Long- and Medium-term Operations Planning and Stochastic Modelling in Hydro-dominated Power Systems Based on Stochastic Dual Dynamic Programming

Anders Gjelsvik, Birger Mo, and Arne Haugstad

Abstract This chapter reviews how stochastic dual dynamic programming (SDDP) has been applied to hydropower scheduling in the Nordic countries. The SDDP method, developed in Brazil, makes it possible to optimize multi-reservoir hydro systems with a detailed representation. Two applications are described: (1) A model intended for the system of a single power company, with the power price as an exogenous stochastic variable. In this case the standard SDDP algorithm has been extended; it is combined with ordinary stochastic dynamic programming. (2) A global model for a large system (possibly many countries) where the power price is an internal (endogenous) variable. The main focus is on (1). The modelling of the stochastic variables is discussed. Setting up proper stochastic models for inflow and price is quite a challenge, especially in the case of (2) above. This is an area where further work would be useful. Long computing time may in some cases be a consideration. In particular, the local model has been used by utilities with good results.

Keywords Energy economics · Hydro scheduling · Stochastic programming

1 Introduction

Finding optimal operational strategies for a large hydrothermal power system with a large fraction of hydropower is a very demanding problem, both theoretically and computationally, since it is stochastic and usually large-scale. One major development in this area is the method of stochastic dual dynamic programming (SDDP) (Pereira 1989; Pereira and Pinto 1991). In this text we shall describe adaptions of this method in a Nordic context.

Norway has about 99% hydropower. In the Nordic countries, Denmark (with no hydropower), Sweden, Finland and Norway have a liberalized common power

A. Gjelsvik (✉)
SINTEF Energy Research, 7465 Trondheim, Norway
e-mail: Anders.Gjelsvik@sintef.no

S. Rebennack et al. (eds.), *Handbook of Power Systems I*, Energy Systems,
DOI 10.1007/978-3-642-02493-1_2,

market, in which hydropower constitutes about 50% of average generation. Because of the high fraction of hydro, market prices may depend very much on the hydrological situation, and taking inflow stochasticity into account is therefore essential. If we consider only a small entity that cannot influence the prices (a price taker), it also becomes necessary to model price stochasticity.

As is well known, it is common to separate the scheduling task in at least three steps: The long-term scheduling, with an horizon of 3–5 years or longer; the medium-term or seasonal scheduling looking 1–2 years ahead; and the short-term scheduling with a horizon of a few days to 1 week. The long- and medium-term scheduling problems are stochastic. The long-term scheduling sets end conditions for the medium-term scheduling, for example in terms of marginal water values, and the medium-term scheduling results are input to the short-term scheduling.

The long-term scheduling problem is frequently approached using some variant of the water value method (Stage and Larsson 1961; Lindqvist 1962), which is based on dynamic programming (see also the dual concept in Scott and Read (1996)). An overview of other models based on stochastic programming is given in Wallace and Fleten (2003). A variety of methods for hydrothermal scheduling (other than SDDP) are also reviewed in Labadie (2004).

In practice, the water value method can only be applied to systems with a very small number of reservoirs. It is therefore often applied to aggregated one-reservoir models of more complicated hydro systems and for simulation purposes supplemented by heuristics. Using SDDP techniques, however, allows stochastic optimization for multi-reservoir systems, which means that more realistic and detailed models can be dealt with.

As examples of application of SDDP and/or related techniques, we mention Tilmant and Kelman (2007), where inflow modelling is also discussed, and Iliadis et al. (2006) and Aouam and Yu (2008). In Philpott and Guan (2008), the convergence of the SDDP-algorithm is discussed, and a theoretical convergence proof is given. In addition to Pereira (1989) and Pereira and Pinto (1991), descriptions of the algorithm can be found in Tilmant and Kelman (2007) and de Oliveira et al. (2002).

In this chapter, we shall deal with two different scheduling models based on SDDP. One is a 'local' model, for a system confined in a geographical area that can be covered by a single power balance equation (without internal transmission bottlenecks), and typically owned by a single power company. It is usually assumed that the system is not large enough to influence market price, so that we have a 'price taker' case. This means that the market price must be dealt with as an exogenous stochastic variable. This model is mainly aimed at medium-term scheduling, but is also used for long-term scheduling.

The other model to be discussed here arises when the SDDP approach is applied to a 'global' system model much similar to the EMPS model (Botnen et al. 1992).

In Sect. 2, we describe elements of a mathematical model of a hydrothermal power system. In Sect. 3, a SDDP-based solution algorithm for the local model is dealt with. To handle the exogenously given price, a combination of SDDP and ordinary stochastic dynamic programming (SDP), originally described in Gjelsvik and Wallace (1996), is used. The method is also described in Gjelsvik et al. (1997), and in more detail in Gjelsvik et al. (1999), and a similar combination of SDDP and SDP was also used in Iliadis et al. (2006). This local scheduling algorithm is the main topic of this chapter. In Sect. 4, we discuss extensions to the basic model, such as an approximate way of handling head variations, and incorporation of risk control.

The global model is briefly outlined in Sect. 6. For this model inflow modelling becomes harder than that in the local case. Although the SDDP approach can deal with many reservoirs, it is not so easy to handle a many-dimensional multivariate inflow process. This is discussed in Sect. 7.

Section 8 deals with some computational issues, and Sect. 9 discusses and sums up some experiences with these models.

2 Basic Power System Model

2.1 Introduction

Most of the material in this section is general for hydropower system modelling; however, price modelling mainly applies to the local model.

The SDDP solution algorithm can be seen as a dynamic programming approach where future costs are represented by hyperplanes (often referred to as cuts), and it relies on linear programming. A fundamental requirement is that the problem must be linear or at least convex. In the model presentation that follows here, we therefore strive to obtain linear or piecewise linear relationships. Fortunately, most relations are close to linear.

It is necessary to use a finite time horizon at time T. For medium-term scheduling, T is usually up to 2–3 years ahead; for long-term scheduling one would use 3–5 years or more. The study period is divided in discrete time steps indexed t, with $t \in [1, \dots, T]$, usually of length 1 week.

2.2 Power Station Model

Let Q be the release of water through a hydropower plant during a time interval, and let P be the corresponding electrical energy generated. It is assumed that

$$P = f(Q)\frac{h}{h_0}\,. \tag{1}$$

We take the function $f(Q)$ as piecewise linear, specific for each power station. For the SDDP algorithm, $f(Q)$ must be a concave function. h is the water head and h_0 a nominal reference head.

It is not possible to handle variable head directly in the SDDP algorithm (at least for a cascaded power system). Therefore, the head correction factor h/h_0 in

(1) must be applied with estimated values of h. In many Norwegian power stations this is a fair approximation, since the head often is quite large compared to the head variations. In principle, optimizing with variable head might lead to a non-convex problem, not suitable for SDDP. A study for a single power station is given in Bortolossi et al. (2002). For an example of a non-linear model for variable head, see Mariano et al. (2008).

In Gjelsvik and Haugstad (2005) a heuristic to deal with head variations was described, as used in connection with a hydropower system with cascaded reservoirs. In Sect. 4.1 we shall review this approach.

Generation in thermal power stations is modelled in a simplified manner, as a set of buying options, each with a fixed marginal cost.

2.3 Reservoir and Inflow

Let q_t be the vector of inflows in time step t and V_t the vector of reservoir volumes at the end of t. Further, let Q_t and s_t be vectors of reservoir releases and spills, respectively. With V_0 given, the water reservoir balances can be written as

$$V_t = V_{t-1} - H_1 Q_t - H_2 s_t + q_t, \tag{2}$$

with the condition that

$$\underline{V}_t \leq V_t \leq \overline{V}_t \tag{3}$$

for $t = 1, \ldots, T$. Here H_1 and H_2 are suitable incidence matrices that describe where releases and spills go, and $\underline{V}_t$ and $\overline{V}_t$ are (possibly time-dependent) limits.

Inflow sequences often show strong sequential correlation, so that q_t depends heavily on q_{t-1}. In stochastic dynamic programming, the stochastic term for the time interval t must depend only on the state at the beginning of the time interval t and not on earlier history. As is well known, this can be arranged by state space enlargement, whereby 'previous' inflows are included in the system state vector.

Historical inflow time series differ in length, but often up to 70 years or more are available. Variations in load etc. are treated as coupled to the inflow; one may speak of 'weather years'. If there are S observed weather years, we construct S parallel inflow scenarios by picking T weeks starting in the first year, then T weeks starting in the second year, and so on. Let q_t^i be the inflow of the ith scenario in week T.

Usually the inflow has strong seasonal variations. One reason for this is the accumulation of snow during winter, with the following spring melt. We try to eliminate the seasonal variations by normalizing computing normalized weekly inflows $\{z_t^i\}$ as

$$z_t^i = \frac{q_t^i - \overline{q_t}}{\sigma_t} \quad \text{for} \quad i = 1, \ldots, S, \quad t = 1, \ldots, T, \tag{4}$$

where $\overline{q_t}$ is the mean inflow in week t averaged over observed years and σ_t is the corresponding sample standard deviation.

A first-order auto-regressive model (AR1) is then used to represent the series $\{z_t^i\}$. Usually there are several series, so that z_t is a vector, whose components have been individually normalized as shown. The model is then

$$z_t = \phi z_{t-1} + \varepsilon_t \quad \text{for} \quad t = 1, \ldots, T. \tag{5}$$

Here ϕ is the transition matrix and the vector ε_t is the model error, or 'noise'. The elements of ϕ and $\text{cov}(\varepsilon_t)$ are estimated by a regression approach, minimizing the sum $\sum_t (z_t - \phi z_{t-1})^T (z_t - \phi z_{t-1})$. It is now assumed that the model error ε_t is independent of z_{t-1}.

The linear AR1 inflow model (5) is not always very good, but it is a compromise between accuracy and computational feasibility. In practice, it has been found that it is best to split the data into different seasons and to fit a separate model for each season.

From the fitting of (5) we have a (usually multivariate) sample distribution of the error term ε_t. For use in the optimization model, we must approximate this distribution by a discrete probability distribution with K discrete values $\varepsilon_t^1, \ldots, \varepsilon_t^K$ and corresponding probabilities ψ_k, where K is not too large.

$$\Pr(\varepsilon_t = \varepsilon_t^k) = \psi_k \quad \text{for} \quad k = 1, \ldots, K \quad \text{and} \quad t = 1, \ldots, T. \tag{6}$$

One way of doing this is to carry out a model reduction by applying principal component analysis (PCA) (Johnson and Wichern 1998) to the sample $\{\varepsilon_t\}$ for $t = 1, \ldots, T$. The principal components are transformed variables constructed so that they are independent, taken over the sample. Only the principal components that contribute most to the total variance are kept (typically 3). After the PCA has been carried out, the distribution of each principal component kept is approximated by a small number of discrete points. Finally the discrete points obtained are transformed back to the axes of the original normalized data and combined.

An example of the use of principal components analysis for inflow modelling, followed by discretization, is da Costa et al. (2006). In Jardim et al. (2001) clustering techniques are used for constructing representative discrete noise terms.

In addition to the above-mentioned method, we have also implemented an approach whereby we obtain the $\{\varepsilon_t^k\}$ by sampling from the collection of sample errors $\{\varepsilon_t\}$ directly.

Inflow modelling will be further dealt with in Sect. 7.

2.4 Power Balance and Objective Function

We deal here with the power balance for a 'local' system. It can be generalized by setting up several such balances and introducing transmission variables. Let u_t denote a vector of decision variables for time step t, containing releases Q_t, spills s_t, thermal generation w_t and transactions outside the spot market. Let c_t be the cost vector associated with u_t. We also write $x_t = [V_t^T, z_t^T]^T$ for (the continuous part of) the system model state vector at the end of time step t. Further we define, for $t = 1, \ldots, T$,

y_t^+ – Sale to the spot market
y_t^- – Purchase from the spot market
y_t – Vector $y_t^T = [y_t^+, y_t^-]$
p_t – Spot price (weekly average)
δ_t – Transmission charge
d_t – Firm power demand

The power balance is then

$$A_t u_t - y_t^+ + y_t^- = d_t \quad \text{for} \quad t = 1, \ldots, T. \tag{7}$$

The cost L_t for one realization in one time step becomes

$$L_t(u_t, y_t, p_t) = c_t^T u_t + (p_t + \delta_t) y_t^- - (p_t - \delta_t) y_t^+. \tag{8}$$

In the power balance (7), the hydro and thermal generations, as well as power transactions outside the spot market, are contained in $A_t u_t$; the matrix A_t contains the power plant conversion factors corresponding to the piecewise linear models in (1).

If a better time resolution in the market description is desired, the hours in a week are grouped into several 'load periods'. Hours with similar prices are lumped together in the same load period, for instance one for night and one for day. Instead of one power balance in (7), there is one for each load period, but the load periods do not follow each other sequentially.

In practice, the transmission cost δ_t is mostly neglected, but we have so far retained it in the model, to make the model more general. We also want to be able to deal with a limited market, and so we introduce limits $\underline{y}_t$ and $\overline{y}_t$ to the y_t vectors; see (19) below. Market limitations may stem from transmission limits, for instance. The model may also be run without a spot market, but with a local load or a local market. In some cases, price elasticity in the spot market may be included by splitting the market into steps with decreasing price for increase in sales. In most cases, though, the market is considered infinite, and so the limits are set at some sufficiently large value and will not be binding.

The firm power demand d_t represents obligations in the local area. However, the firm power demand is zero in the case where all generation is considered sold in the spot market, which is the most common modelling case in the Nordic market. If present, d_t may be considered to vary with the inflow according to the 'weather year'. This causes a difficulty with the SDDP algorithm that we are to use, so that partially we have to use averages; see a remark in Sect. 3.2.

If there is a firm power demand, its influence on the hydro schedules depends on the situation. If market limitations on y_t do not become binding, the hydro schedules will be unaffected by a change in d_t. Otherwise, an increase in d_t will in general lead to higher average storage levels in the reservoirs.

The value of the water remaining in the reservoirs at the horizon must be subtracted from the cost. Let this value be given by a function $\Phi(x_T)$.

Estimating $\Phi(x_T)$ is one of the challenges in the scheduling task. We use results from an aggregated long-term model of the system for this purpose, supplemented by heuristics for distribution between the reservoirs (Johannesen and Flatabø 1989). To a large extent, however, Φ is a function of total storage in the system. Ideally, the horizon T should be set as far away as possible to minimize the influence of errors in $\Phi(x_T)$, but here computing time must also be taken into account.

2.5 Price Modelling

We deal here with the local model, where the spot price is regarded as an exogenous stochastic variable. Price variations can be quite strong, as shown in Fig. 1. Price forecasts can be obtained in several ways. In the Nordic market, forecasts are often obtained by simulations with the so-called EMPS model (Botnen et al. 1992), which is a long-term model covering many areas and several years with the spot price as an internal variable. When using the EMPS model, each price scenario corresponds to a historical weather year.

Time series from such forecasts show that the spot price has a strong sequential correlation. As with the inflow, it is then necessary to include a price state in the system state description, since we intend to use dynamic programming. However, since our objective function (8) contains the product term $p_t y_t$, we cannot expect the future cost functions (see (3.2)) to be convex functions of reservoir and price variables; hence they cannot be represented by cuts, as is necessary when applying SDDP.

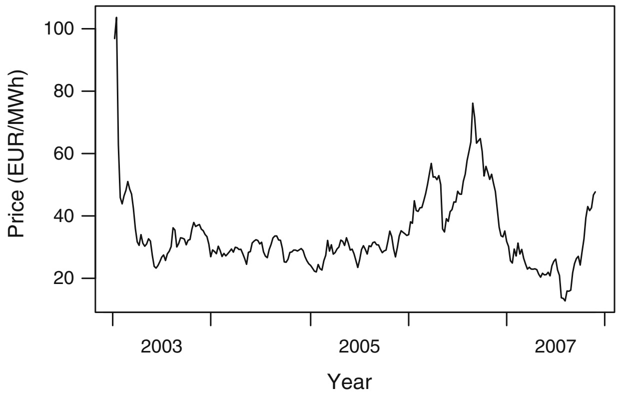

Fig. 1 Weekly average of Nord Pool spot price 2003–2007 (source Nord Pool)

Therefore, while reservoir and inflow states are treated as continuous variables, we combine SDDP with ordinary stochastic dynamic programming with discrete states with regard to p_t, see Sect. 3.1. Thus, p_t is represented by a set of M points $\zeta_t^1, \ldots, \zeta_t^M$.

To establish the probability distribution for price and inflow, we simplify by considering the stochastic processes for inflow and price as independent of each other and using the marginal probability distributions for each.

Broadly speaking, it is reasonable that more water leads to lower prices and the other way round. However, it would be a challenge to include this in the model.

One should also consider that we are dealing with a local part of the total system. To take a Norwegian example: the price will depend on the hydrological state of Norway, Sweden and Finland as a whole, but the system under consideration in the model may be covering only a couple of rivers, where the inflow may not fully follow the trend in the total system.

To describe the transitions between the established discrete price values from 1 week to the next, we use the following Markov model:

$$\Pr(p_t = \zeta_t^j | p_{t-1} = \zeta_{t-1}^i) = \rho_{ij}(t) \quad \text{for all} \quad i, j \in [1, M]. \tag{9}$$

Thus, $\rho_{ij}(t)$ is the probability that $p_t = \zeta_t^j$, given that p_{t-1} was ζ_{t-1}^i.

The numerical values of the transition probabilities are established from a set of price scenarios the following way (Mo et al. 2001b):

First, the price values within each week are grouped in M groups, and ζ_t^i is taken as the mean value of the $N_i(t)$ price values in the ith group in week t. This way, all ζ_t^i are established for all t in the data period. It is recorded which scenarios go into each group in each week, and an estimate $\tilde{\rho}_{ij}$ of ρ_{ij} is then taken as the fraction of the scenarios from the ith group at time $t-1$ that belong to group j at time t. The probabilities obtained this way do not involve the actual values p_t and may not give correct sample conditional means for the price at time t, however. Given that $p_{t-1} = \zeta_{t-1}^i$, $E\{p_t | p_{t-1} = \zeta_{t-1}^i\}$ should be equal to the average price in week t of the scenarios belonging to the ith group at time $t-1$. For this reason, improved values $\{\rho_{ij}\}$ are computed by minimizing a weighted sum of the squared deviations $\{\rho_{ij} - \tilde{\rho}_{ij}\}$ and the squared deviations in the conditionally expected price. One seeks ρ_{ij} so as to find

$$\min \left\{ \sum_{i=1}^{M} \left(\sum_{j=1} \rho_{ij}(t)\zeta_t^j - E\{p_t | p_{t-1} = \zeta_{t-1}^i\} \right)^2 + \omega \sum_{i=1}^{M} \sum_{j=1}^{M} (\rho_{ij}(t) - \tilde{\rho}_{ij}(t))^2 \right\}, \tag{10}$$

subject to the constraints

$$\sum_{j=1}^{M} \rho_{ij}(t) = 1 \quad \text{for all } i \tag{11}$$

$$\sum_{i=1}^{M} \rho_{ij}(t) N_i(t-1) = N_j(t) \quad \text{for all } j \tag{12}$$

$$0 \leq \rho_{ij} \leq 1 \quad \text{for all } i \text{ and } j\,. \tag{13}$$

Here ω is an appropriate weight factor; the second sum of squares in (10) is considered necessary to obtain a unique solution.

2.6 Overall Local Model

From the elements described, the local model can be summarized as follows. Find an operating strategy that gives u_t from x_t, such that

$$\min E \left\{ \sum_{t=1}^{T} L_t(u_t, y_t, p_t) - \Phi(x_T) \right\}, \tag{14}$$

subject to the constraints

$$x_t = F_t x_{t-1} + G_t u_t + \varepsilon_t \tag{15}$$

$$A_t u_t + B_t y_t = d_t \tag{16}$$

$$\underline{x}_t \leq x_t \leq \overline{x}_t \tag{17}$$

$$\underline{u}_t \leq u_t \leq \overline{u}_t \tag{18}$$

$$\underline{y}_t \leq y_t \leq \overline{y}_t \tag{19}$$

for $= 1, \ldots, T$ and x_0 and p_0 given, and with probability distributions given by (9) and (6). L_t is given by (8).

The expectation E is to be taken over both inflow and price. Equation (15) is the transition equation for the states except the price state, and contains (2) and (5). Equation (16) contains the power balances (7), and it may also be generalized to include other constraints that are not coupled in time. A_t, B_t and F_t, G_t are matrices of suitable dimensions. Reservoir limits, equipment ratings etc. are contained in (17)–(19).

3 Solution Method for the Local Model

3.1 Overview

As already mentioned, the solution method that we have chosen is a combination of stochastic dual dynamic programming and ordinary stochastic dynamic programming (SDP) (Dreyfus 1965). The ordinary SDP part is introduced to take care of the

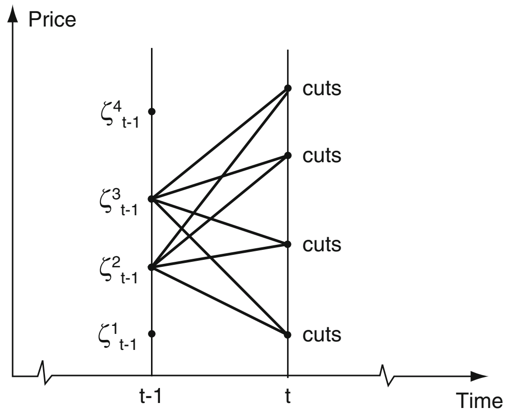

Fig. 2 View of the dynamic programming part of the combined approach, in the time-price plane

price process, which is modelled as described in Sect. 2.5. The reservoir and inflow states are treated as continuous variables and dealt with using hyperplanes, as in the ordinary SDDP algorithm.

A hybrid SDP/SDDP approach was also used in Iliadis et al. (2006).

The approach is visualized in Fig. 2, where the price dimension is shown schematically. In ordinary table-based SDP, there would be a number at each discrete state point, giving the expected future cost going from this state. Here, at each discrete price point there is now instead a set of cuts representing the expected future cost function as a function of the continuous state variables.

As mentioned in Sect. 2.5, the correlation between inflow and price 1 week ahead is neglected.

3.2 Solution by a Dynamic Programming Approach

The solution algorithm for the model established has been described in great detail in Gjelsvik et al. (1999). We outline it here.

As indicated in Fig. 2 we consider a time interval t, with the initial state given by x_{t-1} and p_{t-1}. There are K realizations of the inflow noise $\varepsilon_t = \varepsilon_t^k$, and for each of these M possible price values $p_t = \zeta_t^i$, $i = 1, \ldots, M$. We assume here that we learn ε_t and p_t immediately before the decisions for this time step are to be carried out. One justification for this is that with an interval of 1 week, it is possible to adjust to changing conditions, and a further reason is that usually at a time close to the actual operation more accurate forecasts are available than those assumed in our stochastic models. Let $\alpha_t(x_t|\zeta_t^j)$ be the expected future cost function at the end of time period t, at the system state $[x_t, \zeta_t^j]$. (The expected future cost function is the expected cost in going from the given state at the end of time interval t to an allowed final state using an optimal strategy). Applying the Bellman optimality principle (Dreyfus 1965), we obtain from (14) and (9) and (6) the recursive equation

$$\alpha_{t-1}(x_{t-1}|p_{t-1} = \zeta^i_{t-1}) = \sum_{j=1}^{M}\sum_{k=1}^{K} \rho_{ij}\psi_k \min\left[L_t(u_t, y_t, \zeta^j_t) + \alpha_t(x_t|p_t = \zeta^j_t)\right] \tag{20}$$

for all $t \in [1, T]$ and all $i \in [1, M]$. The constraints (15)–(19) must be satisfied for each transition. For each possible outcome $(\varepsilon^k_t, \zeta^j_t)$ separate decisions (u^{kj}_t, y^{kj}_t) are made, and the final state obtained is x^{kj}_t.

In Pereira (1989), Pereira and Pinto (1991) and Gjelsvik et al. (1999), it is shown that, with a linear model, the expected future cost functions are piecewise linear functions of x and can be represented by hyperplanes in the x state space, which also means that these functions are convex.

We define $\alpha^{kj}_{t-1}(x_{t-1}) = \min[L_t(u_t, y_t, \zeta^j_t) + \alpha_t(x_t|\zeta^j_t)]$ in (20). With the hyperplane representation, (20) then decomposes into single-transition sub-problems of the following form:

With x_{t-1}, $p_{t-1} = \zeta^i_{t-1}$, $p_t = \zeta^j_t$ and $\varepsilon_t = \varepsilon^k_t$ given, find

$$\alpha^{kj}_{t-1}(x_{t-1}) = \min\left[L_t(u_t, y_t, \zeta^j_t) + \alpha\right], \tag{21}$$

with the constraints (15)–(19) and

$$\left.\begin{aligned} \alpha + (\mu^{j1}_t)^T x_t &\geq \gamma^{j1}_t \\ &\vdots \\ \alpha + (\mu^{jR}_t)^T x_t &\geq \gamma^{jR}_t \end{aligned}\right\} . \tag{22}$$

In (22) $\mu^{j1}_t, \ldots, \mu^{jR}_t$ and $\gamma^{j1}_t, \ldots, \gamma^{jR}_t$ define R hyperplanes (cuts) that represent the expected future cost function at the price point $p_t = \zeta^j_t$. For an exact representation, an extremely large number of cuts would usually be required. Therefore, an approximate representation is used, where one starts from zero or few hyperplanes and adds them iteratively to get an improved strategy, as in the ordinary SDDP algorithm.

It is assumed that the single-transition sub-problem described in (21)–(22) has a feasible solution; this is ensured by artificial variables. We note that the price of the previous week, p_{t-1}, does not enter the sub-problem. Therefore, in (20) it is not necessary to solve the sub-problem for all M^2 combinations of i and j on the backward run. One solves for the M different $p_t = \zeta^j_t$; then the initial state $p_t = \zeta^i_t$ enters through averaging with the transition probabilities ρ_{ij}.

One main iteration of the algorithm consists of a forward simulation and a backward recursion:

1. *Forward simulation.* The system is simulated from the initial state, with the given scenarios for price and inflow, using the strategy (hyperplanes) obtained so far. For each week t in inflow scenario s, the operation is found by solving (21) with x_{t-1} and p_{t-1} from the final state of the previous week, and ε_t and p_t from the observed values for this scenario. For values p_t that differ from the defined

points ζ_t, linear interpolation in the hyperplanes of the neighbouring points is used. The cost for each scenario is computed, and the average of these costs gives an upper bound for the operating cost. The reservoir trajectories $\{x_t^s\}$ for all t and s are stored.

2. *Backward recursion.* At the horizon $t = T$ the expected future cost function is given by $-\Phi(x_T)$. Consider a general time step, as indicated in Fig. 2, with the expected cost functions given at time t. For each discrete $p_t = \zeta_t^j, \quad j = 1, \ldots, M$ along the price axis, one solves the single transition sub-problems (21) with $x_{t-1} = x_{t-1}^s$ for the trajectories for all $s = 1, \ldots, S$ and all K inflow transitions in time step t. From this, an improved expected future cost function is constructed at each price point ζ_{t-1}^i at the end of week $t-1$ by adding new hyperplanes computed from the dual variables of (21) and (22) and the transition probabilities $\{\rho_{ij}\}$, as described in Gjelsvik et al. (1999). Proceeding step by step backwards from $t = T$ to $t = 1$, one obtains an updated strategy, and a lower bound for the cost. If converged, then stop, otherwise go to 1.

Convergence means that the simulated expected operating cost from step 1 comes 'close' to the lower bound from step 2. In practice, there is usually a gap, so that convergence is mainly judged by monitoring the maximal change in a reservoir trajectory from one main iteration to the next, prescribing a minimum and a maximum iteration number.

In the procedure outlined, we make a modification for the inflow, in that on the forward run we use the 'observed' inflow-price scenarios. This heuristic is intended to take care of any coupling between inflow and price when averaged over longer periods. It may, however, lead to gaps between upper and lower cost estimates, because the fitted inflow and price models used on the backward run (5) and (9) may not be fully consistent with the observed scenarios. Some numerical values are given in Sect. 9.

In the case with firm power $d_t, t = 1, \ldots, T$, that is modelled as inflow-dependent, we use averages of the firm power demand d_t in the backward run of the algorithm to avoid state dependencies. This may also contribute to a cost gap.

Apart from the 'outer' dynamic programming treatment of the spot price state, the approach is similar to that of the ordinary SDDP algorithm, and the same computational approach can be used for this part. To solve the single-transition sub-problem, a relaxation approach is used for the future cost hyperplanes and the reservoir balances, as in Røtting and Gjelsvik (1992). The LP problems actually solved are quite small, see Sect. 8. There is a limit to the number of hyperplanes allowed for each of the M price values; after reaching that, hyperplanes that are infrequently binding are overwritten. This is a crucial part of the algorithm.

Usually an initial set of reservoir trajectories is available at the start of the solution process, so that one can start with the backward recursion step of the algorithm. The inflow loop is put innermost in the calculations, since this only changes the right-hand side of the single-transition problem of minimizing (21) with constraints. Each problem with a new inflow is started from the solution of the previous one using the dual algorithm of linear programming. When the price p_t changes, both the cost row and the set of cuts changes; in this case, an all slack basis is used for start.

4 Extensions to the Local Model

In this section, we describe a few extensions to the medium-term local scheduling model as described in Sects. 2 and 3.

4.1 Head Variations in Medium-term Scheduling

As mentioned in Sect. 2.2, the SDDP algorithm cannot deal directly with variable head, as the problem then may become non-convex. In this section, we outline how variable head can be taken approximately into account in the local model, using a semi-heuristic approach. The method is described more closely in Gjelsvik and Haugstad (2005). It is based on an expansion around a nominal reservoir operating schedule. Release is considered fixed, and sensitivities of economic gain with respect to small changes in reservoir levels are calculated and added to the cost function.

Consider a hydropower system with n reservoirs. For $i = 1, \ldots, n$ we look at the ith reservoir, with a storage of V_t^i at the end of week t and a water surface elevation of h_t^i, referred to sea level, say. We assume that there is a power plant with output P_t^i immediately downstream of reservoir i. If there is a reservoir below this plant, let j be its number, and let k be the number of any upstream plant. In general, we assume that the outlet of a plant in a downstream reservoir is submerged; otherwise, the contribution to head sensitivity is zero.

As before we assume that generated power depends linearly on water head and is obtained from (1)

$$P_t^i = f^i(Q_t^i)\frac{h_t^i - h_t^j}{h_0^i}, \tag{23}$$

where h_0^i is the nominal head for plant i.

We now consider the situation where the volume V_t^i is changed by a small amount ΔV_t^i, without changing the release q_t^i (the change can be thought of as being brought about by a different operation at earlier stages). The influence of ΔV_t^i on generation in the downstream and upstream plants can be shown to be

$$\frac{\partial P_t^i}{\partial V_t^i} = \frac{\partial P_t^i}{\partial h_t^i}\frac{\partial h_t^i}{\partial V_t^i} = \frac{1}{A_t^i}\frac{P_t^i}{h_t^i - h_t^j} \quad \text{and} \quad \frac{\partial P_t^k}{\partial V_t^i} = -\frac{1}{A_t^i}\frac{P_t^k}{h_t^k - h_t^i}, \tag{24}$$

where $A_t^i = (\partial h_t^i / \partial V_t^i)^{-1}$ is the current surface area of reservoir i.

We take the prevailing market price of power, λ_t, as the marginal value of the generation change.

We now assume that we have available nominal reservoir operation schedule with nominal values of $\{P_t^i\}$, $\{\lambda_t\}$, $\{V_t^i\}$, $\{h_t^i\}$ and $\{A_t^i\}$ for all t and i. Using the above formulas, we may then approximately account for the cost change due to variable head by use of extra cost terms containing ΔV:

$$\sum_{t=1}^{T} \sum_{i=1}^{n} \tilde{c}_t^i \Delta V_t^i , \tag{25}$$

where the $\tilde{c}$- coefficients are to be determined. Using (24) we find

$$\tilde{c}_t^i = \frac{\lambda_t}{A_t^i} \left[-\frac{P_t^i}{h_t^i - h_t^j} + \frac{P_t^k}{h_t^k - h_t^i} \right] \qquad \text{for all } i \text{ and } t \ . \tag{26}$$

A $\tilde{c}$-coefficient may be positive or negative. As expected an increase in reservoir storage decreases cost associated with the downstream plant (i), but increases cost in the upstream plant (k). As seen from (24), this also depends on the nominal generations P_t^i and P_t^k.

Use of the sensitivities derived above has been implemented in the model described in Sect. 3. In this model, the cost functions must be convex, and all state dependency must be contained in the hyperplanes. It is therefore not possible in general to have different model coefficients for various system states. Therefore, the mean values of the sensitivity coefficients above are used, where the mean is taken over the various inflow scenarios.

As indicated earlier, calculations are carried out in two steps. First, the scheduling program is run without head coefficients. From the releases and reservoir and price trajectories obtained from this run, a full set of head sensitivities $\{\tilde{c}_t^i\}$ is calculated for each inflow scenario. For each week, the sensitivities are averaged over the different inflow scenarios. The mean values of the sensitivities are then used to perturb the cost function according to (25) in a second run. The second run is then generally considered as giving the final schedule. Repeated recalculation of the head sensitivities based on rerunning the program with the last calculated sensitivities is possible, but the sensitivities could in principle oscillate from one calculation to the next. Tests indicate, however, that results may converge after a few repetitions of the recalculation procedure. Furthermore, recalculation does not seem necessary for a good result.

The above perturbation is a kind of heuristic. However, simulations with this correction have given reservoir trajectories that look more realistic than without, and the simulated economic result generally improves.

4.2 Risk Control

The local model has also been extended to allow for dynamic hedging using forward delivery contracts and a mechanism for risk control (Fleten 2000; Mo et al. 2001a).

The basic idea is to consider the accumulated profit. The study period is divided into suitable sub-periods, say quarters, with a profit target for each sub-period. There is a penalty for not meeting the profit target. This penalty function is specified by the user, and it can be shown that this is another form of a utility function that defines the user's risk aversion. Mathematically, the accumulated profit is introduced as a state in the state vector x in (15). The profit state is additive and linear, and so goes with the reservoir equations in the basic model, with a zero inflow.

Forward contracts can be bought or sold for every suitable period within the study period. The net amounts of forward contracts for each future week are also defined as state variables, and buying and selling of such contracts are included in the control vector u_t. This increases the computational burden, but the advantage is that the trading in forward contracts is handled dynamically. A closer description can be found in Mo et al. (2001a).

The study (Iliadis et al. 2006) gives some examples of risk management using an SDP/SDDP approach.

4.3 Use of the Results from the Local Model

Results from the local model are used in two ways. First, marginal values or release volumes for the first 1 or 2 weeks are used as input to the short-term scheduling and the daily spot market bidding process. Hyperplanes can be transferred directly to a mathematical model for short-term scheduling if desired (Fosso and Belsnes 2004).

Second, the output from the local model is used for various estimates, such as the hydro generation over the study period, for predicting reservoir levels, in risk analysis and for maintenance scheduling.

5 A Numerical Example

We give here a brief example of application of the local model. The example is taken from Gjelsvik and Haugstad (2005) and is used to illustrate the effect of the correction heuristics for variable head. The system is shown in Fig. 3. The mean annual generation is about 3,270 GWh. There are four plants with significant variations in head (5, 7, 8 and 10). Average market price is 15.4 EUR/MWh in this case. Run with and without head coefficients, the model shows an average increase in income when head coefficients included are of 0.13 EUR/MWh, compared to the schedule without head coefficients. Generation increases by about 1% on average, despite an increase in spilled water.

Inclusion of head coefficients change the strategy for operation of reservoirs considerably. For reservoirs 6 and 7 this is shown in Fig. 4. There is no plant immediately below reservoir 6, and when taking head into consideration, the model transfers more water to reservoir 7 to increase the head.

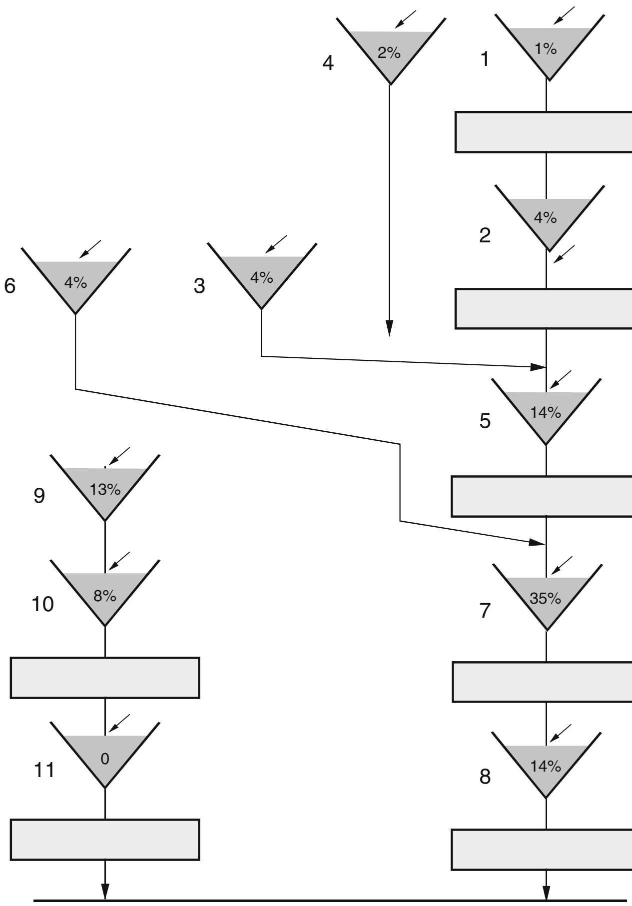

Fig. 3 Case system. *Cones* are reservoirs and *boxes* are power stations. Numbers in reservoir symbols show each reservoir's share of the total reservoir volume. *Boxes* indicate plants. Reservoirs 5, 7, 8 and 10 have a maximum head variation of 45, 70, 17 and 26 m, respectively. Nominal head for all these plants is in the range of 300–500 m

6 A Global Scheduling Model

The SDDP algorithm has also been implemented in an optimization model for a 'global' system, for instance corresponding to the Nord Pool market area and northern Europe. This model is similar to the EMPS system model Botnen et al. (1992), in that the system in each area is lumped together and represented as a single reservoir and a single power station during the optimization phase. An advantage compared to the standard EMPS model is that interconnections between areas are better described in the optimization phase. The model differs from the local model of Sect. 3, in that there are several busbars (one for each area) and that the market price is an internal variable. An external price mechanism has been implemented, though;

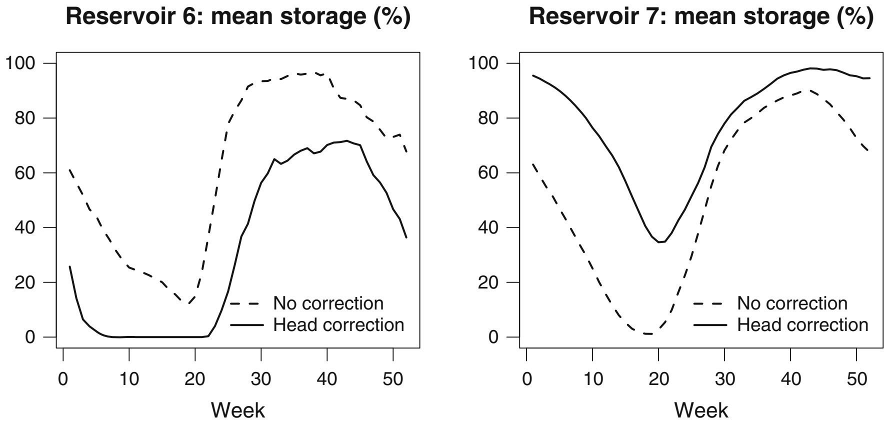

Fig. 4 Mean levels in reservoirs 6 and 7 with and without head correction

this may be useful for describing uncertainties in costs of alternative resources, such as oil or coal.

This model typically has on the order of 20 areas and 40 inflow series. Each area has a power balance similar to (7), with transmission modelled as a transport. The high dimensionality of the inflow process is a difficulty with this model. Since this model is intended to cover a much larger area than the model of Sect. 2, the inflows have more differing characteristics and are not so easily described by a few principal components. This will be further discussed in Sect. 7.

A few extensions have been implemented in the global model, such as internal markets for green certificates (Mo et al. 2005) and CO_2-quotas (Belsnes et al. 2003).

Results from the global model would typically be used for price forecasts and simulation studies from a given initial state. Convergence seems to be slower for this model, and the computing times can be several days on a single processor. Reasons for slower convergence may be that the reservoirs are quite large and less constrained than in the local model, and oscillations in power transfers from one iteration to the next. Also, it is harder to get a good inflow model, as will be discussed in the next section.

7 More on Stochastic Inflow Modelling

Setting up stochastic models for processes involved in stochastic hydropower scheduling usually requires some compromises between accuracy and tractability. In this section, we look at some difficult points of inflow modelling.

In modelling inflow for hydro scheduling, there are several requirements that one should try to meet: The model should give a set of discrete inflow values for each stage, with corresponding probabilities. The model should be as simple as possible to reduce computational burden, but it should also be sufficiently accurate and unbiased. For SDDP applications, the inflow model must preserve convexity.

In Sect. 2.3 we introduced a first order (vector) auto-regressive model (5) for the (normalized) inflow. This is supposed to take care of the sequential correlation. In a geographically dispersed system, it can be difficult to fit a single multivariate model, since different areas may have different inflow characteristics. Fitting an AR1 time series requires that the inflow process is a weakly stationary process, that is, with statistical properties that are independent of time. Even after removal of seasonally varying mean values this is often not the case. The problem can to some extent be avoided by splitting the data into several seasons and fitting the model to each season separately, ensuring proper handling where the seasons join.

If enough data are available, as is frequently the case, it would probably be best to carry out an individual regression analysis for each week.

A difficulty that is sometimes observed is that the residuals for a given week may depend on the initial state. This may happen because a time series is not 'stationary enough', but also when using linear regression directly on a single week. This means that the computations in (21) become biased, since the same distribution of ε_t is used for all initial states. A rather special case is shown in Fig. 5.

The data here come from a 16-area model of Denmark, Sweden, Finland and Norway, intended for use with the global model of Sect. 6. They are for week 24, at the end of the snow melting period. Sample residuals ε_{24} are plotted against normalized inflow values z_{23} at the beginning of the week. Although the distribution for Area 4 is fairly independent of z_{23}, the plot for Area 10 has a conical shape. In this case, it is not correct to use the same distribution of ε_t irrespective of z_{23}; here the model will give too large inflows for higher values of the previous week. One reason for the difficulty is probably that the chosen week 24 is around the average snow melting peak, and so it is reasonable that for the highest inflows there will be a reduction afterwards. It is not clear how this can be dealt with using a linear model.

In the case of several inflow series to the system, one may want to carry out a 'model reduction' to get a noise vector of lower dimensionality. In Sect. 2.3, it was outlined how this can be done by principal components. The success of this procedure varies. In practice, one can retain at most three or four principal components when the total number of noise vectors is to be kept at a reasonable level. (For example, Four principal components represented by 3 points each gives 81 different

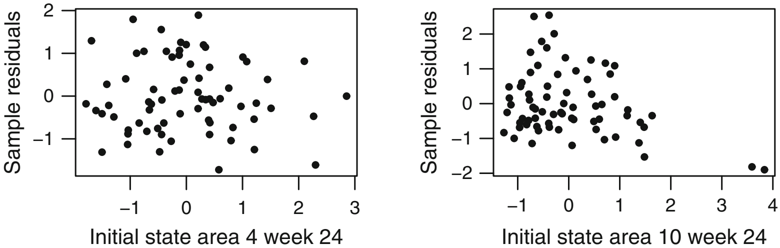

Fig. 5 Residual dependence on initial state. *Left*: Area 4. *Right*: Area 10

inflow cases.) A typical choice can be three principal components, represented by 5, 3 and 2 points, respectively, giving 30 combinations. It turns out that SDDP runs are not always very sensitive to the choice. For a geographically concentrated ('local') system, three principal components may work quite well, as in most cases there is a lot of cross-correlation between the inflows.

For a global model as outlined in Sect. 6, model reduction is more difficult. The inflow series shows less spatial correlation here. For the 16-area case mentioned earlier, it was found that seven principal components were required to cover 90% of the variance. Three principal components cover 75% of the variance. A problem may sometimes arise here, in the case of small areas where most of the variation in the inflow is contained in one or two principal components that are neglected. The inflow variation in such an area may almost disappear. A noise vector with too low variance may give strategies that are too optimistic.

As an alternative in this situation, we have experimented with a direct Monte Carlo approach: We construct the inflow noise vectors by sampling from the sample residuals ε_t available after the fitting of the VAR1-model. Here we avoid systematically cutting off some of the noise space by dropping principal components. On the other hand, it is not clear what a suitable sample size is. One must also check the sampled vectors, so that outliers in the sample of residuals ε_t are not included.

A further difficulty with a linear inflow model such as (5) is that it may generate negative inflows, particularly in weeks where the average inflow is low compared to the standard deviation. We cannot change this to a non-negative value for the scenario in case, since we need consistence to avoid non-convexity. In the SDDP algorithm, this is handled by penalty variables. However, this gives a (further) inconsistency between the forward and backward runs and may slow down convergence.

Negative inflows could be avoided by working with the logarithm of the inflow ($\log q$ instead of q). However, it can be shown that this transformation leads to non-convexity.

In connection with the algorithm description in Sect. 3.2, we mentioned that in the forward SDDP run, we use 'observed' inflow scenarios, which may not be fully consistent with the inflow model used on the backward recursion. This may lead to slower convergence. In a special case, similar to the model of Belsnes et al. (2003), an effect was directly visible in the results. It was with a special version of the global model of Sect. 6, where an internal quota market for CO_2 was modelled as a reservoir. The marginal price of the quota came out with a time variation that was not in line with standard economic theory, due to the inflow inconsistencies. Later simulations of this case showed that if 1,000 consistent inflow scenarios generated from the inflow model (5) were used, the incorrect price behaviour would disappear.

We have not carried out much other direct comparison to study the effect of 'observed' inflows vs. inflows generated by sampling in the stochastic inflow model. In scheduling with the local model, however, we believe that it is most realistic to use observed series of inflow and price, since that helps keep the correct coupling between inflow and price.

Summing up, it is not always easy to find a good inflow representation. Especially in the case of a global model with many inflows that are weakly correlated, the dimensionality of the noise space seems to be a problem. For the local model, the coupling between inflow and price should be further investigated.

8 Computational Issues

The SDDP computations are computer-intensive, and more so with the addition of the price state (Sects. 2 and 3). Depending on the model and the level of detail, the computer time on a Pentium IV computer or similar is in the range half an hour to more than a day. This is the case when weekly time steps are used everywhere. The computations are usually stopped after a given number of main iterations, typically 50–100. The number S of inflow scenarios is typically 50–70, and the number of discrete inflow values at each time step in the backward computations usually is at most 30. The number of hyperplanes stored for each future cost function usually is around 1,000. Solution of the linear programming sub-problems is carried out by general LP software, either a commercial solver or open software.

In a case with 28 reservoirs and 13 power stations, with four load periods, the model for a single transition (a single case of (15) through (19)) has around 460 variables and 33 to some 60 constraints (varying with the number of cuts entered into the active model.) (In this case, relaxation is not used for the reservoir balances but only for the cuts, since most reservoirs are rather small.) The number of simplex iterations may vary from very few (including zero and one), when starting from an advanced basis, up to a few hundreds when an initial basis is not available. One may think that the most usual commercial packages may not be optimal for these problems, since they are primarily intended and tuned for much larger problems. On the other hand, if one wants a finer time resolution within the week, possibly with separate power and reservoir balances for each sub-interval, then the number of rows and columns grows approximately linearly with the number of sub-intervals, and the model for a single transition becomes larger.

Computer time varies much with the size of the system, the length of the study period (T), the number of 'load periods', the number of inflow scenarios and price model levels (M), and the number of discrete inflow values at each stage (K). For the model size mentioned above, with 75 inflow scenarios, 7 price points and $K = 10$ discrete inflow values, the computing time is around 10 h on a single-core Pentium-class processor, using 50 iterations, and solving about 108 million small single-transition LP problems of the kind (21) with the associated constraints. The study period was two and a half year. Smaller and simpler systems can be solved in less than an hour.

Parallel processing has been implemented to reduce computing time. Multicore computers allow this to be easily applied by utilities. The reduction in computation time is almost proportional to the number of cores, but this has not been tested on large-scale parallel computers.

Another way of 'parallel processing' is also used by utilities. The idea is to split a large system with several watercourses into one (sub-)system for each watercourse and optimize each watercourse on a separate computer. This is a perfect decoupling for a price taker with no transmission limits to the market.

9 Discussion

The scheduling algorithm for a local system described in Sect. 3 has been implemented for medium and long-term hydro scheduling. The main advantage of such an algorithm (as with the ordinary SDDP algorithm) is that one can provide stochastic optimization with a detailed model of the hydropower system, so that one obtains reliable incremental water values for each reservoir. The algorithm has been used by power companies for some years and applied to systems with sizes ranging from 4 to over 50 reservoirs.

The global model of Sect. 6 has also been implemented, but is less used.

An obstacle to the use of both implementations is the computer time requirements. Some utilities have started using a parallel version of the local model.

The modelling of the stochastic processes inflow and price seems to be a major area where improvements are wanted. First there is the dimensionality problem. Especially with the global model of Sect. 6, there may be many almost independent inflows to deal with. This requires a relatively high number K of discrete inflow cases to be dealt with in the backward recursion; giving long computing times. An alternative to principal component analysis is sampling from the residuals of the inflow model. A linear regression for the inflow next week may sometimes be problematic, as shown in the example in Sect. 7.

For the fitting of the price model, usually only 50–75 scenarios of 'observed' series are available. This gives rough estimates of the transition probabilities. Some utilities have parametric price models that produce more than one price scenario for each inflow year. The price model can then be generated directly from the parametric model or estimated from a much larger number of price scenarios. As seen in Sect. 2.5, joint modelling of inflow and price remains a challenge.

In computations, we observe gaps between the upper and lower cost bounds. The upper bound, obtained in the forward pass, depends on the inflow scenarios, and is subject to sample variations, so that the 'gap' may even become negative. However, there is usually a positive gap. As mentioned in Sect. 3.2, the inflows used on the forward run are not fully consistent with the inflow model used on the backward run, since we use observed inflows and prices for the forward run. This, combined with the simplifications in price/inflow modelling, probably gives the main contribution to the gap. The size of the gap varies. For small models with a single inflow series, we have sample values in the range 1–2% of the cost, while in larger and more complicated systems with several inflow series, the gap can be in the range 10–15%.

A special problem is that of constructing the final value function $\Phi(x_T)$. If no good estimate for this function is available, the strategy obtained may not be optimal for the last year or so of the study period, and simulation results from this period may be misleading. In our implementation the final value function is based on aggregated water values from a standard SDP model that is distributed to individual reservoirs using heuristics. One should ensure that the study period is long enough.

10 Conclusion

This chapter reviews work on the application of SDDP-based algorithms for hydro scheduling, with some extensions, in the Nordic countries.

It seems that there is room for improvements, particularly in the stochastic models for inflow and price. This is especially the case when there are many independent inflows and in applications to risk management, green certificates or quota modelling. However, the present models work, and in particular the local model gives good results. Parallel processing helps shorten computing time.

Acknowledgements The authors thank International Centre for Hydropower, Trondheim, for permission to use material from Gjelsvik and Haugstad (2005).

References

Aouam T, Yu Z (2008) Multistage Stochastic hydrothermal scheduling. In: 2008 IEEE international conference on electro/information technology, Ames, IA, 18–20 May 2008. IEEE, NY, pp. 66–71

Belsnes MM, Haugstad A, Mo B, Markussen P (2003) Quota modeling in hydrothermal systems. In: 2003 IEEE Bologna PowerTech Proceedings, IEEE, NY

Bortolossi HJ, Pereira MV, Tomei C (2002) Optimal hydrothermal scheduling with variable production coefficient. Math Meth Oper Res 55(1):11–36

Botnen OJ, Johannesen A, Haugstad A, Kroken S, Frøystein O (1992) Modelling of hydropower scheduling in a national/international context. In: Broch E, Lysne DK (eds) Hydropower ’92. A.A. Balkema, Rotterdam

da Costa JP, de Oliveira GC, Legey LFL (2006) Reduced scenario tree generation for mid-term hydrothermal operation planning. In: 2006 international conference on probabilistic methods applied to power systems, vols. 1, 2, Stockholm, Sweden, 11–15 Jun 2006. IEEE, NY, pp. 34–40

de Oliveira GC, Granville S, Pereira M (2002) Optimization in electrical power systems. In: Pardalos PM, Resende MGC (eds) Handbook of applied optimization. Oxford University Press, London, pp. 770–807

Dreyfus SE (1965) Dynamic programming and the calculus of variations. Academic Press, New York

Fleten SE (2000) Portfolio management emphasizing electricity market applications: a stochastic programming approach/Stein-Erik Fleten. PhD thesis, Norwegian University of Science and Technology, Faculty of Social Sciences and Technology Management

Fosso OB, Belsnes MM (2004) Short-term hydro scheduling in a liberalized power system. In: 2004 international conference on power systems technology, Powercon 2004, Singapore, IEEE

Gjelsvik A, Haugstad A (2005) Considering head variations in a linear model for optimal hydro scheduling. In: Proceedings, Hydropower ’05: The backbone of sustainable energy supply, International centre for hydropower, Trondheim, Norway

Gjelsvik A, Wallace SW (1996) Methods for stochastic medium-term scheduling in hydrodominated power systems. Tech. Rep. A4438, Norwegian Electric Power Research Institute, Trondheim, Norway

Gjelsvik A, Belsnes MM, Håland M (1997) A case of hydro scheduling with a stochastic price model. In: Broch E, Lysne DK, Flatabø N, Helland-Hansen E (eds) Procedings of the 3rd international conference on hydropower, Trondheim/Norway/30 June–2 July 1997. A.A. Balkema, Rotterdam, pp. 211–218

Gjelsvik A, Belsnes MM, Haugstad A (1999) An algorithm for stochastic medium-term hydrothermal scheduling under spot price uncertainty. In: 13th power systems computation conference: Proceedings, vol. 2

Iliadis NA, Pereira MVF, Granville S, Finger M, Haldi PA, Barroso LA (2006) Benchmarking of hydroelectric stochastic risk management models using financial indicators. In: 2006 power engineering society general meeting, vols. 1–9, pp. 4449–4456. General Meeting of the Power-Engineering-Society, Montreal, Canada, 18–22 Jun 2006

Jardim D, Maceira M, Falcao D (2001) Stochastic streamflow model for hydroelectric systems using clustering techniques. In: Power tech proceedings, 2001 IEEE Porto, vol. 3, p 6. doi:10.1109/PTC.2001.964916

Johannesen A, Flatabø N (1989) Scheduling methods in operation planning of a hydro-dominated power production system. Int J Electr Power Energ Syst 11(3):189–199

Johnson RA, Wichern DW (1998) Applied multivariate statistical analysis. Prentice Hall, New Jersey

Labadie J (2004) Optimal operation of multireservoir systems: State-of-the-art review. J Water Resour Plann Manag-ASCE 130(2):93–111

Lindqvist J (1962) Operation of a hydrothermal electric system: a multistage decision process. AIEE Trans III (Power Apparatus and Systems) 81:1–7

Mariano SJPS, Catalao JPS, Mendes VMF, Ferreira LAFM (2008) Optimising power generation efficiency for head-sensitive cascaded reservoirs in a competitive electricity market. Int J Electr Power Energ Syst 30(2):125–133. doi:10.1016/j.ijepes.2007.06.017

Mo B, Gjelsvik A, Grundt A (2001a) Integrated risk management of hydro power scheduling and contract management. IEEE Trans Power Syst 16(2):216–221

Mo B, Gjelsvik A, Grundt A, Kåresen K (2001b) Hydropower operation in a liberalised market with focus on price modelling. In: 2001 Porto power tech proceedings, IEEE, NY

Mo B, Wolfgang O, Gjelsvik A, Bjørke S, Dyrstad K (2005) Simulations and optimization of markets for electricity and el-certificates. In: 15th power systems computation conference: Proceedings, PSCC

Pereira MVF (1989) Optimal stochastic operations scheduling of large hydroelectric systems. Electr Power Energ Syst 11(3):161–169

Pereira MVF, Pinto LMVG (1991) Multi-stage stochastic optimization applied to energy planning. Math Program 52:359–375

Philpott AB, Guan Z (2008) On the convergence of stochastic dual dynamic programming and related methods. Oper Res Lett 36(4):450–455. doi:10.1016/j.orl.2008.01.013

Røtting TA, Gjelsvik A (1992) Stochastic dual dynamic programming for seasonal scheduling in the Norwegian power system. IEEE Trans Power Syst 7(1):273–279

Scott TJ, Read EG (1996) Modelling hydro reservoir operation in a deregulated electricity market. Int Trans Oper Res 3(3–4):243–253. doi:10.1111/j.1475-3995.1996.tb00050.x

Stage S, Larsson Y (1961) Incremental cost of water power. AIEE Trans III (Power Apparatus and Systems) 80:361–365

Tilmant A, Kelman R (2007) A stochastic approach to analyze trade-offs and risks associated with large-scale water resources systems. Water Resour Res 43, w06425. doi:10.1029/2006WR005094

Wallace SW, Fleten SE (2003) Stochastic programming models in energy. In: Ruszczyński A, Shapiro A (eds) Handbooks in operations research, vol. 10. Elsevier, Amsterdam