Analyzing the price- and inflow relationships in hydroelectric scheduling

Preface

This paper is prepared as the result of our Project Thesis at the Norwegian University of Science and Technology (NTNU), Department of Industrial Economics and Technology Management during autumn 2007. The paper is a part of the course TIØ4550 Financial Engineering, Specialization Project.

We would specially like to thank our teaching supervisor, Associate Professor Stein-Erik Fleten for good and helpful support and constructive feedback during the semester. In addition Associate Professor Sjur Westgaard has been helpful and also deserves our gratitude.

Trondheim, December 17th 2007-12-17

Tor Olav Seim

Ole Ronny Thorsnes

Abstract

In this paper we have analyzed the relationship between the system price in the Nordic power market and inflow to three different Norwegian power producers. For hydro power producers the objective of the long term hydro power planning problem is to utilize the available generation resources in an optimal way, maximizing revenues under the relevant constraints. The stochastic variables price and inflow are the main drivers of uncertainty in this optimization problem and their relationship needs to be included in order to obtain correct hydro power schedules. This stochastic optimization problem is simplified with a deterministic equivalent and implemented in the optimization program Xpress.

Two different categories of models have been used to describe the dynamics of the inflow and price series. These include the one and two factor models described in Lucia and Schwartz (2001), and state space models presented in Commandeur & Koopman (2007). The differences between these models are discussed and the estimated models are tested both in and out of sample before being used to generate price and inflow scenarios. The program Scenred is then used to generate a scenario tree from the simulated price and inflow scenarios to give a simplified but adequate representation of the different states the stochastic variables can take in each time period. This scenario tree is then used together with information about the different producers as input in the optimization model described above.

Several aspects of the relationship between price and inflow were analyzed, and weekly correlation was found not to be stable over time. Weekly correlation can therefore not be used to create correlated scenarios and other methods were explored. Among these were the inclusion of inflow as an explanatory variable for price in the state space models, and the introduction of different matching methods for the factor models.

The negative correlation that is found between inflow and price in the Nordic system creates a natural hedge for the individual power producers and is illustrated when calculating the variance of the expected income from a producer when incorporating the correlation and not. The revenues for the different scenarios in the reduced scenario tree are calculated for the correlated and the uncorrelated case. The correlation is found to narrow the set of possible outcomes and reduces the risk in uncertain revenues for the power producers. The degree of correlation between a producer's inflow and the system price is therefore found t be an important risk measure for power producers.

Table of Contents

1.0 Introduction

One of the main challenges faced by hydro power producers is the scheduling task. Many factors affect how this task is conducted at company level. Some of these factors are again highly stochastic and uncertain, and this uncertainty needs to be included in the models used in the scheduling procedure. For producers in the Nordic power system, it is mainly the inflow and electricity spot prices that contribute to this uncertainty. Since the share of electricity production that stems from hydro is totally dominating in Norway, these two uncertain factors are expected to be correlated to each other. To be able to hedge cash flows from hydro power plants, it is necessary to have production models which consider the relationship between production/inflow and price. The water is the main resource for a hydro power producer, and due to its limited storability (reservoirs are rarely dimensioned to store several years of inflow) wet years are usually accompanied with low prices and vice versa. This gives rise to what can be characterized as a natural hedge for Norwegian hydro power producers. A natural hedge can be described as a reduction in risk that arises from a company's normal operating procedure.

This paper develops models for inflow and spot price based on two different approaches and is estimated for three different power stations. We use factor models in accordance with Lucia and Schwartz (2001) and state space methodology described in Commandeur and Koopman (2007). An analysis of the correlation between price and inflow is also conducted, and we suggest a set more qualitative and intuitive approaches to match inflow and price in suitable and realistic ways. Fan scenarios are generated based on the factor and state space models, and these are subsequently reduced to scenario trees. A stochastic optimization model is then finally solved as a deterministic equivalent using linear programming. Much of the material in this paper makes use of the master thesis by Bjørnsgard and Hauge (2007). The main purpose of this paper is to obtain and estimate the parameters for the inflow and spot price based on realized data, and further investigate how the correlation between them can be used in the scheduling task faced by hydro power producers. Further we are attempting to verify the natural hedge characteristics in the Norwegian power market.

This report is divided into eleven chapters. Chapter 2 gives an overall description of the hydro power planning procedure, while chapter 3 describes the underlying data for both inflow and price. In chapter 4 the different models used are presented and the estimation process and results are presented in chapter 5. The next chapter then examines the correlation between price and inflow, while chapter 7 describes how the fan scenarios are simulated. The optimization models are described in chapter 8. The three last chapters provide analysis, conclusion and suggestions for further work.

2.0 Hydro Power Planning

Hydro power players, in this paper represented by generating companies, have a wide span of activities and factors affecting their financial performance. Especially the hydro power planning is a comprehensive and complex task requiring resources, data and competence. This chapter describes some of the activities generating companies must undertake.

2.1 The Nordic power market

Nord Pool ASA is the Nordic power exchange and consists of both a physical and a financial market. Nord Pool's role in the power market is to provide a wholesale marketplace for electricity, where the electricity is traded between generators and users (such as industry) or distribution companies (Nord Pool 2006). The consumer market consists of electricity distributors who sell power to consumers. The differences between spot and consumer prices are due to different distribution models in the Nordic countries. It is the most liquid marketplace for electricity in Europe and accounts for 63% of the total value of the Nordic regions power consumption.

The total volume traded in the financial electricity market (traded and cleared) was 2220TWh in 2006, with a value of EUR 79,2billions. Nord Pool has had a substantial growth in traded volume since its beginning in the early nineties as shown by figure 1.

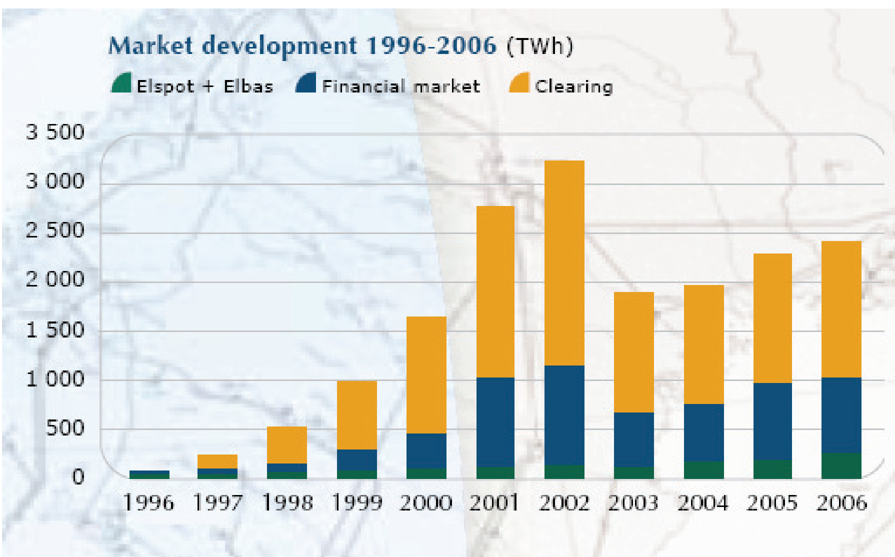


Figure 1. Market development at Nord Pool, 1996-2006

2.1.1 History

From 1971 until 1993 the coordination of power production in Norway was handled by the governmental unit Samkjøringen. The market consisted of vertically integrated units, which handled production and distribution to the individual consumers. In 1991 the Norwegian power market was deregulated as a result of the Energy Act, which introduced free competition in power trading. Statnett Marked was established as a neutral marketplace and changed name later to Nord Pool after the Swedish market opened for free competition and was included. Later also Finland and Denmark was included in this international and neutral electricity exchange which is the first of its kind in the world.

2.1.2 The physical market

The market for physical contracts comprises Elspot and Elbas and is organized by Nord Pool Spot AS. Elspot is an auction based market that trades electrical power contracts for each hour the following day. The physical market forms the basis for all trading in the Nordic power market, and sets the reference price in the financial market. Players at the physical market need to have an agreement with Nord Pool in advance in order to place bids. Bids from the individual producers are prepared and submitted to Nord Pool before 12.00, consisting of tables with the amount of energy wanted bought or sold at different market prices for the coming day. A market cross is calculated from the total demand and sales bids for every hour which then constitute next day's spot prices for the respective hours. Elbas is a physical market where power can be traded up to one hour before consumption/delivery. This market is only available in Finland, Sweden and Denmark.

2.1.3 The financial market

The financial market at Nord Pool is called Eltermin and provides a marketplace where exchange members trade financially settled electricity contracts (futures and forwards). There is no physical delivery in the financial market and contracts can be traded for up to six years. Producers typically use the financial market for risk management and price hedging of production and speculators profit from the volatility in the market and contribute with liquidity and transparency.

2.1.4 Clearing

Nord Pool Clearing offers clearing services where it acts as a contractual counter-party for all the financially settled contracts. The clearing house accepts responsibility for future settlement of the contracts and thereby reduces the financial risk of both buyers and sellers. Nord Pool Clearing undertakes clearing for all standardized contracts traded at Nord Pool in addition to bilateral contracts traded outside the power exchange that is reported for clearing.

2.1.5 Area prices

All the players at the exchange are linked to a certain areas in the Nordic region depending on their geographical location. The different power producers have to report their buy/sales bids in the area where they are connected to the grid. In each area a unique spot price is developed as a reference price for the whole area. The reason why we get different prices in different areas is due to transfer restrictions, which prevents transfer of enough electric power as is demanded from a free market. These bottlenecks in the system create higher prices in deficit areas, and lower prices in production areas. The total system price, which also serves as the reference price in the financial market, is calculated without considering congestions and is the average of the 24 spot prices calculated for the respective day.

2.2 Hydro power scheduling

The objective of power generation scheduling can loosely be defined as “utilizing available generation resources to satisfy the demand for electricity in such a way that the optimal result is obtained and all relevant constraints are satisfied,” Doorman(2007). In a restructured market, such as the Nordic, the optimal result will come from profit maximization, and the constraints typically consisting of generation system constraints (both hydro and thermal), transmission system constraints, environmental constraints and demand characteristics.

2.2.1 Water values

The water streams to the reservoirs at no cost and the variable cost of hydro production is very low (2-5 øre/kWh for Norwegian conditions) but the amount of water available is limited and uncertain, and so the water has an opportunity cost. Production of 1 kWh more in the present period prevents the production of 1kWh in a later period when prices might be higher. The marginal cost of the water, also known as the water value, is therefore dependent on both reservoir volume and production capacity in own system, the demand expectations in future periods, inflow expectations to the reservoir and future spot market prices. Optimal handling of hydropower within each period is achieved by using the water value as the resource cost of hydro power as shown in figure 2 below. The water value can be found using a backward process knowing the water value at the end of the period and optimal production for the given week. The water value is then included in the supply curve to form the correct price cross and the given system price for the respective period. More information about this procedure is given in Doorman (2007).

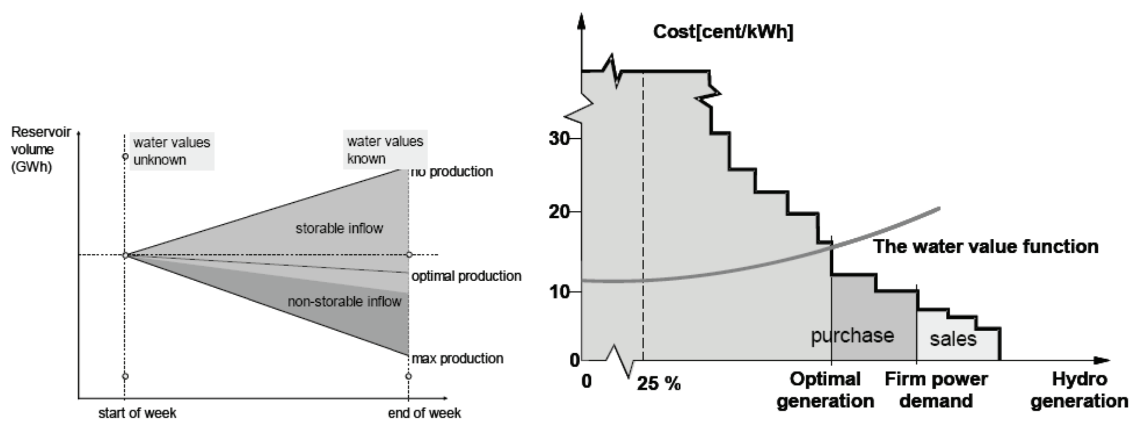


Figure 2. Water Value Method

2.2.2 Scheduling

The Norwegian power system is mainly based on hydro power, but because of strong exchange connections with neighboring countries with extensive thermal production, a hydro-thermal analysis needs to be applied when handling the hydro power scheduling. The scheduling is usually separated into a long, seasonal and short term planning.

The objective in the long term scheduling is to obtain optimal use of resources with a time horizon of up to 5 years. The long term scheduling consist of the strategic management of own resources in interaction with the entire power system. The model description of the whole power system will depend on the model used. A widely used model in Norway is the EMPS[1] model, where the surrounding system is modeled with a market description in a global analysis[2] where among others the price forecast is obtained. This model is usually used by large power producers such as Statkraft. EOPS[3] is a model used typically by smaller producers and consist of a detailed description of the producers own system and models the surroundings through prices. The long term planning models the physical system of reservoirs, power plants topology etc by separating the entire system into geographical areas and aggregating the plants and reservoirs in one area together. This makes the model feasible for computations since a detailed description of the system would lead to unacceptable computational times. An example of such a geographical modeling of the Nordic system in the EMPS model is given below.

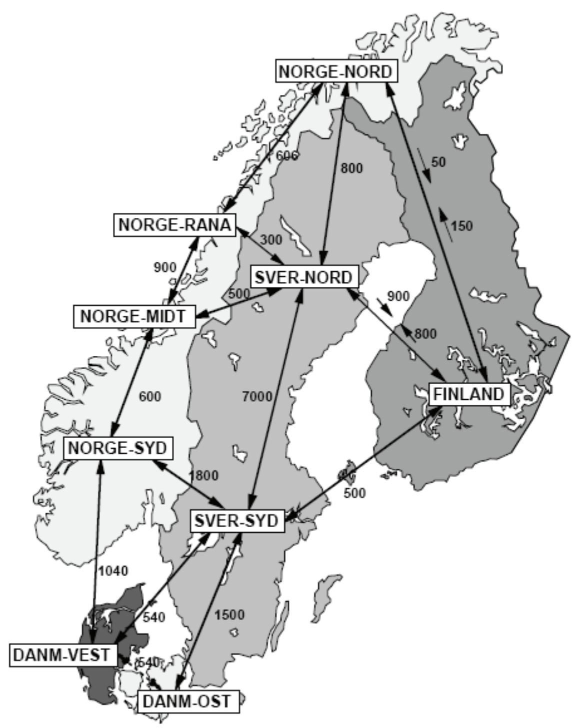


Figure 3. Geographical modeling of the Nordic System in the EMPS model

Seasonal scheduling is used to establish border conditions for the decisions made in the detailed short term planning, and links the long and short term planning together. Historical sets of related time series of price, inflow, temperature, wind etc are used to create scenarios to represent the uncertainty in the seasonal scheduling period.

---

[1] Samkjøringsmodellen, developed by Sintef Energy Research

[2] A global analysis tries to describe the whole system by its physical characteristics (in contrast to a local analysis which describes only a small part by its physical characteristics and the rest by e.g. a fixed price-volume relationship)

[3] Known as Vannsimtap in Norway

The short term planning is used to establish spot market bidding and detailed operational plans, and a detailed description of the physical system is needed for the solution to be feasible. The short term scheduling is represented by one deterministic scenario and so prices and inflow is assumed known.

2.3 The natural hedge in a hydro power portfolio

A hedge is an investment that is performed specifically to reduce or cancel out the risk of another investment. Hedging is thus a strategy designed to minimize exposure to an unwanted business risk, while still allowing the business to profit from an investment activity. It can for example be made to reduce the risk of adverse price movements in an asset, and normally, a hedge consist of taking an offsetting position in a related security, such as a futures contract

A natural hedge is the reduction in risk that arises from a company's normal operations. It is an investment that reduces the undesired risk by matching cash flows i.e. revenues and expenses. A company with significant sales in one country holds a natural hedge to its exposure to currency risk if it also generates expenses in that country, Finance Wise (2007). Another example is a company that opens a subsidiary in another country and borrows money in the local currency to finance the operation. Even though the local interest rate is lower than in the home country; by matching the debt payments to expected revenues in the local currency, the parent company has reduced its foreign currency exposure. A company can thus alter its operational behavior in order to take advantage of a natural hedge, but such hedges are typically less flexible than a financial hedge, Energy Power (2007).

Natural hedges against drought and extreme weather conditions are well known in the Agricultural industry where low yields on corn production are correlated with high prizes, which help to stabilize the farmer's income. This hedge is strongest in the main production areas for corn, since they are closer correlated with the prize. Farmers in peripheral production areas tend to have a higher income risk due to higher corn yield variability and therefore a weaker natural hedge, Harwood (1991)

Hydro power producers are exposed to a substantial risk of inflow and future prices. As in the example from the corn industry, the inflow and prices are negatively correlated in the long run. This result in a natural hedge for the power producers as low inflow will tend to give higher prices, which in turn will stabilize the income for the hydropower producers, Fleten, Wallace, Ziemba (2002). The strong negative correlation we have in Norway is due to the high share of hydro production. This natural hedge will probably diminish as exchange connections with neighboring countries, with significant thermal production, increase.

2.4 Inflow and price

Market players with hydro power production and/or end-user sales need to include the correlations between inflow, end-user sales and prices into their risk management system, Mo, Gjelsvik, Grundt and Kåresen (2001).

In systems with a large share of hydro power, such as the Nordic, the inflow variations are one of the main drivers of uncertainty. The expected annual hydro generation is 119TWh in a normal year, but actual generation may vary between 95TWh and 140TWh, depending on precipitation to the water reservoirs. Figure 4 displays these great variations in inflow, and at the same time shows how the demand are typically spread over the year. Theory suggests that prices and inflow are negatively correlated suggesting that in a year with above average inflow to the reservoirs, the price of electricity should be lower than the average price level.

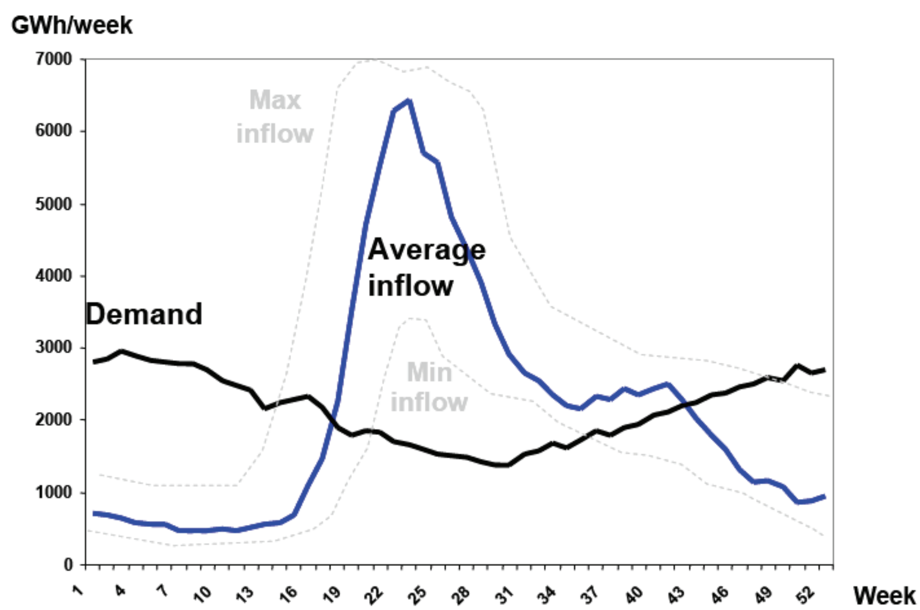


Figure 4. Yearly inflow variation over the year

The figure shows how the inflow can vary over the year in the Norwegian system. Based on these variations, generation in Norway can vary between 95 TWh and 140 TWh depending on the inflow. It can also be seen how the demand is distributed over the year

The demand for electricity and the inflow peaks at different times of the year and so water needs to be stored in reservoirs to be able to supply the consumers throughout the year.

Intuitively the correlation between price and inflow is stronger on an aggregated national level, than between the inflow of a certain power plant and the system price. This because it is the total inflow of water to the power plants that decide if we will have an electric energy surplus or if we need to import energy, use rationing or other energy sources with a higher marginal cost, resulting on average in higher prices. This picture is of course a bit simple since we have reservoirs in the Nordic system that can store water over several years from wet periods to dry. Other factors such as temperature and climate also affect the total demand for energy and further complicate the relationship. Some correlation does however exist between temperature and precipitation. Wet winters are typically warmer than normal, and vice versa, Mo, Gjelsvik and Grundt (2001). Thus during cold winters with little precipitation, demand for electricity is high and further stimulate higher prices.

Theory also implies that a high degree of precipitation one day does not affect prices as much if it is followed by several days of drought. Low inflows to hydro reservoirs are positively correlated with high market prices and low temperatures (resulting in higher end-user sales), Mo, Gjelsvik, Grundt and Kåresen (2001). For hydro producers the correlations between accumulated inflow for the whole season or year and market prices are much more important than the correlations between weekly inflow and market prices because of storage capacity. It is the accumulated precipitation over a period of time that affects prices, and for example less than expected inflow over a long period of time will typically result in higher electricity prices, all else held constant. A presentation made by Kjersti Aas at a conference for Modeling and Measuring Energy Risk on the Nord Pool suggests using a 26 week aggregation of inflow data when modeling the relationship between inflow and price. This is tested against simple weekly correlations for the three selected power plants in this paper which are described in the next chapter. Direct correlations between inflow at local power plants and system prices can however be justified noticing that high inflow at individual power plants is probably strongly correlated with high inflow at an aggregated national level, and so indirectly correlated with the system price.

In the Nordic system especially market prices and inflow are strongly correlated because of the dominating position of hydro power, Doorman (2007). The figure below is taken from Doorman (2007), showing weekly average values of the Nord Pool spot price (left axis) together with the negative of the deviation from normal reservoir in Norway in percentage points (right axis).

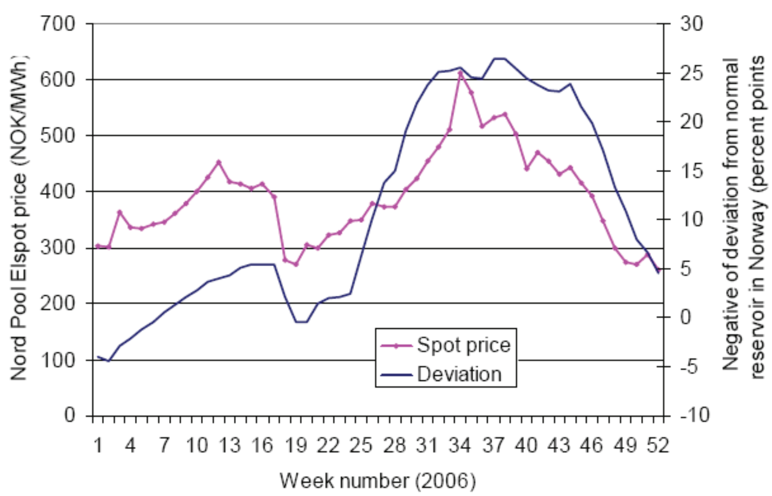


Figure 5. Weekly average values of Nord Pool spot prices and negative deviation from normal reservoir in Norway

The development of fuel prices will also affect the system price to an increasing extent as the European market further consolidates.

3.0 Description of data

3.1 Inflow data

Three power plants will be examined in this paper, and these were chosen based on the characteristics of their inflow time series, quality of input data and geographical location.

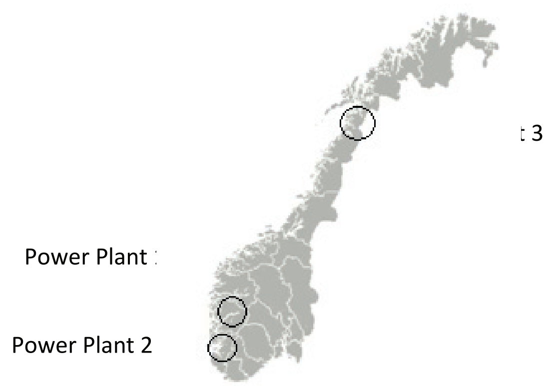


Figure 6. Location of the power plants

The map shows where the power plants are located. The plants are chosen so that they represent different geographical areas in Norway, in addition to that they have different characteristics.

Information about the plants will be given below together with descriptive statistics of their individual inflow series.

3.1.1 Power plant 1

The inflow series of power plant 1 consists of weekly observations in the period 1990(1) – 2006(52). It shows a predictable seasonal pattern with a mean annual inflow of 99,3GWh. The abnormally high observation in week 34 in 2002 is due to the correction of an observation error and will be addressed later when modeling the inflow. The plant has a reservoir size of 177,4Mm^3, a maximum production capacity of 28 MW and an average energy equivalent of 0,67kWh/m^3. The plant has a utilization factor[4] of 49% and a degree of regulation[5] of 1.22, meaning that the reservoir can store more water than the average inflow. This makes this plant relatively flexible in terms of power generation. Figure 7 displays the historic inflow characteristics for this power plant.

[4] Defined as the time it takes to empty a full reservoir when the generator is running at full power (M_{max}/P_{max}) measured in % of hours per year.

[5] Defined as the relationship between the full reservoir and the average annual inflow ($M_{max}/Q_{average}$)

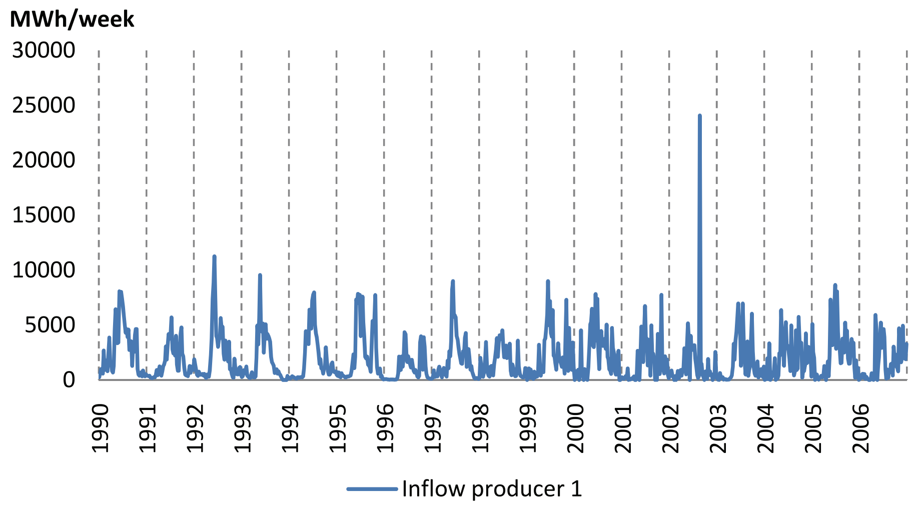


Figure 7. Inflow data producer 1

The figure shows the historic inflow for power producer 1 on a weekly resolution over the period from 1990 to 2006

3.1.2 Power plant 2

The inflow series of power plant 2 consists of weekly observations in the period 1990(1) – 2006(52). The time series show a very seasonally dependent inflow, and has a mean annual inflow of 275,3GWh. The plant has a reservoir size of 204Mm^3, a maximum production capacity of 68MW and an average energy equivalent of 1,25kWh/m^3. The plant has a utilization factor of 59% and a degree of regulation of 1.67, which is the highest among the plants considered in this paper. Figure 8 displays the historic inflow data for this power plant.

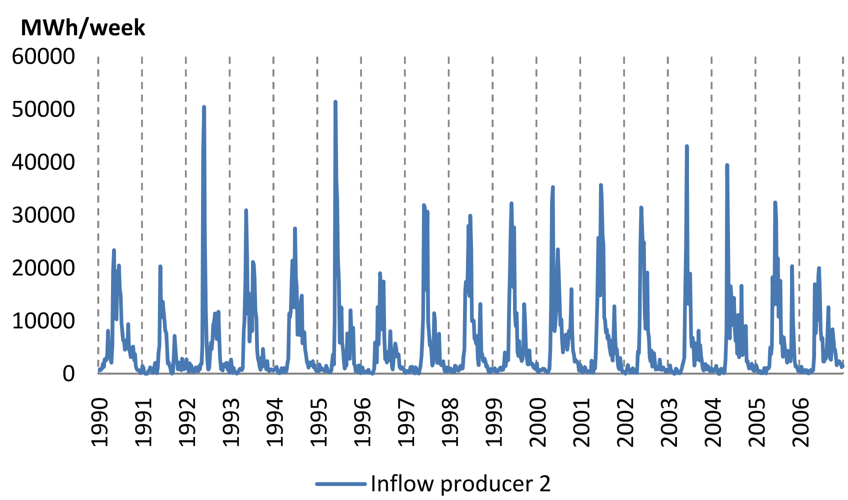


Figure 8. Inflow data producer 2

The figure shows the historic inflow for power producer 2 on a weekly resolution over the period from 1990 to 2006

3.1.3 Power plant 3

The inflow series of power plant 3 consists of weekly observations in the period 2000-2006. The time series has a low degree of seasonal dependence but the highest average annual inflow of 1 247,3GWh. The plant has a reservoir size of 869,4Mm^3, a maximum production capacity of 210MW and an average energy equivalent of 1,46kWh/m^3. The plant has a utilization factor of 47% and a degree of regulation of 0.7. Figure 9 displays the historic inflow data for this power plant.

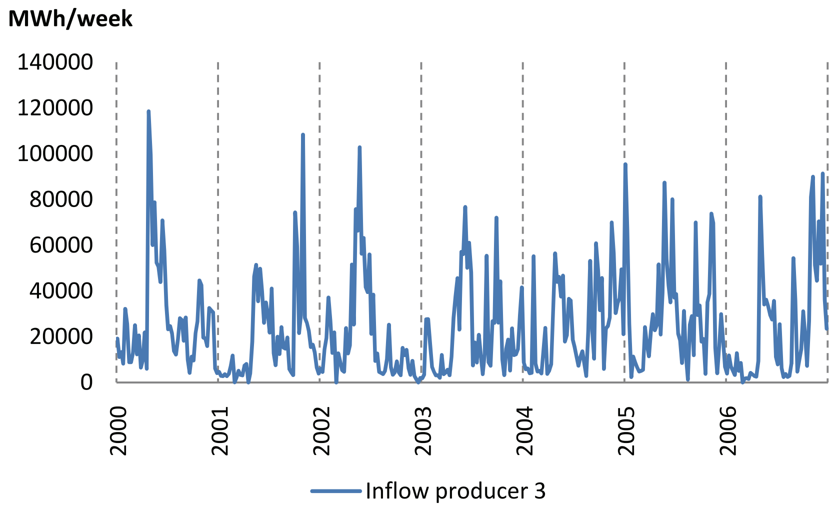


Figure 9. Inflow data producer 3

The figure shows the historic inflow for power producer 3 on a weekly resolution over the period from 1990 to 2006

The first two producers seem to have a more pronounced seasonal pattern than the last, but the total annual inflow to the reservoirs is however stable over time and does not seem to indicate a trend in the amount of inflow (appendix A). This would then have to be explained by changing climate factors or physical change in topology, where a river is lead into the reservoir. The latter was not the case in any of the reservoirs and would not lead to a trend but rather a jump in the inflow series.

3.1.4 Summary of power plant characteristics

Table 1 summaries the characteristics of the power plants dealt with in this paper.

Table 1. Summary of plant characteristics

The table summarizes some of the important characteristics for the three power plants

Power producer	Utilization factor [%]	Degree of regulation	Reservoir size [Mm^3]	Production capacity [MW]	Seasonality in inflow
1	49%	1,22	177,4	28	High
2	59%	1,67	204,0	68	Very high
3	47%	0,70	869,4	210	Low

3.2 Electricity price data

The time series of system prices consist of weekly observations from the period 1993(1) – 2006(52), obtained from Nord Pool's FTP server. System prices are used instead of local prices since these are used in the financial market and to simplify the problem.

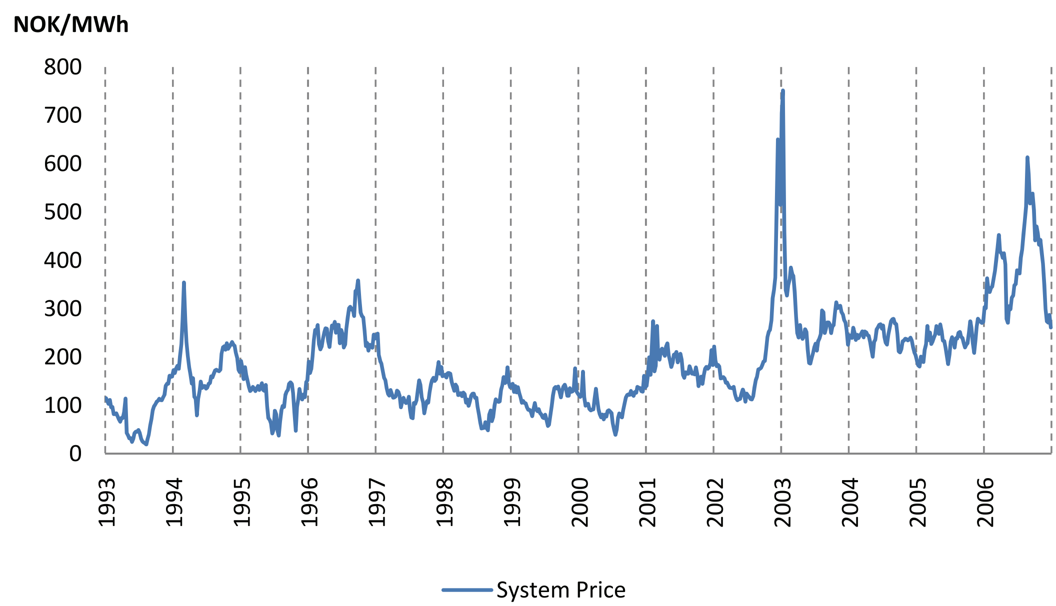


Figure 10 Electricity price data, 1993-2006

The time series show signs of a seasonal trend with relatively lower prices during the filling season in summer compared to the depletion season in winter. The trace also indicates an increasing trend. Certain extreme years can be noticed in the period around 2002/2003 and in 2006, which both were considered as very dry periods with little precipitation.

3.2.1 Electricity price characteristics

Electricity prices have certain characteristics that are different from general commodities prices in the financial market. Existing models used for price modeling and prediction of future prices in the stock market have therefore shown to be inadequate to be directly used to model electricity spot prices (see e.g. Knittel and Roberts (2005) and Seppi (2002)).

Electricity cannot be stored (however, water/coal/gas used for electricity production can be stored) and is used once it is produced. It is not possible for a player in the electric market to buy electricity, store it and then sell it at a later time when prices are higher. For storable commodities the convenience yield represents the benefit of holding the commodity rather than the right to buy it in the future. Since electricity cannot be stored this definition of convenience yield is not sufficient. According to Bøckman and Fleten et.al. (2006), the convenience yield for non-storable commodities can be interpreted as the relative benefit of delivery of the commodity earlier rather than later. Further, the implementation of convenience yield in electricity spot price modeling makes it possible to explain and empirically describe another property of electricity prices, namely random backwardiation and contango. Backwardiation is characterized by a downward sloping forward curve, meaning that the market see early ownership as beneficial. Contango is the opposite situation[6]. The fact that electricity cannot be stored also means that electricity delivered at different points in time can be seen as different products (Lucia and Schwartz (2001)).

Prices are controlled by supply and demand and show seasonal, weekly and daily variations. In the Nordic market prices are usually higher during winter since a lot of the precipitation comes as snow, and the demand for electricity for heating and lighting is higher. There are also weekly and daily variations in the electricity price that can be predicted but this will not be discussed in this paper since we use weekly observations of the electricity spot price.

The spot price time series show a kurtosis of 3,59 (appendix B), indicating a higher probability of incurring extreme prices than in a normal distribution with the same variance. In addition it had a positive skewness of 1,42 indicating that positive extreme values occur more often than negative extreme values. These characteristics indicate existence of unsymmetrical fat tails in the Nordic electricity spot prices.

Electricity prices are very volatile and according to Fleten, Tomasgard and Wallace (2001) this is because electricity have to be used once it is produced and since demand is not particularly price elastic in the short term. Price volatility is stochastic and seasonally dependent (Knittel and Roberts, (2005)) as can be seen from the plot of the Nord Pool system price (figure 10), which display certain extreme periods with high prices and a relatively lower prize during summer. This can also be seen when the time series of spot prices is separated into cold (week 41-52 and 1-18) and warm seasons (week 19-40). The mean spot price is higher during winter but the standard deviation is higher during the warm season. These results can also be found in appendix B. All this points in the direction of heteroscedasticity in the electricity prices in the Nordic market.

Electricity prices also show a mean reversion effect in contrast to stock prices. This means that even if we get shocks in the electricity price, for instance supply or demand shocks, the price tends to revert back to a long run equilibrium level. This and the previous characteristics of the electricity prices will be considered when the electricity spot price is modeled later in this paper.

[6] Consult e.g. Seppi (2002) for a more thorough description of empirical properties of commodity prices

4.0 Models

This paper uses two classes of models when modeling future inflow and price series, namely stochastic factor models and state space models (STAMP).

4.1 Stochastic factor models

The stochastic models used in this paper are mainly based on the ones described in Lucia and Schwartz (2001).

4.1.1 Deterministic component

All the one and two factor models used in this paper to explain the stochastic processes for price and inflow contain a deterministic part. This part attempts to explain predictable components of the time series such as level and seasonality. The level is modeled as a constant and the seasonality is captured in a sinusoidal function. The deterministic function takes the following form:

$$f(t) = \alpha + \gamma \cos\left((t+\tau)\frac{2\pi}{52} \right) \tag{1}$$

Where t is measured in weeks, and hence the cosine function tries to capture annual seasonality. The parameters α, γ and τ needs to be estimated.

4.1.2 One factor model for inflow

The inflow represented by A_t is modeled as the sum of two components. The first being the predictable deterministic function presented in equation 1 and the second a mean-reverting stochastic process.

$$A_t = f(t) + X_t \tag{2}$$

The stochastic term, X_t follow the stochastic process given by:

$$\mathrm{dX_t} = -\kappa \mathrm{X_t dt} + \sigma \mathrm{dZ} \tag{3}$$

where κ>0, $X(0)=x_0$ and dZ represents an increment to a standard Brownian motion Z_t. X_t is said to follow a stationary mean-reverting process, or an Ornstein-Uhlenbeck process with a zero long run mean and a speed of adjustment of K, as explained in McDonald (2006). Equation 2 and 3 can be rewritten as:

$$d(A_t - f(t)) = \kappa(f(t) - A_t)dt + \sigma dZ \tag{4}$$

Showing the mean reverting nature of the process. As A_t deviates from the deterministic part, f(t), it is pulled back at a rate that is proportional to the deviation, and the speed of reversion is given by the mean reverting factor κ.

The distribution of A_t conditional on X_0 is normal with mean and variance equal to (using $X_0 = A_0 - f(0)$):

$$E_0(A_t) = E(P_t / X_0) = f(t) + (P_0 - f(0))e^{-\kappa t} \tag{5}$$

$$Var_0(P_t) = Var(P_t / X_0) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}), \quad \kappa > 0 \tag{6}$$

4.1.3 Two factor model based on the spot price

A two factor model based on Lucia and Schwartz (2001) is also used to describe the price behavior in which the one factor model is expanded with an additional stochastic term. The stochastic price behavior of the spot price is modeled with one short-term mean reverting component and one long-term equilibrium price level component in the equation below:

$$P_t = f(t) + X_t + \varepsilon_t \tag{7}$$

Where

$$dX_t = -\kappa X_t dt + \sigma_x dZ_x \tag{8}$$

$$d\varepsilon_t = \mu_\varepsilon dt + \sigma_\varepsilon dZ_\varepsilon \tag{9}$$

$$dZ_x dZ_\varepsilon = \rho dt \tag{10}$$

The stochastic term X_t is the short run component which follows a mean reverting Ornstein-Uhlenbeck process, and ε_t is the long term equilibrium and follows an arithmetic Brownian motion. The two stochastic processes (dZ_x and dZ_ε) are correlated through equation 10.

In order to use the model for security valuation purposes, we need to use the risk-adjusted process for the stochastic terms in the two factor model. The corresponding risk-adjusted processes are given by:

$$dX_t = \kappa(\alpha^* - X_t)dt + \sigma_x dZ_x^* \tag{11}$$

$$d\varepsilon_t = \mu_\varepsilon^* dt + \sigma_\varepsilon dZ_\varepsilon^* \tag{12}$$

Where: $\alpha^* = -\frac{\lambda_X \sigma_X}{\kappa}$ and $\mu_\varepsilon^* = \mu_\varepsilon - \lambda_\varepsilon \sigma_\varepsilon$

λ_x and λ_ε are the market price of risk for each state variable and assumed to be constant, and dZ* is an increment to a standard Brownian motion, Z_t^*, under the risk-neutral probability measure. λ_x (λ_ε) is the market price per unit risk linked to the state variable X_t (ε_t)and assumed to be constant[7].

It can then be shown that the futures prices are given by:

$$F_0(P_0,T) = E_0^*(P_T) = f(T) + e^{-\kappa T}X_0 + \varepsilon_0 + (1-e^{-\kappa T})\alpha^* + \mu_\varepsilon^* T \tag{13}$$

4.2 STAMP model for inflow and price

In this paper we will also model inflow and price time series using state space methodology. The state space methodology is described in Commandeur and Koopman (2007). The following general model has been used in this paper as a starting point for the analysis:

$$\begin{aligned}
&y_t = \mu_t + \gamma_{1,t} + \beta_t x_t + \lambda_t w_t + \varepsilon_t && \varepsilon_t \sim NID(0,\sigma_\varepsilon^2)\\
&\mu_{t+1} = \mu_t + v_t + \xi_t && \xi_t \sim NID(0,\sigma_\xi^2)\\
&v_{t+1} = v_t + \varsigma_t && \varsigma_t \sim NID(0,\sigma_\varsigma^2)\\
&\beta_{t+1} = \beta_t + \tau_t && \tau_t \sim NID(0,\sigma_\tau^2)\\
&\lambda_{t+1} = \lambda_t + \rho_t && \rho_t \sim NID(0,\sigma_\rho^2)\\
&\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} - \ldots\ldots - \gamma_{51,t} + \omega_t && \omega_t \sim NID(0,\sigma_\omega^2)\\
&\gamma_{2,t+1} = \gamma_{1,t}\\
&\gamma_{3,t+1} = \gamma_{2,t}\\
&\quad\ldots\\
&\gamma_{51,t+1} = \gamma_{50,t}
\end{aligned} \tag{14}$$

Y_t is the dependent variable (price and inflow) that we are modeling. This model have stochastic level (μ), trend (v), explanatory variables(x), intervention variables(w) and seasonal parameters (γ), all of which are normally and independently distributed with zero mean and the respective variances given above. The models are estimated in the software OxMetrics using iterative procedures based on Kalman filtering. The model will be individually adapted to each time series and several models will be explored and tested against each other based on in-sample and out-of-sample tests. In order to capture the correlation between inflow and price, the inflow is used as an explanatory variable describing the price. This way the correlation is directly modeled when estimating the parameters in the model. Both inflow and a 26 week aggregation of inflow is used in the model and compared out of sample to see if any contain superior explanatory power.

When creating scenario trees from these models a time series for inflow is first simulated based on the inflow model. Then a price series is simulated using the inflow series as an explanatory variable. This way we introduce stochasticity in both series and the correlation is introduced through the

[7] See e.g. Hull (2006) chapter 25 for a more thorough description of market price of risk

coefficient of the explanatory variable. With a negative coefficient a simulated “high” inflow series will therefore tend to create a “low” price series since the high value of inflow will reduce the price series by an amount related to the coefficient above. This is further elaborated in the simulation and estimation chapter that follows.

5.0 Estimation

The Nordic market is, as previously mentioned, one of the oldest deregulated power exchanges in the world and thus contains a long and well documented time series of spot prices and other derivatives prices which is necessary for estimating the parameters in the price models. Further, the power producers have provided us with historical inflow data from 1990 to 2006 (2000 to 2006 for power producer 3).

5.1 One factor inflow models

For the purpose of modeling inflow, the one factor model described earlier is used. The parameters in the model are estimated using a numerical nonlinear least squares procedure in Microsoft Excel. Based on Doorman (2007) the Norwegian climate has three periods with different characterizations with regards to inflow.

- Period 1 can be defined from week 1 to week 18. When entering this period the reservoir level has reached its maximum and further precipitation will come mainly as snow and not as direct inflow to the reservoirs. Still this precipitation will be saved as accumulated snow and give rise to inflow during the spring inflow.
- Period 2 can be defined from week 19 to week 40. In this period all precipitation will come as direct inflow to the reservoirs, and at the same time the snow accumulated in period 1 will dissolve.
- Period 3 is defined from week 41 to 52. This is statistically the most “rainful” season of the year.

Based on this we are testing if dividing the year in three sub periods leads to a better description of the realized inflow, compared to using just one period. The models are hereinafter referred to as the Divided Year Model and the Full Year Model respectively. This is done by using data from year 1990 to 2003 (2000-2004 for power station 3) to estimate the parameters, and then using years 2004 to 2006 (2005 to 2006 for power station 3) to run out of sample tests to measure the accuracy of the two models.

Figure 11, 12 and 13 shows out of sample plots for the three power stations investigated, including realized inflow and the forecast based on the estimated parameters. The forecasted estimate here contains only the deterministic term in the one-factor model, since the stochastic term has an expected value of zero.

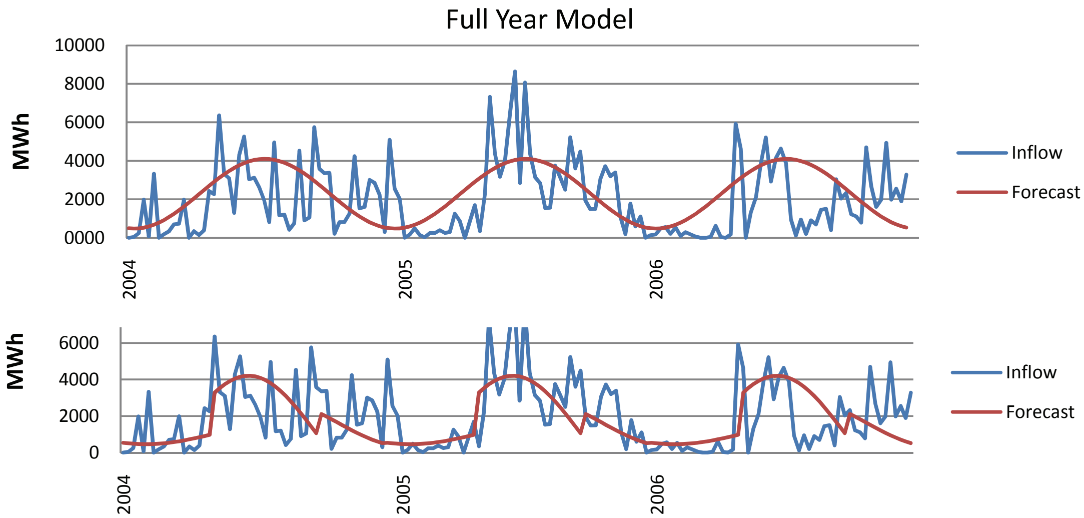


Figure 11. Inflow power plant 1 2004-2006

Upper graph shows the realized inflow compared to the Full Year Model, while the lower shows The Divided Year Model

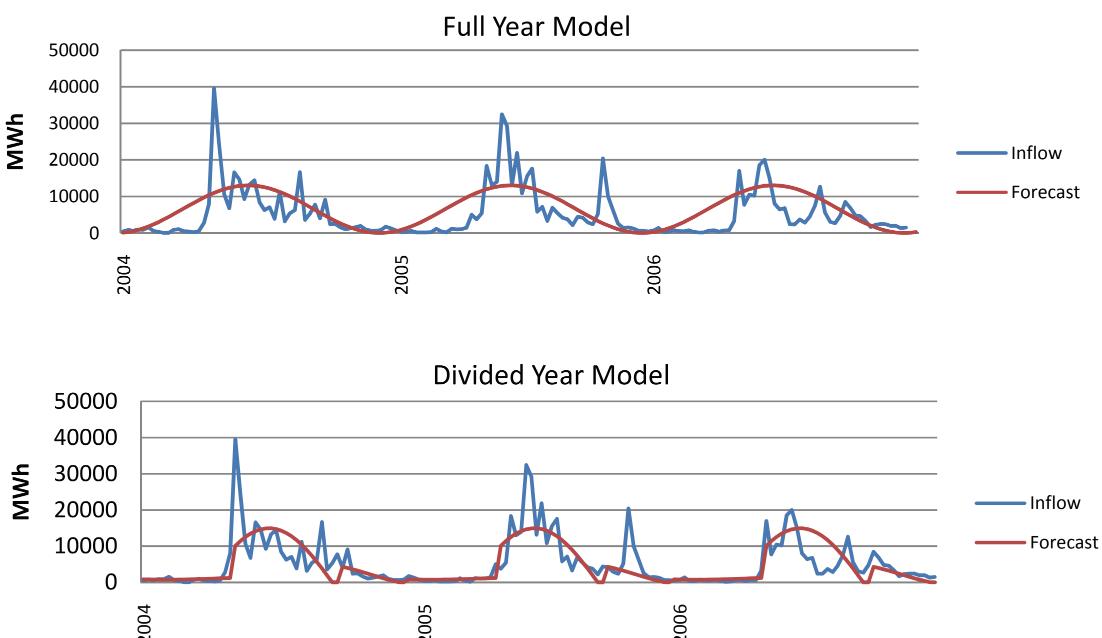


Figure 12. Inflow power plant 2 2004-2006

Upper graph shows the realized inflow compared to the Full Year Model, while the lower shows The Divided Year Model

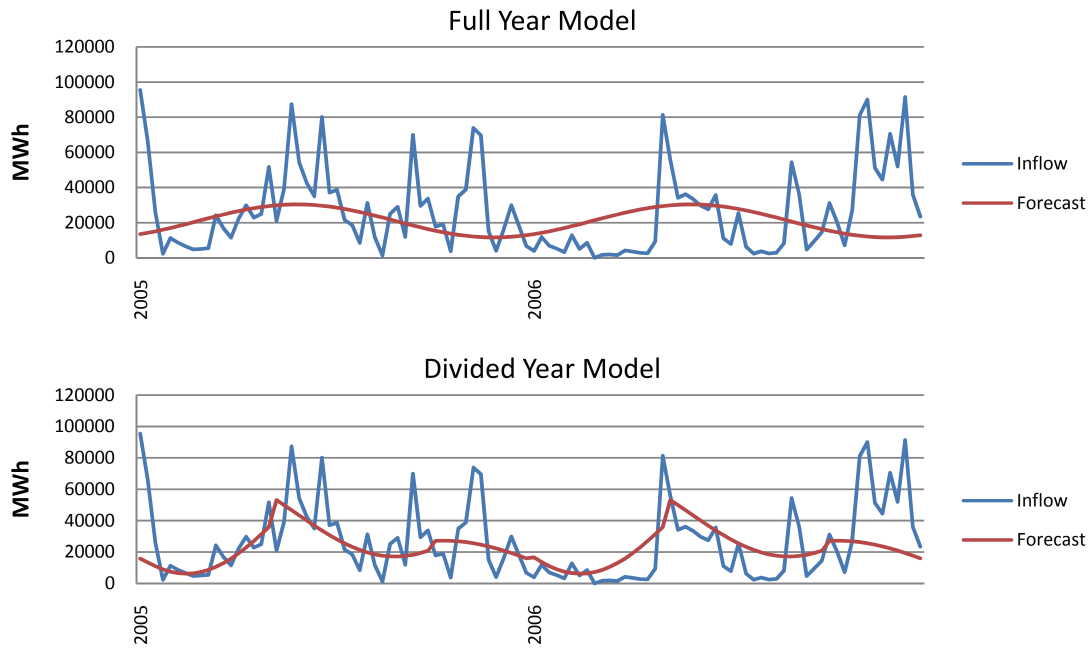


Figure 13. Inflow power plant 3 2005-2006
Upper graph shows the realized inflow compared to the Full Year Model, while the lower shows the Divided Year Model

Intuitively it seems like dividing the year in three yields a better result although the deterministic term is not able to capture the jumps and the following mean reversions. Table NUMBER summarizes common used measures of accuracy for the three power stations[8].

Table 2. Error measures for out of sample tests for inflow models
The table summarizes the results from out of sample tests on the Full Year Model and Divided Year Model for the three power stations

	Mean error	Mean absolute error	Mean squared error	Mean percentage error	Mean absolute percentage error	U	Root mean squared error
Power station 1 Full Year Model	-304,9	1431,8	3186189	-268,9 %	301,0 %	1,52	1785,0
Power station 1 Divided Year Model	129,5	1259,2	2853389	-139,9 %	182,3 %	1,01	1689,2
Power station 2 Full Year Model	-1230,2	3969,4	31068978	-551,0 %	591,1 %	7,64	5574,0
Power station 2 Divided Year Model	409,3	2760,2	23893106	-86,9 %	132,0 %	1,02	4888,1
Power station 3 Full Year Model	5395,4	18929,4	687125575	-151,5 %	196,9 %	1,33	26213,1
Power station 3 Divided Year Model	3887,4	15498,4	500579767	-103,9 %	140,6 %	0,96	22373,7

The error measures used are the ones commonly used when comparing different models to realized data. It could have been an idea to also use error measures penalizing on either model over

[8] See e.g. Cutbertson and Nitzsche (2002) or Brooks (2002) for description on the different error measures used

prediction or under prediction. However, looking at the charts in figure11-13 reveals that none of the models seem to over predict/under predict to a significant extent compared to the others. Hence, error measures penalizing on either model over prediction or under prediction are omitted when evaluating the models. The results from table 2 verify that the Divided Year Model gives less deviation between realized inflow and forecasted inflow in years 2005-2006 for all three power stations. Based on these findings it is chosen to use the Divided Year Model in the further analysis. The estimated parameters for the inflow to the different power station based on the chosen model are given in appendix C.

5.2 One factor aggregate inflow model

Based on findings by Kjersti Aas (as previously mentioned), it is possible that the use of aggregated inflow over 26 weeks better can explain the correlation between inflow and price. Figure 14 shows the 26 weeks aggregated inflow for the three power plants studied.

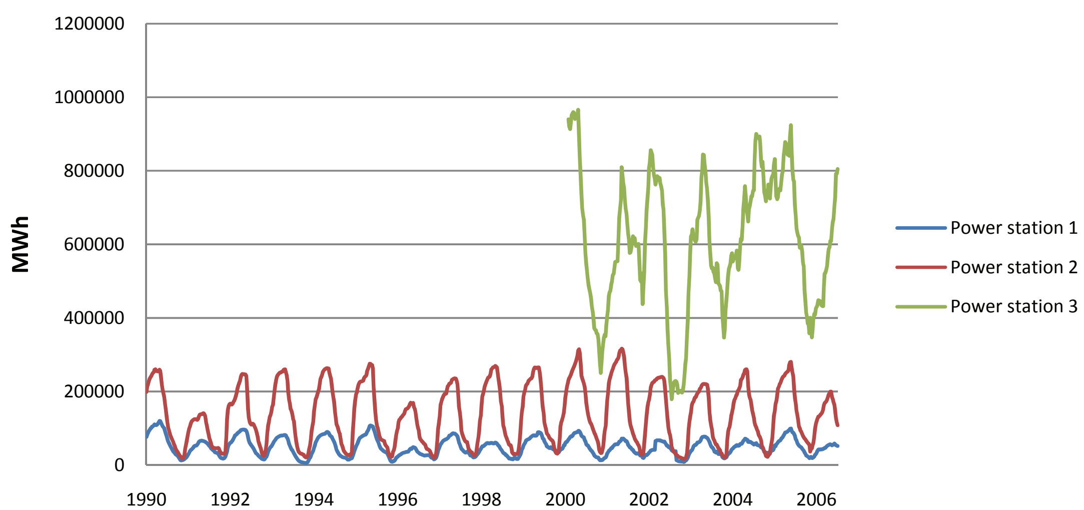


Figure 14. Aggregated inflow 1990-2006

The figure shows the 26 weeks aggregated inflow for the three power stations over the period from 1990 to 2006

It can be seen from the graph that the need to divide the year in three sub periods is not evidently present when working with aggregated inflow (especially for power station 1 and 2), compared to the situation when working with ordinary inflow. Based on this, parameters for the one factor model for aggregated inflow are estimated without dividing the year in three. The estimations are again carried out using data from 1990 to 2003 for power station 1 and 2, and 2000 to 2004 for power station 3. We are then using years 2004 to 2006 (2005 to 2006 for power station 3) to perform out of sample tests on the models. Figure 15-17 shows out of sample plots for the three power stations investigated, including realized aggregated inflow and the forecast based on the estimated parameters.

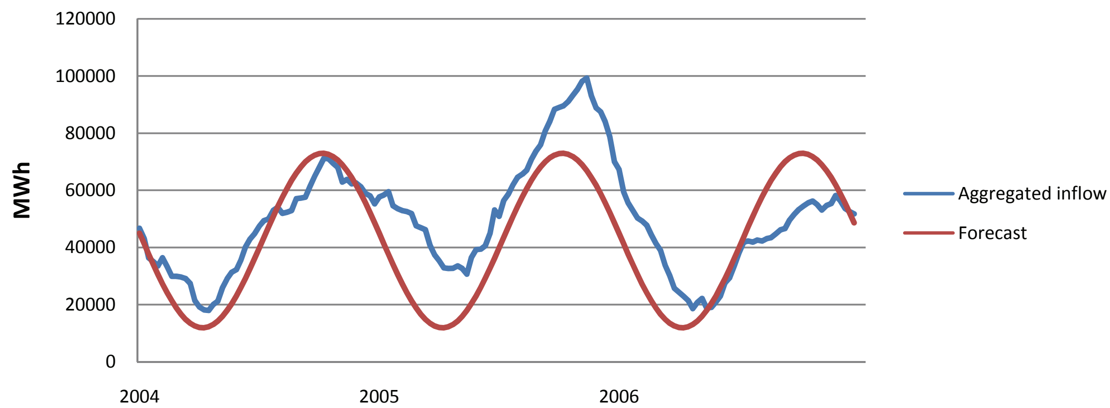


Figure 15. Aggregate inflow power plant 1

The figure shows the 26 weeks aggregate inflow for the years 2004-2006 for power plant 1

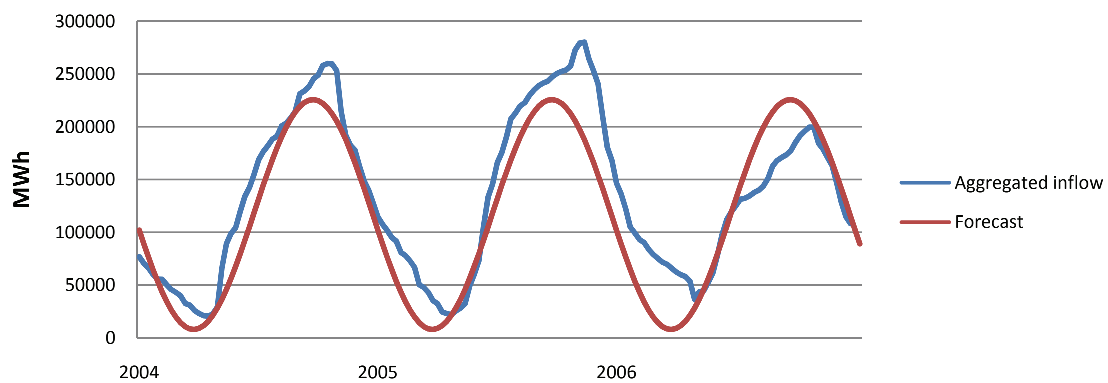


Figure 16. Aggregated inflow power plant 2

The figure shows 26 weeks aggregate inflow for the years 2004-2006 for power plant 2

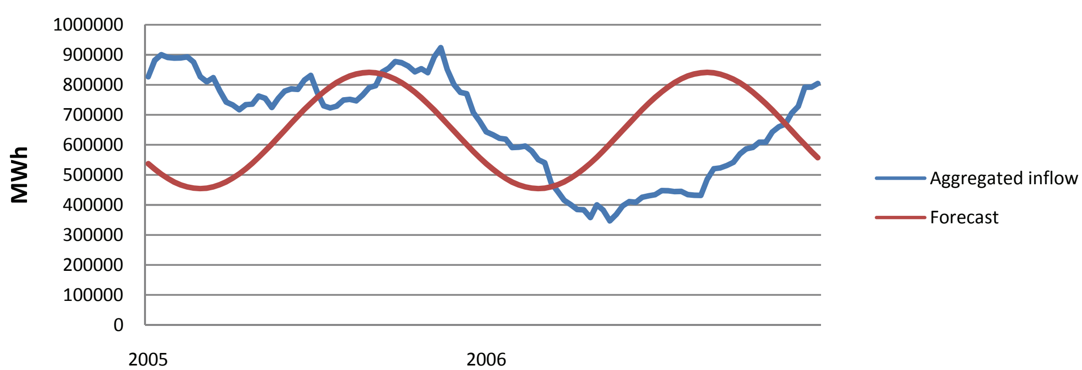


Figure 17. Aggregated inflow power plant 3

The figure shows 26 weeks aggregate inflow for the years 2005-2006 for power plant 3

The models seem to fit well, especially for power station 1 and power station 2. For power station 3 however the fit is less accurate. This can be explained by the fact that power station 3 has the lowest degree of seasonal dependence, which also will have an impact when we try to fit aggregated inflow to a deterministic expression trying to capture seasonal variations. Looking at table 2 reveals the same properties of the inflow data to power station 3, since it has the largest deviation measures from the estimated model. Table 3 displays the previously used measures of accuracy for the three power stations for the aggregated model.

Table 3. Error measures from out of sample tests for aggregate inflow model

The table summarizes the results from out of sample tests on the Full Year Model and Divided Year Model for the three power stations

	Mean error	Mean absolute error	Mean squared error	Mean percentage error	Mean absolute percentage error	U	Root mean squared error
Power station 1	405,4	5739,5	50635359,4	0,5 %	12,1 %	2,10	7115,9
Power station 2	-1446,3	12880,0	269419881,2	-3,4 %	13,7 %	1,26	16414,0
Power station 3	-51359,9	90789,6	12815519554,0	-15,0 %	20,3 %	3,70	113205,7

The error measured for the aggregate inflow model in table 3 quantifies the observations from graphs 15-17. The fit between the model and the aggregated inflow to power station 1 and 2 are quite good. The estimated parameters can be found in appendix C.

5.3 Two factor price model

The two-factor model from Lucia and Schwartz (2001) is used for the purpose of modeling the spot price of electricity. The parameters are estimated using the procedure presented in Lucia and Schwartz (2001) and Cortazar and Schwartz (2002). This is a numerical nonlinear least squares procedure, using data about spot and futures/forward prices from 1996 to 2006. This is considered a more flexible and user-friendly approach than the rather complex Kalman filtering. After running 35 iterations in Excel, the improvements per run are approximately zero and the following parameter estimations are obtained.

As a measure of how accurate our estimated parameters are, a comparison with the estimates obtained in Lucia and Schwartz (2001) and Krossøy and Torgersrud (2003) is presented in table 4. The table shows percentage difference in the estimated parameters.

Table 4. Comparison between estimated parameters

The table displays a comparison between the estimated parameters obtained in this paper and the parameters obtained by Lucia and Schwartz (2001) and Krossøy and Torgersrud (2003). Consult the model chapter for description of the different parameters

	α^*	κ	μ_ε^*	α	γ	τ
Lucia and Schwartz vs. this paper	4 %	-317 %	72 %	0 %	17 %	153 %
Lucia and Schwartz vs. Krossøy and Torgersrud	-37 %	-264 %	354 %	-1 %	10 %	124 %
Krossøy and Torgersrud vs. this paper	30 %	-14 %	111 %	0 %	8 %	-123 %

As can be seen from the table, there are some differences between the parameter values estimated in Lucia and Schwartz (2001) and Krossøy and Torgersrud (2003), and between Lucia and Schwartz (2001) and this paper. The principal reason for this is that Lucia and Schwartz (2001) uses daily data, while this paper and Krossøy and Torgersrud (2003) uses weekly data. The time-frame is also different. Lucia and Schwartz (2001) uses data from 1st January 1993 to 14th December 1998, while Krossøy and Torgersrud (2003) uses data from 1996 to 2002 and this paper uses data from 1996 to 2006. Some of the differences may also stem from the fact that Lucia and Schwartz (2001) includes one extra term in the deterministic part of the model that describes the difference in prices between weekdays and weekends/holidays. There are also some differences between the values estimated in this paper and in Krossøy and Torgersrud (2003). Some of these different parameter estimations may stem from the fact that this paper includes a larger data set than Krossøy and Torgersrud (2003).

5.4 State space time series inflow models

Using state space methodology, the natural log transformation was used to model weekly inflow for all three power producers. This prevents us from obtaining scenarios with negative inflow, since the simulation is done in log transform. When this was not done the lower confidence interval (90%) fall below zero for long periods of the year during winter as shown in the figure below. Setting these scenario values to zero would then skew the distribution upwards resulting on average in scenarios with too high annual inflow.

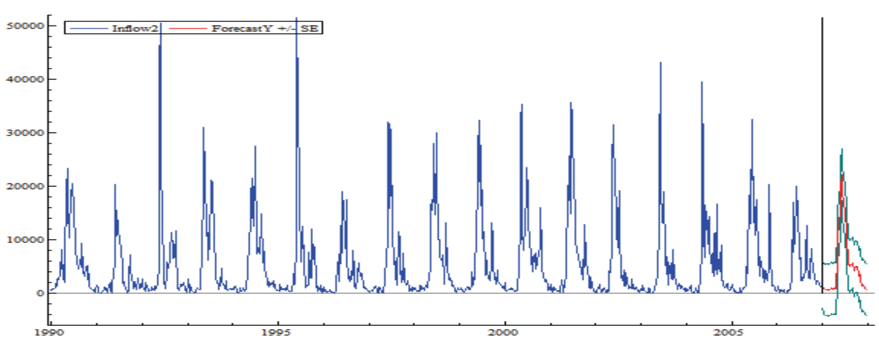


Figure 18. Inflow forecast based on non-log model with 90% confidence intervals

By using the natural log we then avoid the problem with negative inflow scenarios, but the log transform do have some consequences when the model is applied for simulation. The confidence intervals for the log of inflow will be skewed upwards when they are transposed back to inflow by taking exponentials, as shown in the figure below.

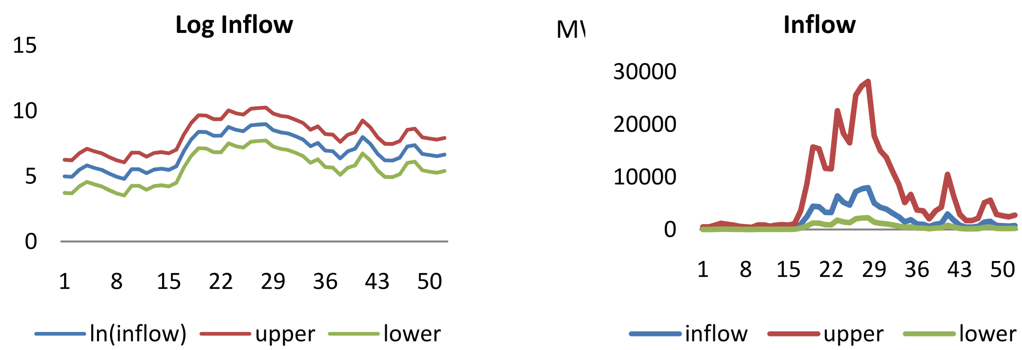


Figure 19. Transformation of confidence interval in log transforms

This can however be justified to an extent since extremely high observations of inflow tend to have a higher absolute deviation from the expected inflow than extremely low observations. Inflow is constrained downward by zero (since we cannot have negative inflow) but there is nothing preventing extremely high observations. These properties were also noticed during the statistical analysis of the inflow series in chapter 3. The scenarios are first simulated in log transform before they are transposed back by taking exponentials. The same problem as described above will then occur and scenarios above the forecast will tend to be overemphasized compared to the scenarios below. So when simulating many scenarios the expected value of the simulations will be a bit higher than the forecast. The annual average inflows to the different producers are fairly stable (appendix A) over time and so it is preferable that the inflow simulations are centered round this mean. This is accomplished by adjusting the forecast so that the model has a forecast that gives a total annual inflow equal to the average total annual inflow for the entire series. The level of inflow is not assumed to have changed much the last 16 years due to climate changes or other hydrological phenomenon.

When adjusting the forecast it is not done equal in all weeks of the year, since it is more likely that inflow will occur during the wet periods. This effect is already captured in the seasonal components. The sum of the seasonal components add to zero with high positive values in typical wet weeks and high negative values during dry winter weeks. By taking the exponential of these seasonal components we get weights that can be used to determine which weeks that typically contributes with most inflow. The adjustment of the forecast is allocated to the different weeks depending on these weights.

Six different models were estimated for the different inflow series and evaluated in sample based on the Akaike information criterion (AIC), defined in Diebold (2007), and other statistics given by the software. The models were also tested out of sample by holding back a part of the time series, in order to determine the forecasting accuracy of the different models. The models differ by which factors and how many stochastic terms that are included when describing the structure of the time series. The components used in the different models are given in table 5 below.

Table 5. Models tested for inflow
The table show the different components used in the 6 models tested for inflow

Model	Level	Slope	Seasonal	Irregular	Explanatory	AR(1)	Interventions
1	Stochastic	Stochastic	Stochastic	Yes	None	None	Yes
2	Stochastic	Fixed	Fixed	Yes	None	None	Yes
3	Stochastic	None	Fixed	Yes	None	None	Yes
4	Fixed	None	Fixed	Yes	None	None	Yes
5	Stochastic	None	Fixed	Yes	None	Yes	Yes
6	Fixed	None	Fixed	Yes	None	Yes	Yes

Slope was attempted but disregarded based on theory suggesting that inflow is stable over time with no trend, and the insignificant values obtained for the slope in the first two models further supported this decision. Interventions are used in the first time series to cope with two irregular measurements.

5.4.1 Producer 1

For producer 1 the different models are estimated using the period 1990(1) – 2003(52), and the remaining 3 years of data is used for out-of-sample testing. All the models account for the two observations in 2002 (week 23 and 34) mentioned by the producer in the dataset, were there was a measurement error and a correction for the error a few weeks later. These were both modeled with irregular intervention components taking the value 1 in the respective week, and zero in all others.
The results of these in- and out-of sample tests are given below[9]:

Table 6. In- and out-of-sample statistics for the inflow models estimated for Producer 1

	Inflow model					
In sample	**1**	**2**	**3**	**4**	**5**	**6**
AIC	-0,119	-0,119	-0,132	0,212	**-0,232**	-0,225
SC	46,96	46,96	51,73	-75,14	90,07	87,68
Q(d)	69,23 (20)	69,23 (20)	69,89 (21)	594,61 (21)	17,48 (20)	23,72 (20)
H(h)	3,30 (216)	3,30 (216)	3,29 (217)	2,29 (217)	3,47 (217)	3,51 (217)
N	33,394	33,395	33,576	17,797	36,417	41,733
DW	1,717	1,717	1,7102	0,77301	1,9458	1,9649
R^2	0,11	0,11	0,11	-0,32	0,20	0,20
Out of sample						
Mean error	-1016	-1016	-868	504	715	510
Mean absolute error	1927	1927	1836	1217	1213	1222
Mean squared error	6916891	6917462	6162938	2913669	3034180	2927104
Mean percentage error	-228 %	-228 %	-212 %	-68 %	-49 %	-73 %
Mean absolute percentage error	256 %	256 %	241 %	123 %	111 %	128 %
U	1,01	1,01	0,98	0,87	0,87	0,87
RMSE	2630	2630	2482	1706	1741	1710

Based on the AIC, model 5 is preferred slightly over model 6. The residual errors of these models are independent based on the Box-Ljung Q statistic of 17,5 and 23,7 which are both smaller than the critical value of $\chi^2_{2,0.05}$ = 31,4. The residuals in all the models are heteroscedastic (H statistic) and

[9] All the test statistics and critical values are thorough described in STAMP user manual

neither have normally distributed residuals (N statistic). Model 5 and 6 seem to outperform the others based on the out-of sample statistics as well, but show little dissimilarities and none of them seem to be dominating the other. Model 5 does however model inflow with a stochastic level resulting in a random walk, which contradicts our assumption of stationary inflow. Model 6 is therefore selected and adjusted to the mean annual inflow before being used to generate scenarios. The adjusted forecast is plotted together with the historical inflow for the period 1996–2006 below.

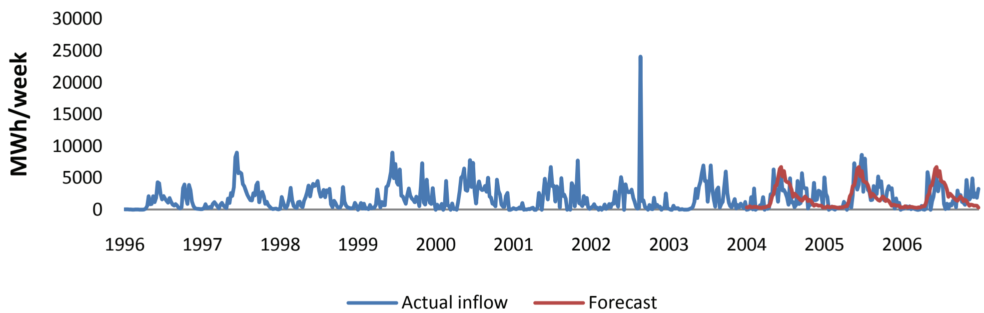


Figure 20. Historical and forecasted inflow for producer 1
The figure show the weekly historical inflow for the last 10 years together with the 3 year out-of-sample forecast

5.4.2 Producer 2

For producer 2 the same procedure was applied as with producer 1 and the different models are estimated using the period 1990(1) – 2003(52), and the remaining 3 years of data is used for out-of-sample testing. The results of the in- and out-of sample tests are given below:

Table 7. In- and out-of-sample statistics for the inflow models estimated for Producer 2

	Inflow model					
In sample	**1**	**2**	**3**	**4**	**5**	**6**
AIC	-0,18	-0,18	-0,19	0,04	-0,36	**-0,37**
SC	69,77	69,77	74,25	-12,88	137,67	137,75
Q(d)	101,52 (20)	101,52 (20)	102,92 (21)	421,19 (21)	34,65 (20)	35,42 (20)
H(h)	0,91 (225)	0,91 (225)	0,92(225)	0,63 (225)	0,88 (225)	0,90 (225)
N	10,08	10,08	10,10	54,24	22,32	22,50
DW	1,83	1,83	1,82	0,80	1,87	1,89
R^2	0,03	0,03	0,03	-0,25	0,19	0,20
Out of sample						
Mean error	-1672	-1672	-1232	945	1609	939
Mean absolute error	3544	3544	3316	2469	2409	2472
Mean squared error	41101657	41101657	35793760	24507462	25824126	24510763
Mean percentage error	-136 %	-136 %	-123 %	-48 %	-26 %	-49 %
Mean absolute percentage error	151 %	151 %	142 %	89 %	78 %	90 %
U	1,35	1,35	1,32	0,89	0,81	0,89
RMSE	6411	6411	5982	4950	5081	4950

Based on the AIC, model 6 is preferred over the others. The residuals of this model show homoscedastisity but are not normally distributed. Neither of the models have independent residuals, but the last two have the least serial correlation, as the Durbin Watson statistic also indicates. From the out-of-sample statistics the last three models show better predictive power, but neither seem to have superior explanatory power. Due to this and the same argumentation as for model 1, the last model is selected to describe the dynamics of the inflow series. It is adjusted to the mean annual inflow and plotted together with the historical inflow for the period 1996(1) – 2006(52) in the graph below.

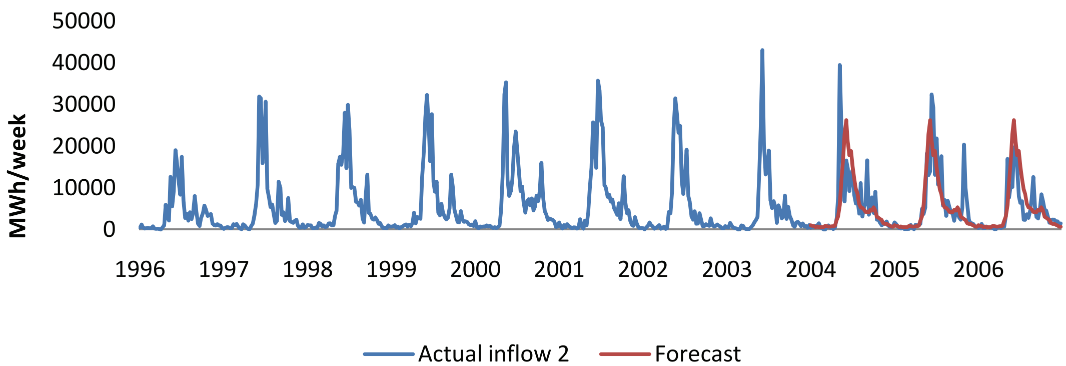


Figure 21. Historical and forecasted inflow for producer 2
The figure show the weekly historical inflow for the last 10 years together with the 3 year out-of-sample forecast

5.4.3 Producer 3

The inflow series for producer 3 starts in 2000 and so we estimate the model using the period 2000(1) – 2005(52), and the data from 2006 is used for out-of-sample testing. A pulse shaped intervention variable was used in week 16 (2001) to cope with a measurement error in the inflow series. The results of the in- and out-of sample tests are given below.

Table 8. In- and out-of-sample statistics for the inflow models estimated for Producer 3

	Inflow model					
In sample	**1**	**2**	**3**	**4**	**5**	**6**
AIC	0,184	0,188	0,175	0,223	**0,158**	0,161
SC	-70,84	-74,29	-69,10	-90,77	-59,91	-61,32
Q(d)	31,94 (10)	34,39 (10)	34,35 (11)	159,93 (11)	8,22 (10)	14,20 (10)
H(h)	1,10 (85)	1,44 (85)	1,43 (85)	1,16 (85)	1,30 (85)	1,29 (85)
N	10,94	10,15	10,62	12,96	17,12	21,97
DW	1,57	1,56	1,54	1,13	1,96	1,95
R^2	-0,06	0,16	0,16	0,00	0,22	0,21
Out of sample						
Mean error	5297	3984	2989	5552	2085	5624
Mean absolute error	11990	12680	12839	12293	13235	12366
Mean squared error	289505798	327054828	323832042	325094270	331833399	329957327
Mean percentage error	-6 %	-18 %	-22 %	-8 %	-28 %	-8 %
Mean absolute percentage error	48 %	57 %	59 %	51 %	62 %	51 %
U	0,73	0,74	0,74	0,75	0,73	0,75
RMSE	17014	18084	17995	18030	18216	18164

Based on the AIC, model 5 is preferred slightly over model 6. The residual errors of these two models are independent based on the Box-Ljung Q statistic. They are both homoscedastic but do not have normally distributed residuals. Neither of the models seems to outperform the others in terms of the out-of sample statistics. Model 6 show low mean percentage- and absolute percentage errors relative to the others. This model is also stationary due to a fixed level component. From examining the time series plot of inflow to Producer 3 it would seem that there was a stochastic seasonal pattern that changed over time, but this was not found to be significant when attempted modeled, and so a fixed seasonal pattern was chosen. Model 6 is therefore selected and adjusted to the mean annual inflow before being used to generate scenarios. The adjusted forecast is given below together with the historical inflow for the period 2000(1) – 2006(52).

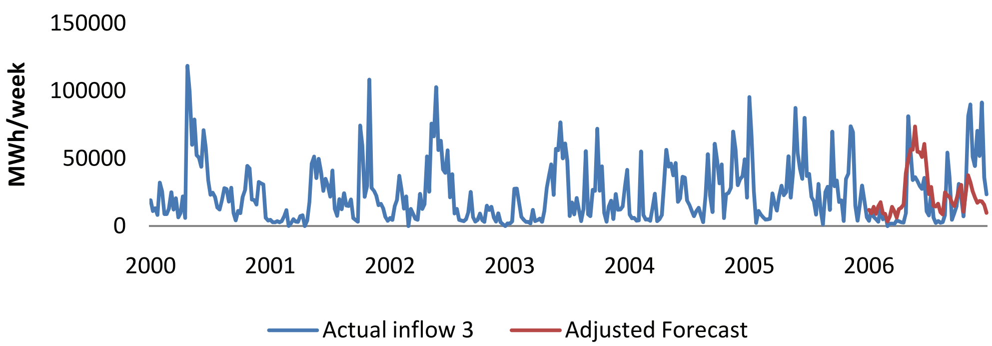


Figure 22. Historical and forecasted inflow for producer 3
The figure show the weekly historical inflow for the last 6 years together with 1 year out-of-sample forecast

5.5 State space time series price models

In this paper we wish to correlate price with inflow and one way of doing this could be to first forecast inflow to the reservoir and then use the forecasted inflow series as an explanatory variable when predicting the price series. If using the inflow series as an explanatory variable does not help increasing the forecasting ability of the model, the price series can be forecasted independently of the inflow series and then matched with a suitable inflow series afterwards. Later in this paper we describe a few such matching rules and their properties. Other factors affecting the System price such as gas/coal prices in Europe, fallouts of lines, surplus of wind energy etc are disregarded in this model. Certain extreme events on prices will however be smoothed out if they are very short, since we use the weekly average system price. A model for price without any explanatory variables is first estimated, which can be used for all the power plants. The following models were attempted in this case

Table 9. Components of models tested for price
The table show the different components used in the different price models estimated

Dependent variable	Level	Slope	Seasonal	Irregular	AR(1)
System Price	Stochastic	Stochastic	Stochastic	Yes	None
System Price	Stochastic	Fixed	Fixed	Yes	None
System Price	Stochastic	None	Fixed	Yes	None
System Price	Stochastic	Fixed	Fixed	Yes	Yes
System Price	Fixed	Fixed	Fixed	Yes	Yes
Ln(System Price)	Stochastic	Stochastic	Stochastic	Yes	None
Ln(System Price)	Stochastic	Fixed	Fixed	Yes	None
Ln(System Price)	Stochastic	None	Fixed	Yes	None
Ln(System Price)	Stochastic	Fixed	Fixed	Yes	Yes
Ln(System Price)	Fixed	Fixed	Fixed	Yes	Yes

The estimation of these models showed which models that were potential candidates when using explanatory variables. The use of log transform in this case was found not to be used since the system price show strong signs of random walk properties. This could result in infeasible high price scenarios due to the exponential transformation and will therefore not be attempted for the individual producer's price models. The results of these models are therefore not shown in the analysis below. This resulted in the following models for the individual producers' price series using functions of inflow as an explanatory variable, and the estimation of these for the individual power producers are given in the following chapters.

Table 10. Components of price models for individual producers
The table show the different components in the price models that were estimated for the individual producers

	Dep. variable	Level	Slope	Seasonal	Irregular	AR(1)	Explanatory
1	System Price	Stochastic	Fixed	Fixed	Yes	None	Inflow
2	System Price	Stochastic	Fixed	Fixed	Yes	None	Agg Inflow
3	System Price	Stochastic	Fixed	Fixed	Yes	Yes	Inflow
4	System Price	Stochastic	Fixed	Fixed	Yes	Yes	Agg Inflow
5	System Price	Fixed	Fixed	Fixed	Yes	Yes	Inflow
6	System Price	Fixed	Fixed	Fixed	Yes	Yes	Agg Inflow

5.5.1 Price Producer 1

The models are estimated using data for the period 1993(1)-2003(52), while the last observations from 2004(1) – 2006(52) are used to test the models out of sample. The actual values for inflow have been used as explanatory variables when generating the out of sample forecast. This is done for all three producers in order to compare the forecasted system price with the actual price. When generating scenarios, the forecasted inflow will be used as the explanatory variable. Results from the in and out of sample tests are given in the table below.

Table 11. In- and out-of-sample statistics for the price models estimated for Producer 1

	Price model					
In sample	**1**	**2**	**3**	**4**	**5**	**6**
AIC	6,428	6,429	6,426	6,423	6,420	6,417
SC	-1836,64	-1836,77	-1832,35	-1831,29	-1832,54	-1831,69
Q(d)	89,88 (17)	96,88 (17)	81,91 (16)	87,10 (16)	83,21 (16)	88,83 (16)
H(h)	3,86 (172)	3,83 (172)	3,76 (172)	3,77 (172)	3,80 (172)	3,81 (172)
N	1398,00	1466,20	1754,20	1768,60	1760,80	1783,70
DW	1,79	1,80	1,74	1,74	1,75	1,75
R^2	-0,08	-0,08	-0,06	-0,06	-0,06	-0,06
t-value explanatory variable	-2,68	-2,70	-2,67	-3,20	-2,67	-3,11
Out of sample						
Mean error	76,09	84,62	30,15	33,84	64,01	70,78
Mean absolute error	83,11	88,85	63,01	58,13	74,85	78,55
Mean squared error	13616	13694	8330	7084	11978	11421
Mean percentage error	21 %	25 %	4 %	7 %	17 %	20 %
Mean absolute percentage error	25 %	27 %	19 %	18 %	22 %	24 %
U	4,27	4,48	3,48	3,19	3,94	3,99
RMSE	116,69	117,02	91,27	84,16	109,45	106,87

Model 6 is preferred slightly over model 5 and 4 based on the in-sample measures. All models have dependent residuals that are heteroscedastic and not normally distributed. Based on the out-of sample testing, model 4 is preferred with the lowest RMSE and u statistic and a relatively low mean- and absolute percentage value. This model uses 26 week aggregated inflow as the explanatory variable and the forecast is given in the graph below together with the historical system price.

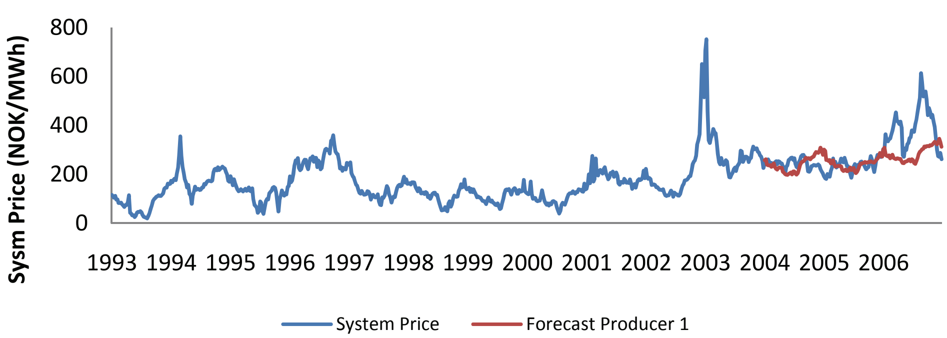


Figure 23. Historical and forecasted price for producer 1

5.5.2 Price Producer 2

The models are estimated using data for the period 1993(1)-2003(52), while the last observations from 2004(1) – 2006(52) is used to test the models out of sample. The actual values for inflow have been used as explanatory variables when generating the out-of sample forecast. Results from the in and out of sample tests are given in the table below.

Table 12. In- and out-of-sample statistics for the price models estimated for Producer 2

	Price model					
In sample	**1**	**2**	**3**	**4**	**5**	**6**
AIC	6,439	6,427	6,430	6,418	6,424	**6,413**
SC	-1837,93	-1836,26	-1833,28	-1830,03	-1833,45	-1830,49
Q(d)	90,14 (17)	94,42 (17)	82,01 (16)	82,47 (16)	83,12 (16)	84,51 (16)
H(h)	3,73 (172)	3,72 (172)	3,63 (172)	3,60 (172)	3,67 (172)	3,65 (172)
N	1355,60	1371,70	1721,10	1722,80	1726,60	1738,20
DW	1,78	1,80	1,74	1,73	1,75	1,74
R^2	-0,08	-0,08	-0,06	-0,05	-0,06	-0,05
t-value explanatory variable	-2,65	-3,22	-2,64	-3,96	-2,65	-3,87
Out of sample						
Mean error	76,79	91,79	31,14	30,41	63,75	59,75
Mean absolute error	83,88	95,61	63,57	59,42	75,10	72,12
Mean squared error	13818	15265	8491	7062	12024	10028
Mean percentage error	21 %	27 %	5 %	5 %	17 %	16 %
Mean absolute percentage error	25 %	29 %	19 %	18 %	22 %	22 %
U	4,31	4,80	3,51	3,29	3,96	3,75
RMSE	117,55	123,55	92,15	84,04	109,65	100,14

Model 6 is preferred slightly over model 4 based on the in-sample measures. All models have dependent residuals that are heteroscedastic and not normally distributed. None of the models seem to fit the data particularly good, but based on the out-of sample testing, model 4 seem to be the best. This uses 26 week aggregated inflow as the explanatory variable and captures some of the variation in the system price as can be seen in 2006, where the model predicts a slightly higher price due to the low aggregated inflow that year (negatively correlated). This can be seen in the figure below where the forecast is given together with the historical system price series.

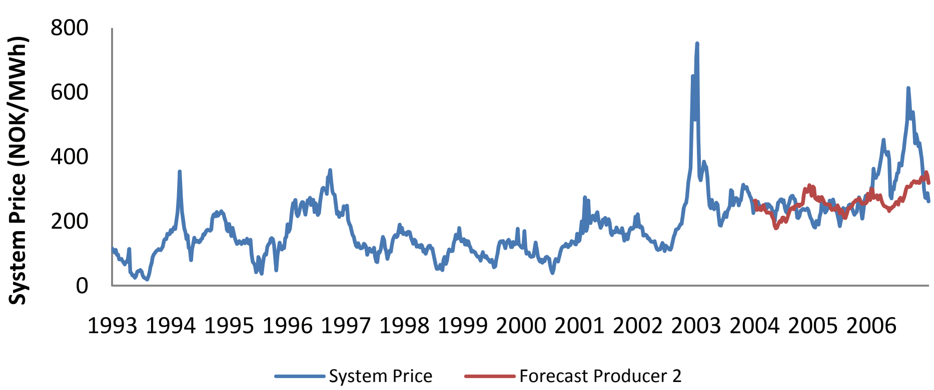


Figure 24. Historical and forecast price for producer 2

5.5.3 Price Producer 3

The models are estimated using data for the period 2000(1)-2004(52), while the last observations from 2005(1) – 2006(52) are used to test the models out of sample. The actual values for inflow have been used as explanatory variables when generating the out-of sample forecast. Results from the in and out of sample tests are given in the table below.

Table 13. In- and out-of-sample statistics for the price models estimated for Producer 3

	Price model					
In sample	**1**	**2**	**3**	**4**	**5**	**6**
AIC	6,651	5,984	6,660	5,974	6,647	5,960
SC	-862,85	-776,23	-860,66	-771,42	-860,68	-771,42
Q(d)	48,76 (8)	47,64 (7)	44,01 (7)	39,09 (6)	44,10 (7)	39,12 (6)
H(h)	0,67 (68)	1,22 (60)	0,74 (68)	1,27 (60)	0,74 (68)	1,27 (60)
N	223,46	205,29	366,48	284,26	365,59	284,28
DW	1,77	1,82	1,70	1,71	1,70	1,71
R^2	-0,21	-0,24	-0,17	-0,17	-0,17	-0,17
t-value explanatory variable	-2,54	-2,93	-2,55	-4,66	-2,55	-4,66
Out of sample						
Mean error	113,77	102,98	16,44	52,16	4,38	51,08
Mean absolute error	117,79	104,80	64,55	60,83	65,22	60,13
Mean squared error	22491	15276	7733	6512	7241	6394
Mean percentage error	31 %	31 %	-1 %	15 %	-5 %	14 %
Mean absolute percentage error	33 %	32 %	19 %	18 %	20 %	18 %
U	4,87	4,52	2,98	3,01	3,09	2,99
RMSE	149,97	123,60	87,94	80,70	85,09	79,96

For this producer all the models using aggregated inflow (2, 4 and 6) as an explanatory variable seem to outperform the ones using weekly inflow on the in-sample tests, with model 6 giving the lowest AIC. All the models have homoscedastic residuals but they are not independent and neither are normally distributed. Based on the out-of-sample tests, model 4 and 6 have the lowest mean average percentage error and RMSE. After examining a plot of these two models together with the actual price series we notice that they do not differ much, and model 4 is chosen instead of model 6 because it contains a stochastic level, which we think coincide with theory. The forecast from this model is given in the graph below together with the historical system prices.

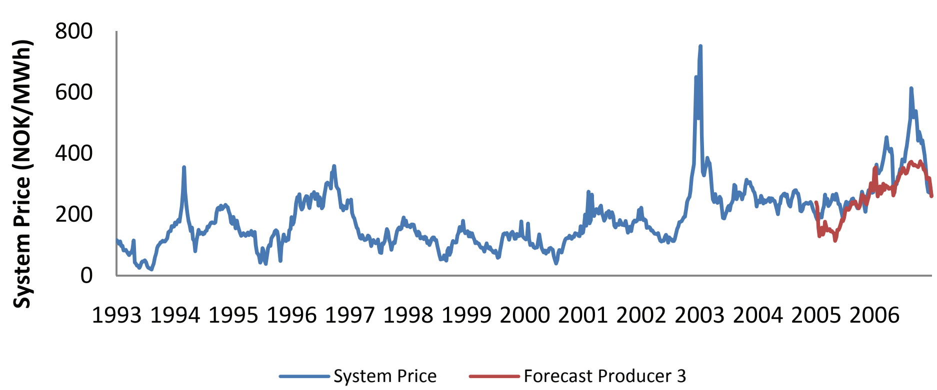


Figure 25. Historical and actual price for producer 3

5.6 Price model comparisons

Two different models have been used to model the system price in this paper and in this section we will discuss some of the differences between the models.

The factor model used in Lucia and Schwartz is calibrated to the futures curve, which contain valuable information about the market's expectations of the system price. It models the seasonality with a trigonometric sine function as can be seen in figure 26. The forecast is dependent on the last deviation between the realized value of the system price and the long run prediction. The speed of reversion back to the long run mean is given by the mean reverting factor. When using the model for forecasting in this paper we have set the last deviations equal to zero in order to have the simulations centered round the long run forecast. The effect of this is to reduce the price forecast notably as can be seen in the figure below, and the simulated price scenarios will also be shifted downward by the same amount.

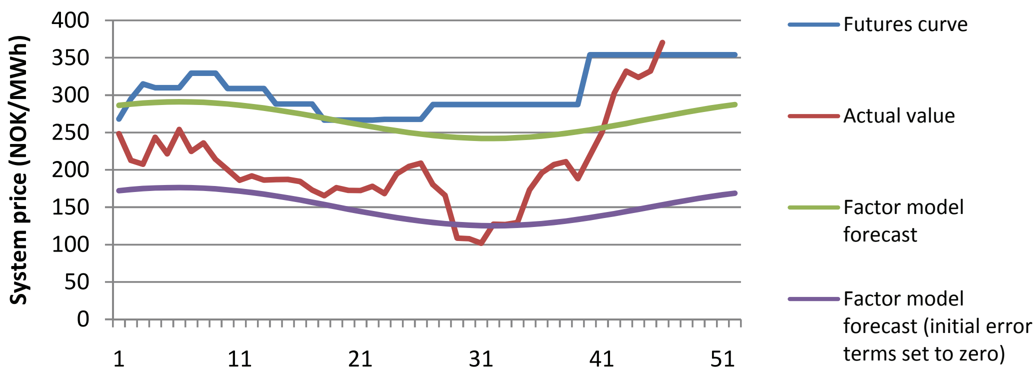


Figure 26. Actual prices, forward curve and Factor price model forecast for 2006

The STAMP model is not calibrated to the futures curve but seem to fit the futures curve pretty good. It uses historical price movements in addition to inflow as an explanatory variable to forecast and create simulations. The estimation of model parameters are fast compared to the more time consuming process of estimating the two factor model based on a Kalman filter or the minimization of least squares iteration procedure used in this paper. The STAMP models weighs information from current periods higher than old observations compared to the equal weighting in our two factor estimation procedure. The STAMP model uses dummy variables to model the seasonality.

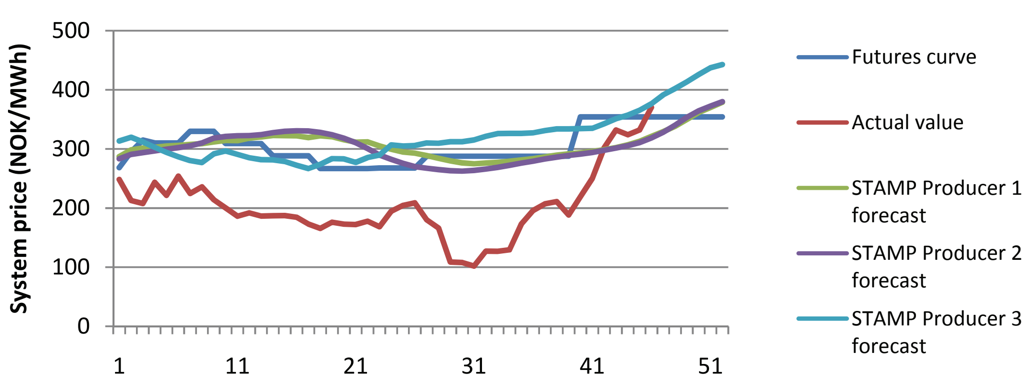


Figure 27. Actual prices, forward curve and STAMP price model forecast for 2006

6.0 Correlation

6.1 Natural correlation testing

Theory suggests negative correlation between system price and inflow at a national level, but for a local power producer we need to examine if there is any correlation between local inflow and the system price. The correlation between local inflow and price and the 26 week aggregated local inflow and price is given in appendix D. In the first two power plants the weekly inflow is stronger correlated with the system price than the aggregated inflow. At power plant 3 however the correlation between the system price and the 26 weeks aggregated inflow is -0.45, about twice as large as for the weekly inflow. This strong negative correlation can also be seen in the graph below which show the indexed[10] system price and 26 week aggregated local inflow to power plant 3. This plot also explains the high electricity prices in 2002/03 and in 2006 due to low aggregated inflow.

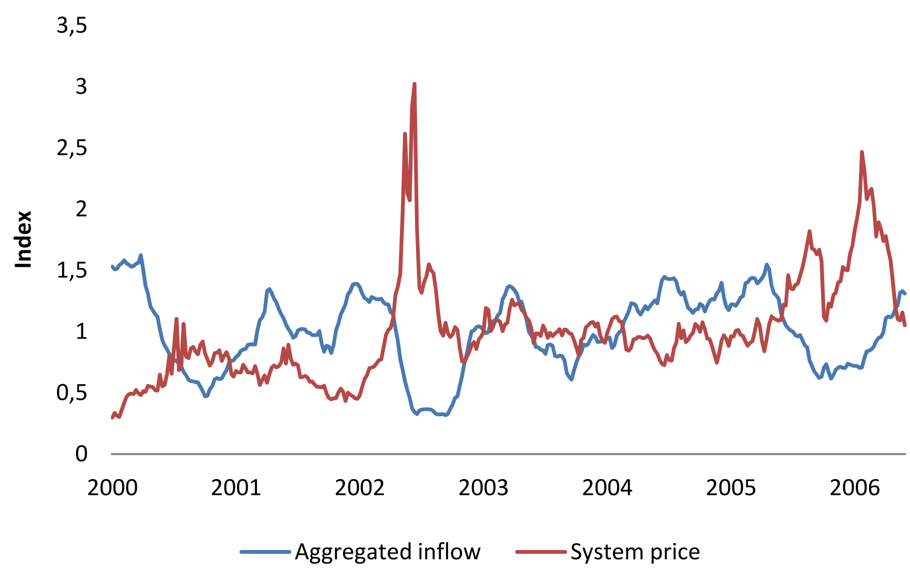


Figure 28. Indexed system price and aggregate inflow
The graph displays the indexed 26 weeks aggregated inflow and the system price for power plant 3

All the correlations are negative (seen in appendix D) as expected, and the strongest correlation was found for the largest power plant (producer 3).

6.2 Rolling window correlation for factor models

Although theory suggests that there seem to be evidence of negative correlation between aggregated inflow and price (especially between price and aggregated national inflow) we still need to see if there is correlation between the error terms of the models we estimate. In Bjørnsgard and Hauge (2007) the weekly correlation is estimated between the error terms of the model and used when simulating price and inflow scenarios. Thus finding a weekly correlation between the error

[10] Inflow and price have been divided by their average over the given time period

terms allows one to correlate simulated scenarios through correlating the random numbers drawn to represent the error terms.
If the correlation between the error terms is not stable over time however the procedure should not be used since the correlation we get is due to randomness and not a structural relationship that can be modeled.

To check for the stability of the correlations we can calculate rolling window correlations between the different error terms of the two factor price model, and the one factor inflow model. Correlations are found for 26 and 40 week correlation lengths. For the 26 week length the correlations are calculated for the first 26 weeks, then for week 2-27, 3-28 etc and the correlation is plotted against time. If the correlation is stable, these graphs should tend to lie close to a constant level. This procedure for checking the correlation stability is used by Hvarnes (2007). A point worth mentioning here is the fact that the jumps in the electricity spot prices may lead to an unstable volatility (and hence correlation) if the rolling windows are too narrow, Seppi (2002). Thus the volatility may be taken to be heteroscedastic even though this is not the fact. Since we use two windows with up to 40 weeks length, and the correlation shows the same irregular tendency independent of the window length, this is not considered to be a problem in the interpretation of our results.

Model correlations

The error terms of the different price and inflow models needs to be found before this procedure can be applied. For the two-factor price model there are two stochastic terms, one which follow a mean reverting process, and one that follow an arithmetic Brownian motion. The short term mean reverting effect follow the stochastic process given by:

$$dx_t = \kappa(\alpha^* - x_t)dt + \sigma dz \tag{15}$$

This is the continuous-time version of a first order autoregressive process in discrete time, and is the limiting case of the following AR(1) process as $\Delta t \to 0$, Dixit and Pindyck (1994).

$$x_t - x_{t-1} = \alpha^*(1-e^{-\kappa}) + (e^{-\kappa}-1)x_{t-1} + \varepsilon_t \tag{16}$$

Where ε_t is normally distributed with mean zero and variance given by:

$$\sigma_\varepsilon^2 = \frac{\sigma^2}{2\kappa}(1-e^{-2\kappa}) \tag{17}$$

The ε_t in equation 16 is the error term of the mean reverting stochastic factor in the two-factor price model and it is equated using the time series we have from the two-factor estimation excel sheet together with the parameters of κ and α* estimated in the same sheet. The error term can therefore be found by rearranging equation 16 to get:

$$\varepsilon_t = X_t - X_{t-1} - \alpha^*(1-e^{-\kappa}) - (e^{-\kappa}-1)X_{t-1} \tag{18}$$

The error of the arithmetic term is a bit easier to calculate by also discretizing the continuous time version of the stochastic process.

$$dE_t = \mu_E dt + \sigma_E dZ_E \tag{19}$$

This gives the following formula for the long term error, where σ_E is the variance of the error and μ_E found in the two-factor price model estimation.

$$\zeta_t = \frac{E_t - E_{t-1} - \mu_E}{\sigma_E} \tag{20}$$

The error term of the inflow series also follow a mean reverting process but this is not transformed to a risk neutral environment, which is needed for the price model. Doing this gives the following stochastic process.

$$dY_t = -\kappa Y_t dt + \sigma dZ \tag{21}$$

And the error term is found by using the same procedure as above giving the following error term:

$$\varsigma_t = Y_t - Y_{t-1} - (e^{-\kappa} - 1)Y_{t-1} \tag{22}$$

All the error terms for X (short term price error), E (long term price error) and Y (inflow error) are then normalized before being used for measuring correlation between each other's errors.

The correlations for the Full Year and the aggregated inflow models for producer 1 with the system price are given below (where X represents the mean-reverting short term variance and E the long term).

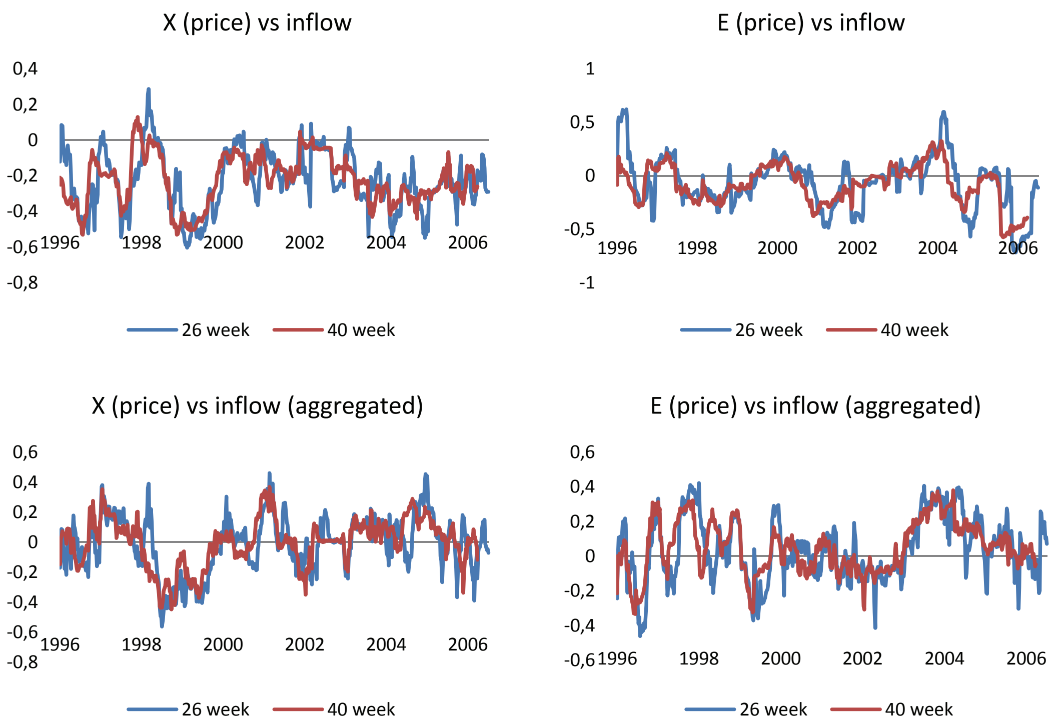


Figure 29. Correlation between error terms of the Factor models for price and inflow
The graphs show the 26 week and 40 week rolling correlations between the different error terms of the Factor model for price and inflow

The correlations between the error terms are very unstable for all the models (appendix E), yet the correlations for the aggregated inflow models seem to be more stable, and centered round zero. Even if they were stable, a correlation of zero would mean that we simulated the price and inflow series independently. These graphs are representative for the other power plants in appendix E. Also the weekly correlation between the error terms in the 2 factor price model seems to be very unstable as seen in the figure below.

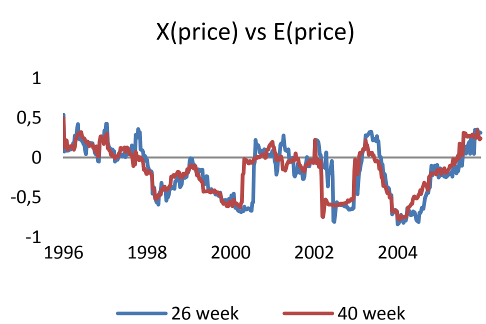


Figure 30. Correlation between the error terms in the two factor model for price
The graph show the 26 week and 40 week rolling correlations between the error terms of the two factor model for price.

Using weekly correlations between the error terms in our models to correlate different scenarios for price and inflow is therefore not suitable. Other methods to correlate price and inflow series then needs to be addressed and a few such are described in the next section.

6.3 Matching price and inflow scenarios

Since weekly correlation is very volatile for the error terms in our models and changes depending on the estimation period and interval length we use, there is no way of correctly correlating them on a weekly basis. When performing simulations the random numbers drawn for each week does not have a constant correlation and so they could be drawn independently, and the scenarios then arranged afterwards based on some sort of matching rule. Descriptions of a few of these methods are given below together with certain intuitive drawbacks, and they are all illustrated in figure 31.

6.3.1 Method 1

A simple and very intuitive form of such a rule is to calculate the average of the price and inflow for the different series, and match the price series with the highest average price with the lowest average inflow series. A problem with this method is that since we are simulating more than two years ahead with this procedure we could end up drawing two scenarios with high inflow and price the first year and then low inflow and low price the second year, and matching these two together since their average match. The method also simplifies the relationship by assuming perfect negative correlation.

6.3.2 Method 2

The problem described in the model above can be solved by simulating price and inflow series for one year at the time and then putting them together to form a longer time series. This way we can use the same procedure above but matching the scenarios for every year and not based on the whole series. Which series that follows each other every year should be random, since there is no reason why a very dry year should have a higher tendency of being followed by another dry year. One could however argue for some kind of correlation between the years since we do have reservoirs that can hold water for more than a year and so a very wet year could tend to reduce prices in the following if it turns out to be a dry year. It is also possible to further separate the year into periods such as winter/summer or the three periods used in this paper for modeling inflow based on the factor model.

6.3.3 Method 3

The two methods above are very strict in terms of their matching, and a year with high inflow does not necessary need to be a year with low prices. If we look at nationally aggregated inflow this is more correct, but for a local power plant this does not necessary have to be true. Occurrences of high inflow and high prices could occur if national aggregated inflow is low that year. In order to capture more of this randomness we could separate either every year as in method 2 or the whole set as in method 1 into equal sized groups of series based on their average inflow and price level. Each year could then be separated into e.g. 5 equal groups (extreme, high, normal, low, dry) based on their average inflow and price, and the different price scenarios be matched randomly with scenarios from the corresponding inflow group.

6.3.4 Method 4

Another procedure that could be attempted is to calculate average price and inflow for the series either for each year or the whole period and then assigning probability distributions to each series. Series with a high average price/inflow would be given a distribution with a high probability of sampling a high number. The price series would then be arranged in a descending order based on one sampling from the probability function of the individual price series. The opposite would be done for the inflow series and the series would then be matched based on the new order. As an example, the price series which drew the highest number from its probability distribution would be matched with the inflow series that drew the smallest number etc. A high probability of sampling a high number is given for series with high average value. The series are then arranged based on one sampling for each series based on their respective distribution. This way high inflow series will tend to be matched with low inflow series but there is still some randomness in how the scenarios are arranged. This would be a time consuming procedure when simulating many scenarios and also include the problem of assigning correct probabilities.

These decision rules should be adjusted to each producer since they might have different correlations with the national aggregated inflow. A producer with strong positive correlation between local inflow and the system price would tend to have high inflow in dry years when the prices are high and so the use of method 1 and 2 for instance would be wrong. They should therefore be adjusted to fit the specific case.

When running the optimization in this paper we have used method 1 to match the scenarios for the factor models for simplicity. The MatLab algorithm is however constructed in such a way that it is easy to include new matching rules/algorithms.

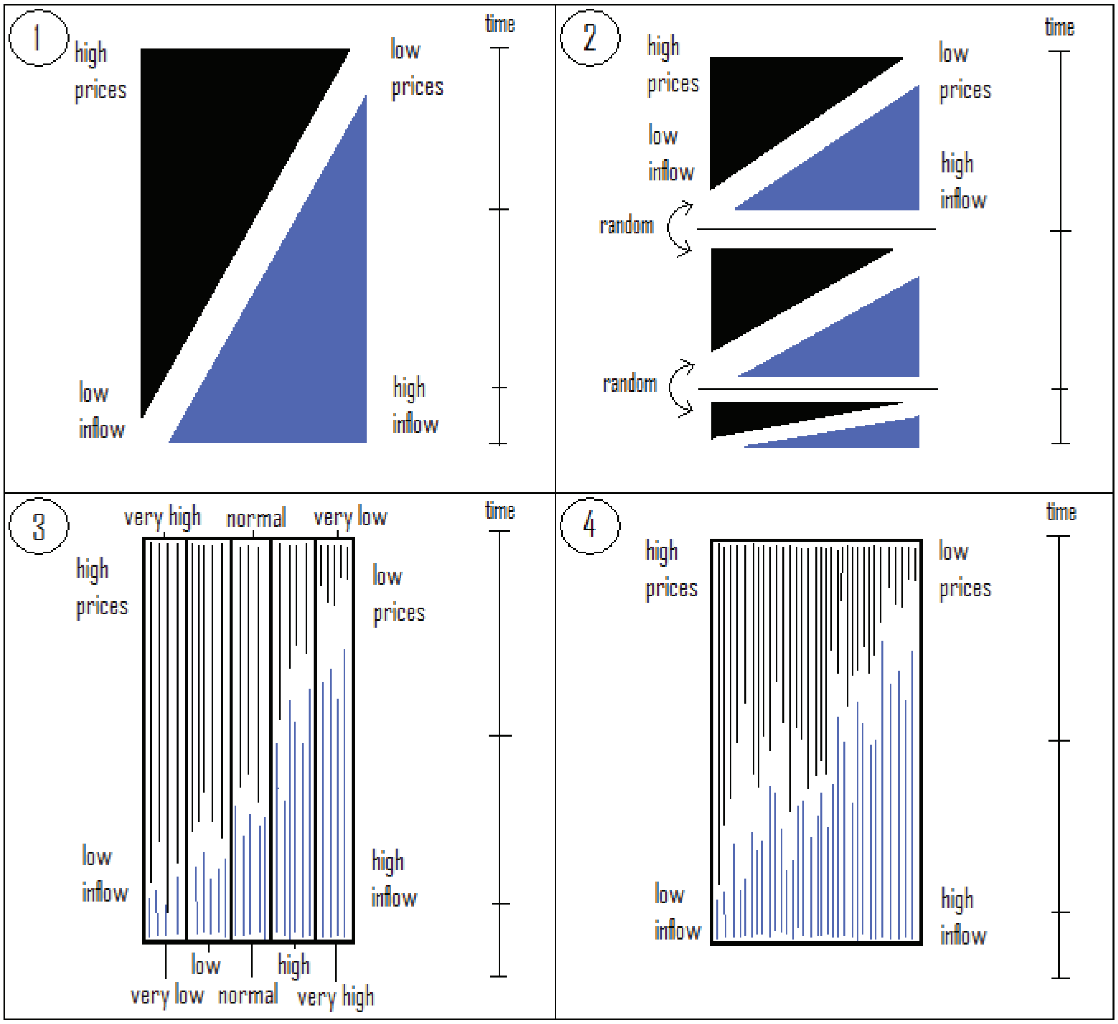


Figure 31. Illustration of the different matching methods described in this chapter

7.0 Simulation

To be able to run the optimization models described later in the paper, it is necessary to be able to generate scenarios for the different variables (inflow and price) and the different models (stochastic factor models and STAMP). This chapter briefly describes the procedure used in the paper for generating these scenarios.

7.1 Stochastic factor models

7.1.1 Inflow

Based on the estimation of the parameters of the model used (Divided Year Model), it is now possible to generate as many scenarios as needed to be able to run a stochastic optimization. This is done using MatLab.

$$P(t) = f(t) + X_t \tag{23}$$

$$f(t) = \alpha + \gamma \cos((t+\tau)\frac{2\pi}{356}) \tag{24}$$

$$X_t = X_{t-1}e^{-\kappa dt} + \varepsilon\sqrt{\frac{\sigma^2}{2\kappa}(1-e^{-2kdt})} \tag{25}$$

The first equation and second equations are presented earlier in the paper and will not be repeated here. The third equation however needs an elaboration. This is the stochastic term which provide the fluctuations around the deterministic term. The factor ε is a standard normally distributed random number (comes from an increment to a standard Brownian motion) and makes the simulations fluctuate around the deterministic term.

Figure 32 shows an example of the result from one simulation over three years using the Divided Year Model for power station 1. For illustrative purpose the realized inflow between 2004 and 2006 is also included.

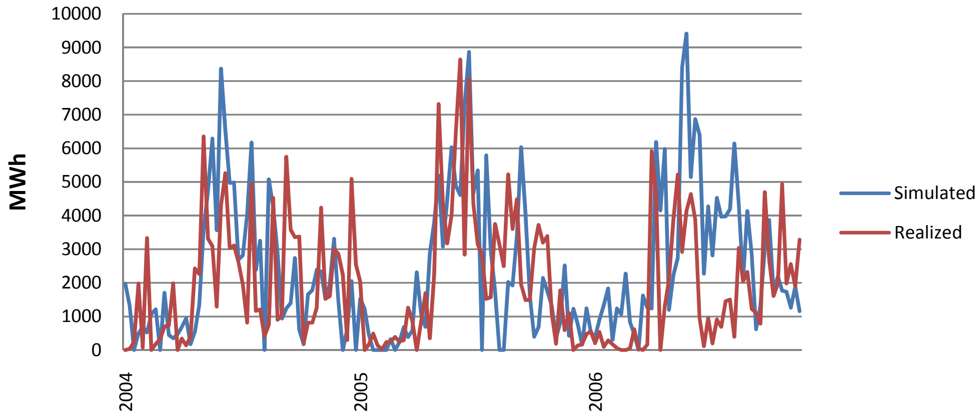


Figure 32. Simulation of inflow, Divided Year Model for power producer 1
The figure shows the result from one simulation over three years using the Divided Year Model and the realized inflow from 2004 to 2006 for power producer 1

It can be observed that the simulated inflow sometimes can be zero over a period of time. The reason for this is that the simulation result is automatically set to zero if the initially simulated result is negative, which can happen since the standard deviation of the simulation is big.

Figure 33 shows the results obtained after 25 simulations over a period of 5 years, also this for the Divided Year Model for power station 1. The chart shows the average simulation result in each week together with the maximum and minimum simulated inflow for that particular week. Again we observe that the minimum simulated inflow is zero over long periods of time. Since the chart is generated using the minimum inflow for each week from a sample of 25 simulated inflow scenarios, the minimum graph is obviously zero over very long periods, especially during the winter.

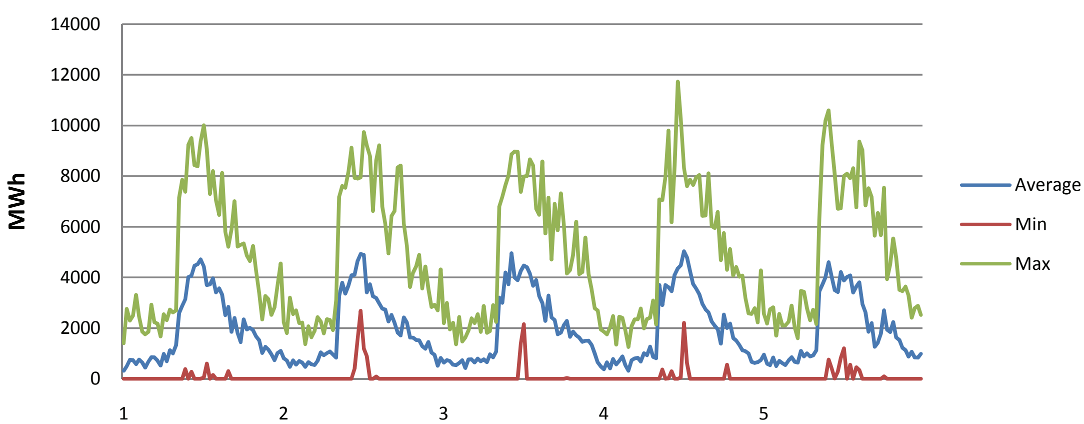


Figure 33. Inflow spread after 25 simulations, Divided Year model for power producer 1
The figure shows the result after 25 simulations over a period of 5 years using the Full Year Model for power producer 1. In addition to the average inflow, the maximum and minimum inflow for each week is also displayed.

7.1.2 Price

It is also necessary to make scenarios from the estimated numbers for the two factor price model. Figure 34 shows an example from running one simulation over three years, in addition to realized prices over three years periods from 1996 to 2006. The single simulation drawn shows that the two factor price model seems to capture some of the seasonalities in the electricity power prices. The abnormal high prices in years 2002 and 2006 do not seem to have too much impact on the model. The reason for this is that we use data from 1996 to 2006, and so the model are not that influenced by the high prices in 2002 and 2006.

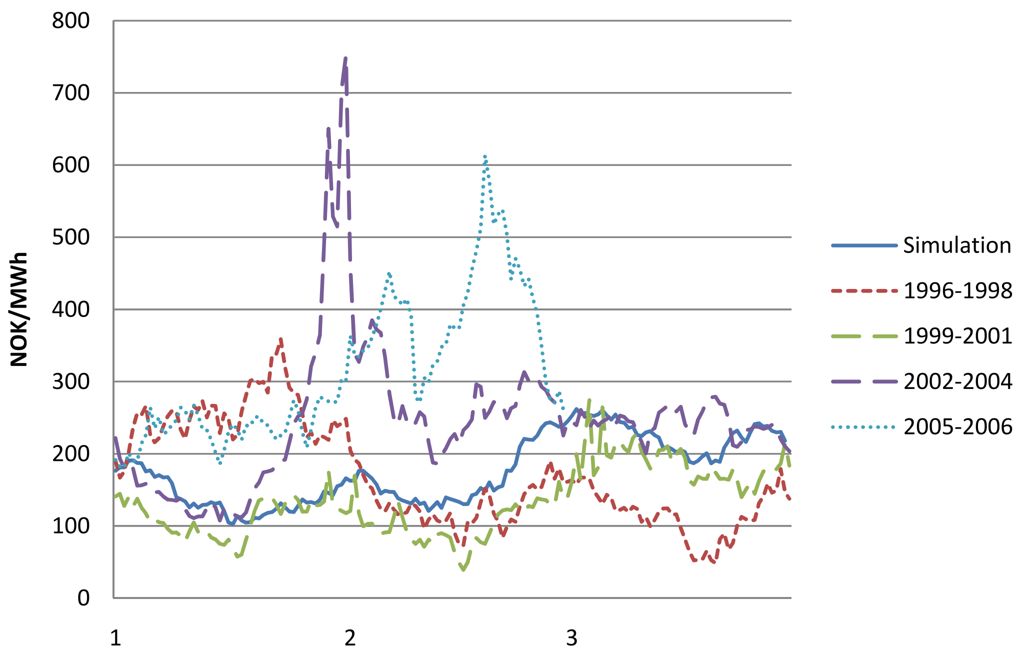


Figure 34. Simulation of price

The figure displays one simulation of the spot price over a period of three years. In addition realized price in the intervals 1996-1998, 1999-2001, 2002-2004 and 2005-2006 are shown

Figure 35 shows the result after 25 simulations for the price model over a three year period. The chart displays the average, maximum and minimum values. As can be seen from the graph, even after 25 simulations, none of the scenarios has a maximum value in magnitude of the realized prices in 2002 and 2006. Further it may look like the average price seems to be on a lower level than the realized data. Even though the graph is a result of only 25 simulations, it reveals that the price model may have some important weaknesses making it not that suitable for price simulations, as discussed in section 5.6.

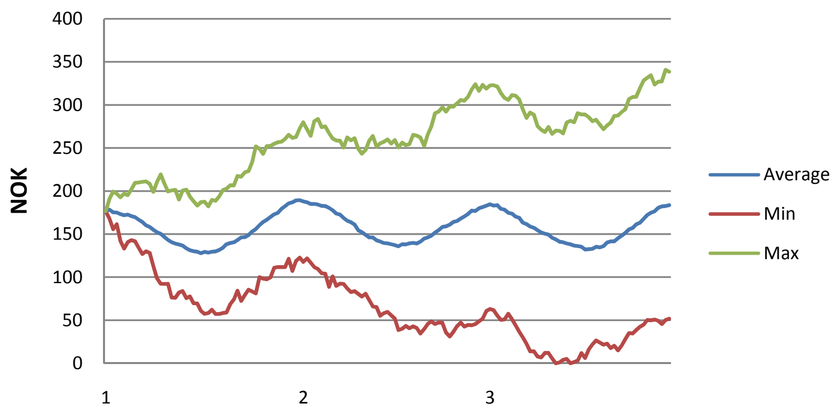


Figure 35. Price spread after 25 simulations
The figure shows the result after 25 simulations over a period of three years for the spot price. In addition to the average spot price, the maximum and minimum spot price for each week is also displayed.

7.2 STAMP models

7.2.1 Inflow

The forecast for inflow is given by the following equations, where the first is the initiation.

$$Y_1 = \exp(\mu + \gamma_1 + \rho_{AR}(\ln(Y_{2003(52)}) - \mu - \gamma_{52})) \tag{26}$$

$$Y_{t+1} = \exp(\mu + \gamma_{t+1} + \rho_{AR}(\ln(Y_t) - \mu - \gamma_t)) \tag{27}$$

The level μ, and the seasonal component γ_t are fixed as are the AR(1) coefficient ρ_A. The previous value of the Inflow in the first period of forecasting is given by the inflow in period 2003(52).
Simulations have been made by simulating the distribution of the irregular term using the standard error of the estimation. The inflow scenarios are first generated for each producer for the given time horizon as shown for 10 scenarios for the three producers in the figure below (Inflow measured in MWh/week).

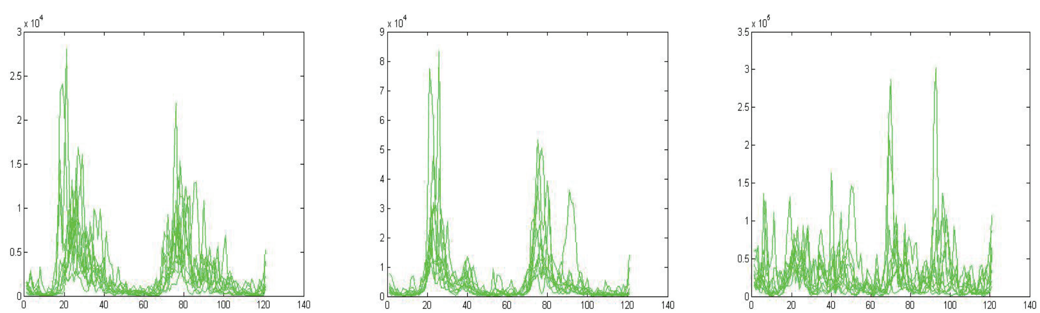


Figure 36. Inflow scenarios for producer 1, 2 and 3

7.2.2 Price

The forecast for price is given by the following equations, the first being the initiation.

$$S_1 = \mu + v + \gamma_1 + \beta x_1 + \rho_{AR}(S_{2003(52)} - (\mu + \gamma_{52} + \beta x_{2003(52)})) \tag{28}$$

$$S_{t+1} = \mu + v \cdot (t+1) + \gamma_{t+1} + \beta x_{t+1} + \rho_{AR}(S_t - (\mu + v \cdot t + \gamma_t + \beta x_t)) \tag{29}$$

The level μ, trend v_t and the seasonal component γ_t are fixed and the explanatory variable x_t is created by aggregating the related simulated inflow over 26 weeks. Price simulations have been made in the same way as above by using the standard error of the estimation to simulate the irregular term. Simulations of 20 fan scenarios for price are given below for producer 1, 2 and 3 respectively (price in NOK per MWh):

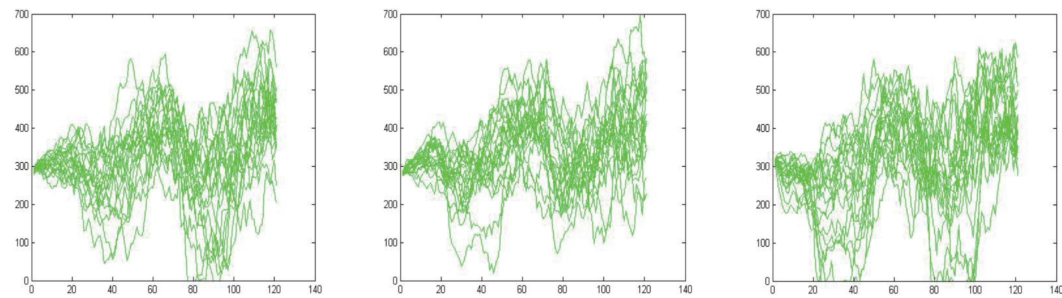


Figure 37. Price scenarios for producer 1, 2 and 3

8.0 Optimization

The main purpose of long term hydropower planning is to utilize the water in the reservoirs in an optimal way, satisfying all relevant constraints. Long term scheduling of the reservoirs represents the strategic management of the individual power plant's own resources in interaction with the whole power system considering uncertainty in inflow, prices, demand etc. After the deregulation of the power market the objective of the producers is to maximize their profit by utilizing their water at the best possible time (when prices are high).

When considering hydro power scheduling there are many relationships that are non-linear such as the plant head to reservoir level, discharge to produced kWh, start/stop costs etc. In long term planning this can be simplified by using linear relationships instead or other techniques to make the problem solvable within reasonable computational times. A constant efficiency coefficient is for instance assumed in this paper, keeping the amount of water needed to produce one kWh fixed. This is a common assumption in long term hydro power planning.

In this paper, only producers with a single reservoir are considered, but the model considerate topology and can incorporate connections between several other power plants. The following optimization models are based on the models used in Winnem (2006) and Pedersen (2006) and Bjørnsgard and Hauge (2007). At the end of the planning period in this model the water in the reservoirs have no value according to this model. To prevent the model from emptying the reservoirs (which would have been done since we are maximizing profits) a minimum ending reservoir is set.

8.1 Deterministic Model

A deterministic model for optimization of water discharge for a power producer will assume that inflow and prices are known with certainty over the whole planning period.

Set:

A : set of planning periods A=(0,1,...,T)

Index:

t : time preiod

Parameters:

Π_t : electricity price in period t

ψ_t : water inflow

η : energy efficiency coefficient

$M_{\max}$: maximum reservoir level

$M_{\min}$: minimum reservoir level

M_0 : initial reservoir level

M_T : ending reservoir level

$Q_{\max}$: maximum water discharge

r : interest rate

Variables:

V_0 : Present value of total production over the planning period

m_t : ending reservoir

s_t : spillage in period t

p_t : energy produced in period t

q_t : discharged water in period t

Objective function:

$$V_0 = \max_{q_t, m_t, s_t} \sum_{t=0}^{T} \frac{\Pi_t}{(1+r)^t} p_t \tag{30}$$

Constraints:

$$\begin{aligned}
&p_t = \eta \times q_t && ,t \in A \\
&m_t - m_{t-1} + q_t + s_t = \psi_t && ,t \in A \\
&M_{\min} \le m_t \le M_{\max} && ,t \in A \\
&q_t \le Q_{\max} && ,t \in A \\
&q_t, s_t \ge 0 && ,t \in A
\end{aligned}$$

The objective function in equation 30 maximizes the discounted income from each time period. The first constraint states the linear relationship assumed between discharge and energy production and the second the reservoir balance where change in reservoir is due to inflow minus discharge and spillage. The following equality then gives the reservoir level constrains. Water discharge is restricted to be lower than max discharge, and spillage and discharge is set bigger than zero in the last constraints.

8.2 Stochastic model

A stochastic model for optimization of water discharge for a power producer will assume that inflow and prices are stochastic variables. The constraints in the stochastic model are equal to the deterministic case only with stochastic variables. The objective function of the optimization is now the expected value of the discounted income from all the periods.

$$V_0 = \max_{q_t, m_t, s_t} \mathrm{E}\left[\sum_{t=0}^{T} \frac{\Pi_t}{(1+r)^t} p_t\right] \tag{31}$$

8.3 Deterministic equivalent

If we assume that the stochastic variables can be represented by discrete probability distributions, we can model the stochasticity with a scenario tree. In this way the different states the stochastic variable can take are represented by nodes in the tree at different time steps. The stochastic model can then be solved by using standard linear programming on the deterministic equivalent. The following modifications have to be applied to the deterministic model

New index:

n : node index

New parameters:

N : number of nodes

n_T : number of nodes in period t

t_n : time period at node n

$\alpha_{(n,k)}$: index of predecessor node n in time period t-k

P_n : probability that the state in node n will occur

Objective function:

$$V_0 = \max_{q_n, m_n, s_n} \sum_{n \in N}^{T} P_n \frac{\Pi_n}{(1+r)^{t_n}} \times p_n \tag{32}$$

Constraints:

$$
\begin{aligned}
& p_n = \eta \times q_n && , n \in N \\
& m_n - m_{\alpha(n,1)} + q_n + s_n = \psi_n && , n \in N \\
& M_{\min} \le m_n \le M_{\max} && , n \in N \\
& q_n \le Q_{\max} && , n \in N \\
& q_n, s_n \ge 0 && , n \in N
\end{aligned}
$$

The objective function in equation 32 is the sum of the discounted income in each possible state multiplied by the probability of being in this state. The reservoir balance in the second constraint makes sure that the change in reservoir level is linked to the reservoir level in the predecessor node.

8.4 Scenario tree

In stochastic programming it is convenient to represent the dynamics of the stochastic variables in a scenario tree. The scenario trees then represent possible future states and realizations of the stochastic variable. When modeling stochastic variables such as inflow and price, many different fan scenarios can be generated based on simulations from the models. The fan scenarios can then be bundled together to create a scenario tree as shown in figure 38 below.

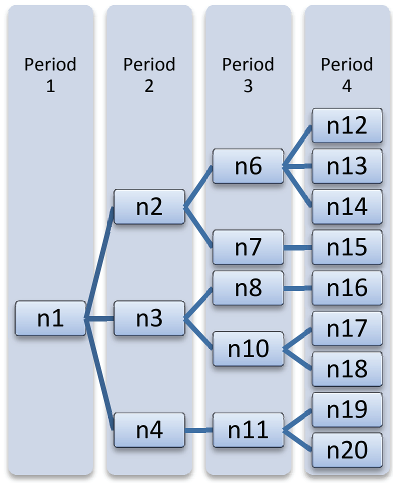


Figure 38. Scenario tree example with four periods and 20 nodes

The scenario tree given in figure 38 above have 4 different time steps and nine different scenarios, given by the total number of end nodes in the final period. The number of nodes at each time step represents all the possible states the stochastic variable can take at that particular time. As time unfolds more uncertainty will be included and the number of nodes in each time step increase. The different length of the time periods can change. The more frequent time steps and nodes the more detailed the description of possible states the variable can have, but the computational time increases quickly to unfeasible dimensions. When bundling the scenarios together several nodes coincide and so it is possible to reduce the number of scenarios given by the fan scenarios and still have good representation of the possible states of the stochastic variable.

Heitsch and Römich (2005) describe two models that can be used to generate scenario trees from a set of fan scenarios. The fan scenarios are bundled together and coinciding nodes are deleted selecting only the most representative nodes, thus generating a scenario tree as described in the article. Two models are presented, one backward and one forward model, and an illustration of the forward method is given below, taken from Pedersen (2006).

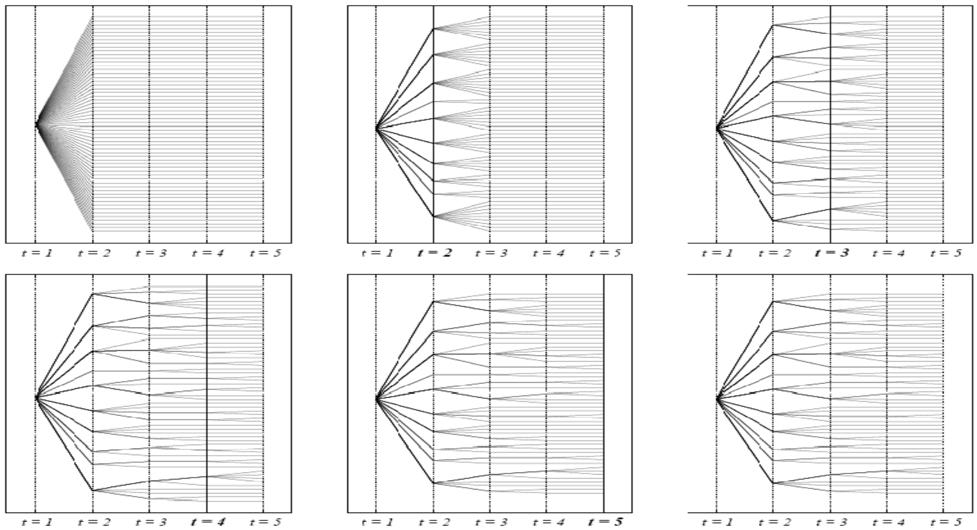


Figure 39. Generation of scenario trees from fan scenarios
The figure shows the procedure developed by Heitsch and Römich (2005) when generation a scenario tree from a set of fan scenarios. The illustration is taken from Pedersen (2006).

The program Scenred is developed by Heitsch, and is a C++ based program constructing scenario trees from sets of fan scenarios. In this paper the program is used to generate scenario trees from inflow and price fan scenarios. When running the program, three different parameters have to be set depending on how the reduction is to be made. The relative probabilistic tolerance ε_p is set equal to 0,8 and measures the distance between the original and the approximated probability distribution. The relative filtration tolerance ε_f is set equal to 0,85 measuring the filtration or information distance, and the tree construction parameter q affecting the tolerance at each branching point is set equal to 0,65. More information on how these parameters are selected is given in the article by Heitsch and Römich. Figure 40 below shows an example of a reduced scenario tree generated from initially 1000 fan scenarios for price and inflow.

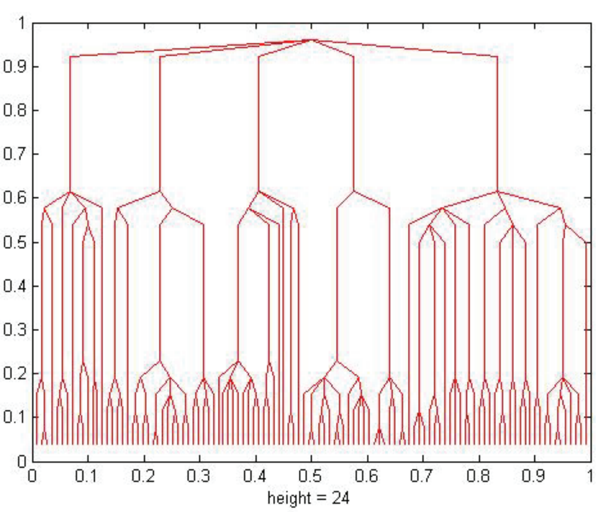


Figure 40. Example of scenario tree generated in Scenred

The use of Scenred to create aseario tree from inflow and price scenarios does have some drawbacks worth mentioning. The mean reverting properties of both the price and inflow series described earlier is not included in a satisfacory way. The figure below show a scenario tree with extreme events at the top and bottom branch. When standing at a node along one of these brances, i.e.node a in the figure, the probability of going to a "less extreme" node should be much higher than going to a state with even higher price due to the mean reverting nature of both electricity price and inflow. By incorporating this effect we would shift the probability density of the different scenarios closer to the center brances of the scenario tree.

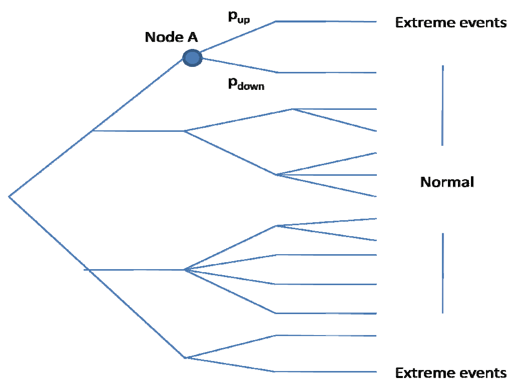


Figure 41. Generation of scenario tree in Scenred
The figure displays how different fan scenarios are selected along "extreme" branches in the scenario tree

When generating fan scenarios in Matlab we will of course end up with more scenarios at the center of the tree, as would happen in any scenario tree, but we argue that we should have an even higher density of scenarios in the scenter of the tree. The program is however still used assuming that the resulting scenario tree is adequate to represent the actual set of scenarios.

8.5 Flow chart optimization process

The flow chart of the computation process used in this paper is given in the figure below. It lists the programs used in the boxes together with a short descriptions of the tasks performed in the respective programs.

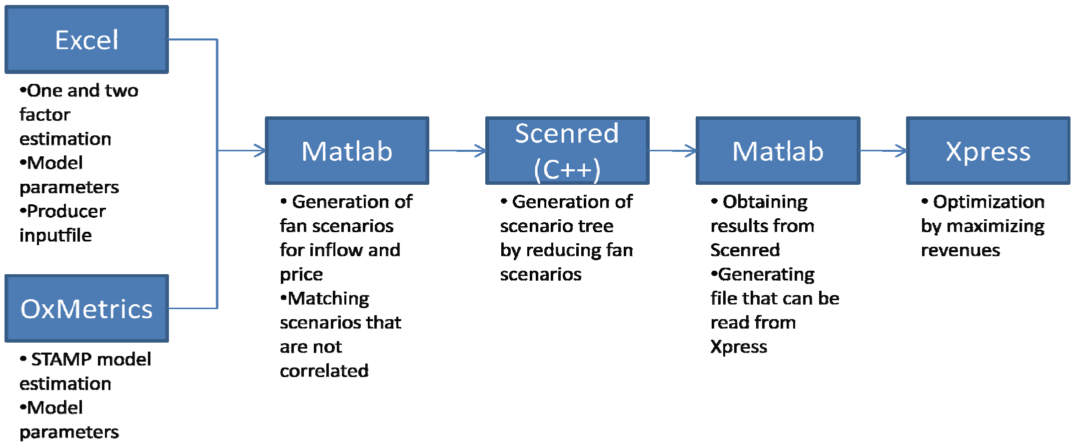


Figure 42. Flow chart of computation process

The figure displays a flow chart of the underlying computation process used in this paper. Each box represent a software program and a description of how the software is used is described

The different models used to describe the dynamics of the inflow and price series are estimated in Excel for the one- and two-factor models from Lucia and Schwartz (2001), and in OxMetrics for the state space models. The model parameters and all the input from the different producers are gathered in an Excel input file. The input file is read by Matlab, and the parameters from the different models are used to simulate price and inflow series resulting in fan scenarios. The inflow and price series are then matched together based on a matching rule described in chapter 6, except from the STAMP models where price and inflow are already linked. Matlab then generates a text file that can be read by Scenred, which reduces the amount of scenarios and creates a scenario tree. The output file is then read by Matlab and a file that can be read by Xpress is generated containing the scenarios from the scenario tree and all the relevant input parameters from the individual producers. The deterministic equivalent of the stochastic optimization problem described in chapter 8 is then run in Xpress MP. The optimization model used is the one developed by Pedersen (2006).

8.6 Value of stochastic solution

Measuring to which extent a stochastic approach is superior to a deterministic one is the underlying principle of the value of stochastic solution (VSS), Wallace (1999).

The value of stochastic solution measures the expected increase in the objective function, when uncertainty is explicitly incorporated in the optimization model. It is found by solving the stochastic model's corresponding deterministic model. More formally VSS is defined by the following formula:

$$VSS = ESS - EMV \tag{32}$$

Here ESS is the expected objective value of stochastic solution, while EMV is the expected objective value of the mean value solution (deterministic).

9.0 Analysis

The flowchart in figure 42 shows how the different chapters in this paper are linked. As can be seen from the figure, all the models developed and estimated are finally run through Xpress MP. Here the scenario tree based on the stochastic factor models and STAMP model is optimized by maximizing revenues. When generating the fan scenarios, inflow and price for 122 weeks are simulated due expected reservoir levels at the end of the simulation period. The initial scenarios all have weekly resolution, but the program procedure aggregates the simulations for the last weeks in to monthly and quarterly time steps to shorten the computational time for the analysis. For the factor models it is decided, based on previous findings, to use only the Divided Year Model when modeling the inflow. For all the analysis conducted, 1000 fan scenarios are generated and this is repeated 20 times. The number of fan scenarios is chosen based on the analysis from Bjørnsgard and Hauge (2007) and the computational time required for the calculation process.

Optimal value

The deterministic equivalent of the stochastic long term planning problem faced by hydro power producers is run. When running the optimization we get different values for the objective function in Xpress due to the simulated price and inflow scenarios that differ each time. The average value and the standard deviation for the different producers and models are given in the table below.

Table 14. Average optimal value and standard deviation for factor and STAMP models

		Producer 1	Producer 2	Producer 3
Factor	Average	1,02E+08	2,31E+08	1,01E+09
Model	St.dev	1,24E+06	2,33E+06	2,03E+7
Stamp	Average	2,34E+08	5,28E+08	2,51E+09
Model	St.dev	3,46E+06	1,23E+06	3,47E+07

The average value of optimal solutions for the two models differs due to the fact that they predict different price and inflow scenarios. Ideally they should be the same. The price forecast from the STAMP model is higher than the Factor model and so the optimal value is always higher for the STAMP model. However the ratio is constant and so this has no effect on the comparison between the different power producers. This drawback is also discussed in section 5.6.

It can also be seen that producer 3 has the highest revenues. This is as expected due to the fact that this power station has the highest average inflow over the year, the largest reservoir and the highest energy equivalent. Producer 2 experiences higher expected revenues than producer 1, even though average inflow to producer 1 is slightly higher than inflow to producer 2 (~15% higher). This can be due to producer 2 having approximately 2,5 times the generation capacity installed, giving this producer the possibility to produce more when prices are high and hence giving a total higher expected revenue. The reservoir is also a bit larger (in terms of GWh) and since we are depleting it down to 20% by the end of the period, this will affect the optimal value.

Natural Hedge

The optimization of the deterministic equivalent is run for all the producers with and without the matching of scenarios as described in chapter 6 (where the lowest inflow scenario is matched with the scenario with highest average value according to method 1). This can only be done for price and inflow based on the factor model, since the STAMP models are already matched. The optimal expected value for the resulting scenario trees was found to be fairly equal, the average being slightly higher for the unmatched case as seen in the table below.

Table 15. Optimal value and standard deviation for matched and unmatched scenarios

Matched scenarios

	Producer 1	Producer 2	Producer 3
Average	1,02E+08	2,31E+08	1,01E+09
St.dev	1,24E+06	2,33E+06	2,03E+07

Unmatched scenarios

	Producer 1	Producer 2	Producer 3
Average	1,05E+08	2,40E+08	1,03E+09
St.dev	6,21E+05	2,50E+06	1,92E+07

For the individual scenarios in the scenario tree we would expect the variation in the resulting revenues to be varying more in the unmatched case. To check this we have found the production and price in each node in the different scenarios, and calculated the revenues in each node. By summing all the revenues in each node for the different scenarios we can calculate the resulting revenues for each inflow/price scenario in the reduced scenario three. The variation in revenues for the different scenarios is calculated for both the matched and the unmatched case and the result from producer 2 is given below.

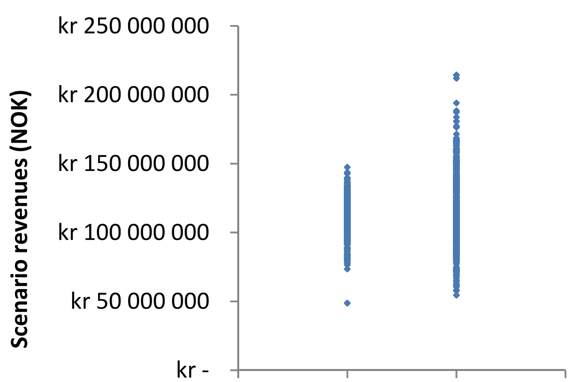


Figure 43. Variation in scenario revenues matched and unmatched scenarios

This power plant show the largest variation in revenues and illustrates best the variation we can get if power producers do not include the correlation of price and inflow in the modeling. This analysis also shows the benchmarks for what can be expected from the natural hedging in hydro power

production. In this example the matching done represents an unrealistically high correlation between price and inflow, and so the realistic set of scenario revenues will probably be somewhere between the two extreme cases in the figure above, depending on the individual power plants.
By running several optimizations for the matched and unmatched case we obtain the following standard deviations for the different scenario revenues.

Table 16. Average standard deviation for matched and unmatched scenario revenues

	Producer 1	**Producer 2**	**Producer 3**
Matched	6,91E+06	1,29E+07	6,19E+07
Unmatched	7,30E+06	2,50+07	6,72E+07
% increase	6 %	94 %	9 %

The results show higher variance for all the unmatched cases as expected, with the highest value for producer 2, once again showing the importance of the natural hedging in hydro power production.

10.0 Conclusion

In this paper we have used two different models to estimate the dynamics of electricity prices and inflow to hydro power producers. The different scenarios have then been matched together based on different methods to create more realistic scenarios in compliance with the correlation we see in the Nordic power market. The long term planning problem faced by hydro power producers in terms of optimal dispatch of the water in the reservoirs, given stochastic inflow and price, is simplified to a deterministic equivalent which is implemented in Xpress. This optimization problem is solved for the different producers and models estimated.

When modeling the relationship between price and inflow, the weekly correlation have been analyzed, both for real values and between the error terms of our models. These correlations were found to be highly unstable for both, and so using the weekly correlation in our models are rejected. The correlation between a 26 week aggregated inflow and price turned out to be stronger, especially for producer 3 and this was also found when estimating the STAMP models, where the 26 week aggregated inflow was used as an explanatory variable to describe the electricity price series. Different ways to correlate the price and inflow scenarios are presented and one model is used in this paper as described in chapter 6.

The negative correlation between price and inflow creates a natural hedge for the different power producers which we have illustrated by calculating the variance of the expected income from a producer when incorporating the correlation and not. The revenues for the respective scenarios in the scenario tree have higher variation when the price inflow scenarios are uncorrelated versus the correlated case. The correlation narrows the set of possible outcome, thus reducing risk in uncertain revenues for the power producers.

The degree of correlation between local inflow and total national inflow, and indirectly the system price at Nord Pool, determines the degree of the natural hedge. Parallels can be drawn to the corn example described in chapter 2. The corn producers in the main production areas are stronger correlated to the corn price since bad crop years and drought tend to affect many producers, thus creating shortage for corn which in term lead to higher prices that will offset some of the financial losses due to less volume produced. The same apply for power producers, where a producer with inflow that is highly correlated to the system price will experience a stronger natural hedge than a producer with less correlation. Based on these findings, the inflow can be used as a benchmark for risk for a hydro power producer. If the hydro power producer's local inflow is highly correlated to the national inflow, this producer can be said to benefit by the natural hedge characteristics in the Nordic power market. If the producer has an inflow with low negative or positive correlation with the national inflow (e.g. due to geographical/climatically differences), this producer faces a risk other producers are hedged against. In years with low prices (due to high national inflow) this produces may experience low inflow, and hence the risk of financial distress. Of course, the absence of natural hedge for one hydro player may also turn out to be profitable in years with low national inflow and high local inflow. However, extreme events happen rarely and the income for hydro producers will usually be clustered around the expected scenarios. This is exemplified in table 15, where it can be seen that the average expected income from the uncorrelated case is only slightly higher than for the correlated situation. Since the natural hedge is a hedge with no direct up-front cost and companies prefer lowest possible risk to a corresponding payoff, it is to be expected that natural hedge is highly preferred characteristic for a power producer.

The presence of natural hedge characteristics for hydro producers in Norway may open for opportunities for electricity producers using e.g. gas or coal. The resource base for such producers is more stable, which is something that they can exploit as long as the Nordic power prices are strongly affected by hydro power producers. As Norway gets more and more interconnected to markets dominated by thermal power producers, some of the natural hedge characteristics faced by power producer may disappear.

11.0 Further work

This paper has explored a very fundamental part of hydro power scheduling for power producers in the Nordic power system, namely the relationship between electricity price and inflow. In relation to this paper there are several other aspects that potentially can give fruitful contributions and should be further explored. Among these are:

- The different matching rules described in chapter 6 can be tested for the different power producers
- The models described in this paper could be used on other power plants
- More analysis could be done on the existing power plants as done in e..g. Bjørnsgard and Hauge (2007)
- Better scenario generation, both in relation to the fan scenarios created and the use of Scenred to create reduced scenario trees
- The importance of inclusion of uncertainty (inflow/prices) in the hydro power scheduling procedure (Value of stochastic solution)
- Analysis of stability of solutions
- Hedging analysis (delta-, gamma-, vega hedging etc.)

Bibliography

Aas, Kjersti, 2007. A model for Simulation of Nordic Electricity Spot Prices. Presentation, conference for Modeling and Measuring Energy Risk on the Nord Pool Market, Oslo.

Bjørnsgard, Mari and Hauge, Linn K., 2007. Case studies in Hydropower Scheduling. Master's thesis, NTNU

Brooks, Chris, 2002. Introductory Econometrics for Finance. Cambridge University Press.

Bøckman, T., Fleten, S.-E., Juliussen, E., Langhammer, H.J., Revdal, I., 2006. Investment timing and optimal capacity choice for small hydropower projects. Working paper NTNU (accepted for publication in European Journal of Operational Research (2008)).

Commandeur, Jacques J. F. and Koopman, Siem J., 2007. An introduction to State Space Time Series Analysis, Oxford University Press.

Cuthberson, Keith, and Nitzsche, Dirk, 2004. Quantitative financial economics, 2nd edition. John Wiley and Sons, New York

Diebold, Francis X., 2007. Elements of Forecasting, 4th Edition, University of Pennsylvania.

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

Doorman, Gerard L., 2007. Hydro Power Scheduling. NTNU.

Energy Power, 2007, http://www.energyandpower.com/

Finance wise, 2007, www.financewise.com

Fleten, S.-E., Tomasgard, A., and Wallace, S. W., 2001. Produksjonsplanlegging og risikostyring I et deregulert kraftmarked med finansielle instrumenter. *Magma: Tidsskrift for økonomi og ledelse*, 4(5), pp. 22-33

Fleten, S.-E., Wallace, S. W., Ziemba, W. T., 2002. Hedging electricity portfolios via stochastic programming. In C. Greengard and A. Ruszczynski (eds.), Decision making under uncertainty: Energy and power, Springer-Verland, New York, pp. 71-94

Harwood, John, 1991. Economic Research Service, Local Government, Community, and Business Working Group Draft Report

Heitsch, Holger and Römich, Werner, 2005. Scenario tree modeling for multistage stochastic programs. Homboldt-university Berlin, Institute of Mathematics.

Hull, John C., 2006. Options, futures and other derivatives, 6th edition. Pearson Prentice Hall.

Hvarnes, Håvard, 2007, Presentation in TIØ01, Autumn, Elkem Energihandel

Knittel, Christoper R., and Roberts, Michael R., 2005. Financial models of deregulated electricity prices: An application to the California Market. *Energy Economics*, 27(5), pp 791-817

Koopman, S. J., Harvey, A. C., Doornik, J. A., Shepard, N., 2006. Stamp 8. Timerlake Consultants Ltd.

Krossøy, K., Torgersrud, P. C. L., 2004. Real options analysis of renewables projects. Project Thesis, NTNU

Lucia, Julio J. and Schwartz, Eduardo S., 2002. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5:5-50.

Mo**,** B., Gjelsvik, A,. Grundt, A., and Kåresen, K., 2001 Optimization of Hydropower Operation in a Liberalized Market with Focus on Price Modeling. Power Tech Proceedings, Porto, Portugal

Mo**,** B., Gjelsvik, A,. Grundt, A, 2001, Integrated Risk Management of Hydro Power
Scheduling and Contract Management. Power Tech Proceedings, Porto, Portugal

Mo, Birger. Meeting, 29.oct 2007

McDonald, Robert L., 2006. Derivatives Markets. Boston, Massachusetts: Pearson Adison Wesley.

Nord Pool, Annual report 2006, www.Nord Pool.com

Nord Pool. Ftp-server. Accessed 19.oct 2007

Pederse, Martin Geving, 2006. Hydroelectric power scheduling and risk management. Master Thesis, NTNU

Seppi, Duane J., 2002. Risk-Neutral Stochastic Processes for Commodity Derivative Pricing: An Introduction and Survey. Carnegie Mellon University.

Wallace, Stein W., 1999. Decision making under uncertainty: The art of modeling. NTNU, Trondheim.

Winnem, Marcs, 2006. Hedging hydroelectric generation. Master Thesis, NTNU

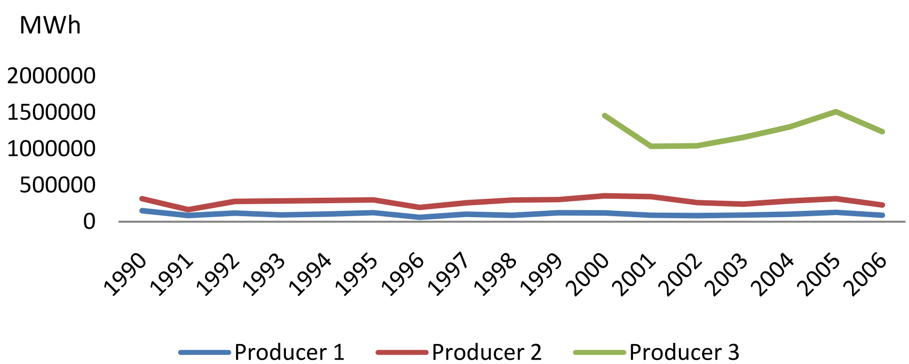
List of figures

List of tables

Appendix

A. Table and graph of annual inflow to the power producers

	Annual inflow in MWh		
Year	Producer 1	Producer 2	Producer 3
1990	146622	313263	
1991	80886	162788	
1992	114047	276350	
1993	90245	283125	
1994	101857	288813	
1995	120248	297300	
1996	56451	193525	
1997	99215	254675	
1998	83974	293775	
1999	116966	300713	
2000	116135	352088	1455600
2001	85873	342163	1033240
2002	79970	258950	1039970
2003	86682	239375	1156490
2004	100179	282013	1300520
2005	123219	313788	1510240
2006	85415	227238	1235110



B. Descriptive statistics for the weekly system price (1993– 2006)

Descriptive statistics	*Total series*	Cold Seasons	Warm Seasons
Mean	189,155	200,580	173,260
Standard Error	3,740	4,670	6,014
Median	168,875	178,852	147,118
Standard Deviation	101,044	96,055	105,551
Sample Variance	10209,858	9226,604	11141,063
Kurtosis	3,585	5,235	2,208
Skewness	1,415	1,727	1,251
Range	732,463	713,689	594,110
Minimum	19,255	38,030	19,255
Maximum	751,719	751,719	613,366
Sum	138083,490	84845,532	53363,990
Count	730	423	308
Confidence Level(95,0%)	7,342	9,180	11,835

C. Estimated parameters for stochastic factor models

The table displays the estimated parameters for Full Year Model and Divided Year Model for the inflow.

	α	γ	τ	κ
Power station 1 (week 1-18)	1008,86	-546,10	-16,11	998,15
Power station 1 (week 18-40)	1185,46	3015,57	-32,49	65,72
Power station 1 (week 40-52)	1603,25	-1255,76	20,91	45,38
Power station 2 (week 1-18)	1235,42	482,95	-13,60	988,99
Power station 2 (week 18-40)	-1183,14	16131,13	-32,49	26,32
Power station 2 (week 40-52)	2172,27	-3435,69	19,88	37,73
Power station 3 (week 1-18)	45934,84	-39733,92	-14,14	73,23
Power station 3 (week 18-40)	44449,39	27337,49	-32,17	65,36
Power station 3 (week 40-52)	12195,98	-15039,17	19,71	61,79

The table displays the estimated parameters for the 26 weeks aggregate inflow model

	α	γ	τ	κ
Power station 1 (week 1-18)	48876,51	-30548,28	1,72	0,94
Power station 1 (week 18-40)	137195,63	108902,10	2,25	1,61
Power station 1 (week 40-52)	561685,04	193743,79	8223,25	0,87

The table displays the estimated parameters for the two factor spot price model

α*	κ	μ_ε*	α	γ	τ
-51,47563	0,0321	-0,0082	151,47	25,179	-2,1137

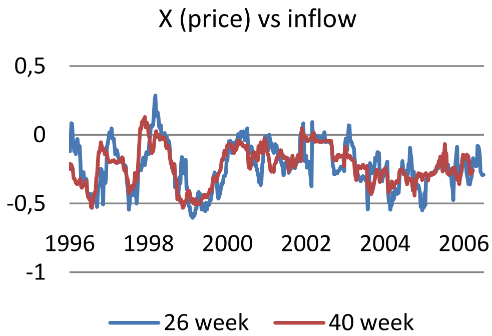
D. Correlation between time series for inflow and System price

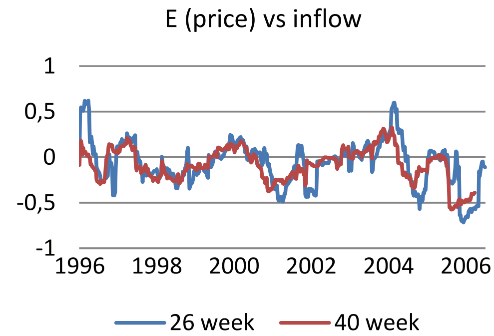
	System Price		
	Total	First half	Last half
Inflow 1	-0,227458339	-0,3999977	-0,233357267
Aggregated inflow 1	-0,109283388	-0,181517657	-0,146309599
Inflow 2	-0,207911041	-0,314570519	-0,2468168
Aggregated inflow 2	-0,126071425	-0,16615801	-0,239053318
Inflow 3	-0,217060664	-0,326581598	-0,489530364
Aggregated inflow 3	-0,447737166	-0,654477445	-0,474232708

E. Rolling correlation window for the different factor models

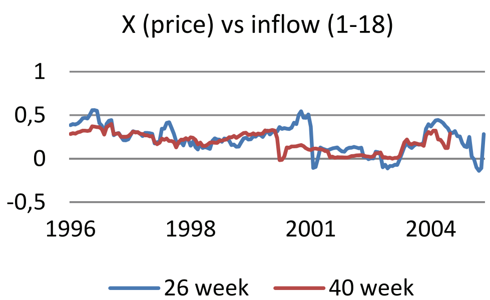
Producer 1

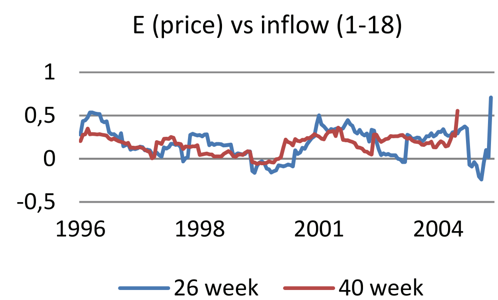
Inflow Full year model

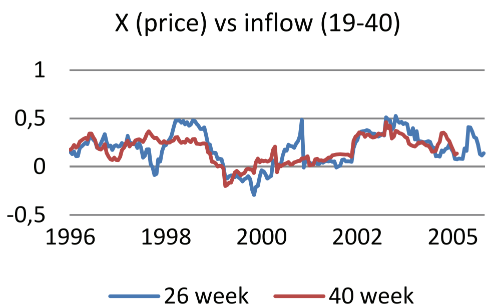


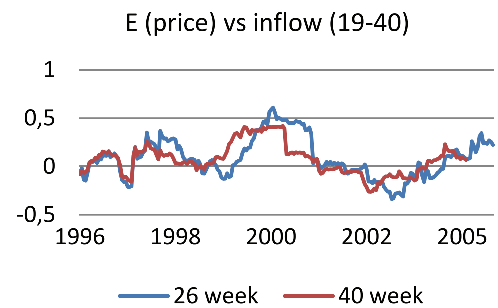


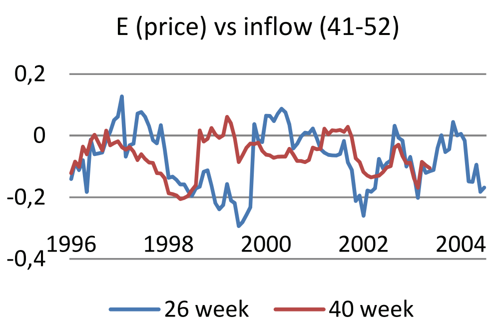
Inflow Divided year model

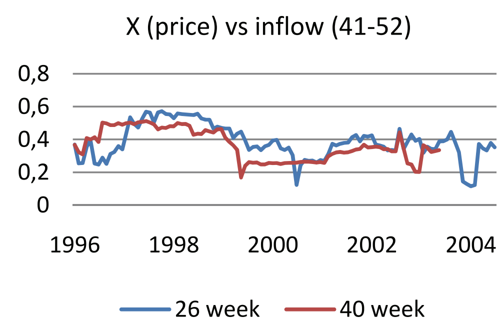




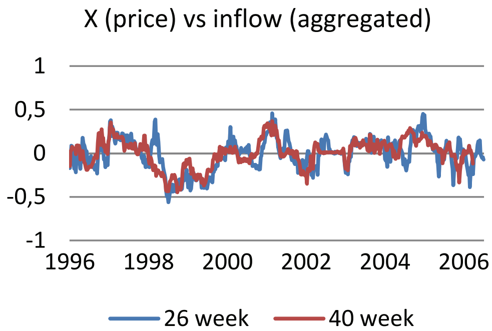


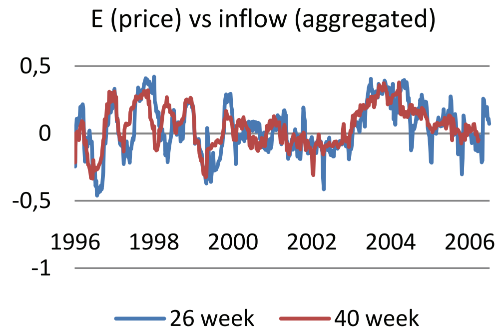






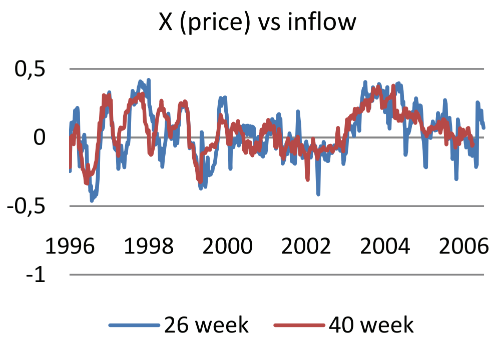
Aggregated inflow model

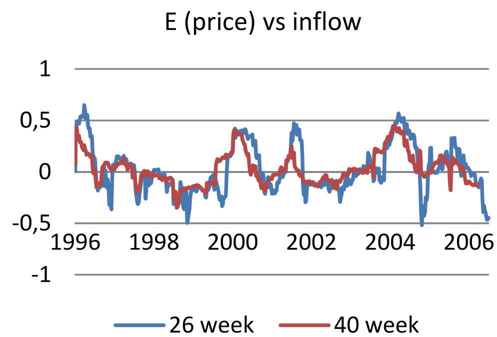




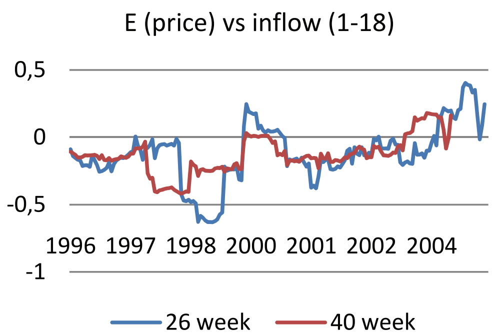
Power station 2

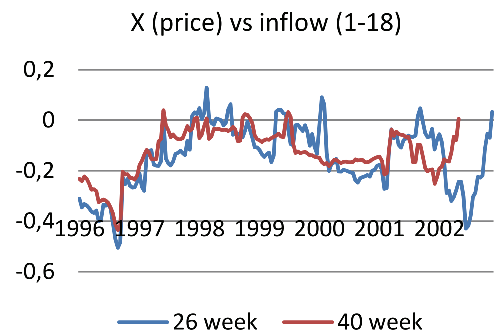
Inflow Full year model

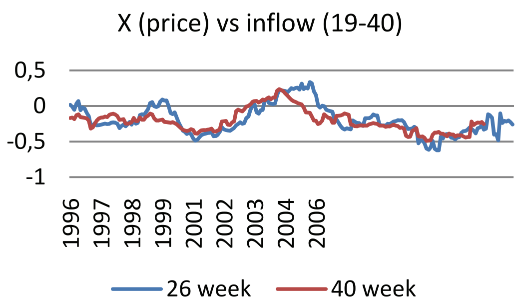


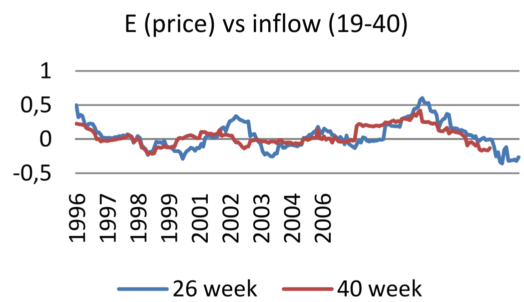


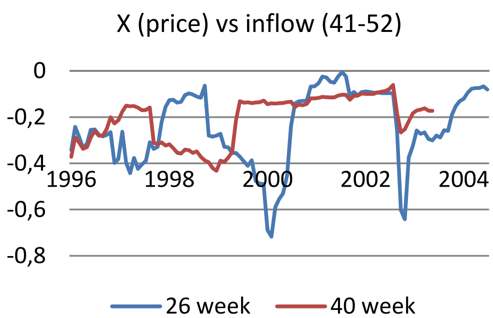
Inflow Divided year model

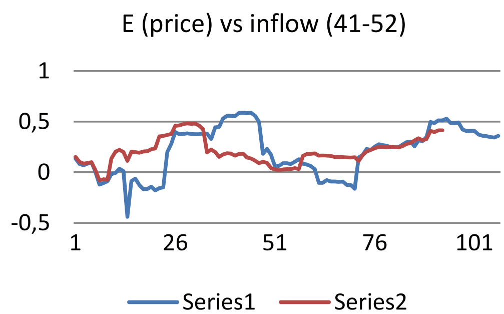




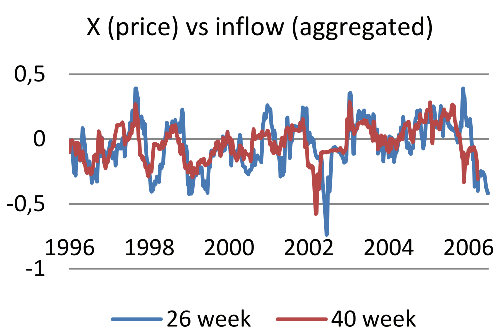


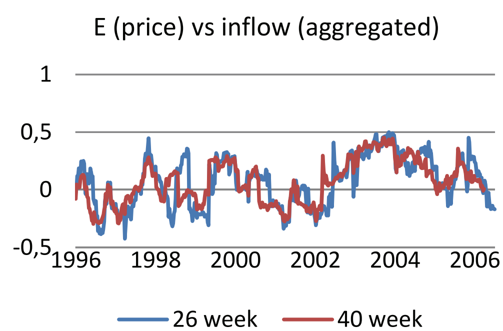






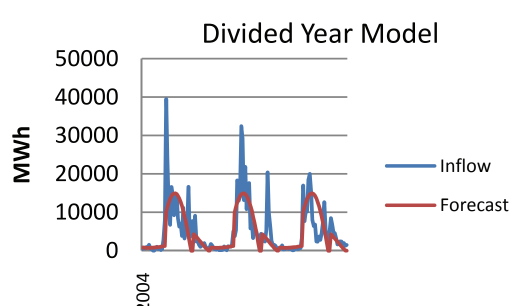
Aggregated inflow model

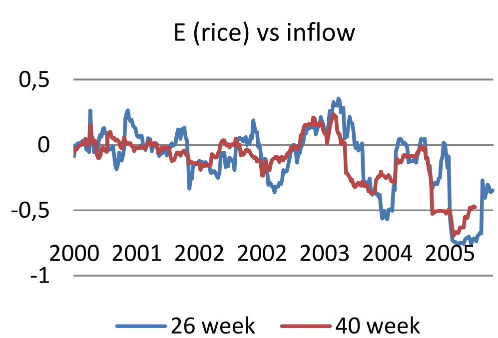




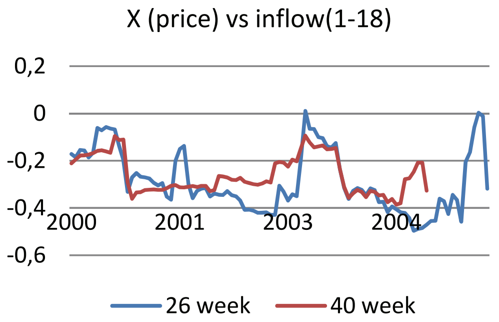
Power station 3

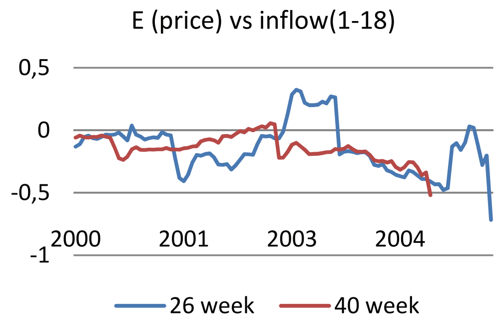
Inflow Full year model

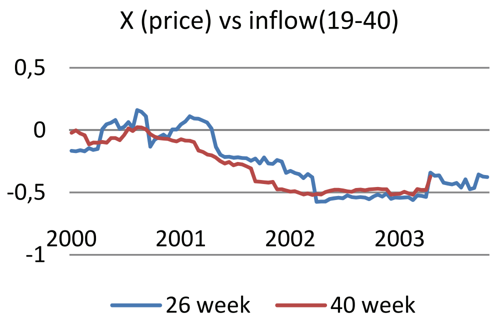


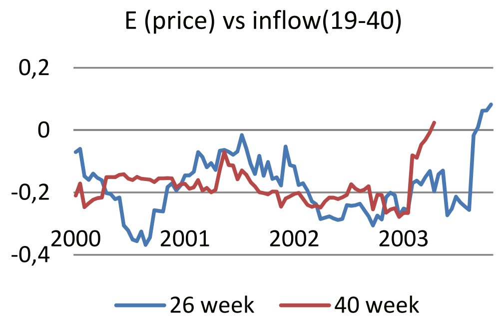


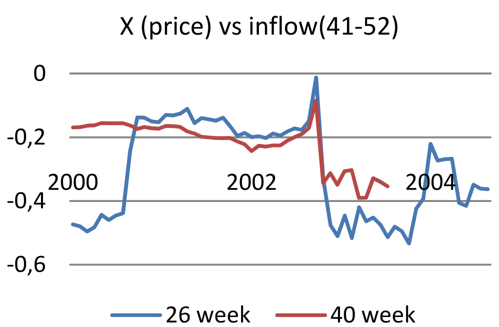
Inflow Divided year model

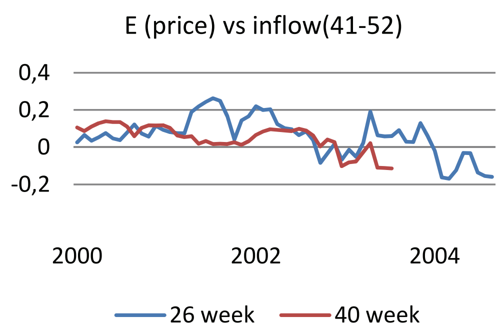




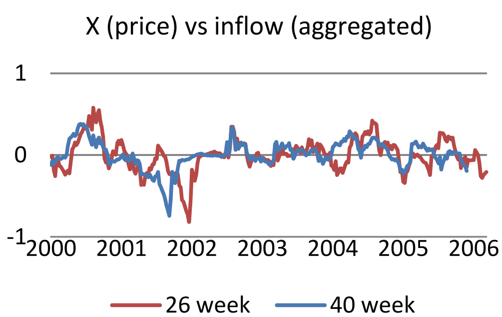


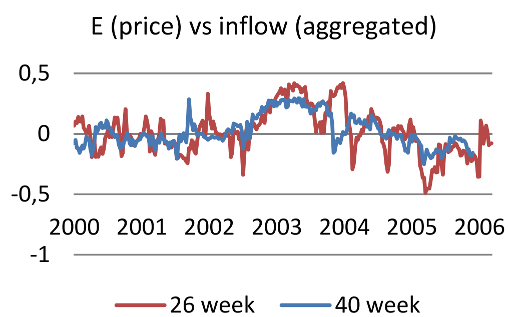






Aggregated inflow model





F. Enclosed CD

Consult enclosed CD for contents and description of files attached.