leading to a lower achieved price compared with the policy accounting for correlation. In figure 4.5 each pair of columns is related to one assumption of correlation. Simulations using the corresponding η is run on an optimized policy for both this η and no correlation.

As expected, the value of using the policies that account for the correlation between inflow and price are higher than the value of using a policy that do not take correlation into consideration on correlated data. This effect increases linearly as η increase. At an η of -0.1 the value of including correlation in the making of the policy is about 2.2 percentage points higher then if not including it. The increase in difference is in accordance with our expectations, since assuming no correlation when generating the policy lead to a higher difference the higher the actual correlation is. We can see that increasing η , implying an increased weighting of the correlation, clearly makes a correlated policy more valuable than an uncorrelated one. Meaning that taking correlation into account when developing a generation policy is more valuable the higher the correlation actually is.

The main finding from this analysis is an observed linear relationship between the achieved price and the mismatch between correlation. The higher the correlation between

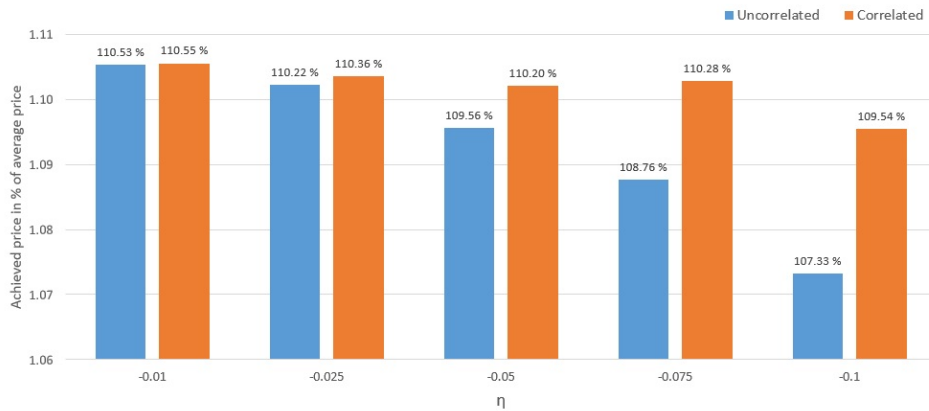


Figure 4.5: Using both correlated and uncorrelated policies on simulations for different η 's

inflow and prices in real life the less the power plant will be able to take advantage of high prices when applying the data on a uncorrelated policy.

Comparing reservoir paths for a correlated and uncorrelated policy

We expect the reservoir level for the uncorrelated policy to be higher than the correlated for high levels of inflow. This can be explained from figure 4.6. Here we observe one correlated (blue) and one uncorrelated (red) production strategy. During high inflow period a correlated strategy recommends an increase in production, compared to an uncorrelated one. This is because the uncorrelated policy operate on lower expected prices during this period compared to the correlated policy. We can observe this effect in practice in figure 4.7. In this figure the average reservoir path is plotted for both an uncorrelated policy, one with η equal to -0.05 and the historical average. In the start of the year the reservoir for uncorrelated producers is lower than what it is for correlated ones. This difference is because the correlated prices are higher here than uncorrelated ones, and when the producers observe higher prices than was expected, they will produce more lowering the reservoir level. From about 150 days onwards the correlated policy produces more than the uncorrelated one, hence have a lower reservoir level. The explanation is that the inflow is high, so the uncorrelated producers see prices that are lower than what they expected, and therefore choose not to produce. At the end of the year inflow is lower, and so by the same logic, an uncorrelated suggest producing more, since the prices are higher than what they expected, which we see in figure 4.6 around day 350.

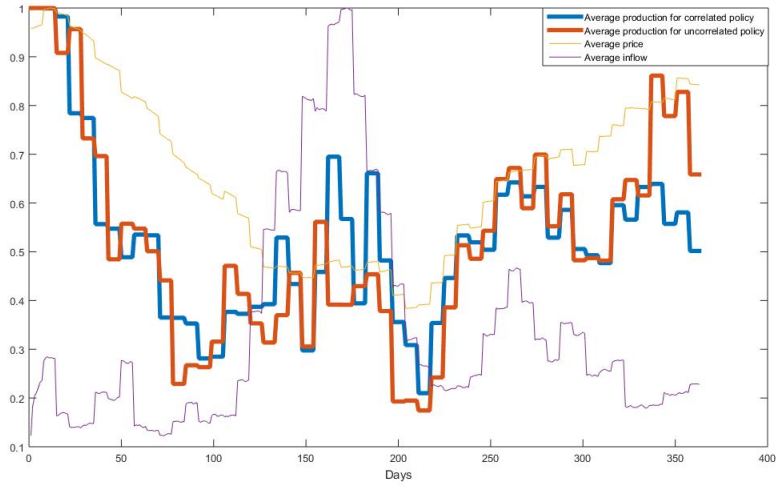


Figure 4.6: Average weekly accumulated production for an uncorrelated and a correlated policy, plotted against average price and average inflow, all in % of maximum observed value

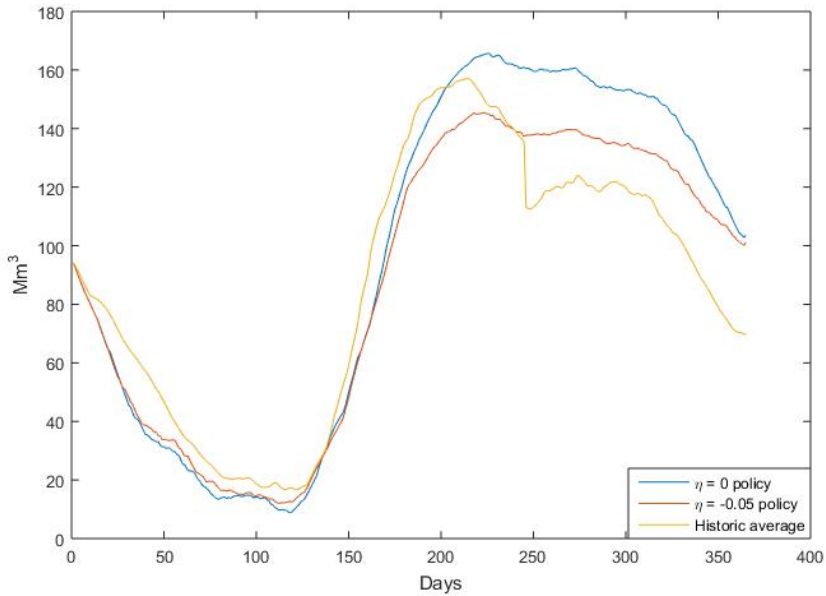


Figure 4.7: Using both correlated and uncorrelated policies on simulations for different η 's

4.3 Shortcomings and improvements for the future

Finding the correct η

From this analysis we observe that having an optimal policy where correlation between price and inflow are taken into consideration when making the optimal policy. Results varies significantly between different assumptions on η . This suggests that in order to make a policy where the value of water are neither overestimated or underestimated, finding the right factor η is important. A good eta will be given by an as realistic weighting of correlation as possible. We see finding the correct η as the element of our analysis to provide the most value for future studies.

Small data sample

We only had eight years of data for the inflow and twelve years of data for the prices. To have a more general representation of the evolution of both inflow and prices over time, we would need more years of data for both time series. Out of these years, only four were overlapping. We can observe how the small amount of data effect our analysis in figure 4.3). In order to be able to use historical data to analyze the effect of correlation between inflow and prices, we need a larger period of overlapping time series. If not our sample space is not large enough to represent the total population.

Alternative way to generate the lattice, and avoid rounding errors

The lattice has opportunities for improvement. Using only 10000 simulations as input can be too low when using 10 price and inflow levels, or nodes, for each stage. With 10 nodes at each stage there will be 100 transition probabilities between two stages, which means that with 10000 simulations there would be, on average, 100 paths going through each transition. This might not be sufficient to explain all the variations possible. Increasing the number of simulations will lead to transition probabilities that is more representative for real values. The problem with increased number of simulation is the increased computation time. The lattice generation is by far the most time consuming part of our model, so we want to avoid making it more demanding. Changing the way we generate the lattice can be a way to cut down on computation time, while also increasing the correctness of the transition probabilities. One way to do this is by keeping the nodes fixed. Then we do not have to calculate these node values, and can simply calculate the transition probabilities between the predefined ones. Computation time would then decline, and we could increase the number of simulations.

Having predefined and fixed price and inflow nodes, opens up for the possibility to let the state values be synchronized with the storage grid, suggested by Nandalal and Bogardi (2007). The two grids would then have the same discretization, and we would no longer need to round S_{t+1} , from restriction 3.3, to the closest grid point in state space S . A drawback of this is that you have to choose values for the nodes initially and keep them throughout the lattice generation, instead of constantly improving the node value, as we do now. Also, Tejada-Guibert and Stedinger (1993) states that synchronization would require a very fine state spaces, which would in turn increase computation time.

For future studies, Markov chain generation could be done by methods like moment matching or conditional sampling (Kaut, 2003). Another possibility is to find node values

based on an averaging principle, like Mo and Kåresen (2001). They first find the maximum and minimum value, for so to divide the space between those extremes into intervals of equal length and the value of each node is given by the average value of all scenario in the respective interval. They later change the size of the intervals to cover a minimum amount of scenarios, while keeping the interval sizes as equal as possible. The transition probabilities are found through a non-linear optimization problem each stage.

Conclusion

This study explains how to develop a release policy for a hydropower producer with one single reservoir, when prices and inflow is correlated and stochastic variables. We then analyze the performance of the policies, when changing the correlation assumed during policy generation.

We are mainly interested in comparing a release policy where the correlation is included, to a policy where it is ignored. This is to see if hydropower producers should take the correlation into account when developing the release policy, to get a better performance from the policy.

We find that using a release policy that accounts for correlation outperforms a policy that does not account for it. The difference in performance seems to increase linearly as the correlation increases.

The results show that a policy with a higher correlation than the data you are using it on, also underperforms compared to a policy with the correct correlation. Our results indicate that the degree of this underperformance can be severe for higher values of assumed correlation. Of this reason it is important to find the correct level of correlation. Assuming either too high, or too low correlation both lead to a performance that is worse than if assuming correct correlation.

This study shows there is improvements to be made in the development of a release policy for hydropower producers, compared to how it is currently done. By including a correct correlation between prices and inflow, the policy will perform better than if ignoring the correlation, leading to higher revenues.

Bibliography

- Bogardi, J.J., B. S. S. D., Nandalal, K., 1988. Effect of state space and inflow discretization on stochastic dynamic programming based reservoir operation rules and system performance. Proceedings, 6th congress, Asian and Pacific Regional Division 1, 429–436.
- Dixit, A. K., Pindyck, R. S., 1994. Investment under Uncertainty. Princeton University Press.
- Doorman, G., 2015. Course ELK15 Hydro Power Scheduling. Department of Electric Power Engineering, NTNU.
- Flatabø, N. e. a., 1988. Efi's models for hydro scheduling. Tech. Rep. 3483, Energiforsyningens Forskningsinstitutt A/S.
- Fosso, O.B., e. a., 1999. Generating scheduling in deregulated system. the norwegian case. IEEE Transactions on Power Systems 14 (1), 75–81.
- IEA, 2015. Renewable energy essentials:hydropower.
URL http://www.iea.org/publications/freepublications/publication/hydropower_essentials.pdf
- Kaut, M., W. S. W., 2003. Evaluation of scenario-generation methods for stochastic programming.
- Kolsrud, C. W., Prokosch, M., 2010. Reservoir hydropower: The value of flexibility. Master thesis TIØ4900, Norwegian University of science and Technology.
- Maidment, D., Chow, V., 1981. Stochastic state variable dynamic programming for reservoir systems analysis. Water Resources Research 17 (6), 1578–1584.
- Miranda, M. J., Fackler, P. L., 2002. Applied Computational Economics and Finance. The MIT Press.
- Mo, B., G. A. G. A., Kåresen, K., 2001. Optimisation of hydropower operation in a liberalised market with focus on price modelling.

Nandalal, K. D. W., Bogardi, J. J., 2007. Dynamic Programming Based Operation of Reservoirs: Applicability and Limits. Cambridge University Press.

NVE, 2015. Norwegian water resources and energy directorate.

URL <http://vanmagasinfylling.nve.no/Default.aspx?ViewType=AllYearsTable&Omr=NO>

Näsäkkälä, E., Keppo, J., 2008. Hydropower with financial information. Applied Mathematical Finance 15 (5-6), 503–529.

Pflug, G. C., Pichler, A., 2015. Dynamic generation of scenario trees.

Su, S. Y., Deininger, R. A., 1972. Generalization of white's method of successive approximations of periodic markovian decision processes. Operations Research 20 (2), 318–326.

Tejada-Guibert, J. A., J. S. A., Stedinger, J. R., 1993. Comparison of two approaches for implementing multireservoir operating policies derived using stochastic dynamic programming. Water Resources Research 29 (12), 3969–3980.

Wolfgang, O., H. A. M. B. G. A. W. I., Doorman, G., 2009. Hydro reservoir handling in norway before and after deregulation. Energy 34, 1642–1651.

Appendix

5.1 Power plant descriptive data

For our power plant we have

a) Inflow	
Average yearly inflow	481,9 $Mm^3/year$
b) Reservoir volume	
For the largest reservoir alone	158,5 Mm^3
For aggregated reservoirs at the power plant	179,3 $Mm^3/year$
c) Degree of regulation (DoR)	
For the largest reservoir alone	$158,5/481,9 = 0,329$
For the aggregated reservoir	$179,3/481,9 = 0,372$
d) Degree of regulation in flooding season (19.04-20.07)	
For the largest reservoir alone	$158,5/261,5 = 0,606$
For the aggregated reservoir	$179,3/261,5 = 0,686$
e) Production capacity	
Yearly production	504 GWh
Maximum daily production	3072 MWh
Energy equivalent	1,2385 kWh/m^3
3072000 $[kWh]/1,2385 [kWh/m^3]$	$= 2480419,9 m^3/day$
f) Capacity factor	
$481,9 [Mm^3/years]/2480419,9[m^3/day]*365[days]$	$= 53,23 \%$

5.2 Power plant capacity calculations, \bar{U}

The given inflow and reservoir level data sets are already converted to the energy equivalent, how much energy stored in each m^3 of water. For our model we have \bar{U} given from given data as a constant weekly capacity, whereas in reality it would have depended on variations in plant head, H , and discharge, Q . For our case the differences between the highest and lowest head is below 10 percent. Because of that effects of the changes in head is negligible for power plant calculations.

- Q : discharge capacity in m^3/s
- η : power plant efficiency
- γ : water density in kg/m^3
- H : plant head in m

Table 5.1: Regression results for inflow and price auto-regressive base processes and the overall reservoir level in Norway

Inflow				
	Estimate	SE	tStat	pValue
Intercept	0.0001	0.0153	0.0082	0.9935
X_{t-1}	0.9282	0.0184	50.4404	0.0000
X_{t-2}	-0.2775	0.0247	-11.2463	0.0000
X_{t-3}	0.1143	0.0184	6.2050	0.0000
R-Squared: 61.62%				

Price				
	Estimate	SE	tStat	pValue
Intercept	0.0155	0.0077	2.0197	0.0442
X_{t-1}	0.9670	0.0141	68.5417	0.0000
R-Squared: 92.83%				

Overall reservoir level				
	Estimate	SE	tStat	pValue
Intercept	0.0225	0.0219	1.0272	0.3187
Δr_{t-1}	0.9343	0.1133	8.2443	0.0000
Δi_t	0.0795	0.0332	2.3935	0.0285
R-Squared: 86.10%				

Average energy equivalent in kWh/m^3

$$e = \frac{1}{3.6 \cdot 10^6} \cdot \gamma \cdot g \cdot H \cdot \eta \quad (5.1)$$

Power plant capacity in MW . The equation is divided by 1000 to get the capacity in MW and multiplied with 3600 for converting from seconds to hours.

$$\bar{U} = e \cdot Q \cdot \frac{3600}{1000} \quad (5.2)$$

5.3 Regression analysis

Price and inflow models

In section 3.4 and section 3.5 we assumed that both inflow and price had a seasonal as well as an auto-regressive part. The aim of this part is to analyse the regression results of these models.

Season function

In figure 5.1 we can see the seasonal components of the models plotted against the historical average, for inflow and price respectively. Even though the R-squared values are low, we can see in the plots that the functions captures the seasonality of the historic average.

Table 5.2: Regression results for inflow and price seasonal dummy functions

	Inflow				Price			
	Estimate	SE	tStat	pValue	Estimate	SE	tStat	pValue
Intercept	0.8482	0.1801	4.7091	0	2.9758	0.0547	54.3699	0
D1	0.0998	0.2476	0.4029	0.687	-0.074	0.0759	-0.9756	0.3293
D2	0.3178	0.2515	1.2636	0.2065	-0.0414	0.0759	-0.5457	0.5853
D3	-0.3888	0.2547	-1.5262	0.1271	-0.0493	0.0772	-0.6388	0.523
D4	-0.5653	0.2547	-2.2189	0.0266	-0.0678	0.0774	-0.8765	0.3808
D5	-0.5395	0.2547	-2.1179	0.0343	-0.0718	0.0774	-0.928	0.3535
D6	-0.1152	0.2547	-0.4521	0.6513	-0.1081	0.0774	-1.3967	0.1626
D7	-0.2018	0.2547	-0.7923	0.4283	-0.1072	0.0774	-1.3852	0.1661
D8	0.2258	0.2547	0.8864	0.3755	-0.1586	0.0774	-2.0484	0.0406
D9	-0.5939	0.2547	-2.3312	0.0198	-0.1518	0.0774	-1.9607	0.05
D10	-0.6528	0.2547	-2.5625	0.0104	-0.1658	0.0774	-2.1427	0.0322
D11	-0.7235	0.2547	-2.84	0.0045	-0.2108	0.0774	-2.7233	0.0065
D12	-0.5179	0.2547	-2.0329	0.0422	-0.2451	0.0774	-3.1663	0.0016
D13	-0.2767	0.2547	-1.0861	0.2775	-0.2536	0.0774	-3.2758	0.0011
D14	-0.5191	0.2547	-2.0377	0.0417	-0.2624	0.0774	-3.3903	0.0007
D15	-0.4102	0.2547	-1.6104	0.1074	-0.2864	0.0774	-3.6995	0.0002
D16	-0.4336	0.2547	-1.702	0.0889	-0.2554	0.0774	-3.2998	0.001
D17	0.036	0.2547	0.1414	0.8876	-0.3063	0.0774	-3.9577	0.0001
D18	0.7563	0.2547	2.9688	0.003	-0.411	0.0774	-5.3095	0
D19	1.5002	0.2547	5.8891	0	-0.4823	0.0774	-6.2306	0
D20	1.9862	0.2547	7.7968	0	-0.472	0.0774	-6.0983	0
D21	1.6553	0.2547	6.4978	0	-0.5	0.0774	-6.4592	0
D22	2.5924	0.2547	10.1767	0	-0.5213	0.0774	-6.7346	0
D23	2.4874	0.2547	9.7644	0	-0.4665	0.0774	-6.027	0
D24	3.1905	0.2547	12.5243	0	-0.4606	0.0774	-5.9501	0
D25	3.2979	0.2547	12.9459	0	-0.4932	0.0774	-6.372	0
D26	2.598	0.2547	10.1987	0	-0.513	0.0774	-6.6272	0
D27	1.9868	0.2547	7.7993	0	-0.4856	0.0774	-6.2733	0
D28	1.6459	0.2547	6.4609	0	-0.5339	0.0774	-6.8971	0
D29	1.0047	0.2547	3.9438	0.0001	-0.6556	0.0774	-8.4697	0
D30	0.4847	0.2547	1.9028	0.0572	-0.7279	0.0774	-9.4037	0
D31	0.1993	0.2547	0.7824	0.434	-0.7163	0.0774	-9.2542	0
D32	-0.0421	0.2547	-0.1652	0.8688	-0.6118	0.0774	-7.904	0
D33	-0.0686	0.2547	-0.2693	0.7877	-0.496	0.0774	-6.4076	0
D34	-0.0431	0.2547	-0.1692	0.8657	-0.382	0.0774	-4.9357	0
D35	0.0933	0.2547	0.3664	0.7141	-0.3971	0.0774	-5.1304	0
D36	0.5315	0.2547	2.0863	0.037	-0.3113	0.0774	-4.0222	0.0001
D37	0.794	0.2547	3.1167	0.0018	-0.2505	0.0774	-3.236	0.0012
D38	1.1558	0.2547	4.5372	0	-0.2282	0.0774	-2.948	0.0032
D39	0.8429	0.2547	3.3089	0.0009	-0.2294	0.0774	-2.9633	0.0031
D40	0.4907	0.2547	1.9264	0.0542	-0.1947	0.0774	-2.5151	0.0119
D41	0.266	0.2547	1.044	0.2966	-0.2068	0.0774	-2.6723	0.0076

D42	0.637	0.2547	2.5006	0.0125	-0.1885	0.0774	-2.4351	0.0149
D43	0.5241	0.2547	2.0575	0.0397	-0.2367	0.0774	-3.0577	0.0022
D44	0.1051	0.2547	0.4127	0.6799	-0.1966	0.0774	-2.5406	0.0111
D45	0.1585	0.2547	0.6223	0.5338	-0.1533	0.0774	-1.981	0.0477
D46	0.2575	0.2547	1.011	0.3121	-0.1209	0.0774	-1.5621	0.1183
D47	-0.2908	0.2547	-1.1414	0.2538	-0.0739	0.0774	-0.955	0.3396
D48	-0.2836	0.2547	-1.1131	0.2657	-0.0728	0.0774	-0.9411	0.3467
D49	-0.2925	0.2547	-1.1482	0.251	-0.0518	0.0774	-0.6687	0.5037
D50	-0.1404	0.2547	-0.551	0.5817	-0.0401	0.0774	-0.5184	0.6042
D51	-0.1149	0.2547	-0.4511	0.652	0.0132	0.0774	0.1707	0.8645

R-squared Inflow: 36.72%, R-squared Price: 13.28%

Our main focus is not on inflow and price modelling, and so we see this as good enough for our purpose. One thing we could improve upon is removing the sudden drops during the flooding season in the inflow model. These are not realistic since we expect there to be a continuous increase during this period, and is probably a case of over-fitting. We can therefore smooth this section out to get a better result.

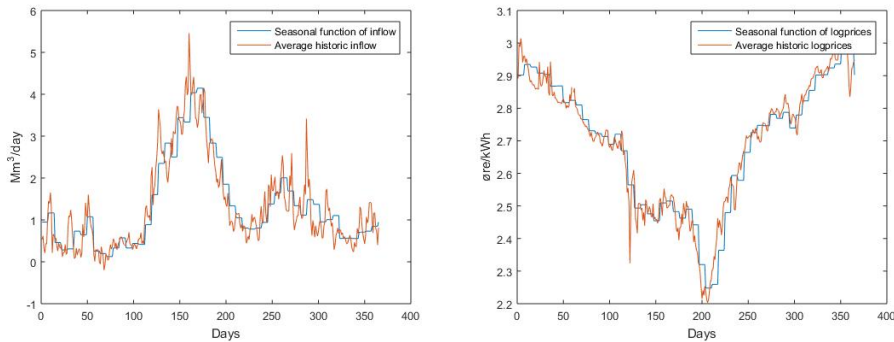


Figure 5.1: Seasonal function of (a) the inflow model and (b) the logprice model plotted against average historical values

Base processes

As mentioned in section 3.4 and 3.5 expect both inflow and price to be serial correlated. In the regression results in table 5.1 we observe from the p-values that for both inflow and price at least one X_{t-p} is significant. The assumption of serial correlation is then correct according to these results. Based on number of significant coefficients we choose an AR(1) model for the price, and an AR(3) model for the inflow. An AR(3) for the inflow makes sense since weather the past couple of days is descriptive of what the weather will be today. In retrospect, using an AR(3) process might not be necessary, since we are only using a Markov-I chain for our lattice. In the Markov-I chain only the values yesterday have an effect on the current value. In the AR(3) process, the value today is explained by the past three days. Transforming the AR(3) to an Markov-I chain would then mean we

lose the dependence on two and three days ago. We could ignore the dependence on two and three days prior, even though they are significant according to the regression results, and only model the inflow base process as an AR(1) process.

Figure 5.2 shows a simulated path against the historical data series for both inflow and price. We observe that the spikes in the historical data are not captured in the model. Spikes can be included through a spike filtering process like similar to what's done by Kolsrud and Prokosch (2010). By searching for spikes and removing them until the variance of the remaining process reaches a lower bound, we would get the distribution of the spike together with the sizes of them. We could use the knowledge of the spike distribution and size to model a separate spike process to include in the overall price process. For now we see our models as sufficient, since they are not the main focus of our study.

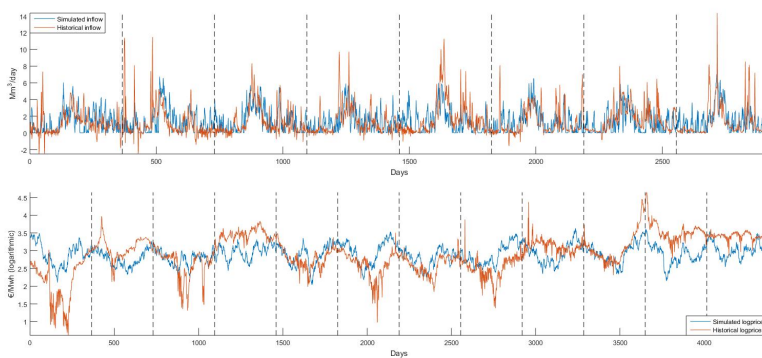


Figure 5.2: Simulated of (a) the inflow model and (b) the logprice model plotted against average historical values

Overall reservoir level

Our expectations of the deviation from overall reservoir level in Norway is that it is heavily serial correlated with the value last week and also correlated with the accumulated inflow in our area to some degree, as is stated in section 3.6. From the regression results in table 5.1 we observe that this assumption is reasonable according to the significance of the coefficients. Given the R-squared value of 86% we deem the model sufficient.

5.4 Lattice analysis

Figure 5.3 show both the inflow and price part of the lattice. We are using a lattice with 10 nodes at each stage through the year. To generate the lattice we have used 10000 simulations as input. The reason for this choice is mainly due to computation time. To analyze our main research questions we need to run the lattice program multiple times, hence it is important to have the computation time as low as possible. On the other hand, it is also important to get a lattice that is a good representation of reality, and having a higher number of input simulations will improve upon that.

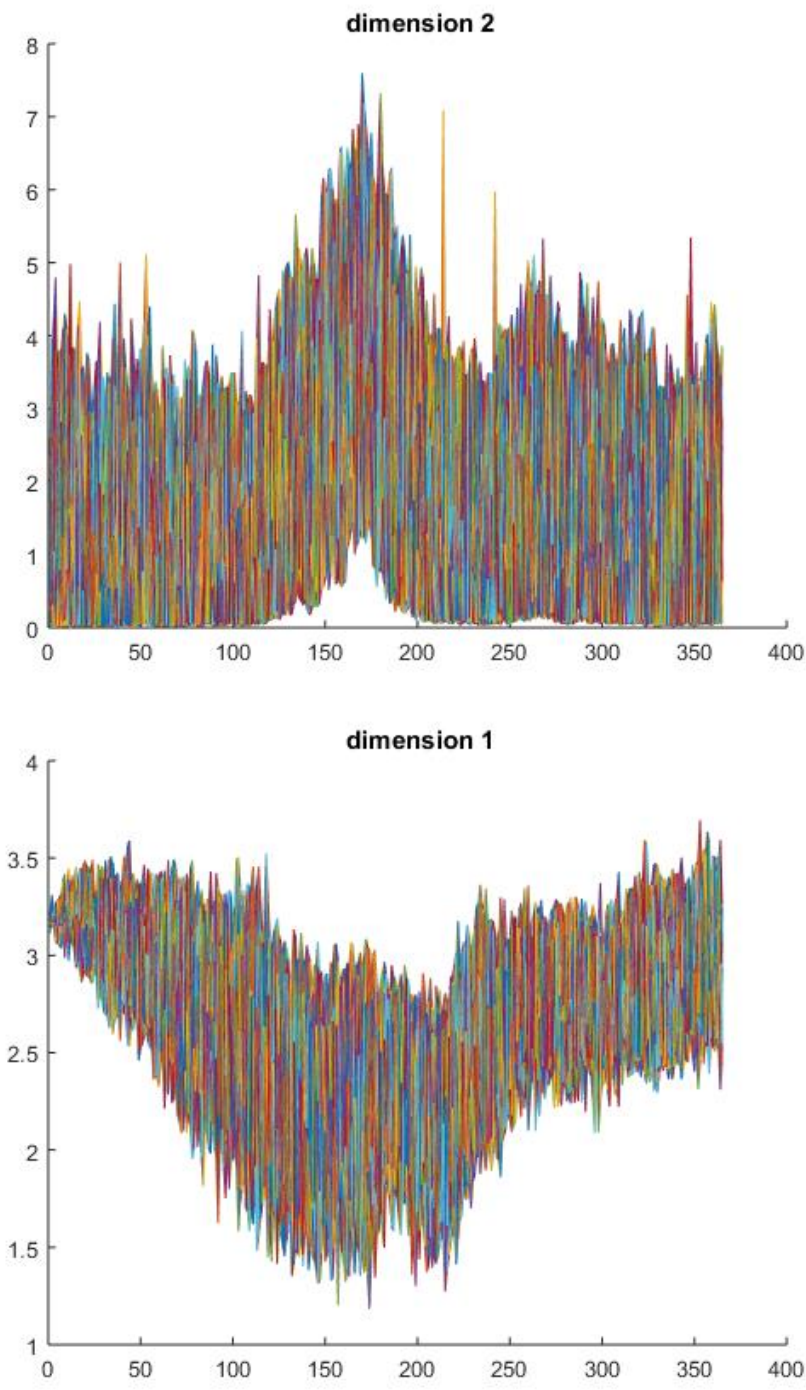


Figure 5.3: Lattice for (a) the inflow and (b) the logprice

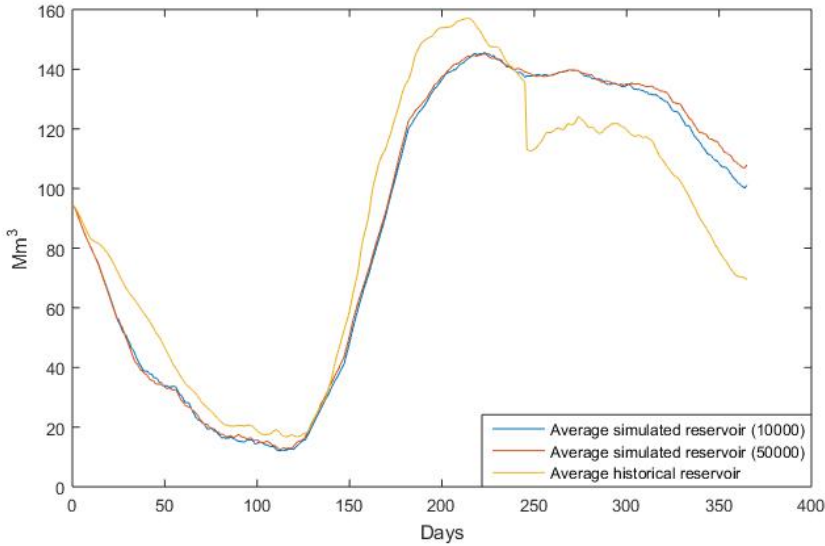


Figure 5.4: Average simulated reservoir path for a lattice with 10000 input simulations, and a lattice with 50000 input simulations, plotted against the historical average

Figure 3.7 shows a reservoir path for a lattice with 10000 simulations and another lattice with 50000 simulations. The two paths align almost the whole year. They only break apart near the end of the year, because the lattice with 50000 simulations has a slightly higher continuation value than the lattice with 10000 simulations. Seeing as that the paths are so similar, and the difference in continuation value is relatively small, we consider our choice of only using 10000 simulations as good enough for our purpose, given the much shorter computation time. We could however expect the continuation value to be even higher when increasing the number of simulations further. The effect of a further increase in simulations is left for future studies.

5.5 Effect of running until steady state

We assume that running til steady state, using the method explained in the convergence criteria check in chapter 3.2, will effect the end of planning horizon reservoir level. We see from 5.5a that the model are favoring emptying the reservoir at the end of the season, instead of saving until next year. We are referring to this as the end of horizon effect. The reason for this is that we have assumed the water has no value at the end of the planning period, which means there is nothing to gain from having water left in the reservoir at this point. The convergence criteria check in the description of the algorithm in chapter 3.2 is utilized in order to obtain a continuation value for the power plant, and thereby accounting for the future value of water. The effect of obtaining a steady state convergence can be seen in the difference between figure 5.5a and 5.5b. Here the end year value represents a

continuation value for running the plant forever.

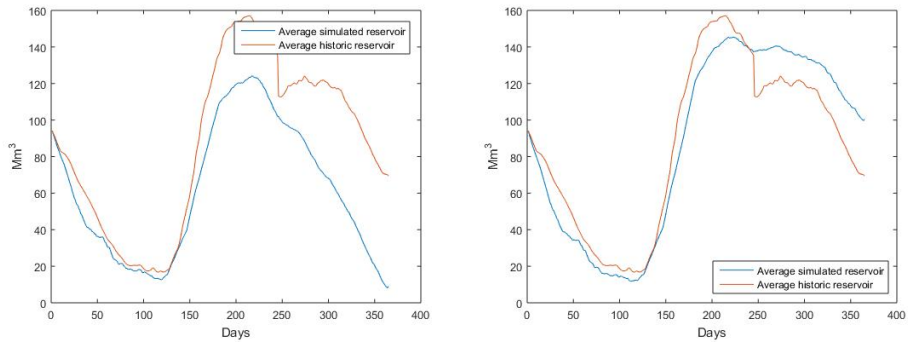


Figure 5.5: Average reservoir path given optimal policy plotted against historical average, for (a) initial and (b) converged continuation value