Hydropower generation management under uncertainty via scenario analysis and parallel computation

L.F. Escudero
UITESA, grupo Iberdrola
Dept. of Statistics and Operations Research
Univ. Complutense, Madrid

J.L. de la Fuente
Iberdrola
Dept. of Applied Mathematics
Industrial Engineering School
Univ. Politecnica, Madrid

C. Garcia
UITESA, grupo Iberdrola
Spain

F.J. Prieto
Dept. of Statistics and Econometrics
Univ. Carlos III, Madrid

Abstract—We present a modeling framework for the robust solution of hydroelectric power management problems with uncertainty in the values of the water inflows and outflows. A deterministic treatment of the problem provides unsatisfactory results, except for very short time horizons. We describe a model based on scenario analysis that allows a satisfactory treatment of uncertainty in the model data for medium and long-term planning problems.

Our approach results in a huge model with a network submodel per scenario plus coupling constraints. The size of the problem and the structure of the constraints are adequate for the use of decomposition techniques and parallel computation tools. We present computational results for both sequential and parallel implementation versions of the codes, running on a cluster of workstations. The codes have been tested on data obtained from the reservoir network of Iberdrola, a power utility owning 50% of the total installed hydroelectric capacity of Spain, and generating 40% of the total energy demand.

I. INTRODUCTION

The problem of planning the operation of a hydrothermal system is usually formulated through the coordination of the thermal and hydraulic parts of the system via the use of the thermal generating prices to obtain the cost of the system operation, and the addition of coupling constraints on the satisfaction of the demand—see [1], [2], [3], [4].

Due to the complexity of the problem and the uncertain nature of several components in the operation of the system, this planning process is normally divided into two stages: the long-term planning problem, on a horizon of one year or longer, where uncertainty plays a crucial role; and the short-term planning problem, on a one-day to one-week horizon, where uncertainty is treated through the results obtained from the long-term model.

We are specifically interested in the efficient treatment of uncertainty within the long-range planning problem; this uncertainty is fundamentally due to the stochastic nature of the water inflows and outflows to the system. Decomposition approaches have been used in the past to solve this problem with remarkable success—see [5], [6], [7], [8] and [9]. Most of these approaches are based on Benders decomposition; we wish to explore the feasibility of using other alternative decomposition-based techniques on a model of the hydrothermal coordination problem where uncertainty is treated via a scenario approach. These techniques are based on Augmented Lagrangian approaches, and from our experience in the solution of general stochastic linear network problems, reported in [10], we believe they may prove more efficient regarding their storage requirements, running times for large-scale problems and flexibility to be extended to the general nonlinear non-convex case.
Our goal is to be able to solve problems large enough to allow a sufficiently detailed representation of uncertainty in the data. An additional advantage of this approach is the possibility of achieving significant reductions in running times, by performing the computations in a distributed environment.

In this paper we have chosen to study the hydraulic part of the system in isolation, in order to better concentrate on the issues mentioned above, as it is the part of the system that is most directly affected by uncertainty in the environment.

The paper is structured as follows: in Section II we introduce the model for the hydraulic subsystem. Section III discusses several issues associated with uncertainty in the data of the problem. Section IV presents the decomposition framework. In Section V we give the implementation results from both the sequential and parallel versions of the code. Finally, Section VI is devoted to the analysis of the conclusions obtained from the results reported in the preceding Section.

II. PROBLEM DESCRIPTION

Our formulation of the problem of maximizing the hydropower generated along a time horizon by a multireservoir power system starts with the consideration of a basic network obtained from a river, where the nodes represent the reservoirs and the arcs correspond to the sections of the river that connect the reservoirs. This basic network is then replicated on a number of segments, corresponding to the time periods (usually, weeks) included in the planning horizon; these segments are linked by special arcs on some of the nodes.

The decision variables in the optimization problem are the amount of water to be released from each reservoir to its direct downstream reservoirs in a given period (corresponding to the arcs in the basic network), or the amount of water to be stored in the reservoirs from one period to the next one (the special arcs linking the segments). This amount becomes available to be released or stored in the next period.

The constraints are of two kinds: linear equations that ensure flow balance in each reservoir, and simple bounds on the variables. These simple bounds have the following purposes: first, to ensure that the water released serves the desired flood control, irrigation and navigational purposes; second, to ensure that the amount of water released from a given reservoir to any of its directly downstream reservoirs does not exceed the canal capacity; third, to penalize the amount of stored water that exceeds a safety capacity for any given reservoir; and fourth, to force the amount of water stored in the reservoirs to remain below a given upper bound.

The objective consists of the maximization of a nonlinear function, typically the generation of hydroelectricity over the planning horizon, or alternatively, the savings in the cost of thermal generated power.

We now introduce a formal model for the problem described above. Let \( J \) denote the set of reservoirs, \( T \) be the time periods set, \( W \subset J \) corresponds to the water storage reservoirs (so that \( J \setminus W \) is the set of run-of-the-river plants), \( E \subset J \) is the set of reservoirs that are not used for hydropower generation, and \( F \) \((Q_j)\) is the set of reservoirs directly upstream (downstream) from reservoir \( j \). The decision variables will be denoted by \( r_{ij}, s_{ij} \), the amount of water released from reservoir \( j \) to reservoir \( i \in Q_j \) in period \( t \), and \( s_{ij} \), the amount of water stored in reservoir \( j \) at the beginning of period \( t \).

The flow balance equations in each reservoir for each time period are as follows

\[
- \sum_{i \in F_j} r_{ij} - s_{ij} + \sum_{i \in Q_j} r_{ij} + s_{i+1,j} = b_j \quad \forall i \in T, j \in J \tag{1}
\]

where \( b_j \) is the exogenous inflow to reservoir \( j \) in time period \( t \), that could be negative if we have a net outflow during period \( t \); \( s_{ij} \) is fixed for \( t = 1 \) and \( t = |T| + 1 \). If we denote by \( b \) the vector of all decision variables, and \( b \) is the vector of exogenous inflows, system (1) can be written as

\[
Ax = b,
\]

where \( A \) is the node-arc incidence matrix.

The bounds on the arcs of the replicated network are given by

\[
l_{ij} \leq r_{ij} \leq u_{ij} \quad \forall i \in Q_j, j \in J, t \in T \tag{2}
\]

\[
m_{ij} \leq s_{ij} \leq M_{ij} \quad \forall j \in W, t \in T(1) \tag{3}
\]

The hydropower generated in reservoir \( j \) at time period \( t \) can be expressed as follows

\[
h_{ij} = K_{ij} r_{ij} \quad \forall i \in T, j \in J \setminus E, i \in Q_j \tag{4}
\]

where \( K_{ij} \) may be a linear or nonlinear function of the \( s_{ij} \)-variables for the reservoirs in set \( J \setminus E \), and it can be approximated by

\[
K_{ij} \approx f_{ij}(s_{ij}, s_{i+1,j}) \tag{5}
\]

Let \( R_{ij} \) denote the power generation water flow capacity of canal \((j,i)\) for \( j \in J \setminus E \), where \( l_{ii} \leq
$R_{ij} \leq u_{ij}$, and let $T_{ij}$ be an upper bound on the amount of water that can be safely stored in reservoir $j$ for $j \in W$ at time period $t$. Since the water overflow $r_{ijj} - R_{ij}$ cannot be used for hydropower generation, and we would be interested in penalizing the excess water stored $s_{ij} - T_{ij}$, the objective function can be written as

$$\max_{i \in P} \left( \sum_{j \in S} h_{ijj} \min(r_{ijj}, R_{ij}) \right)$$

where $h_{ijj}$ is now defined as

$$h_{ijj} = K_{ijj} \min(r_{ijj}, R_{ij})$$

and $U_{ij}$ is either a unit penalty or a nonlinear penalty function of the excess water stored.

The focus of this paper is on methodologies to deal with uncertainty on the right-hand side of (1). These methodologies can be easily extended to treat uncertainty in other parts of the problem, such as the objective function; this would be the case if we would want to model uncertainty in the thermal generation cost savings.

In order to facilitate the presentation of the solution procedures we will introduce some simplifications on the model described in (1)–(6); we will assume that the objective function (6) is approximated by some polynomial function $f$, and we will consider all bounds on the variables to be just non-negativity conditions, so that we obtain the following representation of the problem

$$\min_{z} \quad f(z)$$

s.t. $Ax = b$

$$z \geq 0.$$  

(7)

III. Uncertainty

A. Modelling uncertainty via scenarios

The model described in Section II must be modified in order to deal properly with uncertainty on the values of some of the parameters—in this case the water inflows and outflows. It is important to model adequately the availability of information over time and state, as it will restrict the decisions that can be made at each of the various stages. Also, in order to compute an optimizer any proposed solution will have to be compared with other candidate solutions; in the stochastic setting, the criteria on which this comparison can be based are not uniquely defined.

A traditional approach to represent uncertainty in this setting is to make distributional assumptions, estimate the parameters from historical data and then develop a stochastic model to take uncertainty into account. Such approach may not be appropriate if only limited information is available. On the other hand, in many situations it is often necessary and possible to take into account information that is not reflected in the historical data. In these cases we may employ a technique called "scenario analysis", where uncertainty is modeled via a set of scenarios, say $S$.

In our network model the right-hand side vector may take different values for different scenarios, $b^s$ for $s \in S$, corresponding to different exogenous water inflows and outflows. We also introduce weights $w^s$ representing the likelihood that the decision maker (modeller) associates with each scenario $s \in S$.

B. Implementable policies.

In order to modify model (7) to include uncertainty in the right-hand side, it is important to take into account the use intended for the information provided by the model. Some formulations do anticipate decisions in $z$ that for multistage applications may not be needed at stage $r = 1$. However, frequently the decisions for stage $r = 1$ are the only decisions to be made, since at stage $r = 2$ one may realize that some of the data has been changed, some scenarios vanish, etc. In this case, the models will usually be reoptimized in a rolling planning horizon mode. When only spot decisions (i.e., decisions for stage $r = 1$) are to be made, the information about future uncertainty is only taken into account for a better spot decision making. This type of approach is termed full recourse. In this full recourse model the variables may take different values in different scenarios.

Let $z^s$ denote the arc flows under scenario $s$, for $s \in S$. The following nonanticipativity principle has been stated in [11].

"If two different scenarios $s$ and $s'$ are identical up to stage $r$ on the basis of the information available about them at stage $r$, then the values for the $z$-variables must be identical up to stage $r'".$

This condition guarantees that the solution obtained by the model is not dependent at stage $r$ on information that is not yet available.

Let $N$ denote the set of solutions that satisfy the so-called nonanticipativity constraints. A full re-
course model may have the following form

\[
\begin{align*}
\min & \quad \sum_{s \in S} w_s f(x^s) \\
\text{s.t.} & \quad Ax^s = b^s \quad \forall s \in S \\
& \quad x \in N \\
& \quad x^s \geq 0 \quad \forall s \in S.
\end{align*}
\]

\( (8) \)

C. Scenario immunization

We still have to address the problem of defining the meaning of a solution for a stochastic optimization problem. One way to address this difficulty is to obtain a solution that best tracks each of the scenarios, so that this computed solution is hedged against the occurrence of low-probability scenarios with unfavorable consequences for the operation of the system—see [12]. We can achieve this by obtaining a solution that minimizes the weighted difference between the proposed solution value and the optimal solution value for each scenario.

Our full recourse (FR) model with scenario immunization is as follows

\[
\begin{align*}
\min & \quad \sum_{s \in S} \omega(s) f(x^s) - z^*_s \\
\text{s.t.} & \quad Ax^s = b^s \quad \forall s \in S \\
& \quad x \in N \\
& \quad x^s \geq 0 \quad \forall s \in S.
\end{align*}
\]

\( (9) \)

where \( \omega = \max(0, r) \), \( z^*_s \) denotes the optimality tolerance,

\[
z^*_s = z^*_s + \epsilon \quad \text{for } \epsilon \geq 0,
\]

\( (10) \)

\( z^*_s \) is the optimal solution value for model (7) under scenario \( s \) for \( s \in S \), and \( d \) is a positive parameter. See in [10] the description of a two-stage procedure to deal with the nonlinearity caused by the \( d \)-norm in the objective function of (9).

Model FR (9) has a nice structure that we may exploit. Two approaches can be used to represent the nonanticipativity constraints \( x \in N \). Approach A1, also known as the deterministic equivalent of the stochastic problem—see [13], is based on a compact representation, where some variables are directly eliminated from the formulation. This approach reduces the model size, but the network structure of the constraints is destroyed. Approach A2, the one presented in the remainder of the paper, is based on a splitting variable representation where condition \( x \in N \) is part of the model. Dualizing \( x \in N \) results in independent network submodels.

IV. THE DECOMPOSITION ALGORITHM

A. General framework.

The nonanticipativity constraints \( x \in N \) constitute the major obstacle to the decomposition of model (9). An Augmented Lagrangian method partially overcomes this difficulty by moving the nonanticipativity constraints into the objective function.

Using the notation introduced in Section II, the nonanticipativity constraints for period \( t \) can be written as

\[
s'_t = s'^t
\]

\( (11) \)

for all scenarios \( s \) and \( s' \) that are identical up to period \( t \).

Consider now one such pair of scenarios, \( s \) and \( s' \). The corresponding constraint (11) can be handled by removing it and adding to the objective function the term

\[
w^s \left( (\pi_t)^s (s^t - s') + \frac{p_{ss}}{2} \| s^t - s' \|^2 \right),
\]

\( (12) \)

where \( \pi_t \) is the vector of multipliers for the constraint (11) and \( p > 0 \) is a penalty parameter.

A rough algorithm to solve the modified problem is as follows:

Algorithm A-1

1. For a given multiplier vector \( \pi^t \) available at iteration \( t \), solve the modified problem. Let \( x^t \) be the solution of this problem.

2. If for some tolerance parameter \( \epsilon > 0 \),

\[
\| (s^t) - (s') \| \leq \epsilon \quad \forall s, s', t,
\]

holds, then stop; the optimal solution for the original problem (9), has been found.

3. Otherwise, reduce by an adequate amount the penalty parameter \( p \), and update the dual multipliers \( \pi \) according to

\[
\pi_{t+1} = \pi_t - \beta(s^t - s') \quad \forall s,
\]

\( (13) \)

where \( \beta \geq 0 \) is an appropriate steplength.

A potential disadvantage of the Augmented Lagrangian approach presented above is that the additional terms introduced in (12) are not linear but quasi-separable quadratic.

If the quadratic terms of the form \( \| s^t - s' \|^2 \) in (12) are expanded, and the cross-product terms \( (s^t)^s s'^t \) are approximated using a suggestion from [14], it results that for the variables \( s^t \) and \( s'^t \), and some particular values of these variables \( s^t \) and \( s'^t \),
such that \(\|s_t^j - \tilde{s}_t^j\|\) and \(\|s_t^i - \tilde{s}_t^i\|\) are small, we can approximate the quadratic terms by

\[
\|s_t^j - \tilde{s}_t^j\|^2 \approx\]

\[
\|s_t^i\|^2 + \|s_t^i\|^2 + 2(s_t^j)^T s_t^i - 2(s_t^j)^T \tilde{s}_t^i - 2(s_t^i)^T \tilde{s}_t^i.
\]

The submodels obtained from this procedure are separable QP networks—see [15], [16], [17]. Our rough algorithm for solving the extended problem via separable quadratic approximations is as follows:

1. Let \(i\) be a given iteration of algorithm A-1 and \(\pi_i\) and \(\tilde{\pi}_i\) be the dual multipliers and current point, respectively.

Algorithm A-2

1. Let \(\pi = \pi_i\), \(\tilde{\pi}_m = \tilde{\pi}_i\) and \(m = 1\).
2. Solve the extended problem with \(\tilde{x} = \tilde{\pi}_m\), to obtain a new point \(x_{m+1}\).
3. If \(\|x_{m+1} - \tilde{x}\| \leq \epsilon\), then stop. Otherwise, set

\[
x_{m+1} = x_{m+1} + \gamma(x_{m+1} - \tilde{x}_m),
\]

where \(\gamma\) is an appropriate steplength, increase \(m\) by one and go to step 2.

V. Implementation results.

In this section we present the results obtained from the implementation of the algorithm described in the preceding Section.

A relevant property of the decomposition approach is that the algorithm described above must solve a problem that can be trivially decomposed into the solution of a collection of independent subproblems. In this case, a reasonable procedure to improve the efficiency of the algorithm is to solve these subproblems in parallel. Along this line, we are working on the development of a C code that implements the Augmented Lagrangian algorithm A-1 and A-2, to run on a distributed computation environment.

To solve the quadratic networks in the subproblems, we have developed a C code, QPNET, based on the preconditioned variable reduction truncated Newton approach described in [15] and [16].

The computational experiments were conducted on a network of three HP 9000/715-50 workstations running HP-UX 9.01. These workstations are rated at 13 Mflops. The codes were written in C using the HP-UX “cc” compiler, and the parallel code was prepared using PVM version 3.1.3.

The problems in the set are based on the reservoir network of Iberdrola, a spanish power utility, on the river Tajo; this network is composed of 10 reservoirs. Another subset (larger) problems has been obtained from an artificial network with 35 reservoirs.

The generation functions for each reservoir, and the bounds on all the variables of the network formulation have been taken from the data provided by the utility (currently, the functions included in the tests are only linear functions). On the other hand, the exogenous water inflows and outflows have been randomly generated.

The problems differ in their time horizons and number of scenarios associated to each problem. The naming conventions used in Tables 1 and 2 are as follows: the notation “redx-yy-zz” indicates a problem having a basic network with “xx” reservoirs and “yy” links between reservoirs, that is solved for a time horizon of “zz” time periods. The number of scenarios and the characteristics of the resulting problems are shown in Table 1. Table 2 shows the computational results for the algorithm in its sequential and distributed computation versions; the distributed computation results have been obtained using a straightforward static load balancing scheme (this part of the code is still under development) that minimizes the communication costs, but does not achieve an optimal load balance between processors. The times in this table are shown in the format hh:mm:ss.

The influence of the improvement in the running times obtained
<table>
<thead>
<tr>
<th>Problem</th>
<th>SLD Iter.</th>
<th>Time</th>
<th>PLD Iter.</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>red10-9-6</td>
<td>910</td>
<td>2:42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>red10-9-7</td>
<td>726</td>
<td>5:42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>red10-9-14</td>
<td>1430</td>
<td>8:49:23</td>
<td>4:54:38</td>
<td></td>
</tr>
<tr>
<td>red10-9-30</td>
<td>322</td>
<td>24:08</td>
<td>12:17</td>
<td></td>
</tr>
<tr>
<td>red10-9-50</td>
<td>337</td>
<td>45:08</td>
<td></td>
<td>22:34</td>
</tr>
<tr>
<td>red10-10-14</td>
<td>254</td>
<td>1:33:16</td>
<td>45:48</td>
<td></td>
</tr>
<tr>
<td>red35-52-12</td>
<td>610</td>
<td>4:31:57</td>
<td>2:32:32</td>
<td></td>
</tr>
<tr>
<td>red35-60-6</td>
<td>407</td>
<td>5:47</td>
<td></td>
<td>2:34</td>
</tr>
<tr>
<td>red35-60-7</td>
<td>319</td>
<td>11:21</td>
<td></td>
<td>5:18</td>
</tr>
<tr>
<td>red35-60-10</td>
<td>463</td>
<td>47:39</td>
<td></td>
<td>26:18</td>
</tr>
<tr>
<td>red35-60-12</td>
<td>488</td>
<td>2:51:24</td>
<td>1:31:31</td>
<td></td>
</tr>
</tbody>
</table>

Legend: SLD : Sequential code
        PLD : Parallel code
        Iter : Number of iterations

when executing the parallel version of the code is reasonably close to the number of workstations used (3) for all problems, given the current state of code development. These results confirm the promising properties of the Augmented Lagrangian decomposition algorithm in a parallel computation environment.

VI. Conclusions

In this paper we have presented a formulation of a general hydropower generation management problem as an instance of network flow problems with uncertainty. We have proposed a methodology based on “scenario analysis” to deal with uncertainty. It results in a huge model with network-like constraints and complicating constraints. We have introduced an Augmented Lagrangian approach based on the dualization of the complicating constraints in the splitting variable formulation for the solution of the problem, giving rise to a model that can be decomposed into separable quadratic network submodels. Furthermore, the decomposition algorithm lends itself easily to being implemented in a parallel computation environment, with reductions in time very close to the optimal values.

These results are part of an ongoing research effort, in which our current lines of work include improving certain aspects of the Augmented Lagrangian algorithm, with particular emphasis in the study of more efficient methods for the update of the different parameters in the algorithm, such as the penalty parameter, left unmodified at the moment, the procedure indicated in (13) to update the multipliers, or the procedure in step 2 of Algorithm A-2 to update the reference values \( \hat{x} \). We are also continuing work on more efficient load balancing schemes for the distributed version of the code, and on extensions to problems with nonlinear and non-convex objective functions.

VII. References


VIII. Biographies

Laureano F. Escudero received his Ph.D. degree in Operations Research from Deusto University, Spain in 1974. He has worked at several IBM research centers and is a member of the IBM Academy of Technology. Currently, he is Director of Technology at UTESA, a consulting company of the Iberdrola group, and a Professor in the Statistics and Operations Research Department of the Universidad Complutense, Madrid, Spain. His research interests include strategic planning, logistic and investment analysis, and investment selection and control, with a special interest on their applications to the energy sector.

Jose Luis de la Fuente received his Ph.D. in Electrical Engineering from the Polytechnic University of Madrid, Spain in 1982. He is head of the Planning and Cooperation Area in the Research and Development Department at Iberdrola, and an Associate Professor at the Polytechnic University, Madrid, Spain. His current research interests are linear and nonlinear optimization for large-scale problems, decision theory and simulation of large-scale systems for investment planning and control, specifically regarding their application to problems in the energy and electricity generation sectors.

Cristina Garcia received her B.S. degree in Electrical Engineering from the Polytechnic University, Madrid, Spain in 1991. She currently works for UTESA, a consulting company in the Iberdrola group. Her research interests are linear and nonlinear optimization for large-scale problems, parallel computation and simulation.

Francisco J. Prieto received a Ph.D. degree in Operations Research from Stanford University in 1989. He is currently an Assistant Professor in the Department of Statistics and Econometrics of University Carlos III, Madrid, Spain. His research interests include large-scale optimization, planning under uncertainty and robust multivariate estimation.