STOCHASTIC DUAL DYNAMIC PROGRAMMING APPLIED TO NONLINEAR HYDROTHERMAL MODELS

Santiago Cerisola, Jesús M. Latorre, Andrés Ramos
Instituto de Investigación Tecnológica (IIT)
Escuela Técnica Superior de Ingeniería ICAI
Universidad Pontificia Comillas
Alberto Aguilera 23, 28015, Madrid Spain
santiago.cerisola@iit.upcomillas.es

Abstract:

In this paper we apply stochastic dual dynamic programming decomposition to a nonlinear multistage stochastic hydrothermal model where the nonlinear water head effects on production and the nonlinear dependence between the reservoir head and the reservoir volume are modeled. The nonlinear constraints that represent the production function of a hydro plant are approximated by McCormick envelopes. These constraints are split into smaller regions and the McCormick envelopes are used for each region. Binary variables are used for this disjunctive programming approach which complicates the application of the decomposition method. We resort to a variant of the L-shaped method for solving the MIP subproblem with binary variables at any stage inside the stochastic dual dynamic programming algorithm. A realistic large-scale case study is presented.

I Introduction

The hydrothermal coordination problem (HCP) can be defined as the problem of deciding the operation of thermal units as well as reservoir levels and water released from the upstream reservoirs to the hydro plants (El-Hawary and Christensen 1979). These decisions are taken with the purpose of minimizing the total operation cost while satisfying the load demand. Typically, the HCP problem is defined for several time periods (months, weeks, days, hours) connected by intertemporal restrictions such as the equations of reservoir inventory. Modeling of thermal units usually depends on the time scope of the problem. For medium and long-term models (up to several years), the variable cost of thermal units can be modeled considering a constant heat rate with continuous unit commitment decisions. For short-term models (up to one week), this approach is improved with the addition of startup and shutdown costs and binary unit commitment decisions. For very short-term unit commitment models, minimum startup and shutdown time constraints can also be considered (Nowak and Romisch 2000).

The power output of a hydro unit is a function of the water discharge, the net head and the unit efficiency (El-Hawary and Christensen 1979), (Finardi, Silva et al. 2005). Thus, it can be considered as the result of multiplying a transforming coefficient by the water discharge. Let us denote this coefficient as production function. In simplified models, a constant production function can be considered. With this approach, the water value of different reservoirs is the same, and the optimization model suffers from multiple solutions because the substitution of thermal production by hydro production can be done indifferently by any hydro plant of the system. The multiplicity of solutions of the HCP may sometimes derive in inadequate management of reservoir levels. The use of a constant production function does not take into account the nonlinear water head effect that increases the production of the hydro plant with the reservoir head. In this paper we model this relation in a simplified manner by using a bilinear relation. This allows considering the effect that it has on the optimal operation of hydrothermal systems with multiple basins and multiple cascaded reservoirs.

The stochastic nature of HCP must be considered. These models typically incorporate stochasticity in natural water inflows or in load demand, which can be modeled with separate scenarios, scenario trees or recombining scenario trees. In this paper we model the stochasticity of natural hydro inflows by means of a recombining tree. Thus, for any period of the planning horizon, we determine a fixed number of scenarios to describe this multivariate random variable of such period and create the recombining tree assigning transition probabilities among states of consecutive periods. The selection of the scenarios is carried out with a clustering technique based on the neural gas algorithm as described in (Latorre 2003), (Latorre, Cerisola et al. 2007).

Introducing stochasticity in hydrothermal models greatly increases the difficulty of solving the optimization problem. When the number of scenarios is kept small enough, the deterministic equivalent problem (DEP) can be formulated and written algebraically, and it can be solved with the use of commercial solvers (ILOG 2009). Typically, this DEP adopts the structure of a scenario tree and can be solved directly or by means of decomposition methods like Benders decomposition (Benders 1962) or Lagrangean Relaxation (LR) method (Geoffrion 1974).

When the number of scenarios exceeds the capability of formulating the DEP, the solution of the problem has to rely upon sampling methods for stochastic programming (SP) (Shapiro 2003). Sampling techniques for SP are classified in the literature as "interior" and "exterior" methods. The former methods resort to sampling inside the algorithm during its evolution. For example, in SP interior sampling is done whenever an slope or an intercept of a cut needs to be calculated (Infanger 1994), (Linderoth, Shapiro et al. 2006). The later methods approximate the problem by estimating the objective function with a reduced number of scenarios and optimizes a *sample average approximation* of the objective function (Morton 1993) (Shapiro 2003). The stochastic dual dynamic programming technique (SDDP) developed by Pereira (Pereira and Pinto 1991) falls among these methods and it is the technique applied in this paper.

This technique is a Benders-type method that solves a multistage problem by solving smaller subproblems for each stage. Each subproblem incorporates the part of the objective function corresponding to that stage together with the future cost function as a function of the decision taken in that stage. The fundamental property that enables the application of such an algorithm is the convexity of this future cost or recourse function. The convexity of the recourse function is obtained immediately for linear problems from the theory of duality of linear programming (Minoux 1986).

Unfortunately, for nonlinear problems, the convexity of the recourse function is no longer guaranteed. It can be proven that the recourse function is convex provided that the second stage problem is also convex (Geoffrion 1972). However, if the nonlinear water head effect is included in the HTC model, the recourse function becomes nonconvex and Benders-type decomposition methods cannot be immediately applied. We will show this nonconvexity of the recourse function in the paper.

An alternative to apply a Benders or SDDP type algorithm when nonlinear nonconvex head effects are included in a model consists of convexifying the recourse function using LR techniques. A similar convexification technique to the one we present in this paper was successfully applied to solve weekly stochastic unit commitment models (Cerisola, Baíllo et al. 2009).

One of our main contributions in the current paper is to take advantage of the bilinear modeling for the production function of hydro plants and to replace this bilinear relation with its McCormick envelope, (McCormick 1976), (Liberti and Pantelides 2006), (Iyer and Grossman 1997), (Quesada and Grossmann 1995). This substitution underestimates the nonconvex recourse function by a convex function that can be used as an approximation of the future cost function. A further step can be done at improving the accuracy of that approximation. With this purpose, we create a grid for the variables of the bilinear function and create the McCormick envelope of the bilinear function over each cell of the grid. We force the model to select one cell for every bilinear function for each period using disjunctive programming (Williams 1993; Williams 1999). Binary variables are thus included in each stage in the HCP. By doing this we obtain an upper function and a lower function that bound the bilinear surface. Those functions are neither convex nor concave, but their description by using MIP gives the possibility of using branch and bound techniques and commercial solvers for solving the problem. We denote this approximation as multiple-cell McCormick approximation and, in a parallel manner, the McCormick envelope linearization as single-cell McCormick approximation. The nonconvex function that approximates the future cost function can now be convexifed by using LR techniques (Frangioni 2005; Lemacheral 2007). The SDDP decomposition technique is extended to multistage stochastic integer problems where binary variables appear at any stage. This extension is our second main contribution and is developed in detail in section 3 of the article.

The paper is organized as follows. Section 2 summarizes the hydro subsystem model considered in this work, describing the simplifications as well as the taken assumptions. This section also presents the approximation for the bilinear relation as well as the disjunctive programming model. Section 3 describes the extension of the SDDP technique to the case where binary variables appear in any stage of the model. In section 4 we present a real-sized problem of a HCP with multiple basins and cascaded reservoirs. Section 5 is devoted to a numerical analysis of the method for the previous described problem. Figures showing convergence and reservoir level evolutions for the stochastic case are presented. We conclude the paper suggesting future lines of research.

II Hydro Subsystem Modeling

The power output P_h, in MW, of a hydro unit can be expressed as

$$P_h = qhg\eta_t\eta_g \tag{1}$$

where g stands for the gravity constant in m/s^2, q is the water release in m^3/s, h the net water head in m and η the efficiency of the unit, which comprises the turbine efficiency η_t and the generator efficiency η_g (El-Hawary and Christensen 1979).

The net head of a plant is defined as the difference of the forebay height y, the tailrace height y_T and the head losses in the penstock, which can be modeled as a quadratic function of the water release, kq^2.

$$h = y - y_T - kq^2 \tag{2}$$

The tailrace height accounts for the height of its discharging level and the downstream reservoir. It is also a function of the release as well as the spillage. In our model we neglect the losses in the penstock and the elevation due to the own water flow discharge. We explicity consider the tailrace level as the maximum value of the downstream reservoir forebay height and the drainage level of the hydro plant. Thus, we simplify the net head modeling as

$$h = y - y_T \tag{3}$$

The production function is usually given by level curves that relate the power output of a plant with the net head for a known amount of water released through the turbine. Figure 1 shows a typical hill diagram. (A Mathematical Model for the Efficiency Curves of Hydroelectric units A. L. Diniz, P. P. I. Esteves, C. A. Sagastizábal). It may be observed that given a net head (vertical dashed line) for the reservoir, there is an optimum power output (thick line).

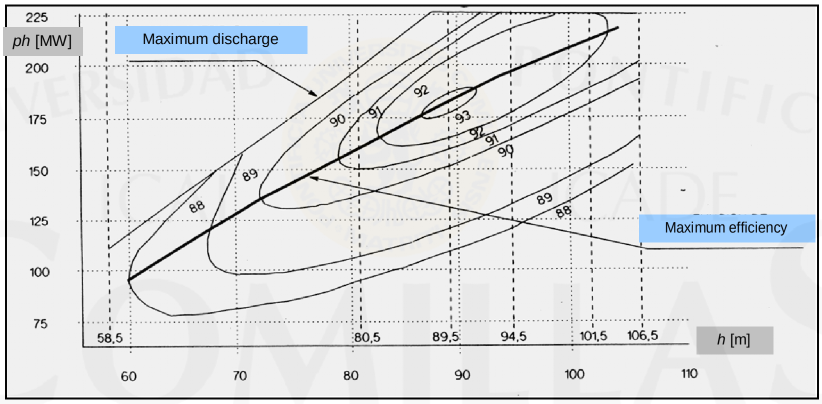


Figure 1. Hill diagram of a hydro plant.

Even more, under the rational assumption that given a net head, the outflow through the turbine will be the optimal one, there is a relation between the net head of the hydro plant and the efficiency of the unit. This relation can be modeled as a piecewise linear function or as a polynomial function (Finardi, Silva et al. 2005).

$$\eta = \phi\left(h\right) \tag{4}$$

Merging previous equations, we have the power output of the hydro unit modeled as

$$P_h = qh\eta = qh\phi\left(h\right) \tag{5}$$

The right part of above equation, $h\phi\left(h\right)$, formulates a relation that only depends on the net head. We approximate this relation with an affine function, and denote it as production function

$$C_e\left(h\right) = h\phi\left(h\right) \approx \alpha + \beta h \tag{6}$$

Thus we finally have a bilinear expression that relates the power output of each hydro unit to its net head and the water released through the turbine.

$$P_h = qC_e(h) \approx q(\alpha + \beta h) \tag{7}$$

For the multistage model we present in this paper, we assume that the production function $C_e(h)$ as well as the water release q remains constant for each period (that will be one week). The above equation expressed in terms of power is used in terms of energy considering the time duration Dur of each period. Thus, we consider

$$Dur \cdot P_h = Dur \cdot q(\alpha + \beta h) \tag{8}$$

and replace in the equation the water release by time unit with the water release w in the period

$$Dur \cdot P_h = w(\alpha + \beta h) \tag{9}$$

The introduction of the production function in the HCP incorporates the signal to the model that the higher the reservoir level, the greater the hydro production and the greater the saved thermal cost. We present firstly an academic model to show this idea.

II.1 Example 1

For the shake of simplicity let us consider a two period case. We assume to have one thermal unit and one hydro plant associated with a reservoir. We note:

Variables

x_1, x_2	Thermal output in periods 1 and 2
y_1, y_2	Hydro output in periods 1 and 2
w_1, w_2	Water release in periods 1 and 2
r_0, r_1, r_2	Initial reserve level and reserve levels at the end of periods 1 and 2

Parameters

c_1, c_2	Constant production function in periods 1 and 2
i_1, i_2	Natural inflows in periods 1 and 2
d_1, d_2	Load demand in periods 1 and 2

Typically, the constant production functions will be equal for both periods. However, we keep the difference for the later comparison with the example 2.

We assume a normalized cost for thermal output. Thus, a linear problem that minimizes the total operation cost can be formulated as:

$$\begin{array}{lll} \min x_1 + x_2 & & \\ x_1 + y_1 = d_1, & x_2 + y_2 = d_2 & \text{Demand balance equations} \\ y_1 = c_1 w_1, & y_2 = c_2 w_2 & \text{Hydro production constraints} \\ r_1 = r_0 + i_1 - w_1, & r_2 = r_1 + i_2 - w_2 & \text{Reserve balance equations} \end{array} \tag{10}$$

Neglecting variable bounds and after same algebra we obtain $w_1 = r_0 + i_1 - r_1$, so that $y_1 = c_1\left(r_0 + i_1 - r_1\right)$ and consequently $x_1 = d_1 - c_1\left(r_0 + i_1 - r_1\right)$. Similarly for the second period, $w_2 = r_1 + i_2 - r_2$, so that $y_2 = c_2\left(r_1 + i_2 - r_2\right)$ and consequently $x_2 = d_2 - c_2\left(r_1 + i_2 - r_2\right)$. The objective function becomes

$$\min d_1 - c_1\left(r_0 + i_1 - r_1\right) + d_2 - c_2\left(r_1 + i_2 - r_2\right) \tag{11}$$

In hydrothermal models, usually the initial and final reserve levels are known values. In our example, the unique decision variable is just r_1. Gathering terms we obtain the objective function

$$\min d_1 - c_1\left(r_0 + i_1\right) + d_2 - c_2\left(i_2 - r_2\right) + \left(c_1 - c_2\right) r_1 \tag{12}$$

which for the case of identical and constant production functions, $c_1 = c_2$, leads to

$$\min d_1 - c_1\left(r_0 + i_1\right) + d_2 - c_2\left(i_2 - r_2\right) \tag{13}$$

Note that this is an objective function that does not depend on the decision variable. For this reason, the optimization problem presents multiplicity of solutions. A large range of first-period final reserves gives the same objective value and no evolution profile is decided for the reserve. We now compare the above case with the next one where the net head effect of the hydro plant is considered in the model.

II.2 Example 2

We consider a linear relation between the production function and the reserve level. Let's the new parameters be:

Parameters

α, β Intercept and slope of the production function

We incorporate the next relation into the optimization model

$$c_1 = \alpha + \beta r_1$$
$$c_2 = \alpha + \beta r_2 \tag{14}$$

Now, the outflow-energy function is nonlinear and the optimization problem also becomes nonlinear. We again neglect variable bounds and after some algebra we obtain $y_1 = (\alpha + \beta r_1)(r_0 + i_1 - r_1)$ and $y_2 = (\alpha + \beta r_2)(r_1 + i_2 - r_2)$. The objective function gets the form

$$\min d_1 - (\alpha + \beta r_1)(r_0 + i_1 - r_1) + d_2 - (\alpha + \beta r_2)(r_1 + i_2 - r_2) \tag{15}$$

Notice that r_1 is the unique decision variable of the problem. The optimality condition is

$$-\beta(r_0 + i_1 - r_1) + (\alpha + \beta r_1) - (\alpha + \beta r_2) = 0 \tag{16}$$

This leaves the reserve decision at the end of the first period as

$$r_1 = \frac{r_0 + r_2}{2} + \frac{i_1}{2} \tag{17}$$

and the water released in the periods

$$w_1 = \frac{r_0 - r_2}{2} + \frac{i_1}{2},\ w_2 = \frac{r_0 - r_2}{2} + \frac{i_1}{2} + i_2 \tag{18}$$

Note that if initial and final reserve levels are equal, half of the natural inflows for the first period are saved for the second period. The optimization problem decides to keep the reserve at a high level, so that the thermal cost of the first period offsets the cost of the second period.

Although this is a small academic example it presents the overall behavior of the reserve levels under an efficient economic dispatch for the hydro units. For realistic models as the one we present in this paper, the multiplicity of constraints and variable bounds together with multiple basins and cascaded reservoirs, and with the additional complexity of stochasticity in the natural inflows, an analytic rule about the behavior of the reservoir is impossible to obtain. However, it is expected that the inclusion of the nonlinear water head effect increases the reservoir levels in comparison with the profiles that a constant production function would produce.

III Problem Formulation

We consider a time scope divided into periods p (weeks) and these periods into load levels n (typically peak and off-peak) and a generation portfolio with different thermal units and hydro plants. Hydro plants are assigned to reservoirs, taking into account that some reservoirs can be shared by different hydro plants and they can be located in cascade. We first list the indexes, variables and parameters to ease the exposition of the equations of the model formulation. Although the model is stochastic, we explicitly avoid showing the scenario index in the mathematical formulation.

Indexes

P, p	Period
N, n	Load level (peak, off-peak)
T, t	Thermal unit
H, h	Hydro plant
E, e	Reservoir
$H(e)$	Hydro plants sharing reservoir e
$Up(e)$	Upstream hydro plants of reservoir e
$Down(e)$	Downstream hydro plants of reservoir e

Parameters

$D_{p,n}$	Load demand of period p and load level n	[MWh]
$Dur_{p,n}$	Duration of period p and load level n	[h]
CV_t	Variable cost of thermal unit t	[€/MW]
Cdr_e, Cer_e	Penalty for the slack and surplus of final level of reservoir e	[€/MWh]
η_h	Efficiency of pumped storage hydro plant h	[p.u]
$\underline{PT}_t$, $\overline{PT}_t$	Minimum and maximum output of thermal unit t	[MW]
$NI_{e,p}$	Natural inflows in reservoir e in period p	[MWh]
α_h, β_h	Intercept and slope of the linear production function of hydro plant h	[MW/m^3/s, MW/m^4/s]
a_e, b_e, c_e	Parameters of the quadratic reservoir volume curve	
T_h	Drainage level of a hydro plant	[m]
R^*_{eP}	Final reservoir level of hydro plant h	[GWh]

Variables

$P_{t,p,n}$	Power output of thermal unit t in period p and load level n	[MW]
$P_{h,p,n}$	Power output of hydro plant h in period p and load level n	[MW]

$B_{h,p,n}$ Consumption of hydro plant h due to pumping in period p and load level n [MW]

$R_{e,p}$	Level of reservoir e at the end of period p	[MWh]
$S_{e,p}$	Spillage of reservoir e in period p	[MWh]
$W_{h,p}$	Water release of hydro plant h in period p	[hm^3]
$H_{h,p}$	Net head of hydro plant h in period p	[m]
$F_{e,p}$	Forebay level of reservoir e in period p	[m]
$T_{h,p}$	Tailrace level of hydro plant h in period p	[m]
$I_{e,p}$	Inflows in reservoir e in period p	[MWh]
$C_{h,p}$	Production function for hydro plant h in period p	[MW/m^3/s]
$dr_{e,P}$, $er_{e,P}$	Slack and surplus of final level of reservoir e	[MWh]

Equations

The energy balance equation guarantees the production of the required demand

$$D_{p,n} = \sum_{t\in T} P_{t,p,n} + \sum_{h\in H} P_{h,p,n} - \sum_{h\in H} B_{h,p,n} / \eta_h \quad \forall p,n \tag{19}$$

The water inventory equation establishes that the reserve level at the end of a period is obtained as the initial reserve level plus the period inflows minus the own hydro reservoir release. In this equation, the period inflows comprise the natural inflows as well as their upstream hydro plant water releases or spillages.

$$\begin{aligned} R_{e,p} &= R_{e,p-1} + I_{e,p} - \sum_{h\in H(e)} W_{h,p} - \sum_{n,h\in H(e)} Dur_{p,n} B_{h,p,n} - S_{e,p} \quad \forall e,p \\ I_{e,p} &= NI_{e,p} + \sum_{\substack{h\in H(e') \\ e'\in Up(e)}} W_{h,p} + \sum_{\substack{n,h\in H(e') \\ e'\in Down(e)}} Dur_{p,n} B_{h,p,n} + \sum_{e'\in Up(e)} S_{e',p} \quad \forall e,p \end{aligned} \tag{20}$$

The hydro plant output is related to the water release $W_{h,p}$ and to the production function $C_{h,p}$, which we approximate as an affine function of the net head of the plant $H_{h,p}$. Thus

$$C_{h,p} = \alpha_h + \beta_h H_{h,p} \quad \forall h,p \tag{21}$$

$$\sum_n Dur_{p,n} P_{h,p,n} = W_{h,p} C_{h,p} \quad \forall h,p \tag{22}$$

The net head of the plant is computed as the forebay level of the upstream reservoir $F_{e,p}$ minus the tailrace level of the hydro plant $T_{h,p}$

$$H_{h,p} = F_{\substack{e,p \\ h \in H(e)}} - T_{h,p} \qquad \forall h, p \tag{23}$$

The reserve level of the reservoir $R_{e,p}$ is modeled as a quadratic function of the forebay level, see Figure 2,

$$R_{e,p} = a_e + b_e F_{e,p} + c_e F_{e,p}^2 \qquad \forall e, p \tag{24}$$

and the tailrace level of the hydro plant is modeled as the maximum between the forebay level of the downstream reservoir $F_{e',p}$ or the drainage level of the plant T_h

$$T_{h,p} = \max\left(F_{e',p}, T_h\right) \qquad \forall h, p \tag{25}$$

where e' indicates the downstream reservoir of hydro plant h.

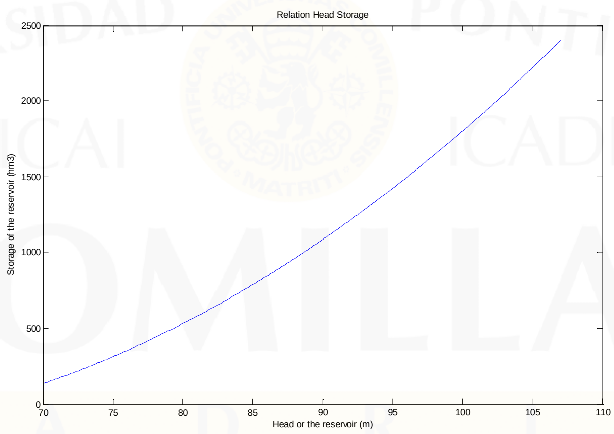


Figure 2. Volume of a reservoir as a quadratic function of its head.

With the purpose of obtaining a realistic profile, the model forces the production in peak hours, load level n_1, to be greater than the production in off-peak hours, load level n_2. On the contrary, it also forces the energy consumption for pumping to satisfy the opposite inequality.

$$\begin{aligned} P_{t,p,n_1} &\geq P_{t,p,n_2} \qquad \forall t, p \\ P_{h,p,n_1} &\geq P_{h,p,n_2} \qquad \forall t, p \\ B_{h,p,n_1} &\leq B_{h,p,n_2} \qquad \forall t, p \end{aligned} \tag{26}$$

The model also incorporates a final reservoir level that ought to be achieved by each reservoir. This final reserve is determined by a long-term model and integrated in this medium-term model by means of using slack and surplus variables.

$$R_{e,P} = er_{e,P} - dr_{e,P} - R^{*}_{eP} \quad \forall e \tag{27}$$

The objective function minimizes the total variable operating cost and penalizes the deviation with respect to the final reserve level.

$$\min \sum_{t,p,n} Dur_{p,n} CV_t P_{t,p,n} + \sum_{e} \left(Cdr_e dr_{e,P} + Cer_e er_{e,P} \right) \tag{28}$$

III.1 Linear Approximation

We approximate the above nonlinear HCP by a linear problem. Firstly, the head-reserve quadratic relation (24) is replaced with a piecewise linear outer approximation, that is not worthwhile to detail it.

On the other hand, the bilinear relation of the hydro plant production (22) is replaced by its convex envelope, which is known in the literature as McCormick approximation [McCormick 1973].

The McCormick approximation considers a bilinear function $z = xy$ with $x \in \left[\underline{x}, \overline{x}\right]$ and $y \in \left[\underline{y}, \overline{y}\right]$ and replaces it by the next family of linear constrains.

$$\begin{aligned} z &\geq x\overline{y} + \overline{x}y - \overline{xy} \\ z &\geq x\underline{y} + \underline{x}y - \underline{xy} \\ z &\leq x\overline{y} + \underline{x}y - \underline{x}\overline{y} \\ z &\leq x\underline{y} + \overline{x}y - \overline{x}\underline{y} \end{aligned} \tag{29}$$

For this particular hydro scheduling problem, we replace equation (22) with

$$\begin{aligned} \sum_{n} Dur_{p,n} P_{h,p,n} &\geq W_{h,p} \overline{C}_{h,p} + \overline{W}_{h,p} C_{h,p} - \overline{W}_{h,p} \overline{C}_{h,p} \\ \sum_{n} Dur_{p,n} P_{h,p,n} &\geq W_{h,p} \underline{C}_{h,p} + \underline{W}_{h,p} C_{h,p} - \underline{W}_{h,p} \underline{C}_{h,p} \\ \sum_{n} Dur_{p,n} P_{h,p,n} &\leq W_{h,p} \overline{C}_{h,p} + \underline{W}_{h,p} C_{h,p} - \underline{W}_{h,p} \overline{C}_{h,p} \\ \sum_{n} Dur_{p,n} P_{h,p,n} &\leq W_{h,p} \underline{C}_{h,p} + \overline{W}_{h,p} C_{h,p} - \overline{W}_{h,p} \underline{C}_{h,p} \end{aligned} \tag{30}$$

The surface defined by the bilinear constraint together with the McCormick envelope (single-cell approximation) that replaces it is depicted in the next figures. Both images represent the same function and the axes have been rotated to appreciate the nonlinearity of the function.

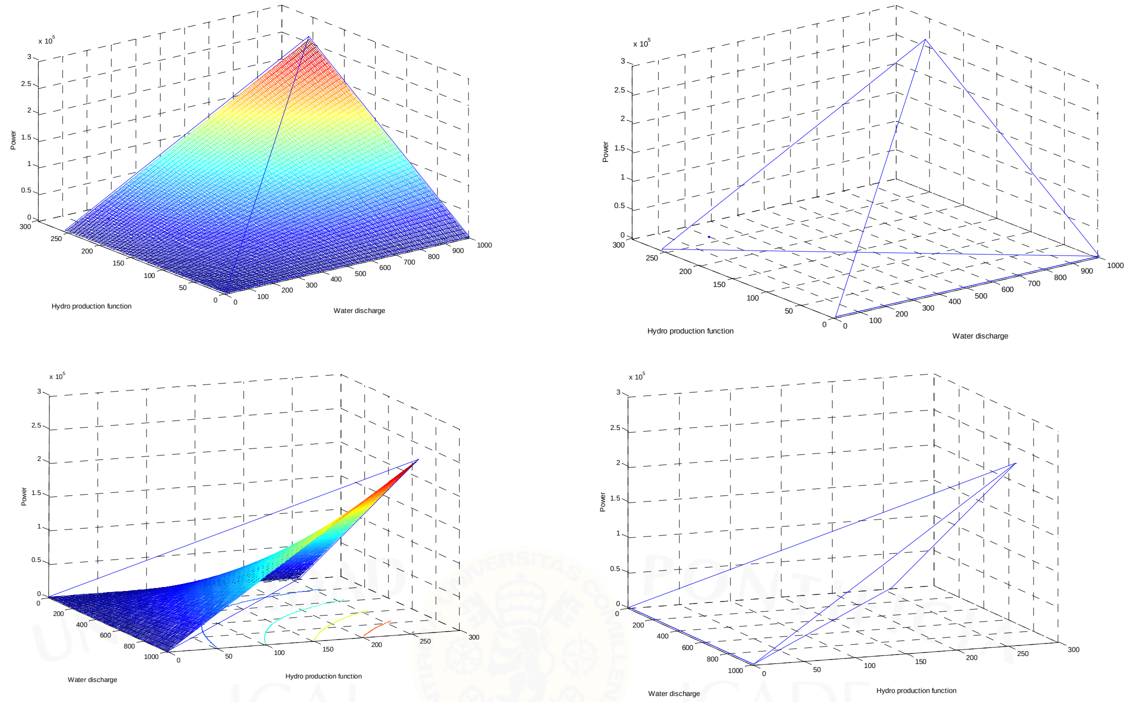


Figure 3. Bilinear surface and McCormick approximation.

It must be outlined that the precision of the McCormick approximation increases with the precision of the variable bounds. Thus, when modeling the HCP the variable bounds should be as tight as possible.

III.2 MIP Approximation

With the purpose of increasing the accuracy of the approximation, we divide the domain of the bilinear function into smaller rectangles, and build the McCormick approximation for each individual cell. We use disjunctive programming techniques to force the model to select one cell out of the total and therefore we modify equations (30) by introducing binary variables.

For the mathematical formulation of this approach we create a grid for the variables of the bilinear relation (water release and production function) $W^{i}_{h,p}$, $i = 1,\dots,I$ and $C^{j}_{h,p}$, $j = 1,\dots,J$, that accounts for a total of $I \cdot J$ cells. Note that $W^{I}_{h,p}$, $W^{0}_{h,p}$ and $C^{J}_{h,p}$, $C^{0}_{h,p}$ coincide with the upper and lower bounds of variables $W_{h,p}$ and $C_{h,p}$.

The lower bounds for the variables will be used when needed.

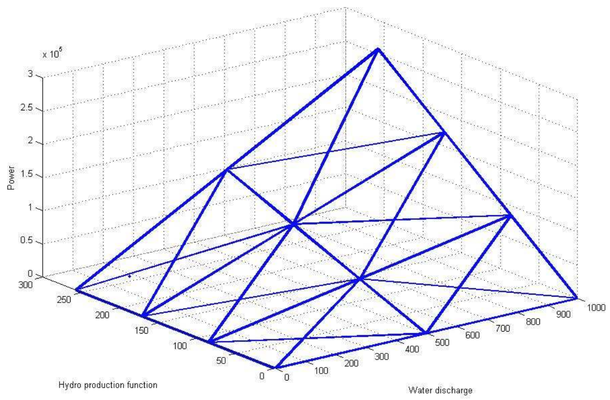


Figure 4. Example of approximation of the bilinear hydro output function with a multiple-cell McCormick approximation for a 3x2 grid. 3 slots for the production function variable and 2 slots for the water release variable.

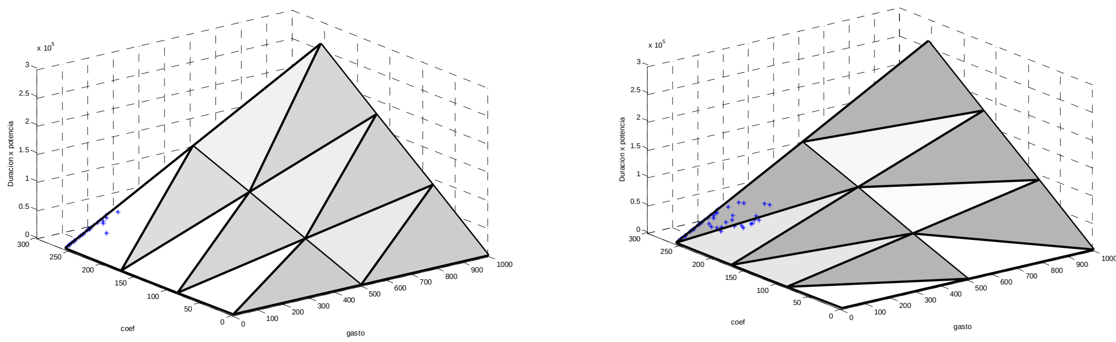


Figure 5. Upper (left) and lower (right) approximation of the bilinear hydro output function with the McCormick multiple cell approximation.

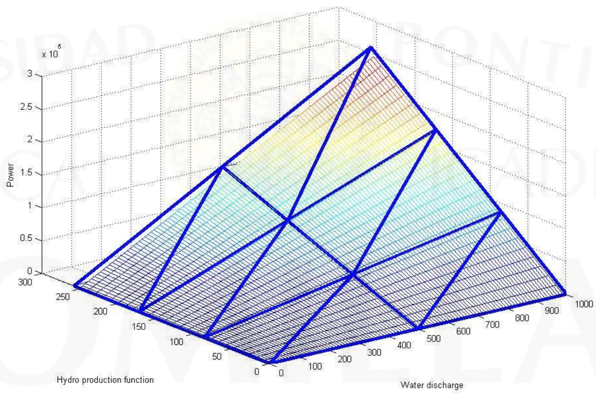


Figure 6. Bilinear surface of hydro output.

The convex envelope of the grid cell corresponding to the (i,j) slot is given by the constraints

$$\begin{aligned}
&\sum_n Dur_{p,n} P_{h,p,n} \geq W_{h,p} C^{j}_{h,p} + W^{i}_{h,p} C_{h,p} - W^{i}_{h,p} C^{j}_{h,p} \\
&\sum_n Dur_{p,n} P_{h,p,n} \geq W_{h,p} C^{j-1}_{h,p} + W^{i-1}_{h,p} C_{h,p} - W^{i-1}_{h,p} C^{j-1}_{h,p} \\
&\sum_n Dur_{p,n} P_{h,p,n} \leq W_{h,p} C^{j}_{h,p} + W^{i-1}_{h,p} C_{h,p} - W^{i-1}_{h,p} C^{j}_{h,p} \\
&\sum_n Dur_{p,n} P_{h,p,n} \leq W_{h,p} C^{j-1}_{h,p} + W^{i}_{h,p} C_{h,p} - W^{i}_{h,p} C^{j-1}_{h,p}
\end{aligned} \tag{31}$$

Additionally, we need to impose the condition that just one of the McCormick envelopes needs to be active. In this is disjunctive programming approach we reformulate above equations by introducing a binary variable $u^{i,j}$ that indicates whether the (i,j) cell is selected $u^{i,j} = 1$ or not $u^{i,j} = 0$. We choose a big-M method (Williams 1993) leading to the next equations

$$
\begin{aligned}
&\sum_n Dur_{p,n} P_{h,p,n} \geq W_{h,p} C^{j}_{h,p} + W^{i}_{h,p} C_{h,p} - u^{i,j}_{h,p} W^{i}_{h,p} C^{j}_{h,p} - \left(1 - u^{i,j}_{h,p}\right) M1^{i,j}_{h,p} \\
&\sum_n Dur_{p,n} P_{h,p,n} \geq W_{h,p} C^{j-1}_{h,p} + W^{i-1}_{h,p} C_{h,p} - u^{i,j}_{h,p} W^{i-1}_{h,p} C^{j-1}_{h,p} - \left(1 - u^{i,j}_{h,p}\right) M2^{i,j}_{h,p} \\
&\sum_n Dur_{p,n} P_{h,p,n} \leq W_{h,p} C^{j}_{h,p} + W^{i-1}_{h,p} C_{h,p} - u^{i,j}_{h,p} W^{i-1}_{h,p} C^{j}_{h,p} - \left(1 - u^{i,j}_{h,p}\right) M3^{i,j}_{h,p} \\
&\sum_n Dur_{p,n} P_{h,p,n} \leq W_{h,p} C^{j-1}_{h,p} + W^{i}_{h,p} C_{h,p} - u^{i,j}_{h,p} W^{i}_{h,p} C^{j-1}_{h,p} - \left(1 - u^{i,j}_{h,p}\right) M4^{n,m}_{h,p}
\end{aligned}
\tag{32}
$$

We now force the model to select just one cell over the grid for every hydro plant and every time period.

$$
\sum_{i,j} u^{i,j}_{h,p} = 1 \quad \forall h, p \tag{33}
$$

This set of equations allows the formulation of the following models that describe the HCP:

- A nonlinear nonconvex hydrothermal coordination problem (HCPN), build up with equations (19) to (28).
- A linear hydrothermal coordination problem (HCPL), where the production function (21) is replaced by a constant. Besides, the reserve level equation as a function of the head (24) is outer approximated by a piecewise linear function.
- A linear approximation (HCPM) of the nonlinear problem, where the bilinear function (22) is replaced with its McCormick convex envelope (single-cell approximation), equation (31). Same as above for equation (24).
- A mixed integer problem (HCPT), where we refine the McCormick approximation with a collection of tetrahedrons (multiple-cell approximation), equations (32), and employ disjunctive programming to force the selection variables to lay in one of them, equation (33). Same as above for equation (24).

Next table summarizes the characteristics and presents the objective function for a realistic deterministic case study that we will present in section 5.

Model	Type of Mathematical Programming problem	Hydro output function representation	Objective function
HCPN	NLP	Bilinear function	7824.912
HCPL	LP	Constant	7837.094
HCPM	LP	Single-cell McCormick approximation	7818.214

HCPT	MIP	Multiple-cell McCormick approximation[1]	7824.474

We may observe in the above table the achieved results for the objective functions of the different models. With a constant hydro production value (HCPL) we obtain an objective function greater than the bilinear model one. On the othe hand, the relaxed formulation with just one cell (HCPM model) obtains an objective value smaller than the optimun, an indicator that this approximation can be inaccurate. Finally the approximated problem with multiple cells gives a precise value for the objective function at the same time that returs adecuate profiles for the hydro reserves evolution.

IV Stochastic Dual Dynamic Programming

We now focus our attention in the resolution of the nonlinear hydrothermal stochastic problem. This multistage problem can be formulated in a general form as (Ermoliev and Wets 1988)

$$\begin{aligned}
&\min z = c_1 x_1 + E_{\xi_2}\left[\min c_2 x_2 + E_{\xi_3}\left[\min c_3 x_3 + \cdots\right]\right] && \\
&A x_t \le b_t && t:1,\ldots,T \\
&B\left(x_t\right) = d_t && t:1,\ldots,T \\
&T x_{t-1} + W x_t \le h\left(\xi_t\right) && t:1,\ldots,T-1
\end{aligned} \tag{34}$$

where t are the stages, the constraints $Ax_t \le b_t$ represent intraperiod constraints, the constraints $B\left(x_t\right) = d_t$ are the bilinear constraints and $Tx_{t-1} + Wx_t \le h\left(\xi_t\right)$ the interperiod constraints, also denoted coupling or tender constraints. In HCP those constraints account for the water inventory equations for the reservoirs.

Typically multistage models are represented by considering the recourse function, which carries the future operation cost as a function of the current value of the decision variable. Thus, the multistage problem is formulated as

$$\begin{aligned}
&\min z = c_1 x_1 + E_{\xi_2}\left[Q_1\left(x_1, \xi^2\right)\right] \\
&A x_1 \le b_1 \\
&B\left(x_1\right) = d_1
\end{aligned} \tag{35}$$

with the recourse function $Q_{t-1}\left(x_{t-1}, \xi_t\right)$ defined in a recursive manner as

[1] This value corresponds to a grid of 10 slot for the production function and 1 slot for the water release and it has been solved with a relative optimality tolerance of 0.08 %.

$$
\begin{aligned}
Q_{t-1}\left(x_{t-1},\xi_t\right) = \quad & \min c_t x_t + E_{\xi_t}\left[Q_t\left(x_t,\xi_{t+1}\right)\right] \\
& Ax_t \le b_t \\
& B_t\left(x_t\right) = d_t \\
& Wx_t \le h\left(\xi_t\right) - Tx_{t-1}
\end{aligned}
\tag{36}
$$

The above formulation is the starting point of a large number of algorithms that solve the multistage problem by solving smaller problems associated to individual stages. The key point in the development of those methods is the convexity of the recourse function because it enables to approximate it by means of outer hyperplanes. Typically these methods iterate between stages proposing values for the decision variables x_t and obtaining approximations for the recourse functions Q_t. The convexity property of the recourse function is guaranteed in the case of pure linear problems or in the case of convexity of the second and further stages (Minoux 1986).

Unfortunately, the HCP of this paper does not fall into the category of problems having convex second stages. The bilinear hydro output function breaks this property. For our problem, neither the recourse function is convex. We now show this nonconvexity of the recourse function by computing it with the next academic example. It is a very reduced problem that comprises the structure of the HCP problems. We have a single period problem with a demand that need to be satisfied by using two thermal units and one hydro plant.

IV.1 Example 3

Consider the next problem given as

$$
\begin{aligned}
& \min 6x_1 + 2x_2 \\
& x_1 + x_2 + y = 500 \\
& y = cw \\
& c = -0.5 + 0.02h \\
& r = 100 + 4h \\
& r \le r^0 + 10 - w \\
& 0 \le x_1 \le 250 \quad 0 \le x_2 \le 250 \\
& 0 \le w \le 600 \\
& 50 \le h \le 130 \\
& 300 \le r
\end{aligned}
\tag{37}
$$

where we have replicated the case of a HCP with nonlinearity in the hydro output. We keep the notation of examples 1 and 2. The initial water reserve is indicated by r^0. We modify the initial reserve level from 300 to 900 and solve the above nonlinear problem. The result is a relation that gives the objective function value of the nonlinear problem as a function of the initial value of r^0 considered. This function is depicted in the next figure.

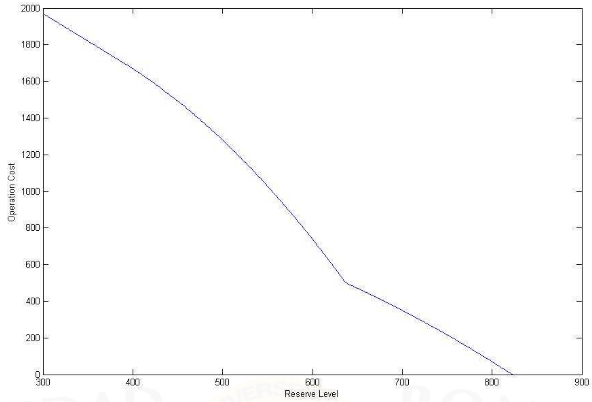


Figure 7. Example of a nonconvex recourse function (in blue) with bilinear hydro output function.

We clearly appreciate the nonconvexity of the recourse function. It can be observed that, the more water, the greater the substitution of thermal production. Although this is also a feature of general hydrothermal models, the above example goes a step further on and shows that, the more water, the higher the net head and consequently the greater the production function. As a consequence, the greater the thermal substitution but with a greater substitution ratio, which causes the loss of the convexity.

The above academic example alerts about the impossibility of applying the multistage algorithms, which approximate the recourse function by tangent lines, to the current case. Our approach to solve this difficulty is to create the convexification of the recourse function. However, with the intention of avoiding the solution of the nonlinear models, we first analyze the recourse function that appears when the bilinear function is replaced with a single-cell McCormick approximation or with a multiple-cell McCormick approximation.

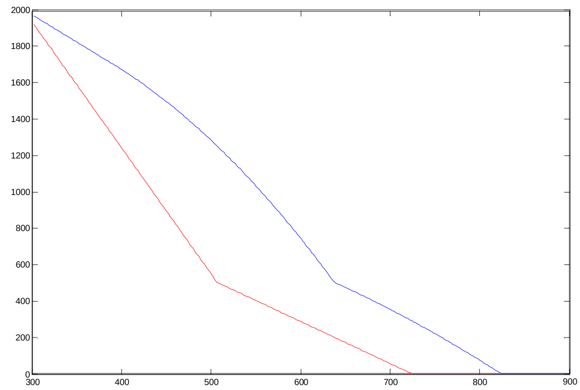


Figure 8. Convex recourse function (in red) when using the single-cell McCormick approximation.

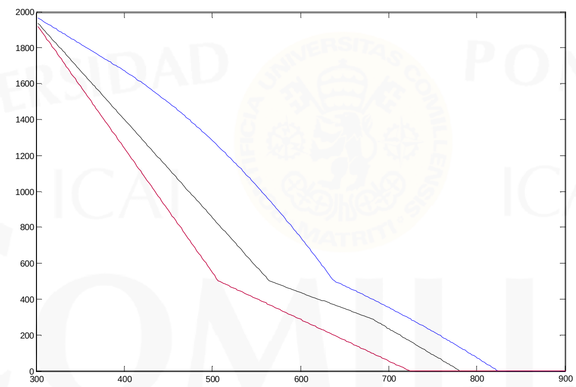


Figure 9. Nonconvex recourse function (in black) when using the multiple-cell McCormick approximation.

So far we have presented three models for the hydrothermal coordination problem and we have analyzed the recourse functions obtained with them. In the construction of a Benders type algorithm we are forced to obtain outer approximations for the convexifications of the recourse function. This convexification is performed by means of the LR technique (Lemarechal and Renaud 2001), (Lemacheral 2007). We now discuss different models under that point of view and argument our selection of the MIP one for carrying out the computations.

The model HCPN is a NLP one. The LR method faces the solution of a dual problem and an NLP inner problem. Any algorithm that could stop in a local but non global minimum would destroy the convergence of the method. An alternative can be the use of spatial branch and bound techniques (Smith and Pantelides 1999), (Sahinidis 1996). Nevertheless, we reject the use of NLP models because our intention is to take advantage in the decomposition method of the capabilities of robust and powerful LP and MIP solvers.

The model HCPM is a LP one but, as shown, the recourse function that it generates may be not a good approximation of the nonconvex recourse function, as seen in Figure 8. Thus, we disregard the use of this model.

Finally, the HCPT model is a MIP one. Although the recourse function that it generates is nonconvex, the dimensions of the grid can be tunned so that an accurate approximation of the recourse function is obtained. The convexification via the application of the LR method iterates between a dual problem and a MIP inner problem. The best bound value that most MIP solvers return may be used to obtain valid Benders cuts that do not eliminate any part of the recourse function. We dive into these ideas in the next section.

IV.2 Construction of the Lower Convex Envelope

The HCPT is a MIP problem with binary variables at any stage. It is also a stochastic problem where stochasticity is considered within this paper via a recombining tree. We focus next in the description of the convexification procedure and later in its extension to the stochastic case. An intermediate stage of problem HCPT for a given scenario ξ_t can be formulated as:

$$\left(HCPT_t^{\xi_t}\right) \qquad \begin{aligned} Q_{t-1}\left(x_{t-1},\xi_t\right) = \quad & \min c_t x_t + E_{\xi_{t+1}}\left[Q_t\left(x_t,\xi_{t+1}\right)\right] \\ & Ax_t \le b_t \\ & Mx_t + Nu_t \le m_t \\ & Wx_t \le h\left(\xi_t\right) - Tx_{t-1} \end{aligned} \tag{38}$$

where the expectation is considered over the transition probabilities from the current node of the scenario tree. The constraints $Mx_t + Nu_t \le m_t$ aggregate those constraints that define the McComick multiple-cell aproximation for the hydro output and $Ax_t \le b_t$ are the thermal operating limits. The remaining equations are those that connect stages and consider the inventory of the water reserves.

In the development of our L-shape type algorithm, the *t*-stage recourse function $Q_{t-1}\left(x_{t-1},\xi_t\right)$ is approximated with outer hyperplanes obtained form the convexification of the recourse function. We keep the convention of denoting those approximations as Benders cuts. Thus, at any stage t a subproblem $(RP_t^{\xi_t})$ can be solved that takes the next form

$$(RP_t^{\xi_t}) \qquad \begin{aligned} Q_{t-1}\left(x_{t-1},\xi_t\right) = \quad & \min c_t x_t + E_{\xi_{t+1}}\left[\theta_t^{\xi_{t+1}}\right] \\ & Ax_t \le b_t \\ & Mx_t + Nu_t \le m_t \\ & Wx_t \le h\left(\xi_t\right) - Tx_{t-1} \\ & \theta_t^{\xi_{t+1}} \ge G_{t,k}^{\xi_{t+1}} x_{t-1} - g_{t,k}^{\xi_{t+1}} \qquad k:1,\dots,K_t^{\xi_{t+1}} \end{aligned} \tag{39}$$

where $K_t^{\xi_t}$ indicates the number of outer approximations to the recourse function that have been built so far. The convexification can be obtained through the maximization of the dual function, defined for each multiplier as the solution of the Lagrangean subproblem:

$$(LP_t^{\xi_t}) \qquad \begin{aligned} \omega_t^{\xi_t}\left(\lambda_t^{\xi_t}\right) = \quad & \min c_t x_t + E_{\xi_{t+1}}\left[\theta_t^{\xi_{t+1}}\right] + \lambda_t^{\xi_t}\left(Tx_{t-1} + Wx_t - h\left(\xi_t\right)\right) \\ & Ax_t \le b_t \\ & Mx_t + Nu_t \le m_t \\ & \theta_t^{\xi_{t+1}} \ge G_{t,k}^{\xi_{t+1}} x_{t-1} - g_{t,k}^{\xi_{t+1}} \qquad k:1,\dots,K_t^{\xi_{t+1}} \end{aligned} \tag{40}$$

It can be proven (Geoffrion 1972) that

$$(DP_t^{\xi_t}) \qquad Q_{t-1}^{*}\left(x_{t-1},\xi_t\right) \triangleq conv\left(Q_{t-1}\left(x_{t-1},\xi_t\right)\right) = \max_{\lambda_t^{\xi_t}} \omega_t^{\xi_t}\left(\lambda_t^{\xi_t}\right) \tag{41}$$

Assume we want to approximate the lower convex envelope of the recourse function at a given decision value $\widehat{x}_{t-1}$ and for a chosen scenario ξ_t. Let $\dot{\lambda}_t^{\xi_t}$ be the optimal multiplier of problem $(DP_t^{\xi_t})$. By definition we would have for every x_{t-1}

$$\begin{aligned} & Q_{t-1}\left(x_{t-1},\xi_t\right) \ge Q_{t-1}^{*}\left(x_{t-1},\xi_t\right) = \omega_t^{\xi_t}\left(\dot{\lambda}_t^{\xi_t}\right) = \\ & = c_t\dot{x}_t + E_{\xi_{t+1}}\left[\dot{\theta}_t^{\xi_t}\right] + \dot{\lambda}_t^{\xi_t}\left(Tx_{t-1} + W\dot{x}_t - h\left(\xi^t\right)\right) \end{aligned} \tag{42}$$

where the dot letters indicate the values where the optimum has been achieved. We particularized the above expression for the $t-1$ stage value where we are interested in building the approximation. After some algebra we have

$$\begin{aligned} & c_t\dot{x}_t + E_{\xi_{t+1}}\left[\dot{\theta}_t^{\xi_t}\right] + \dot{\lambda}_t^{\xi_t}\left(Tx_{t-1} + W\dot{x}_t - h\left(\xi_t\right)\right) = \\ & = c_t\dot{x}_t + E_{\xi_{t+1}}\left[\dot{\theta}_t^{\xi_t}\right] + \dot{\lambda}_t^{\xi_t}\left(T\widehat{x}_{t-1} + Tx_{t-1} - T\widehat{x}_{t-1} + W\dot{x}_t - h\left(\xi_t\right)\right) = \\ & = c_t\dot{x}_t + E_{\xi_{t+1}}\left[\dot{\theta}_t^{\xi_t}\right] + \dot{\lambda}_t^{\xi_t}\left(T\widehat{x}_{t-1} + W\dot{x}_t - h\left(\xi_t\right)\right) + \dot{\lambda}_t^{\xi_t}\left(Tx_{t-1} - T\widehat{x}_{t-1}\right) = \\ & = \omega_t^{\xi_t}\left(\dot{\lambda}_t^{\xi_t}\right) + \dot{\lambda}_t^{\xi_t} T\left(x_{t-1} - \widehat{x}_{t-1}\right) \end{aligned} \tag{43}$$

thus obtaining an approximation in the form of a Benders cut given as

$$\theta_t^{\xi_t} \geq \omega_t^{\xi_t}\left(\dot{\lambda}_t^{\xi_t}\right) + \dot{\lambda}_t^{\xi_t} T\left(x_{t-1} - \widehat{x}_{t-1}\right) \tag{44}$$

and is used to augment problem $(RP_t^{\xi_t})$. Gathering the terms of the above expression we can obtain the coefficients expression in the form of $G_t^{\xi_t}$ and $g_t^{\xi_t}$.

The difficulty in the above development appears in the optimization of the dual function. Proximal bundle methods (Fentelmark and Kiwiel 1999), sequential refinement (Jiménez and Conejo 1999), ascending directions (Beltran-Royo 2009), outer-approximations (Cerisola, Baíllo et al. 2009) are some of the techniques that have been used for optimizing it. In general, this method requires a huge number of iterations to converge. The overall procedure for the multistage problem together with the necessary number of required iterations of a LR procedure at every stage and scenario of the tree can transform the task of finding a solution of the HCP into a cumbersome one. We avoid the optimization of the dual function by employing the next two heuristics.

Heuristic 1.

We solve the linear relaxation of problem $\left(HCPT_t^{\xi_t}\right)$ and obtain the dual variable for the coupling constraint $\dot{\pi}_t^{\xi_t}$. We then consider the multiplier given by $\dot{\lambda}_t^{\xi_t} = -\dot{\pi}_t^{\xi_t}$. With this multiplier we perform a single evaluation of the dual function and obtain the value $\omega_t^{\xi_t}\left(-\dot{\pi}_t^{\xi_t}\right)$, which is used to create a new cut.

Heuristic 2.

The construction of the Benders cut has the goal of approximating the implicit recourse function defined by the t-stage problem. This recourse function is employed in the $t-1$ stage subproblem. For this reason, when optimizing this $t-1$ stage subproblem, we can identify the active cuts. After doing this, we determine the next multiplier as a linear combination of the coefficients of the active cuts.

Consider the recourse function constructed in example 3. Assume we want to optimize a two stage problem with a first stage objective function of $3r^0$ and a second stage objective function given by the recourse function. The overall objective function is then

$$3r^0 + Q(r^0) \tag{45}$$

If we replace the bilinear hydro output function by the McCormick multiple-cell aproximation we face the problem of optimizing the function

$$3r^0 + Q_M(r^0) \tag{46}$$

We depict both functions in the next figures. The figure 10 shows the highly nonconvex function that has to be optimized. In figure 11.a we assume that two cuts have already been generated that approximate the recourse function. We assume those two cuts to be active in the solution of the first stage problem. We take the coefficients of those cuts and propose a new multiplier to evaluate the Lagrangean subproblem. This Lagrangean subproblem returns a value for the dual function that is used to create a new cut. Finally, figure 11.b shows all the cuts and the convergence of first stage decision to the solution that minimizes the objective function.

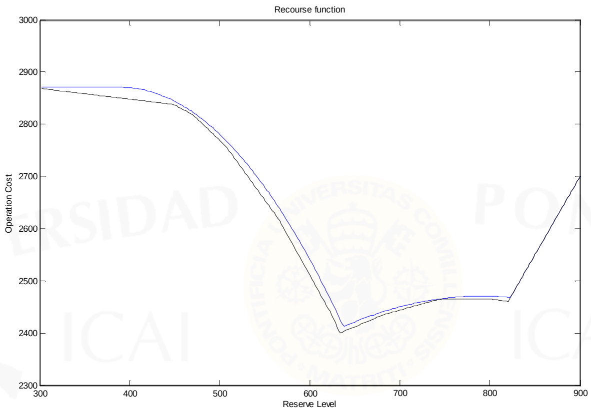


Figure 10. Nonconvex recourse function.

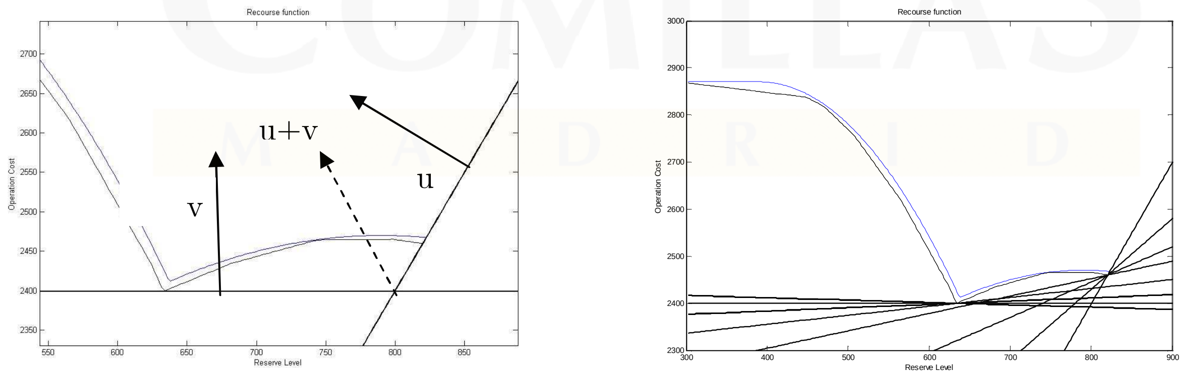


Figure 11. Nonconvex recourse function with two cuts (left) and with all the cuts (right).

IV.3 Stochastic Algorithm

We consider the stochasticity of the problem to be given by means of a recombining tree as the one presented in the figure 12. A recombining tree assumes a discrete representation of the uncertainty at each stage. It also assumes that the distribution depends of the previous stage realization of uncertainty. Thus we consider transition probabilities between the nodes of the recombining tree. In this way, a scenario is constituted by a realization of each stage random variable, being the probability of the scenario the product of the transition probabilities between stages.

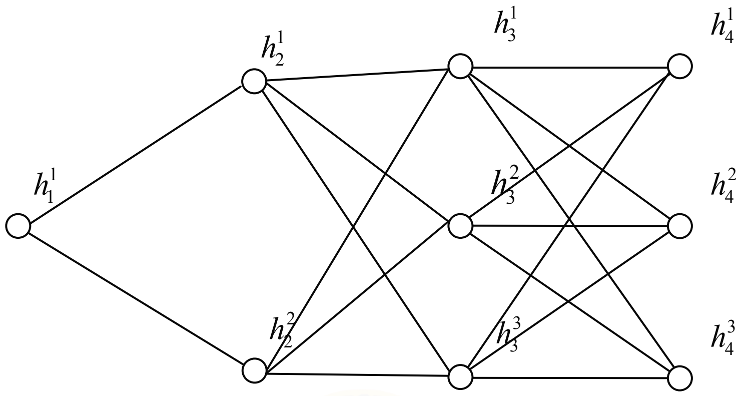


Figure 12. Recombining tree.

With this uncertainty representation, the number of resulting scenarios grows exponentially. It is usually impossible to solve the DEP and the solution of the multistage problem has necessarily to be solved by the application of sampling methods.

The solution of the stochastic problem is based on the approximation of the recourse functions. Those are functions of the reserve levels and the stochastic parameters. In our approach, we consider a different recourse function for every node of the scenario tree and approximate each one independently. With this goal, primal decisions need to be proposed for evaluating the recourse function. The SDDP algorithm samples a scenario by conditional sampling of the discrete distribution function of each stage. After that, a forward pass is performed, each stage problem is solved and primal decisions values proposed. Notice that in the forward pass it is only solved one node of the scenario tree at each stage. Secondly, a backward pass is performed solving every node of the scenario tree and creating a new cut to approximate the recourse function. The construction of the new cut involves the computation of the LR algorithm, or the application of the heuristics previously described.

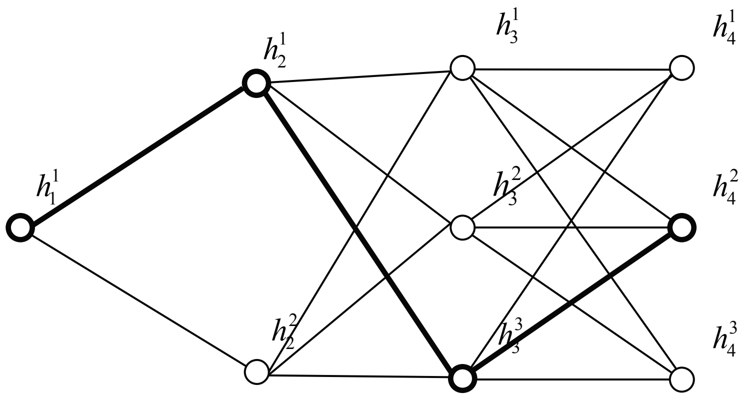


Figure 13a. Scenario solved in a forward pass of SDDP.

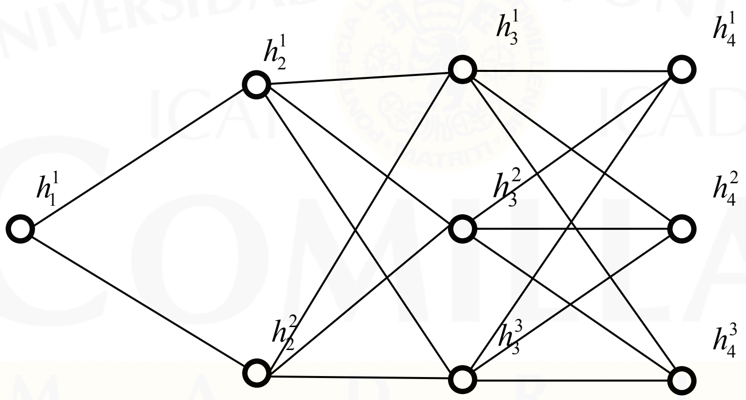


Figure 13b. Nodes solved in the backward pass of SDDP.

The forward and backward passes of the SDDP algorithm may be summarized as follows. We assume a relative complete recourse for every stage of the problem. Thus no infeasible subproblem will appear in the application of the algorithm.

Step 0	Set $K_t^{\xi_t} = 0$. Set $\theta_t^{\xi_t} \equiv 0$ at the initial iteration.
At iteration n	
Step 1	Sample one scenario $(h_t^{\xi_t})^n$, $t = 1,\ldots,T$
	Forward pass:
	Repeat for $t = 1,\ldots,T$
	Solve the node ξ_t of stage t subproblem $(RP_t^{\xi_t})$
	Obtain solution $x_t^{\xi_t}$
	If $t = 1$ obtain lower bound $\underline{z} = v(RP_1^{\xi_1})$
Step 2	**Backward pass**
	Repeat for $t = T,\ldots,1$
	Repeat for each node ξ_t of stage t
	Solve $(RP_t^{\xi_t})$
	Obtain objective $\theta_{t,k}^{\xi_t} = v(RP_t^{\xi_t})$ and dual values $\pi_{t,k}^{\xi_t}$
	Augment $K_t^{\xi_t} = K_t^{\xi_t} + 1$
	Go to step 1

Description of the traversing strategy in SDDP.

When applying the L-shaped method of (Van Slyke and Wets 1969), a termination criterion is considered that stops the algorithm when the relative tolerance between a lower and an upper bound is reached. In the setting of sampling within a multistage stochastic program, a lower bound is also available after the resolution of the first stage problem. However, the construction of the upper bound needs to be revisited. In two-stage problems (Shapiro 2003), the upper bound is computed once a first stage solution is available. With that purpose, a collection of independent scenarios is sampled and the objective function evaluated for the sample average function. When this evaluation is repeated for T independent batches of samples of size N, a confidence interval for the upper bound of the problem is available. The sampling method for two stage stochastic problems is stopped when the lower bounds falls within this confidence interval. The case for multistage SP is more complicated that the two-stage problem (Linderoth, Shapiro et al. 2006). We propose the next stopping rule for the multistage case.

Our experimental results show that in the evolution of the traversing strategy, there is an iteration from which the values of the first stage decision variables are virtually repeated. From that iteration, we store the objective function values and compute the sample mean of those values together with the confidence interval. The algorithm stops when the lower bound falls within that interval.

Let n_0 be the iteration where the first stage solution is repeated. After N additional iterations we have stochastic values for the objective function given by $\sum_{t=1}^{T} c_t \left(\dot{x}_t^{\xi_t}\right)^n$, $n = 1,...,N$. We estimate the sample mean as well as the standard deviation

$$\begin{aligned} \overline{z} &= \frac{1}{N}\sum_{t=1}^{T} c_t \left(\dot{x}_t^{\xi_t}\right)^n \\ \sigma_n &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\sum_{t=1}^{T} c_t \left(\dot{x}_t^{\xi_t}\right)^n - \overline{z}\right)^2} \end{aligned} \tag{47}$$

and create a 95 % confidence interval for the upper bound

$$\left(\overline{z} - 1.96\frac{\sigma_n}{\sqrt{N}}, \overline{z} + 1.96\frac{\sigma_n}{\sqrt{N}}\right) \tag{48}$$

Now, we describe the SDDP algorithm including the extension to MIP subproblems and the stopping rule based on the confidence interval.

Step 0 Set $K_t^{\xi_t}=0$. Set $\theta_t^{\xi_t}\equiv 0$ at the initial iteration

At iteration n

Step 1 (**Forward pass**)

Sample one scenario $(h_t^{\xi_t})^n$, $t=1,\dots,T$

Repeat for $t=1,\dots,T$

Solve the node ξ_t of stage t subproblem $(RP_t^{\xi_t})$

Obtain solution $x_t^{\xi_t}$

If $t=1$ obtain lower bound $\underline{z}=v(RP_1^{\xi_1})$

Step 2 (**Upper bound computation**)

If the variation for first stage variable is less than a given tolerance:

Get the mean and the standard deviation for the last objective function values of the complete objective function evaluated for the primal solutions so far obtained.

$$\overline{z}=\frac{1}{N}\sum_{t=1}^{T}c_t\left(\dot{x}_t^{\xi_t}\right)^n$$

$$\sigma_n=\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\sum_{t=1}^{T}c_t\left(\dot{x}_t^{\xi_t}\right)^n-\overline{z}\right)^2}$$

Update the upper bound confidence interval

$$I=\left(\overline{z}-1.96\frac{\sigma_n}{\sqrt{N}},\overline{z}+1.96\frac{\sigma_n}{\sqrt{N}}\right)$$

Step 3 (**Stopping rule**)

If $1.96\frac{\sigma}{\sqrt{N}}$ is less than a given tolerance and $\underline{z}\in I$ stop

$x_1^{\xi_1}$ is optimal solution,

else go to Step 4

Step 4 **Backward pass**

Repeat for $t=T,\dots,1$

Repeat for each node ξ_t of stage t

If heuristic rule 1 is used

Solve linear relaxation of $(RP_t^{\xi_t})$

Get dual variable and create the multiplier $\lambda_{t,k}^{\xi_t}=-\pi_{t,k}^{\xi_t}$

Solve Lagrangean subproblem $(LP_t^{\xi_t})$

Obtain objective $\theta_{t,k}^{\xi_t}=\omega(DP_t^{\xi_t})$

If heuristic rule 2 is used

Construct the multiplier $\lambda_{t,k}^{\xi_t}$ as a linear combination of the coefficients of the active cuts for $t-1$ stage subproblem

Solve Lagrangean subproblem $(LP_t^{\xi_t})$

Obtain objective $\theta_{t,k}^{\xi_t} = \omega(DP_t^{\xi_t})$

If no heuristic rule is used

Optimize the dual function

Obtain objective $\theta_{t,k}^{\xi_t} = v(DP_t^{\xi_t})$ and dual values $\pi_{t,k}^{\xi_t} = -\lambda_{t,k}^{\xi_t}$

Create new cut and augment $K_t^{\xi_t} = K_t^{\xi_t} + 1$

Go to step 1

Description of the traversing strategy in SDDP for MIP subproblems.

We next apply the SDDP to a realistic large-scale HCP extracted form the Spanish system.

V Case Study

We consider a real size problem for a generation company over a time horizon of one year. We consider the hydraulic year which starts in October and finishes in September and consider a deterministic demand profile. We divide the planning horizon into 52 periods for weeks and each period into two load level subperiods (peak, off-peak). We consider uncertainty in natural water inflows. We classified historical data for the last 50 years and use a clustering technique based in the neural gas algorithm to create the recombining tree that represents uncertainty. The generating portfolio is made of 100 thermal plants (nuclear, coal, fuel, gas) with an installed capacity of 43375 MW and 51 hydro plants with an installed capacity of 11835 MW.

We consider hydro plants to be located in three different hydro basins and implicitly include in the mathematical modeling the physical constraints that describe the cascaded connections among hydro reservoirs and hydro plants.

We have represented our HCP model with GAMS 23.1 (Brooke, Kendrick et al. 1996) and solve the subproblems with CPLEX 11.2 in a Pentium V 1.8 GHz.

For a practical implementation of the algorithm, we perform the decomposition algorithm in three phases.

In phase 1 we solve in the forward pass and in the backward pass the LP relaxation of the t-stage subproblems. In doing so we underestimate the recourse functions but obtain valid approximations that guide the primal decisions towards the optimal ones. This phase has the obvious advantage of being computationally cheaper than the forthcoming ones.

In phase 2 we used the heuristic 2 described in the previous section. In the forward pass we solve the LP relaxation of the subproblem of each stage and store the dual variables of the active Benders cuts in that iteration. In the backward pass we solve the Lagrangean subproblem for every node of the tree, proposing as multiplier the linear combination on the dual variables previously stored.

In phase 3 we solve in the forward pass the MIP subproblem of each stage. This provides us with upper bound estimations that can be used to create an upper bound confidence interval. In the backward pass we use heuristic 1 and solve the LP relaxation of the subproblems for each scenario, propose the multiplier as the opposite of the dual variable of the coupling constraints and solve the Lagrangean subproblem. We obtain different approximation cuts because they are created for primal decisions different than those of phase 1 and phase 2. Notice that in this phase 3 the primal decisions satisfy the integer constraints and thus the hydro output variables lie within the approximating cells of the bilinear function.

We have solved three instances of the stochastic problem: i) a first one with a binary branching structure every four weeks of the planning period, accounting for a total of 2^{12} scenarios, ii) a second instance with a binary branching structure every two weeks, accounting for a total of 2^{25} scenarios, iii) and a final instance where the binary ramification takes place every week. The total number of scenarios is 2^{51}. The next table summarizes the computational results after 500 iterations of the decomposition algorithm.

Branching	Scenarios	Lower Bound	Upper Bound	Confidence Interval	Relative Tolerance
Every 4 weeks	2^{12}	7821.0	7817.7	[7803.5,7831.9]	0.0036
Every 2 weeks	2^{25}	7828.0	7839.3	[7826.4,7852.2]	0.0033
Every 1 weeks	2^{51}	7839.0	7850.3	[7831.6,7868.9]	0.0047

The evolution of the lower bound as well as the upper bound is depicted in the next figure. Notice that the confidence interval is computed starting in the 400^{th} iteration, when the forward pass solves the MIP subproblem for each stage and we have estimations for the upper bound of the objective function.

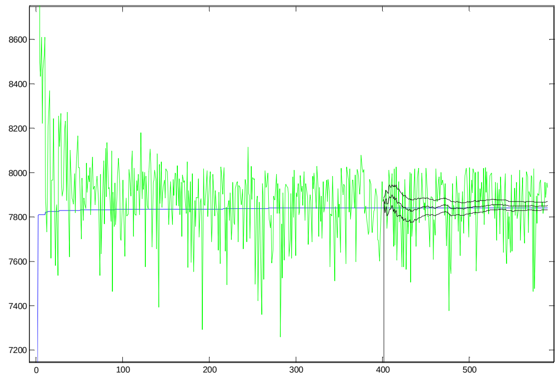
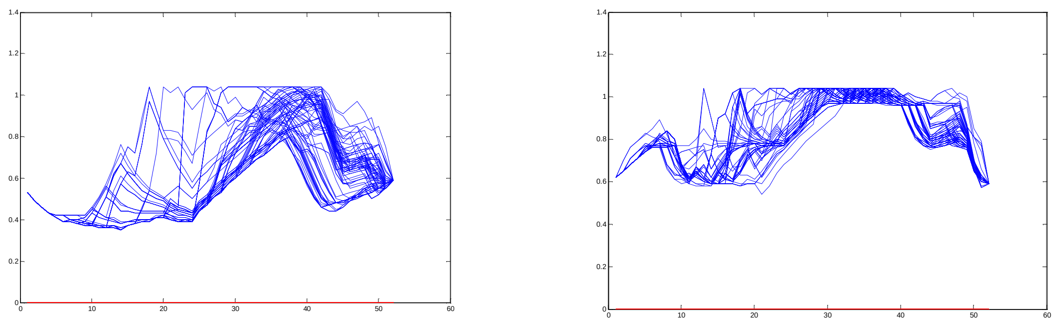


Figure 14. Converge of SDDP with MIP subproblems.

We show next the evolution of some reservoirs of the hydro system. We want to outline that this was the initial objective of the current research. To include the nonlinear head effect of main reservoirs to appreciate the differences in the evolution of the reserve level profiles compared with the case where a constant production function is considered. We show now this effect for five of the main reservoirs of the system. We have tested the evolution of the water profiles for 100 simulated scenarios after the solution of the multistage stochastic problem. The solution is carried out by performing similar forward passes to those of the decomposition method. In the left column we plot the results when a constant production function was used. On the right column we plot the equivalent evolution profiles (same scenario tree and same simulated scenarios) when the net head nonlinear effect is considered.



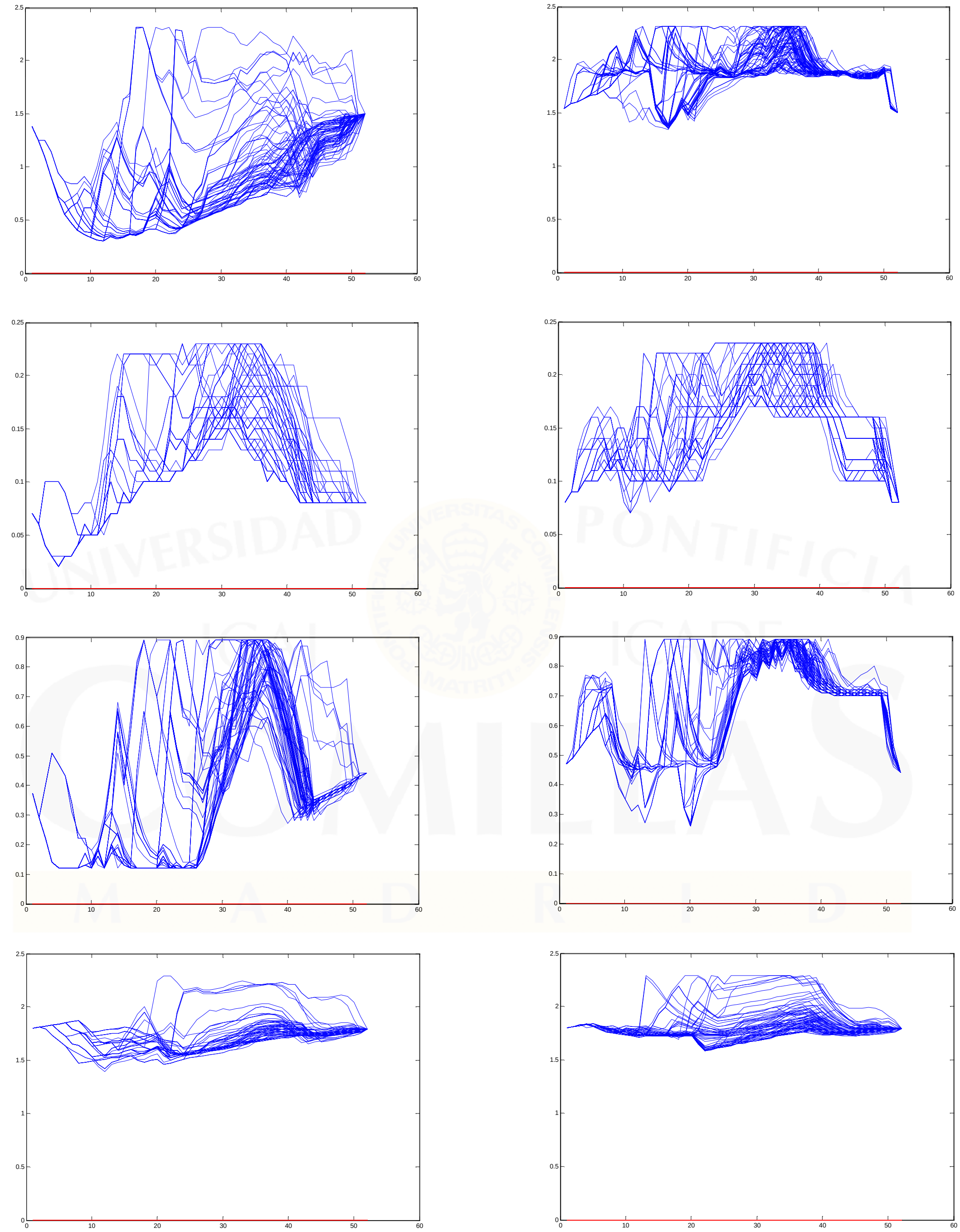


Figure 14. Evolution of the reserve levels for five reservoirs Left: constant production function. Right: McCormick multiple-cell approximation.

We appreciate that the reserve evolution when the nonlinear head effect is considered differs from the case when a constant production function is used. The model gives higher reserves with the goal of increasing the production function and replacing more thermal generation.

VI Conclusion

We have presented a hydrothermal model that considers the net head nonlinear effect in hydro output. We consider within this model the stochasticity given by means of a recombining scenario tree. The main contribution of the paper is the extension of the SDDP to the case where bilinear constraints define the feasible region. Our main approach consists of replacing the bilinear constraints with the convex envelope given by the McCormick approximation of the bilinear function, transforming the nonlinear problem into a LP one. When more precision is required, we replace the bilinear function with an upper and a lower McCormick surfaces. Those surfaces appear when applying the McCormick envelope to every cell of a grid created for the two variables of the domain of the bilinear function. Then the problem becomes MIP.

We have presented a successful application of the technique here developed to a real size case extracted from the Spanish system. The hydrothermal model has a one year time scope divided into 52 periods for every week of the year. We have solved instances with 2^{12}, 2^{25} and 2^{51} scenarios and shown the results. The reserve level of the reservoirs increases with the solution of the proposed model. The production function increases and therefore the operating cost is minimized as more thermal production is replaced.

We suggest some lines of research for the model presented. The first two lines are oriented to accelerate the convergence of the method. The third line is oriented to incorporate the spillage control in the stochastic model.

The immediate idea to accelerate the convergence is to use an importance sampling method within the decomposition method. A second possibility of accelerating the convergence is to solve the nodes of the recombining tree in an aggregate manner. Thus, the computation time could be reduced by solving in the forward and in the backward pass small recombining subtrees bunching the subproblems for different stages.

Finally, reservoir management can be improved by considering some spillage risk measure in the stochastic HCP to avoid keeping reservoirs at their very upper bounds and therefore reducing the risk of spillage.

VII References

Beltran-Royo, C. (2009). "The radar method: An effective line search for piecewise linear concave functions." Annals of Operations Research **166**(1): 299-312.

Benders, J. F. (1962). "Partitioning Procedures for Solving Mixed Variables Programming Problems." Numerische Mathematik **4**: 238-252.

Brooke, A., D. Kendrick, et al. (1996). GAMS: A User's Guide. Whashington, DC, GAMS Development Corporation.

Cerisola, S., Á. Baíllo, et al. (2009). "Stochastic power generation unit commitment in electricity markets: a novel formulation and comparison of solution methods." Operations Research **57**(1): 32-46.

El-Hawary, M. E. and G. S. Christensen (1979). Optimal Economic Operation of Electric Power Systems. New York, ACADEMIC PRESS, INC.

Ermoliev, Y. and R. Wets (1988). Numerical Techniques for Stochastic Optimization. Berlin.

Fentelmark, S. and K. C. Kiwiel (1999). "Dual Applications of Proximal Bundle Methods, Including Lagrangian Relaxation of Nonconvex Problems." SIAM Journal on Optimization **10**(3): 697 - 721.

Finardi, E. C., E. L. D. Silva, et al. (2005). "Solving the unit cmmitment problem of hydropower plants via Lagrangean Relaxations and Sequential Quadratic Programming." Computational and Applied Mathematics **24**(3): 317-341.

Frangioni, A. (2005). "About Lagrangean Methods in Integer Optimization." Annals of Operations Research **139**: 163-193.

Geoffrion, A. M. (1972). "Generalized Benders Decomposition." Journal of Optimization thoery Applications (JOTA) **10**: 237-259.

Geoffrion, A. M. (1974). "Lagrangean relaxation for integer programming." Mathematical programming study **2**(2): 82–114.

ILOG, I. (2009). http://www.ilog.com.

Infanger, G. (1994). Planning Under Uncertainty: Solving Large Scale Stochastic Linear Programs, Boyd and Fraser Publising Co. Danvers, MA.

Iyer, R. R. and I. E. Grossman (1997). "Optimal multiperiod operational planning for utility systems." Computers & Chemical Engineering **21**(8): 787-800.

Jiménez, N. and A. J. Conejo (1999). "Short-term hydrothermal coordination by Lagrangian relaxation: solution of the dual problem." IEEE Transactions on Power Systems **14**(1): 89-95.

Latorre, J. M. (2003). ARBOLES, A Tool for Managing Scenario Trees. Madrid, Instituto de Investigación Tecnológica, Universidad Pontificia Comillas.

Latorre, J. M., S. Cerisola, et al. (2007). "Clustering algorithms for scenario tree generation: Application to natural hydro inflows." European Journal of Operational Research **181**(3): 1339-1353.

Lemacheral, C. (2007). "The omnipresence of Lagrange." Annals of Operations Research **153**: 9-27.

Lemarechal, C. and A. Renaud (2001). "A geometric study of duality gaps, with applications." Mathematical Programming **90**: 399-427.

Liberti, L. and C. C. Pantelides (2006). "An Exact Reformulation Algorithm for Large Nonconvex NLPs Involving Bilinear Terms." Journal of Gobal Optimization **36**(2): 161-189.

Linderoth, J., A. Shapiro, et al. (2006). "The empirical behaviour of sampling methods for stochastic programming." Annals of Operations Research **142**: 215-241.

McCormick, G. (1976). "Computability of Global Solutions to Factorable NonConvex Programs: Part I - Convex Underestimating Problems." Mathematical Programming **10**: 146-175.

Minoux, M. (1986). Mathematical Programming. Theory and Algorithms. New York, John Wiley and Sons.

Morton, D. (1993). Algorithmic Advances in Stochastic Programming, Systems Optimization Laboratory, Department of Operations Research, Stanford University.

Nowak, M. P. and W. Romisch (2000). "Stochastic Lagrangian Relaxation applied to Power Scheduling in a Hydro-Thermal System under Uncertainty." Annals of Operations Research.

Pereira, M. V. F. and L. M. V. G. Pinto (1991). "Multi-Stage Stochastic Optimization Applied to Energy Planning." Mathematical Programming **52**: 359-375.

Quesada, I. and I. E. Grossmann (1995). "Global Optimization Of Bilinear Process Networks With Multicomponent Flows." Computers & Chemical Engineering **19**(12): 1219-1242.

Sahinidis, N. V. (1996). "BARON: A general purpose global optimization software package." Journal Of Global Optimization **8**(2): 201-205.

Shapiro, A. (2003). "Monte Carlo Sampling Approach to Stochastic Programming." ESAIM: Proceedings **13**: 65-73.

Smith, E. M. B. and C. C. Pantelides (1999). "A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs." Computers & Chemical Engineering **23**(4-5): 457-478.

Van Slyke, R. M. and R. Wets (1969). "L-Shaped Linear Progams with Applications to Optimal Control and Stochastic Programming." SIAM Journal on Applied Mathematics **17**(4): 638-663.

Williams, H. P. (1993). Model Building in Mathematical Programming, John Willey & Sons.

Williams, H. P. (1999). Model building in mathematical programming. Chichester, Wiley.