Handling Infeasibilities when Applying Benders Decomposition to Scheduling Optimization

C. Yuen
Corporate Research
ABB Switzerland Ltd
Switzerland

A. Poncet
Corporate Research
ABB Switzerland Ltd
Switzerland

X.Y. Chao
ABB Consulting
ABB Inc
USA

Abstract: - This paper describes the application of Benders decomposition to hydro-thermal optimal scheduling problems. In particular, Benders decomposition combined with dynamic programming techniques can handle multiple-stage scheduling problems which are common for both short-term and long-term planning. The main advantage of the decomposition approach is the ability to solve linear programming (LP) problems which would be otherwise too large or too time-consuming when standard LP algorithms are employed. Although similar topics were covered in existing literature, the techniques required to handle overall problem and single-stage sub-problem infeasibilities have not been discussed in prior work. The paper will address this issue in detail.

Key-Words: - Benders decomposition, Hydro-thermal scheduling, Dynamic programming, Infeasibilities handling.

1 Introduction

The optimal utilization of hydro energy resources has become more important than ever, due to competitive market environment arising from energy deregulation and the promotion of renewable energies. The problems presented in this article address systems of mixed hydro-thermal energy resources with small to large proportion of hydro energy. The key point of hydro scheduling is the transitional effect of decision from one time stage to another. For example, the exploitation of hydro energy in the early stages can lead to the unnecessary use of expensive resources (thermal) in the later stages when the demand is high. Therefore, given a fixed horizon, the optimal use of hydro energy must be worked out taking into account water inflows, energy prices and demand of all stages in this horizon, while respecting all the system and operating constraints. The techniques presented in this paper can be applied to both short-term and long-term scheduling problems. Even though only deterministic cases are considered in this paper, the techniques can easily be extended to cope with stochastic cases. The latter relate to long-term problems, in which accurate prediction of stochastic variables such as inflow and demand cannot be guaranteed.

The following section gives an introduction to Benders decomposition, followed by a section with the problem formulation for the decomposition approach. Afterwards, we present techniques of handling infeasibilities encountered while solving the single-stage sub-problems under the decomposition approach, which have not been addressed in prior literature. Then the results of some case studies are shown to illustrate the robustness and accuracy of the various approaches. The last section of this paper is reserved for conclusions.

2 Benders Decomposition

One of the first technical articles on Benders decomposition appeared in the early 60s [1]. The initial intended application was to solve mixed-integer programming problems. The main principle is to divide the problem into two parts: one with integers only and the other with continuous variables only. The advantage is that the divided problems are easier to tackle than their original mixed counterpart. The master problem is the pure integer problem with the associated constraints plus the so called Benders cuts. These cuts are formulated through iterations with the dual variables found in solving the sub-problem, in this case the pure continuous variable problem. There are two types of Benders cuts: optimality and feasibility cuts. The first ones are to enhance optimality while the latter to avoid infeasibilities of the sub-problem. The logic of Benders decomposition is depicted in the flowchart of Figure 1.

This decomposition technique was elaborated in the 80s to solve multi-stage stochastic optimization problems for multi-reservoir hydroelectric systems [2][3]. This is done through a nested approach of the original Benders decomposition based on the principle of dynamic programming. The deterministic case is called Dual Dynamic Programming (DDP) and the stochastic case Stochastic Dual Dynamic Programming (SDDP), respectively.

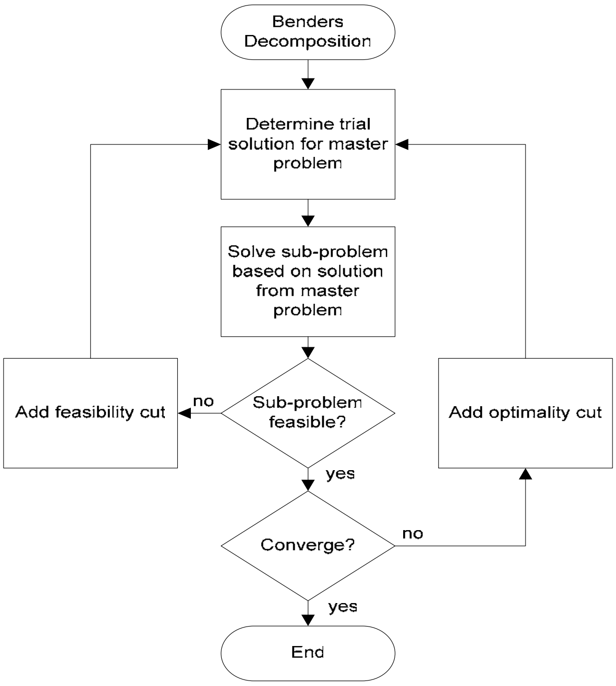

Figure 1 Logic Flowchart of Benders Decomposition

3 Problem Formulation

In this paper the problems aim at system cost minimization, i.e. the role played by the system/market operator. However, it should be noted that the same principles can equally be applied to profit maximization, i.e. for individual generation companies. Thus consider a problem with T stages, for which the objective function can be expressed as:

$$\min \ f(\tilde{x}_1,\ldots,\tilde{x}_T)=\sum_{t=1}^{T}\tilde{c}_t{}'\,\tilde{x}_t \ , \tag{1}$$

where $\tilde{c}_t$ and $\tilde{x}_t$ are column vectors of the same length with cost coefficients and optimization variables respectively.

For each stage t = 1, 2, … T, the following constraints must be satisfied:

$$\sum_{i}^{G} x_t^i = v_t^1 \tag{2}$$

and, for each storage plant,

$$\kappa\, x_t^p + x_t^o + x_t^s - x_{t-1}^s = v_t^2 \tag{3}$$

where

x_t^i = the energy output from all thermal and hydro plants

G = the total number of plants

v_t^1 = the energy demand in stage t

$\kappa\, x_t^p$ = the amount of water through turbine in stage t

x_t^o = the amount of water spilt in stage t

x_t^s = the amount of water in reservoir at end of stage t

v_t^2 = the amount of inflow into the reservoir in stage t

Equation (2) stands for the energy balance equation while equation (3) is for water conservation. Note that more complicated hydraulic constraints can be added to model upstream / downstream reservoir relationships. For demonstration purposes we only show the simplest model here. Besides, the maximum and minimum allowed values for each variable must be respected:

$$\tilde{x}_t^{\min} \le \tilde{x}_t \le \tilde{x}_t^{\max} \tag{4}$$

These form a set of linear equations and can be generalized as:

$$A_t\tilde{x}_t \ge \tilde{b}_{1t}\ ,$$
$$E_{t-1}\tilde{x}_{t-1} + A_t\tilde{x}_t \ge \tilde{b}_{2t} \tag{5}$$

Based on the decomposition approach, the problem is solved stage by stage and each single-stage master problem is formulated as:

$$\begin{aligned} \min \quad & \tilde{c}_t{}'\tilde{x}_t + \hat{\alpha}_t(\tilde{x}_t) \\ s.t. \quad & A_t\tilde{x}_t \ge \tilde{b}_1 \end{aligned} \tag{6}$$

where represents the approximate future cost function of stages t+1 to T. Based on duality theory [5], this approximate future cost function is in fact a convex piecewise linear function and is expressed as:

$$\hat{\alpha}_t(\tilde{x}_t) = \min \alpha$$
$$\begin{aligned} s.t. \quad & \alpha \ge \tilde{\pi}^1(\tilde{b}_2 - E_{t-1}\tilde{x}_t) \\ & \alpha \ge \tilde{\pi}^2(\tilde{b}_2 - E_{t-1}\tilde{x}_t) \\ & \vdots \\ & \alpha \ge \tilde{\pi}^{n_o}(\tilde{b}_2 - E_{t-1}\tilde{x}_t) \end{aligned} \tag{7}$$

where the $\tilde{\pi}^j$ are the dual variables of the sub-problem for stage t+1 and no is the total number of iterations. For multi-stage problems, the real future cost for the current solution is compared with the approximate future cost to determine if convergence is reached. Each iteration consists of forward simulation where trial solutions are found and backward recursion where the approximate future cost functions are updated. Further explanation of the approximate future cost function and details of derivation can be found in [3].

4 Handling Infeasibilities

In this article we put emphasis on how infeasibilities should be handled since most other articles focus on how optimality cuts are derived but not on feasibility cuts [2][3][6][8]. The handling of infeasibilities is important while searching for the optimal solution since convergence cannot be achieved when infeasibilities are not dealt with properly.

In this article two possible approaches are proposed. The first one is the traditional approach where feasibility cuts are added when necessary. When infeasibilities are encountered while solving the single-stage sub-problem at stage t, a feasibility check sub-problem has to be set up [7]. This feasibility check sub-problem is to minimize all the violations while respecting the same set of constraints as the original sub-problem. Mathematically, these cuts are formulated as follows:

$$s(\hat{\tilde{x}}_{t-1}) - (\tilde{x}_{t-1} - \hat{\tilde{x}}_{t-1})' E'_{t-1} \tilde{u} \leq \tilde{0} \qquad (8)$$

where

$s(\hat{\tilde{x}}_t)$ = the optimal solution of the feasibility check sub-problem

$\tilde{u}$ = dual variables found from the feasibility check sub-problem

$\hat{\tilde{x}}_{t-1}$ = the original trial values of $\tilde{x}_{t-1}$.

The principle of the feasibility cuts is that a more restrictive constraint should be applied to the previous stage problem in order to ensure the feasibility of the next stage. In nested Benders decomposition (i.e. multi-stage problems), optimality cuts for the approximate cost function cannot be updated until infeasibilities are cleared from all stages t = 1 to T. Moreover, the constraints from one stage must be carried over to as many previous stages as necessary in order to clear the infeasibilities. This is depicted in Figure 2.

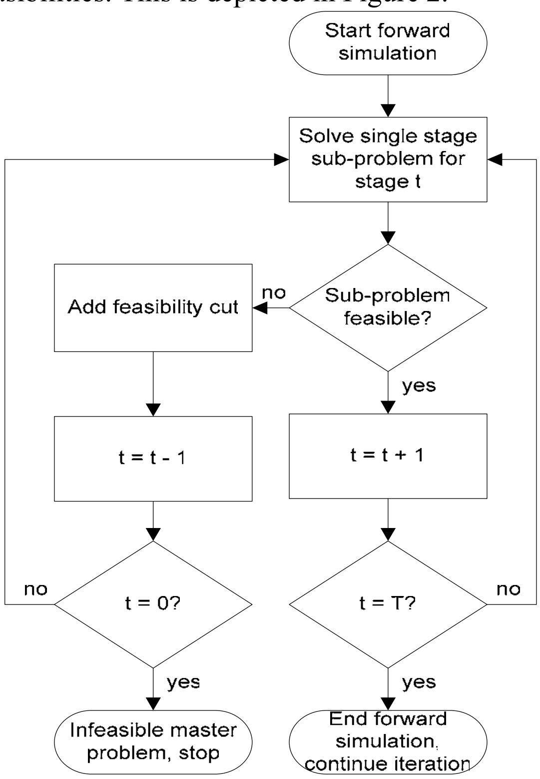

Figure 2 Feasibility Cuts' Implementation for multi-stage problems

The above-mentioned approach is the standard approach of handling sub-problem infeasibilities when implementing Benders decomposition. However, it suffers from the disadvantage that the iteration has to come to a stop when the overall problem is infeasible, i.e. when t = 0 as shown in Figure 2. In view of this, a second approach is presented.

In this second approach, a set of so-called "violation" variables $\tilde{\varepsilon}$ are added to the original optimization variables to form a new set of optimization variables. The objective function of this new formulation is

$$\min f(\tilde{x}_1, \ldots, \tilde{x}_T, \tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_T) = \sum_{t=1}^{T} (\tilde{c}' \tilde{x}_t + \tilde{d}_t' \tilde{\varepsilon}_t) \qquad (9)$$

where the new components are:

$\tilde{d}_t$ = penalty cost vector

$\tilde{\varepsilon}_t$ = violation variables for constraints.

The purpose of using violation variables is to make sure that any violations of the constraints can be expressed in terms of positive ε. In order to achieve that, it is necessary that

$$\tilde{d}_t > \tilde{c}_t$$

and

$$\tilde{\varepsilon}_t \geq 0. \qquad (10)$$

Note that the penalty costs should be bigger than all the other original costs so that the amount of violation is minimized. These violation variables are included into the various constraints as follows:

For equality constraints, e.g. the energy balance equation (2):

$$\sum_{i}^{G} x_t^i + \varepsilon_t^1 - \varepsilon_t^2 = v_t^1 \qquad (11)$$

For inequalities, e.g. soft upper bounds:

$$\tilde{x}_t - \tilde{\varepsilon}_t^{\max} \leq \tilde{x}_t^{\max}$$

The main advantage of this approach is that no infeasibilities are encountered during iterations and therefore the feasibility cuts are not needed. However, although there are no additional optimization problems (the feasibility check sub-problems) that have to be solved each time infeasibilities are encountered, the overall computational time is not necessarily reduced. This is because the decomposed problems have additional variables and therefore take more time to solve. Moreover, it can be shown that the penalty costs need to be above a certain threshold so that optimal solution can be obtained [4]. In other words, convergence can only be guaranteed when the penalty costs are chosen with care.

5 Case Studies

To illustrate the performance of the various proposed approaches, a prototype has been set up in Matlab to run several small test cases. The linear programming engine is provided by the Optimization Toolbox of Matlab. Each case is solved using both the feasibility cuts approach and the violation variable approach. Furthermore, for each of these cases, the results are compared with the corresponding equivalent LP formulation, i.e. multi-stage variables and constraints are aggregated to form one large-scale LP problem. The solution from this equivalent LP provides a benchmark, which can be used to check the accuracy of the other methods.

5.1 Case Description

For demonstration purposes, only small cases are studied. However, the same algorithm is under implementation in an energy market simulation tool, GridView (please see Acknowledgement), and the preliminary results for large-scale cases are encouraging. The base case consists of two thermal plants with different costs and five hydro-storage plants. The constraints for each stage are those stated in equations (2), (3) and (4). For more realistic results, the exogenous variables, i.e. inflow and demand, are scaled historical data. The total system demand and the inflow for one of the reservoirs are plotted in Figure 3.

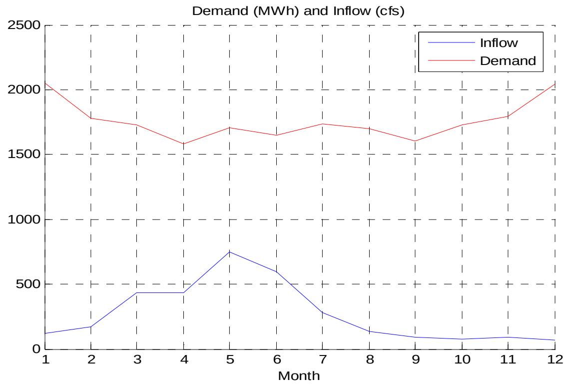

Figure 3 Annual Inflow and Demand Values

It should be noted the relative values of the demand and the inflow in different months of the year are typical. For many places where there is snow in winter, the large quantity of inflow comes when the snow starts melting in spring. However, the demand is usually relatively low this time of the year since the highest demand is during the coldest months of the year, i.e. in winter. This explains the important role of storage plants in a yearly scheduling horizon: water is saved in the reservoir when inflow is high and later on this water is used for electrical energy generation when the demand is high.

5.2 Case 1a: 2-thermal, 5-hydro, overall feasible

In this case there are five hydro plants and the exogeneous variables and constraints are set so that the overall problem is feasible. However, sub-problem infeasibilities are still encountered during iterations, especially the initial ones. The reason is because of the requirement to fill up the reservoirs at the end of the year to the same level at the beginning of the year. Some of the outputs are plotted in the following diagrams. Figures 4 to 8 show the outputs from the violation approach. Figure 9 shows the convergence trend from the feasibility cut approach. Since the results from the latter method are similar to those from the former method, the results are not plotted.

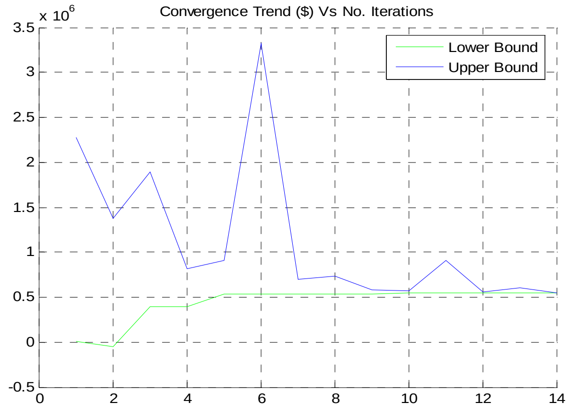

Figure 4 Convergence Trend (violation approach)

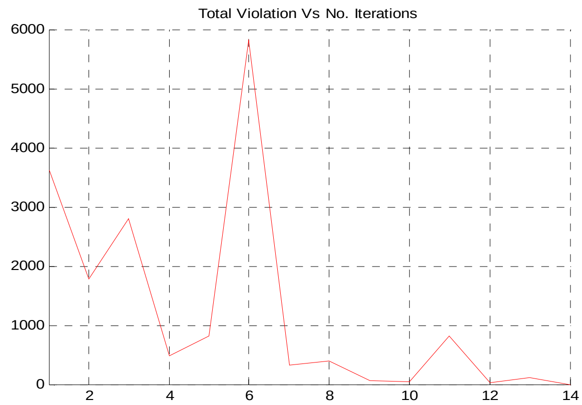

Figure 5 Total Amount of Violation (violation approach)

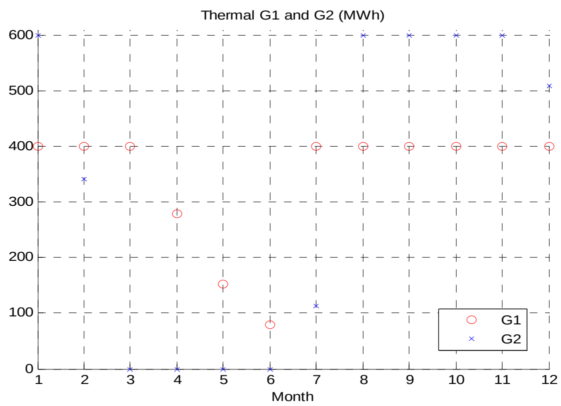

Figure 6 Thermal Outputs of 2 Thermal Plants (violation approach)

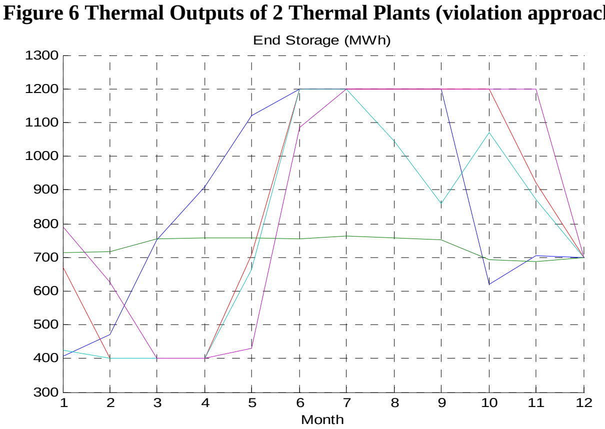

Figure 7 End Storage of 5 Reservoirs (energy equivalent; violation approach)

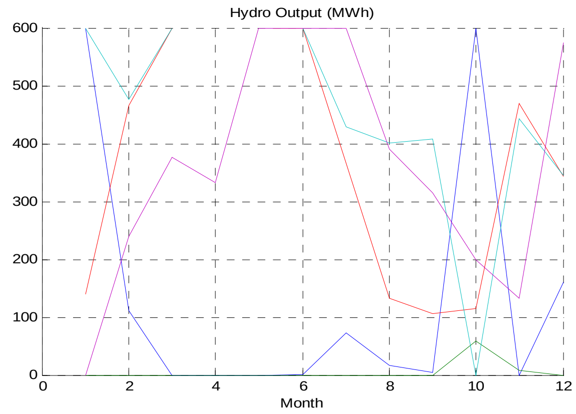

Figure 8 Hydro Outputs of 5 Hydro Plants (violation approach)

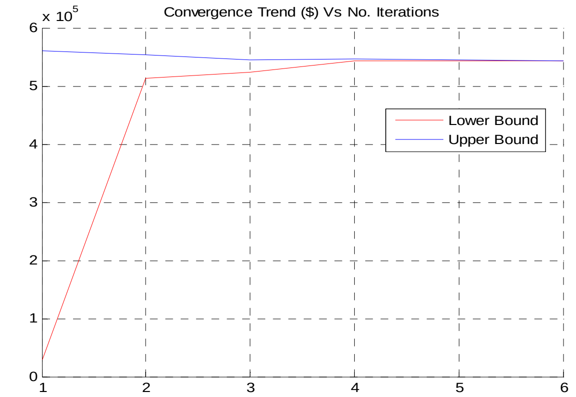

Figure 9 Convergence Trend (feasibility cut approach)

Comparison of the final objective values from the two different approaches with the equivalent LP results show that they are identical, meaning that global optimum is reached in both cases. However, in terms of computational time, the feasibility cut approach takes only 6 seconds while the violation approach takes 41 seconds. Moreover, the first approach takes 6 iterations while the latter one requires 14 for convergence. Besides, it is also interesting to look at Figures 4 and 5 which show the convergence trend (upper bound and lower bound) and total amount of violations plotted against the number of iterations for the violation approach. The total amount of violations is defined as the sum of the values of ε's after each iteration. The upper bound follows the same trend of the total amount of violations while the lower bound increases slowly. When total amount of violation reaches zero at 14th iteration, the upper bound decreases to the same value of the lower bound, i.e. convergence is reached.

5.3 Case 1b: 2-thermal, 5-hydro, overall infeasible

This case is the same as case 1a, except that some exogeneous variables are changed to make the overall problem infeasible. It is therefore obvious that the approach using feasibility cuts will not work, since the iteration process will come to a halt when violations happen in the first stage. In this particular case, it is enough to increase the amount of demand for all the months to create infeasibilities for the overall problem. Some of the results of this case are plotted in the following diagrams.

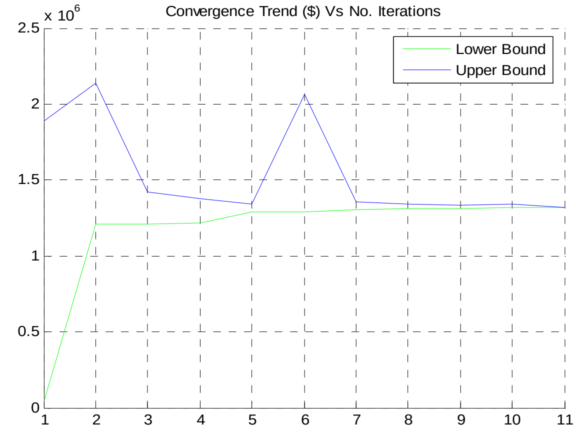

Figure 10 Convergence Trend

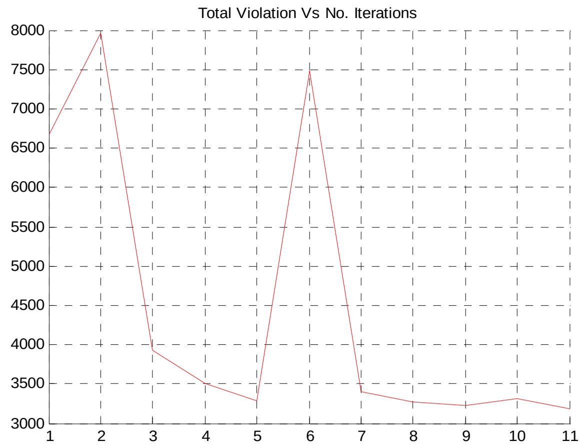

Figure 11 Total Amount of Violation

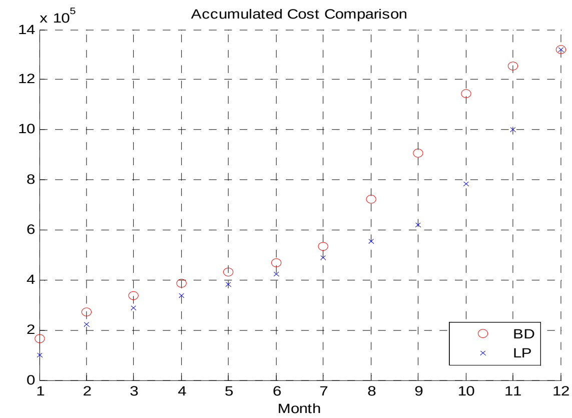

Figure 12 Accumulated Cost BD Vs LP

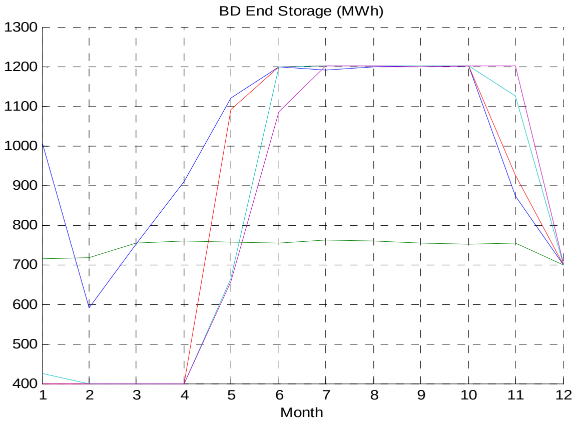

Figure 13 End Storage of 5 Reservoirs (energy equivalent; BD)

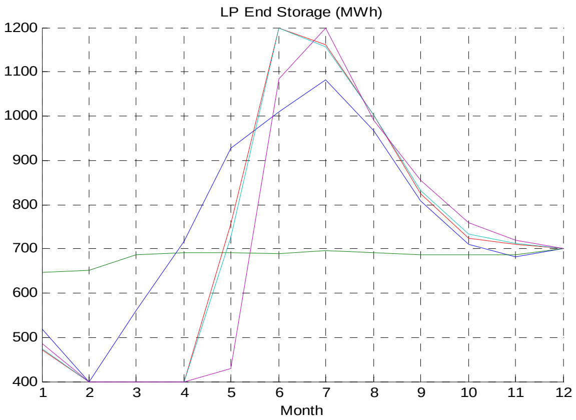

Figure 14 End Storage of 5 Reservoirs (energy equivalent; LP)

Figures 10 and 11 show the convergence trend and the total amount of violation plotted against the number of iterations. Similar to the previous case the upper bound has a similar shape as the total amount of violation, while in this case the amount of violation does not reach zero when convergence is reached. This is expected since the overall case is infeasible in reality and therefore non-zero ε's are needed in the optimum solution.

Figures 12 – 14 show the comparison of results from the decomposition approach and the equivalent LP approach. Note that even though the total costs summed up for the whole year are the same in both cases, the values of the optimization variables are not necessarily identical. Figure 12 displays the accumulated costs along the year. It can be seen that, for this particular case, the total cost from the decomposition approach is higher than the LP approach except the very last month when they reach the same value. Figures 13 and 14 show the reservoir levels along the year for both approaches. Clearly the use of hydro energy is different, which implies the different use of thermal energy, therefore the different operating costs in the two approaches. Indeed, depending on the data and the setup of the cases, it is possible that more than one global optimum exists. As in this example: decomposition approach finds one, the equivalent LP finds another.

5.4 The Importance of the Penalty Costs

This section refers to the exact same case studied in the previous section. In Section 5.3 the results from the decomposition approach are obtained from a case which converges, with careful selection of the penalty costs, i.e. the costs associated to the ε's in the objective function. In other words, they are the penalty costs caused by one unit of violation. In fact these costs can have a very important effect on the convergence of the problem. As a rule of thumb they should be larger than any of the other costs in the objective function. In our case, the costs associated to the thermal plants are 50$/MWh and 65$/MWh, those for all the hydro plants are 5$/MWh, while spillage and storage have zero costs. The penalty costs for the ε's used in Section 5.3 are 190$/MWh. Other values of the penalty costs are used and the results are summarized in Table 1.

TABLE 1 THE EFFECTS OF DIFFERENT PENALTY COSTS ON CONVERGENCE

Penalty Cost ($/unit violation)	Total Violation	No. Iterations
150	4384	20
170	3180 (optimal)	16
190	3180 (optimal)	11
200	3201	20
300	3456	20

The table shows the total amount of violation at convergence or after 20 iterations. Figure 15 shows the "oscillating" pattern of the total violation plotted against the number of iterations for the case with 150$/unit violation as penalty cost.

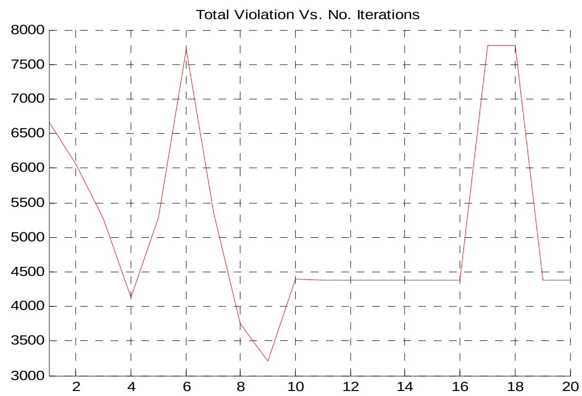

Figure 15 Total Amount of Violation (Non-convergent case)

6 Conclusions

This paper describes the application of Benders decomposition to hydro-thermal scheduling problems. In particular the issue of how to handle single-stage sub-problem or overall-problem infeasibilities is discussed and two possible approaches are described. The use of feasibility cuts is more efficient when the problem is overall feasible. However, the use of violation variables is indispensable for cases that are overall infeasible.

Nevertheless, the penalty costs must be chosen carefully so as to ensure convergence in the latter case. In general, both approaches can be combined to achieve the required performance of the optimization process. For example, violations such as load shedding or storage requirement are foreseeable and therefore violation variables should be included to take care of the relevant constraints. On the contrary, the violations arising from other constraints that occur in single-stage sub-problems during iterations should be taken care of by the feasibility cuts.

7 Acknowledgement

The authors would like to thank John Fazio of Northwest Power and Conservation Council, Mary Johannis and Dennis Phillips of BPA and the ABB Consulting unit in Raleigh, North Carolina, USA for their valuable contribution to the project, in particular for their contribution of practical knowledge and requirements on the development of algorithms for the optimal hydro-thermal scheduling functionalities for the ABB energy market simulation tool GridView.

References:

[1] J. F. Benders, "Partitioning Methods for Solving Mixed Variables Programming Problems," *Numerische Mathematik*, vol. 4, pp. 238-252, 1962.

[2] M. V. F. Pereira and L. M. V. G. Pinto, "Stochastic Optimization of a Multireservoir Hydroelectric System: A Decomposition Approach," *Water Resources Research,* vol. 21, no. 6, pp. 779-792, 1985.

[3] M. V. F. Pereira and L. M. V. G. Pinto, "Multi-stage Stochastic Optimization Applied to Energy Planning" *Mathematical Programming,* vol. 52, pp. 359-375, 1991.

[4] X. Cai, D. C. McKinney, L. S. Lasdon and D. W. Watkins, Jr, "Solving Large Nonconvex Water Resources Management Models Using Generalized Benders Decomposition", *Operations Research, INFORMS,* vol. 49, No. 2, , pp. 235-245, Mar-Apr 2001.

[5] R. J. Vanderbei, Linear Programming: Foundations and Extensions, Springer (International Series in Operations Research & Management Science, Vol. 37), 2nd Edition, 2001.

[6] G. B. Dantzig and G. Infanger, "Multi-Stage Stochastic Linear Programs for Portfolio Optimization", Department of Operations Research, Stanford University, Technical Report, SOL 91-11, Sep 1991.

[7] Mohammad Shahidehpour and Yong Fu, "Benders Decomposition in Restructured Power Systems", IEEE Techtorial, Apr 2005.

[8] M. Pereira, N. Campodonico, R. Kelman, "Long-term Hydro Scheduling based on Stochastic Models", EPSOM'98, Zurich, Sep 1998.