Comparing stochastic optimization methods to solve the medium-term operation planning problem

RAPHAEL E.C. GONÇALVES*, ERLON C. FINARDI,
EDSON L. DA SILVA and MARCELO L.L. DOS SANTOS

Electical Systems Planning Research Laboratory – LabPlan – EEL – CTC – UFSC
Campus Universitário, Trindade, Caixa Postal 88040-900, Florianópolis, SC, Brazil
E-mail: raphael@labplan.ufsc.br / erlon@labplan.ufsc.br /
edsonls@labplan.ufsc.br / mllsantos@labplan.ufsc.br

Abstract. The Medium-Term Operation Planning (MTOP) of hydrothermal systems aims to define the generation for each power plant, minimizing the expected operating cost over the planning horizon. Mathematically, this task can be characterized as a linear, stochastic, large-scale problem which requires the application of suitable optimization tools. To solve this problem, this paper proposes to use the Nested Decomposition, frequently used to solve similar problems (as in Brazilian case), and Progressive Hedging, an alternative method, which has interesting features that make it promising to address this problem. To make a comparative analysis between these two methods with respect to the quality of the solution and the computational burden, a benchmark is established, which is obtained by solving a single Linear Programming problem (the Deterministic Equivalent Problem). An application considering a hydrothermal system is carried out.

1   Introduction

The Medium-Term Operation Planning (MTOP) problem of hydrothermal systems consists in defining a generation strategy to minimize the production cost over the planning horizon, usually ranging from two months to one year ahead, taking into account constraints associated with the system and the generation plants.

Depending on the regulatory framework, this problem can be solved either by the Independent System Operator (ISO) or by the generation companies that own a mix of hydro and thermal plants, which need to submit bids (price and quantity) to the ISO. Particularly, in Brazil, a similar model is used by the ISO in order to define the dispatch and the spot price[1]. This model [1] has being improved continuously, aiming to produce a satisfactory response, which opens room for contributions, such as the proposal of this paper.

This problem is particularly complex owing to some characteristics specially related to randomness of water inflows to the reservoirs [2]. Thus, solutions obtained by models that do not recognize this uncertainty produce unsatisfactory results. Additionally, the availability of hydro energy in the future depends on future inflows, which are uncertain, and the reservoirs operation. For instance, assuming that we discharge today as much as possible followed by drought period in the future, it may be necessary to use expensive thermal generation in the future. On the other hand, if the reservoir level is preserved by means of a more intensive use of thermal generation in the present, and wet period occur in the future, spillage may occur (implying a waste of energy). For this reason, the hydro system operation is a problem coupled in time, that is, there is a connection between an operating decision in a given stage and the future consequences of this decision [3].

Besides, the MTOP usually has the following operating characteristics: coordination with the long-term problem by means of a future cost function; the water released from one hydro plant affects the production of other plants downstream; the hydro plants are located in different basins, each one with a particular hydrology pattern [4].

---

[1]In fact, the spot price is set by the Short Run Marginal Cost obtained as a byproduct of an optimization model.

In general, some simplifications are introduced into the model with purpose to eliminate some nonlinearities related to the operation cost of thermal power plants and the production function of hydro plants and, so, allowing it to be solved within a reasonable cpu time[2]. For example, currently in Brazil, the non-linear hydro plant production function is approximated by a piecewise linear convex model, which depends on the reservoir volume, discharged outflow and, in some cases, spillage from the reservoir [5].

Like other multistage stochastic optimization problems, the MTOP problem can be classified as a difficult one, given that its size grows exponentially according to the number of stages and scenarios considered in the problem. Thus, MTOP requires a high computational effort to be solved.

So far, Nested Decomposition (ND) [6] has been used exhaustively to solve similar problems [7, 8] by using algorithms based on a stochastic extension of Benders decomposition [9]. By this approach, the original problem is decomposed into several subproblems which are easier to solve. The communication between these subproblems is made by the optimality constraints. Then, these constraints are iteratively added to the problem in order to enhance the modeling and, in turn, the solution quality.

Recent advances in the theory of stochastic programming make it possible to develop new methods for solving multistage stochastic programs of remarkable sizes. So, for the multistage stochastic linear programs, the decomposition framework based on Augmented Lagrangian (AL) [10] has properties that make it promising for large-scale problems. In this method, some constraints are moved to the objective function generating a simpler but equivalent problem. The resulting subproblems of the decomposition scheme are quadratic, leading to a differentiable problem. When comparing to other methods with the same features, such as the Lagrangian Relaxation (LR) [11], this method has the advantage of obtaining a feasible primal solution.

Progressive Hedging (PH) is an AL based decomposition method that has been broadly applied to similar problems of others fields of knowledge [12, 13]. Due to its solution features it can be an alternative method to solve the MTOP problem.

---

[2]In this case, the stochastic characteristic is prioritized in relation to other characteristics of the problem.

In this context, the main purpose of this paper is to present a comparative study about the performance of different multistage stochastic optimization methods applied to the MTOP. Therefore, the quality of the solution and the CPU time will be shown. The first method considers the representation of the model through a single Linear Programming (LP) problem, the Deterministic Equivalent problem (DE)[3] [14]. The main idea of this approach is to establish a benchmark for a comparison with the other methods, which make use of decomposition strategies: the ND and the PH. Both decomposition algorithms break the MTOP into smaller subproblems and therefore greatly reduce memory requirements. To obtain reliable results, a realistic hydrothermal configuration extracted from the Brazilian hydrothermal power system was used. Additionally, sensitivity analyses[4] were carried out considering different piecewise linear models that describe the hydro plant production function.

The remainder of the paper is organized as follows. In Section 2, a brief description of the multistage stochastic optimization problem features is shown. The list of symbols used in this paper is presented in Section 3. The three problems, related to each solution strategy, are detailed in Section 3. The test problem and the results on solving this problem are shown in Section 4. Finally, conclusions are presented in Section 5.

2 Multistage Stochastic Optimization Problems

The MTOP is a multistage stochastic problem [16] and, therefore, a difficult problem to be solved. In general, computational methods for multistage stochastic programming problems can be divided into two main groups [17]. The first group takes advantage of the special features of stochastic problems to improve data structures and solution strategies [18]. The second group uses

---

[3]There are two ways of writing DEs: the implicit and the explicit [15]. The difference is the way in which the constraints (nonanticipativity) are handled. The explicit DE has more variables and more constraints than the implicit DE. There are no practical reasons to solve a stochastic optimization problem by its explicit DE. However, this modeling can be useful once there are several strategies that explore that particular structure. In this paper, the explicit DE will be used by the Progressive Hedging method, while the implicit DE will be used to solve the problems in a single linear programming.

[4]These studies are important to analyze the trade-off between the solution quality and the computational burden.

special decomposition methods which exploit the problem structure to split it into manageable pieces and coordinate their solution [19].

As presented in the previous section, in this work three methods are used to solve the MTOP Brazilian problem: (i) the implicit DE: a large-scale LP is given to a commercial package to solve which belongs to the first group of the computational methods [20]; (ii) the ND: a Benders decomposition method; (iii) the PH: an AL-based method.

In a multistage stochastic optimization problem, decisions are to be made in stages, and the uncertainties can be modeled by means of a scenario tree that can be generated by sampling techniques [21]. Figure 1 gives an example of a scenario tree for a three stage problem, in which the water inflow to the hydro plants is uncertain. In this structure, each node (filled circle) represents a specific random realization for the water inflow (system state). A branch represents the relationship between two water inflow realizations (state transition). Thus, a scenario consists of a complete path from the node at stage one to a node at stage three. As a consequence, this tree has nine possible scenarios and 13 nodes.
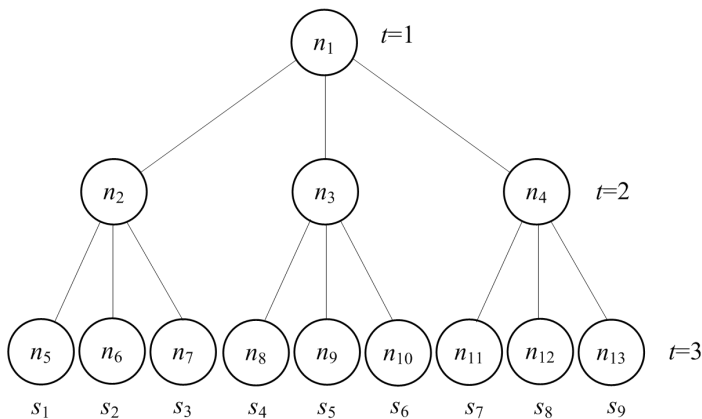

Figure 1 – Scenario tree.

Now, consider that a decision must be taken for some node at stage t in Figure 1. This decision should take into account two available information on the scenario tree: the full path up to stage t (from the past) and the possible realizations for the following nodes (to the future). Since it is not possible to anticipate which scenario will happen, this decision must be unique for all scenarios passing by this node. This condition is called nonanticipativity [22]. For instance,

the decisions of nodes n_1 and n_2 must be identical for the scenarios s_1, s_2 and s_3.

The Figure 2 shows a different representation for the same scenario tree as in Figure 1. At this representation, the scenarios are represented by the full lines while the dotted lines link the decisions for different scenarios, representing the nonanticipativity concept.
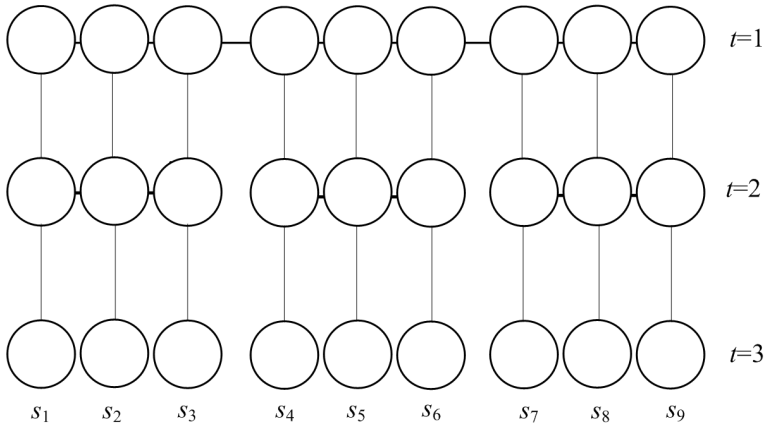

Figure 2 – Sequences of decisions and nonanticipativity (horizontal dotted lines).

The problem can be modeled in different ways according to the chosen solution method. Therefore, the ND decomposes the problem into stages of decision, in which subproblems correspond to nodes and are linked through the temporary joining constraints, so that the mathematical model corresponds to that shown in Figure 1. Given that we intend to use the PH method, introducing scenario decomposition, the mathematical model must use the nonanticipativity condition to link the subproblems, like in Figure 2.

The mathematical modeling of the problem for each solution method will be detailed on the next sections.

3   Problem formulation

In this section, the problem formulation model and some theoretical properties of the methods studied in this paper are presented. Aiming to make easier reading of the remaining of this paper, the notations are shown bellow. The notations describe fixed parameters of the thermal and hydro plants, indices and variables.

T total stages;

t index of stage, so that $t = 1, \ldots, T$;

E total number of subsystems;

e index of subsystems, so that $e = 1, \ldots, E$;

Ω_t set of nodes (decisions) on stage t;

ω a specific node in the stage t, so that $\omega \in \Omega_t$;

S total number of scenarios;

s index of scenario, so that, $s = 1, \ldots, S$;

I total number of thermal plants;

i index of thermal plants, so that, $i = 1, \ldots, I$;

R total number of hydro plants;

r index of the hydro plants, so that, $r = 1, \ldots, R$;

Γ_e set of subsystems linked to the subsystem e;

a_ω^1 immediate ancestor node from node ω;

k_ω successor node from the node ω;

M_r set of upstream reservoirs from reservoir r;

m index of upstream reservoirs;

N number of linear constraints used in piecewise linear function of the hydro plant;

n index associated with the piecewise linear function of the hydro plant, so that, $n = 1, \ldots, N$;

J number of linear constraints used in the piecewise future cost function;

j index associated with the piecewise future cost function;

K_ω set of successor nodes from the node ω;

$X_s t$ set of all scenarios related to scenario s at stage t by the nonanticipativity, including itself;

$iter$ iteration index;

pt_i power output in the i-th thermal plant [MWh];

d_e energy deficit in the i-th subsystem [MWh];

α expected value of the operation cost (scalar) from the stage $T+1$ on;

ph_r power output in r-th hydro plant [MWh];

Int_{le} power interchange from subsystem l to subsystem e [MWh];

v_r final volume of the r-th reservoir [hm^3];

Q_r discharged outflow in the r-th reservoir [m^3/s];

sp_r spillage in the reservoir of the hydro plant r [m^3/s];

$\bar{v}$ average volume of reservoir [hm^3];

p_ω probability associated to the each node ω, such that, $\sum\limits_{\omega \in \Omega_t} p_\omega = 1$;

ct_i incremental operation cost of the i-th thermal plant [$\$/MWh$];

cd_e incremental penalty associated to the energy deficit in the e-th subsystem [$/MWh];

L_e energy demand at subsystem e [MWh];

y_r incremental inflow into the r-th reservoir [m^3/s];

γ_r^j linear function segment at the r-th reservoir of the system;

π Lagrange multiplier [$/MWh] vector associated with the stream-flow balance constraint;

λ Lagrange multiplier vector associated with the nonanticipativity constraint;

ρ positive penalty parameter used in the Augmented Lagrangian method;

σ zero-spillage adjusting factor of the piecewise production function;

τ minimal total square difference factor of the piecewise production function considering the spillage.

3.1   Deterministic equivalent

In this strategy, the stochastic linear problem is solved by means of a standard Linear Program solver. Given that the size of the problem increases sharply as

more stages and scenarios are considered, it is important to exploit the particular matrix structured and its sparsity as much as possible. In this paper, commercial software, CPLEX, was used [23], which uses an advanced optimization algorithm to improve the performance. However, for problems in the real life with the similar structure of the MTOP problem, the usage of CPLEX may not be feasible, given that memory requirements. Indeed, according to the literature [17], multistage stochastic optimization problems, in general, need to be solved by using decomposition techniques, even considering the recent advances in the computing technology.

Therefore, given that we are interested in developing efficient decomposition algorithms to solve MTOP problem in an acceptable time, we make use of the implicit DE approach applied to a small problem[5] in order to obtain a benchmark solution to other methods.

The problem formulation is presented as follows.

- Objective function

   The objective function aims to minimize the system's operation cost. This cost is composed by thermal fuel costs over the planning horizon plus α future cost, which depends on the reservoir's volume at the end of this horizon (at the end of stage T). Then, it can be written as:

$$\text{Min } F = \sum_{t=1}^{T} \sum_{\omega \in \Omega_t} p_\omega \sum_{e \in E} \left(\sum_{i \in I_e} ct_i \, pt_{i\omega} + cd_e d_{e\omega} \right) + \alpha. \quad (1)$$

- The demand supply constraints

$$\sum_{i \in I_e} pt_{i\omega} + \sum_{r \in R_e} ph_{r\omega} + \sum_{l \in \Gamma_e} Int_{l e\omega} + d_{e\omega} = L_{e\omega}, \ e = 1, E. \quad (2)$$

Notice that the hydrothermal system is composed of subsystems that are interconnected. In this context, power plants are located in different subsystems defined by the indexes I_e and R_e, respectively.

Additionally, the demand was considered constant through all stages.

---

[5]A problem within a suitable size is build to demonstrate the method's limitation, considering our computational resources.

- Stream-flow balance constraints

$$v_{r\omega} - v_{ra_\omega} + Q_{r\omega} + sp_{r\omega} - \sum_{m \in M_r} (Q_{m\omega} + sp_{m\omega}) = y_{r\omega}, \ r = 1, R. \quad (3)$$

Here, it is important to discuss the successors and the ancestor nodes in the scenario tree. So, observe the Figure 3.
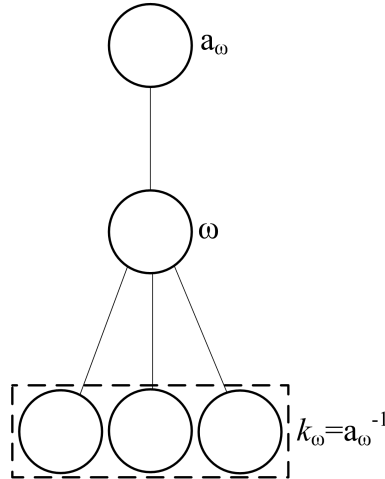

Figure 3 – Structure tree (successor and ancestor node).

Observe that the node a_ω denotes the immediately ancestor node from node ω. This idea can be expanded for a bigger scenario tree (Fig. 1). Otherwise, a_ω^1 or k_ω denotes the successor node from the node ω.

- Piecewise linear function of the hydro plant

$$ph_{r\omega} \leq ph_{r\omega}^n \left(v_{r\omega}^n, Q_{r\omega}^n, sp_{r\omega}^n\right) + \sigma_r^n \frac{\partial ph_{r\omega}}{\partial v_{r\omega}} \left(v_{r\omega}^n, Q_{r\omega}^n, sp_{r\omega}^n\right) \left(v_{r\omega} - v_{r\omega}^n\right)$$

$$+ \sigma_r^n \frac{\partial ph_{r\omega}}{\partial Q_{r\omega}} \left(v_{r\omega}^n, Q_{r\omega}^n, sp_{r\omega}^n\right) \left(Q_{r\omega} - Q_{r\omega}^n\right) + \tau_r^n sp_{r\omega}, \ n = 1, N, r \in R. \quad (4)$$

The hydro production function depends on the discharged outflow, the volume of water in the reservoirs and the spillage. Furthermore, this function may be neither convex nor concave, which requires some a careful approach in the linearization process, similar to the presented in [5], with the purpose to guarantee the construction of a convex function. Therefore, the following steps, based on [5], are required:

i. For all plants, choose N, which is dependent on the precision required. Set $sp = 0$ (ignoring, at the first moment, the spillage).

ii. For each plant, $r = 1, \ldots, R$, set $n = 0$.

Set $Q^0 = Q^{\max}$.

Let (Q^0, v^0) be a feasible point, set $n = n + 1$.

iii. While $n < N$.

If $v^n = v^{n-1}$, calculate $\frac{\partial ph_r^n}{\partial Q_r} (v_r^n, Q_r, 0)$.

Let $Q^n < Q^{n-1}$.

If $\frac{\partial ph_r^n}{\partial Q_r} (v_r^n, Q_r^n, 0) > \frac{\partial ph_r^n}{\partial Q_r} (v_r^n, Q_r^{n-1}, 0)$.

Return to step 3 and set $n = n + 1$.

Else return to step 3.

Else return to step 3 and set $n = n + 1$.

iv. For $n = 0, \ldots, N$, find a first order Taylor approximation of the hydro production function.

$$ph_{r\omega} \leq ph_{r\omega}^n \left(v_{r\omega}^n, Q_{r\omega}^n, 0\right) + \frac{\partial ph_{r\omega}}{\partial v_{r\omega}} \left(v_{r\omega}^n, Q_{r\omega}^n, 0\right) \left(v_{r\omega} - v_{r\omega}^n\right)$$
$$+ \frac{\partial ph_{r\omega}}{\partial Q_{r\omega}} \left(v_{r\omega}^n, Q_{r\omega}^n, 0\right) \left(Q_{r\omega} - Q_{r\omega}^n\right).$$

v. Based on the first order Taylor approximation, a more realistic approximation is obtained by applying a correction scaling factor σ that minimizes the average square deviation between the real function and approximation function.

$$ph_{r\omega} \left(v_{r\omega}, Q_{r\omega}, 0\right) = \sigma ph_{r\omega} \left(v_{r\omega}, Q_{r\omega}, 0\right).$$

vi. Finally, in order to consider the spillage, for $n = 0, \ldots, N$ use secant cuts to approximate the hydro production on the dimension sp, similar to the explained in [5].

$$ph_{r\omega}^n \left(v_{r\omega}, Q_{r\omega}, 0\right) = \sigma ph_{r\omega}^n \left(v_{r\omega}, Q_{r\omega}, 0\right) + \tau_r^n sp_{r\omega}.$$

As a result of this approach, the hyperplanes 4 form a wrap function tangent to the hydro plant production, maintaining, in this way, the problem's convexity.

- Future cost function

$$\alpha - \sum_{\omega \in \Omega_T} \sum_{r \in R} \gamma_r^{(j)} v_{r\omega} \geq \alpha_0^{(j)}, \ j = 1, \ldots, J. \tag{5}$$

This function is given by the longer-term planning model, such as [24] and estimates the expected future cost. It is a piecewise linear function depending on the volume of water in the reservoir at the end of the planning horizon, T. Therefore, it represents the expected future cost from $T + 1$ on. It is worth to notice that these constraints 5 are only included in the subproblems associated with the last stage.

- Power interchange limits between subsystems

$$-Int_{le\omega}^{max} \leq Int_{le\omega} \leq Int_{le\omega}^{max}, \ e \in E, \ l \in \Gamma_e. \tag{6}$$

The variable that describes the power interchange can assume a negative value if the interchange occurs from subsystem e to subsystem l.

- Maximum and minimum volume in the reservoir r

$$v_r^{min} \leq v_{r\omega} \leq v_r^{max}, \quad v_r^{min} \leq v_{ra_\omega} \leq v_r^{max}, \ r \in R. \tag{7}$$

- Limit of the spillage in the reservoir r

$$sp_{r\omega} \geq 0, \ r \in R. \tag{8}$$

- Limits of the discharged outflow in the reservoir r

$$0 \leq Q_{r\omega} \leq Q_r^{max}, \ r \in R. \tag{9}$$

- Thermal and hydro production limits

$$0 \leq pt_{i\omega} \leq pt_i^{max}, i \in I, 0 \leq ph_{r\omega} \leq ph_r^{max}, r \in R. \tag{10}$$

Then, the MTOP problem can be represented by the Linear Problem model:

$$\begin{aligned} \min F &= \ (1), \\ s.t. : \ &(2) - (10). \end{aligned} \tag{11}$$

3.2 Nested decomposition

The ND is a decomposition method based on the Bender's decomposition principle. This method solves the first stage problem and deals with the remaining stages as other subproblems, solving them recursively. In other words, this solves Problem 11 in a recursive manner.

In summary, ND is an iterative process divided into two steps: *forward*, in which the operative cost for each stage is calculated and passed forward to later stages as input to the right hand side; and *backward*, in which the approximations costs are fulfilled and passed back from later stages in the form of optimality cuts, also called a Bender's cut, to the ancestor problem.

In this way, by the application of the single-cut version [8], the Problem 11 is decomposed into subproblems by nodes, which represent a possible water inflow in each stage of the planning horizon. The objective function for the subproblems has the expression:

$$F_\omega = \sum_{e \in E} \left(\sum_{i \in I_e} ct_i \, pt_{i\omega} + cd_e d_{e\omega} + B_\omega \right). \tag{12}$$

The subproblems's constraints are similar to those presented in the previous section; however, in the Eq. 3, the time coupling constraint, is changed for:

$$v_{r\omega} + Q_{r\omega} + sp_{r\omega} - \sum_{m \in M_r} \left(Q_{ma_\omega^{tv}} + sp_{ma_\omega^{tv}} \right)$$
$$= y_{r\omega} + v_{ra_\omega}^{(0)}, \, r = 1, R. \tag{13}$$

In this decomposition scheme, the current stage decision depends on the previous stage decision and, therefore, $v_{ra_\omega}^{(0)}$ denotes the ancestor node solution.

As aforementioned, the cuts created are passed back to the ancestor node subproblem to represent an outer linearization of the recourse function. The expected value of the Lagrange multipliers (dual prices) is used to form the optimality cuts:

$$B_\omega - \sum_{r \in R} \sum_{k \in K_\omega} \left(\frac{p_k}{p_\omega} \right) \pi_{rk} v_{r\omega} \geq H_{r\omega}, \tag{14}$$

where,

$$H_{r\omega} = -\sum_{r \in R} \sum_{k \in K_\omega} \left(\frac{p_k}{p_\omega}\right) \pi_{rk} v_{r\omega}^{(0)} + \sum_{k \in k_\omega} \left(\frac{p_k}{p_\omega}\right) F_k^{(0)}, \qquad (15)$$

where $v_{r\omega}^{(0)}$ denotes the solution of the previous forward simulation and $F_k^{(0)}$ represents the expected future cost on taking the $v_{r\omega}^{(0)}$ decision.

Thus, as the subproblems size is increased iteration after iteration and there are finitely many optimality cuts, the ND method is finitely convergent. Therefore, the initial iterations of this method are very inefficient.

3.3   Progressive hedging

An alternative algorithm to the ND is the PH, which is based on the AL [25]. The fundamental difference between ND and PH is the way that the two algorithms address the nonanticipativity constraints. As shown, ND handles these constraints by having a master problem generating proposals to the subproblems further down the tree scenario; proposal are affected by "futures" nodes by optimality cuts. In PH a different approach is taken: nonanticipativity constraints are relaxed by expressing the large-scale problem in terms of smaller subproblems that are discouraged from violating the original constraints.

Each scenario subproblem is a deterministic problem and has a separate set of variables. These subproblems are coupled by the nonanticipativity constraints which, as discussed above, stipulate that nodes sharing the same stochastic history up to and including that stage must make the same sequence of decisions.

More precisely, the PH models the nonanticipativity by the average value of these scenarios decisions, such as shown in 16.

$$\bar{v}_{rst} = \sum_{x=1}^{X_{st}} p_x v_{rxt}, \; s = 1, S_t, \; t = 1, T - 1. \qquad (16)$$

Observe that in this work the nonanticipativity constraints are represented by the water stored in the reservoirs. This is possible because the storage volumes are the decision variables, which couple constraints in different time steps. Thus, once the water volume is known, the values of all other decision variables are easily obtained.

Applying the AL to the objective function from the DE model, Eq. 1, and considering the nonanticipativity requirement, the subproblems have the following structure:

$$\min \; \Lambda_s = p_s \left[\sum_{t=1}^{T} \sum_{e \in E_t} \left(\sum_{i \in I_e} ct_i \left(pt_{ist} \right) + cd_e \left(d_{est} \right) \right) + \alpha_s \right]$$

$$+ \; p_s \left[\sum_{t=1}^{T-1} \lambda_{rst} \left(v_{rst} - \bar{v}_{rst} \right) \right]$$

$$+ \; p_s \left[\frac{1}{2} \rho \sum_{t=1}^{T-1} \left(\| v_{rst} - \bar{v}_{rst} \|^2 \right) \right],$$

$$s.t. : \sum_{i \in I_e} pt_{ist} + \sum_{r \in R_e} ph_{rst} + \sum_{l \in \Omega_e} Int_{lest} + d_{est} = L_{et},$$

$$v_{rst} - v_{rs,t-1} + Q_{rst} + sp_{rst} - \sum_{m \in M_r} \left(Q_{ms,t-tv} + sp_{ms,t-tv} \right) = y_{rst},$$

$$ph_{rst} \leq ph_{rst}^n(.) + \frac{\partial ph_{rst}}{\partial v_{rst}}(.) \left(v_{rst} - v_{rst}^n \right)$$

$$+ \; \frac{\partial ph_{rst}}{\partial Q_{rst}}(.) \left(Q_{rst} - Q_{rst}^n \right) + \frac{\partial ph_{rst}}{\partial sp_{rst}}(.) \left(sp_{rst} - sp_{rst}^n \right),$$

$$\alpha_s - \sum_{r \in R} \gamma_r^{(j)} v_{rsT} \geq \alpha_0^{(j)}, \; j = 1, \ldots, J,$$

$$- Int_{lest}^{\max} \leq Int_{lest} \leq Int_{lest}^{\max},$$

$$v_r^{\min} \leq v_{rst} \leq v_r^{\max}, \; 0 \leq ph_{rst} \leq ph_r^{\max},$$

$$0 \leq Q_{rst} \leq Q_r^{\max}, \; 0 \leq pt_{ist} \leq pt_i^{\max}, \; sp_{rst} \geq 0,$$

$$e \in E, \; l \in \Gamma_e, \; r \in R,$$

$$s = 1, \ldots, S_t, \; t = 1, \ldots, T, \; n = 1, \ldots, N.$$

(17)

Notice that in the AL function there exists an additional parameter, $\bar{v}$, for each iteration, which is used to uncouple these subproblems for all scenarios [26].

The solution algorithm of the PH consists in an iterative process which solves the subproblems 17 and the dual problem 18:

$$\max \; \psi(\lambda) = \sum_{s=1}^{S} \psi_s(\lambda_s),$$

$$s.t. : \quad \lambda \in \mathbb{R}^m.$$

(18)

Since the dual problem 18 is differentiable, it can be solved by applying of the gradient method:

$$\lambda_{r,s,t}^{iter+1} = \lambda_{r,s,t}^{iter} + \rho \left(v_{r,s,t} - \bar{v}_{r,s,t} \right). \tag{19}$$

This method has two parameter sets that are iteratively updated: the additional parameter $\bar{v}$ and the Lagrange multipliers. Besides, the gradient method only uses information from the last iteration. Owing to these features, the PH is also sensitive to the use of warm start [27].

This class of methods has some advantages when compared to other methods: the primal solution feasibility is guaranteed and it has a weak link between scenarios. The weak link means that in case of computational parallelization, only little information needs to be shared by the processors. The only communication requirement is a share of scenario solutions to enforce nonanticipativity constraints. Thus, this algorithm is suitable for paralleling processing.

4   Computational results

To observe the efficiency of the methods, a performance analysis of each method facing test systems with different complexities are shown. For that, a real hydrothermal system is modeled into two different LP problems, differing from each other by the precision of the hydro production model[6]. In the first case, the hydro plant production is modeled by a piecewise linear function only depending on the discharge outflow, resulting in three linear functions. In the second case, the piecewise linear function is more accurate and it depends on the reservoir storage, discharged outflow and spillage.

In this paper, with the purpose of presenting a coherent analysis of the methods, the computational time was used as the stop criterion.

Before showing the study cases, a brief description of the hydrothermal system is presented.

---

[6]It is important to state that the models of hydro production function are not the focus of this paper. The precision of the hydro production function is only used to evaluate the method's performance. For further details on different manners to model the hydro production function see [28, 29].

4.1   Test problem

The hydrothermal system is based on the Brazilian system and made up of 21 hydro and 20 thermal plants. The total installed capacity is about 57 GW and the demand is about 66% of this value.

The hydro plants are physically connected in cascades, as depicted in Figure 4. This hydro configuration has approximately 47 GW installed. In the Figure 4, the indexes SE, S, NE and N represent the subsystems Southeast, South, Northeast and North, respectively, in which the Brazilian system is subdivided.
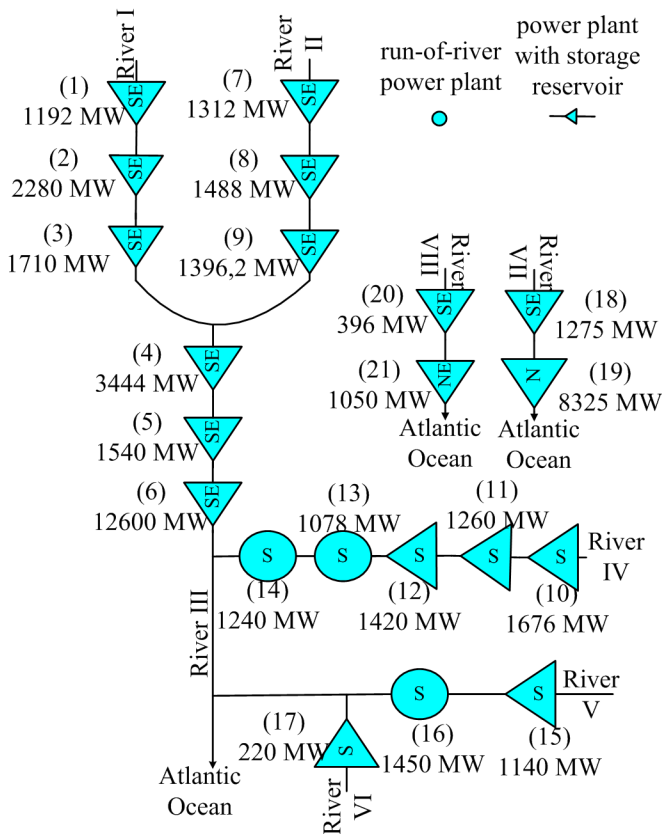

Figure 4 – Test system of hydro plants.

The thermal configuration is composed by 20 plants using different fuels. These plants are located in three different subsystems: NE, SE and S. The total thermal installed capacity is equal to 9.5 GW.

The stochastic model has a study period made up of five weeks plus two months, as depicted in Figure 5, similar to the Brazilian MTOP. The first three stages (three weeks) are modeled as deterministic and the remaining stages as stochastic, each one with seven possible realizations. This scenario model was generated by the periodic autoregressive (PAR) model [30] that is made available by the ISO.
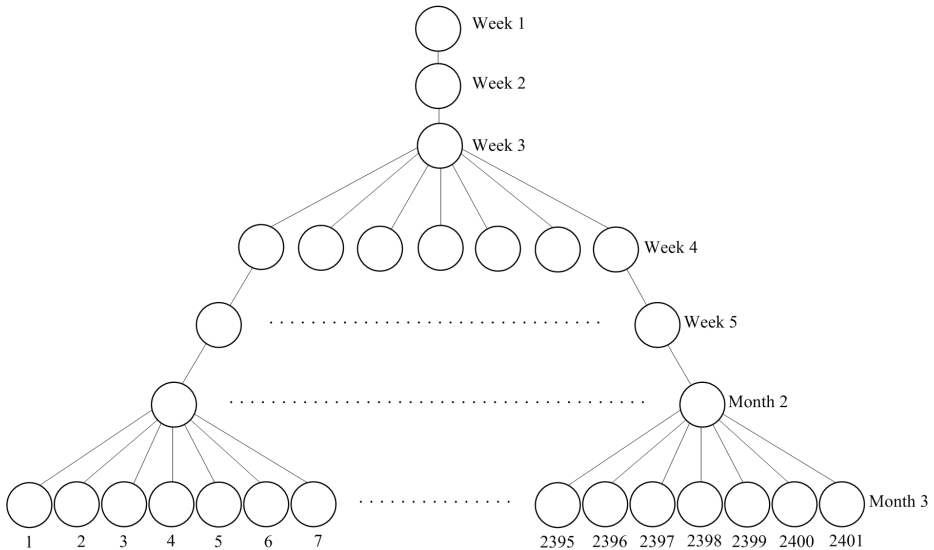

Figure 5 – Inflow scenario tree.

All models of LP (DE and ND) as well as the quadratic programming problems resulting from the PH were solved by the commercial software ILOG CPLEX 7.1. Tests were performed on an Intel Core2 Duo 2.33GHz with 2GB of RAM.

4.2   Case I

In this case, the hydro plant production is represented by a piecewise linear function with three constraints. In addition, this function depends only on the discharged outflow of the hydro plant. This is a simplified model that does not consider the impact of the water head on the production function.

In this case, no warm start techniques were used, and the value of the PH's penalty parameter ρ is 0.001. Additionally, the maximum numbers of iterations

was fixed in five, once the cpu time requested is significantly higher when compared with ND.

Thus, the comparative study is presented in the Table 1. It is important to say that in this case the DE problem has 441,492 variables and 348,323 constraints (101,582 are equality constraints and 246,741 are inequality constraints).

Methods	Iteration	Computation time (h)	Objective function ($ millions)	Deviation
Deterministic Equivalent	—	1.277	96.80	—
Nested Decomposition	21	1.517	96.86	0.070%
Progressive Hedging	5	3.254	96.42	0.391%

Table 1 – Methods comparison – Case I.

The deviations shown in refer to the solution deviation of each methodology (objective function) with respect to the DE solution, which is the exact solution. Notice that the DE method presents the best cpu performance when compared with other methods. Additionally, it can also be observed that ND has reached a better solution faster than PH.

4.3 Case II

The aim of this case is to demonstrate that problem's size increases, the use of DE can be infeasible and, at the same time, to analyze the performance of the other methods against this benchmark.

In contrast to Case I, the hydro plant production is modeled by a piecewise linear function depending on the reservoir storage, the discharged outflow and the spillage. Moreover, this function has five linear constraints. In others words, five points are used to construct the piecewise linear approximation.

In a different way to the Case I, due to preliminary tests, the value of the PH's penalty parameter ρ is 0.1. For both decomposition methods, the computational time was considered as the stop criterion and both should be similar for propitiating a suitable comparison. Additionally, the PH uses the solution of the expected value problem [19] to set additional parameter[7]; therefore this approach works as a warm start.

---

[7]Lagrange multipliers are initialized with value zero.

The comparative analysis between each algorithm is shown in Table 2. The DE problem has 441,492 variables and 492,282 constraints, with 390,700 inequality constraints. Each ND subproblem (node subproblem) in its first iteration has 113 variables and 126 constraints, which 100 are inequality constraints, except in the last stage where it has 114 variables and 726 constraints due to the expected future cost function, summing up 700 inequality constraints. It is worth pointing out that in the ND method, the number of constraints increases at each iteration. On the other hand, each quadratic programming subproblem resulting from the PH decomposition (scenario subproblem) has a fixed set of 792 variables and 1482 constraints, which 1300 are inequality constraints.

Methods	Iteration	Computation time (h)	Objective function ($ millions)	Deviation
Deterministic Equivalent	—	7.03	234.14	—
Nested Decomposition	43	3.76	234.70	0.239%
Progressive Hedging	3	3.75	231.01	1.354%

Table 2 – Methods comparison – Case II.

First of all, notice that the production cost resulting from this case is approximately two times greater than the obtained in the previous case. This occurs because in Case II we used a more realistic model regarding the volume variation[8] and spillage[9]. It is important to model this characteristic because the volume variation in some reservoirs is significant and the spillage is quite common in some plants.

Also, according to Table 2, observe that ND and PH have obtained satisfactory results in a small cpu time in comparison to the DE approach[10], although

---

[8]In the Case I, although it has been calculated three linear functions, the reservoir volume is considered constant in each of them.

[9]The gross head of a reservoir is given by the difference between the forebay and tailrace levels. The net head is defined by the gross head minus the penstock head losses. The forebay level depends on the reservoir volume while the tailrace level is dependent on the total discharged outflow and, in some cases, is also dependent on the spillage. So, considering that a reasonable precision is desired, the modeling of the net head is a critical aspect of problem. For more detail, see [5].

[10]As we mentioned in Section 3, for large scale problems the DE approach is not competitive.

the PH did not already find a feasible solution[11].

Figure 6 illustrates the convergence process of Case II. Observe that if both algorithms were limited to almost one hour of processing, the smallest deviation would be given by the PH. On the other hand, if the time limit was close to two hours and thirty minutes, the PH presents a deviation slightly bigger than the ND. Finally, if the stop criterion was extended to close to four hours, the ND presents the best solution. The Table 3 illustrates these comments.
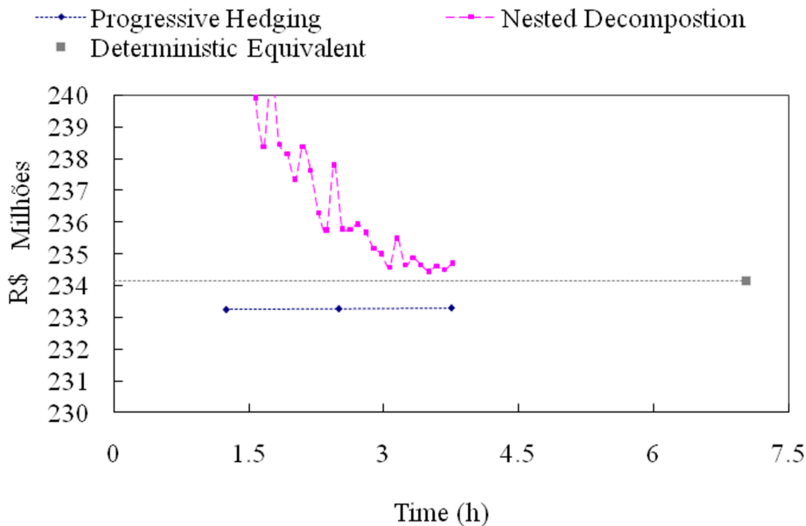

Figure 6 – Convergence process.

Computation time (h)	Progressive Hedging		Nested Decomposition	
	Objetive function ($ millions)	Deviation	Objetive function ($ millions)	Deviation
1.2	230.94	1.384%	246.26	4.922%
2.5	230.98	1.369%	235.78	0.695%
3.7	231.01	1.354%	234.70	0.239%

Table 3 – Methods evolution.

In agreement with Table 2, the number of iterations of ND is bigger than PH.

[11] Given that ND converged at the cpu time of 3.7 hours, we used this reference as the stop criterion.

The ND subproblems are linear and refer to nodes while the PH subproblems are quadratic and refer to scenarios. Therefore, each iteration in the PH method requires more time than the ND method.

These results show that the precision of the mathematical modeling can interfere with the method's performance. In other words, the increase of the problem size (which in this paper depends on the hydro plant's production accuracy) affects the convergence for each method. Additionally, operation cost in Case II is higher than in Case I (as expected), given that a more accurate modeling is used for the hydro plant production. As a consequence, in Case II a larger computational effort is required.

Also, although the ND and PH methods require more computational time than the Case I, now they require less time than the DE to reach lower deviations. This proves that as the size of the problem increases the DE requires more computational cpu time, making its application prohibitive for large systems[12]. We can also observe that ND is more sensitive than PH with respect to the problem's size; this may be an indicative that for bigger problems PH will work better than ND.

5   Conclusions

The objective of the hydrothermal systems MTOP is to calculate an operation strategy for the purpose of minimizing the operation cost over a time horizon. Owing to the uncertainties associated with the inflows scenarios, this problem is quite complex and addressed as a multistage stochastic programming problem. In this paper we compare three solution methods: (a) Deterministic Equivalent, (b) Nested Decomposition, and (c) Progressive Hedging.

Although the Deterministic Equivalent can offer the exact solution, its application in a large-scale problem may be infeasible due to the huge memory requirements and high computational burden. Thus, alternative approaches like Nested Decomposition and Progressive Hedging methods constitute interesting paths to be investigated.

---

[12]It is important to say that limitation in memory occurs when the size of the problem increases. For example, when we tried to simulate a larger case than Case II, the solver gives an error (out of memory). This fact confirms that the solution of large-scale problem by DE approach can be impractical.

Additionally, it is important to emphasize that the accurate representation of the hydro plant production inserts an additional complexity into this particular problem but, as benefit, it improves the quality solution. Certainly the results obtained in the Case I cannot implemented in real life. This fact can be observed when it is compared the computational time and the operation cost of the cases I and II.

Studies presented demonstrated that the Equivalent Deterministic is not useful to solve huge problems and that the Progressive Hedging is competitive when compared to the Nested Decomposition. Moreover, the Progressive Hedging is more stable than the Nested Decomposition, allowing for good solutions with less computational time and for bigger problems PH may respond better than ND.

REFERENCES

[1]   M.E.P. Maceira, L.A. Terry, F.S. Costa, J.M. Damázio and A.C.G. Melo, *Chain of Optimization Models for Setting the Energy Dispatch and Spot Price in the Brazilian Syste*, presented in: 14th PSCC Proceedings, Seville, Spain (2002).

[2]   L.F. Escudero, *WARSYP: a robust modeling approach for water resources system planning under uncertainty*. Annals of Operations Research, **95** (2000), 313–339.

[3]   E.C. Finardi and E.L. da Silva, *Solving the Hydro Unit Commitment Problem via Dual Decomposition and Sequential Quadratic Programming*. IEEE Transactions on Power Systems, **21**(2) (2006), 835–844.

[4]   M.V.F. Pereira and L.M.V.G. Pinto, *Application of Decomposition Techniques to the Mid- and Short-Term Scheduling of Hydrothermal Systems*. IEEE Transactions on Power Systems, **PAS-102**(11) (1983), 3611–3618.

[5]   A.S.L. Diniz and M.E.P. Marceira, *A Four-Dimensional Model of Hydro Generation for Short-Term Hydrothermal Dispatch Problem Considering Head and Spillage Effects*. IEEE transactions on power systems, **23**(3) (2008), 1298–1308.

[6]   M.A.H. Dempster and R.T. Thompson, *Parallelization and aggregation of nested Benders decomposition*. Annals of Operations Research, **81** (1998), 163–187.

[7]   D.P. Morton, *An enhanced decomposition algorithm for multistage stochastic hydroelectric scheduling*. Annals of Operations Research, **64** (1996), 211–235.

[8]   M.V.F. Pereira and L.M.V.G. Pinto, *Stochastic Optimization of a Multireservoir Hydroelectric System: A Decomposition Approach*. Water Resources Research, **21**(6) (1985), 779–792.

[9]   J.F. Benders, *Partitioning Procedures for Solving Mixed Variables Programming Problems*. Numerische Mathematik, **4** (1962), 238–525.

[10] A. Ruszczynski, *Augmented Lagrangian Decomposition for Sparse Convex Optimization*. Annals of Operations Research (1992).

[11] R. Fuentes-Loyola, V.H. Quintana and M. Madrigal, *A Performance Comparison of Primal-Dual Interior Point Method Vs Lagrangian Relaxation to Solve the Medium Term Hydro-Thermal Coordination Problem*. presented at IEEE Power Engineering Society Summer, Seattle, **4** (2000), 2255–2260.

[12] R.T. Rockafellar and R.J.B. Wets, *Scenarios and Policy Aggregation in Optimization under Uncertainty*. Mathematical of Operations Research, **16** (1991), 119–147.

[13] J.P. Watson, D.L. Woodruff and D.R. Strip, *Progressive Hedging Innovations for a Stochastic Spare Parts Support Enterprise Problem*. Journal Article – Naval research Logistics (2007).

[14] K.K. Lau, *Multistage quadratic Stochastic Programming*. Ph.D. dissertation, Dept. Mathematics, Univ. New South Wales (1999).

[15] R. Fourer and L. Lopes, *A Management System for Decomposition in Stochastic Programming*. Annals of Operations Research, **142**(1) (2006), 99–118.

[16] G. Infanger, *Planning under uncertainty: Solving Large-Scale Stochastic Linear Programs*. Ed. Boyd & Fraser Publishing Company (1994).

[17] C. Rosa and A. Ruszczynki, *On Augmented Lagrangian Decomposition Methods for Multistage Stochastic Programs*. Annals of Operations Research (1996).

[18] V.R. Sherkat, K. Moslehi, E.O. Sanchez Lo and J.G. Diaz, *Modelar and Flexible Software for Medium and Short-Thermal Scheduling*. IEEE Transactions on Power Apparatus and Systems, **3**(2) (1988), 1390–1396.

[19] J.R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. 1st ed., New York: Springer (1997).

[20] R. Fuentes-Loyola and V.H. Quintana, *Medium-Term Hydrothermal Coordination by Semidefinite Programming*. IEEE Transactions on Power Systems, **18**(4) (2003), 1515–1522.

[21] J. Dupaèová, G. Consigli and S.W. Wallace, *Scenarios for Multistage Stochastic Programs*. Annals of Operations Research, **100** (2000), 25–53.

[22] A.J. Berger, J. Mulvey and A. Ruszczynski, *An Extension of the DQA algorithm to Convex Stochastic Programs*. SIAM Journal of Optimization, **4** (1994), 735–753.

[23] ILOG, *ILOG CPLEX 7.1*. User's Manual (2001).

[24] E.L. da Silva and E.C. Finardi, *Parallel Processing Applied to the Planning of Hydrothermal Systems*. IEEE Transactions on Parallel and Distributed Systems, **14**(8) (2003), 721–729.

[25] D.P. Bertsekas, *Nonlinear Programming*. 2nd ed., Belmont: Athena Scientific (1999).

[26] A. Ruszczynki, *Decomposition methods in stochastic programming*. Mathematical Programming, **79** (1997), 333–353.

[27] M.L.L dos Santos, E.L. da Silva, E.C. Finardi and R.E.C. Gonçalves, *Solving the Short Term Operating Planning Problem of Hydrothermal Systems by Using the Progressive Hedging Method*, presented at 16th PSCC, Glasgow, Scotland (2008).

[28] J. García-González and G.A Castro, *Short-term hydro scheduling with cascaded and head-dependent reservoirs based on mixed-integer linear programming*. In: Power Tech Proceeding, 2001 IEEE Porto, **3** (2001).

[29] G.W Chang, M. Aganagic, J.G. Waight, J. Medina, T. Burton, S. Reeves and M. Christoforidis, *Experiences with mixed integer linear programming based approaches on short-term hydro scheduling*. IEEE Transactions on Power Systems, **6**(4) (2001), 743–749.

[30] M.E.P. Maceira and J.M. Damázio, *The use of PAR(p) model in the Stochastic Dual Dynamic Programming Optimization Scheme used in the operation planning of the Brazilian Hydropower System*. In: 8th International Conference on Probabilistic Methods Applied to Power Systems, Ames-Iowa, (2004), 397–402.