Long-term electric power planning in a competitive market using the Bloom and Gallant procedure and a modeling language

N. Nabona and A. Pagès
Dept. of Statistics and Operations Research, UPC

DR 2002/24
27–12–2002

Postcript copies of this report may be obtained via anonymous *ftp* at `ftp-eio.upc.es` in directory `pub/onl/reports`. It is the file named `dr200324.ps.Z`. It needs decompressing before printing (e.g. via `uncompress dr200324.ps.Z` to get a postscript file).

corresponding author: Narcís Nabona
Dept. Statistics and Operations Research,
UPC Campus Sud, edifici U,
c. Pau Gargallo 5, 08028 Barcelona,
tel. +34 934017035 fax +34 934015855
email:narcis.nabona@upc.es

Long-term electric power planning in a competitive market using the Bloom and Gallant procedure and a modeling language.[1]

Narcís Nabona and Adela Pagès

Abstract: *Bloom and Gallant have proposed an elegant model for finding the optimal thermal schedule subject to matching the load-duration curve and general linear constraints. Their method is based on a linear program with some linear equality constraints and many linear inequality constraints. There have been applications of this procedure to multi-interval problems using the active set method and the Dantzig-Wolfe column generation method, and through the direct application of linear optimization packages using an available modeling language. This work describes the long-term electric power planning model adapted to the Bloom and Gallant procedure. It presents a new quadratic formulation of the maximum profit problem, and solves a number of long-term power planning problems of various sizes using a modeling language and available linear and quadratic programming solvers. Several remarks are made regarding its implementation.*
Keywords: *Linear Programming, Active set optimization methods, Dantzig-Wolfe column generation, Long-term power generation scheduling, Stochastic processes, Thermal power generation.*

1 Introduction and Motivation

Long-term generation planning is a key issue in the operation of an electricity generation company. Its results are used both for budgeting and planning fuel acquisitions and to provide a framework for short-term generation planning.

The long-term problem is a well-known stochastic optimisation problem because several of its parameters are only known as probability distributions (for example: load, availability of thermal units, hydrogeneration and generations from renewable sources in general).

A long-term planning *period* (e.g., a natural year) is normally subdivided into shorter *intervals* (e.g., a week or a month), for which parameters (e.g., the load-duration curve) are known or predicted, and optimized variables (e.g., the expected energy productions of each generating unit) must be found.

Predicted load-duration curves (LDC's) — equivalent to cumulative probability load distributions — for each interval are used as data for the problem, which is appropriate since load uncertainty can be suitably described through the LDC. The probability of failure for each thermal unit is assumed to be known.

Bloom and Gallant [3] proposed a linear model (with an exponential number of inequality constraints) and used an *active set* methodology [10] to find the optimal way of matching the LDC of a single interval with thermal units only, when there are load-matching and other operational non-load-matching constraints. These could be, for example, limits on the availability of certain fuels, or environmental maximum emission limits. The optimal *loading order* obtained with Bloom and Gallant's method may include permutations with respect to the *merit order* and *splittings* in the loading of units [3, 8]. In this way the energies generated satisfy the limitations imposed by the non-load-matching constraints while having the best possible placement, with respect to generation cost, in the matching of the LDC. (Changes in loading order bring about discrete changes in energy generation, while splittings cause a continuous variation in the energy generated depending on where the split starts [3].)

When the long-term planning power problem is to be solved for a generation company operating in a competitive market, the company has not a load of its own to satisfy, but it bids the energies of its units to a *market operator*, who selects the lowest-price among biding companies to match the load. In this case, the scope of the problem is no longer that of the generation units of a single generation company but that of all units of all companies biding in the same competitive market, matching the load of the whole system. This makes planning problems much larger than before and is a reason for developing more efficient codes to solve them.

[1] Partially supported by CICYT Grant TAP99-1975-C02-01 of *Ministerio de Educación y Cultura*, Spain

The Bloom and Gallant model has been successfully extended to multi-interval long-term planning problems using either the active-set method [10], or the Dantzig-Wolfe column-generation method [14, 11].

In this work, the multi-interval Bloom and Gallant model has been coded using the modeling language *AMPL* [6] and has been solved with a linear programming package *Cplex 7.5* [4] as carried out in [12] for a single interval. A quadratic model is put forward here to formulate the long term profit maximization of generation companies in a competitive electricity market, and the *AMPL* model developed has been solved with quadratic optimization packages. This last model is refined by introducing an external nonbidding energy source.

A companion report by the same authors [13] compares the performance of the active-set and the Dantzig-Wolfe solution procedures applied to the solution of minimum cost long-term power planning problems, which are linear. The smaller test problems solved in that report are also dealt with here, so a comparison of results and efficiency with *AMPL* plus *Cplex 7.5* is possible.

2 The load-duration curve

The LDC is the most sensible way to represent the load of a future interval in an integral way (the load depends on random factors such as weather in several geographical areas, human decisions, social events, etc.). The main features of an LDC (corresponding to the $i^{\underline{\text{th}}}$ interval) can be described through 5 characteristics:

- ⋆ the duration T^i
- ⋆ the peak load power $\widehat{P}^i$
- ⋆ the base load power $\underline{P}^i$
- ⋆ the total energy $\widehat{E}^i$
- ⋆ the shape, which is not a single parameter and is usually described through a table of durations and powers, or through a function.

The LDC for future intervals must be predicted. For a past interval, for which the hourly load record is available, the LDC is equivalent to the load over time curve sorted in order of decreasing power (see Fig. 1).

It should be noted that in a *predicted* LDC, random events such as weather, shifts in consumption timing, etc., that cause modifications of different signs in the load tend to cancel out, and that the LDC keeps all the power variability of the load.

2.1 Power and energy constraints imposed by the load-duration curve

The loading of thermal units in an LDC was first formulated in [1] and practical procedures to compute the covering can be found in [16].

Analytically, given the probability density function of load $p(x)$, the cumulative load distribution function $L_0(x)$ is calculated as follows:

$$L_0(x) = 1 - \int_0^x p(y)\,dy$$

3 Thermal Units

As far as loading an LDC is concerned, the relevant parameters of a thermal unit are:

- ⋆ *power capacity*: (C_j for the $j^{\underline{\text{th}}}$ unit) maximum power output (MW) that the unit can generate

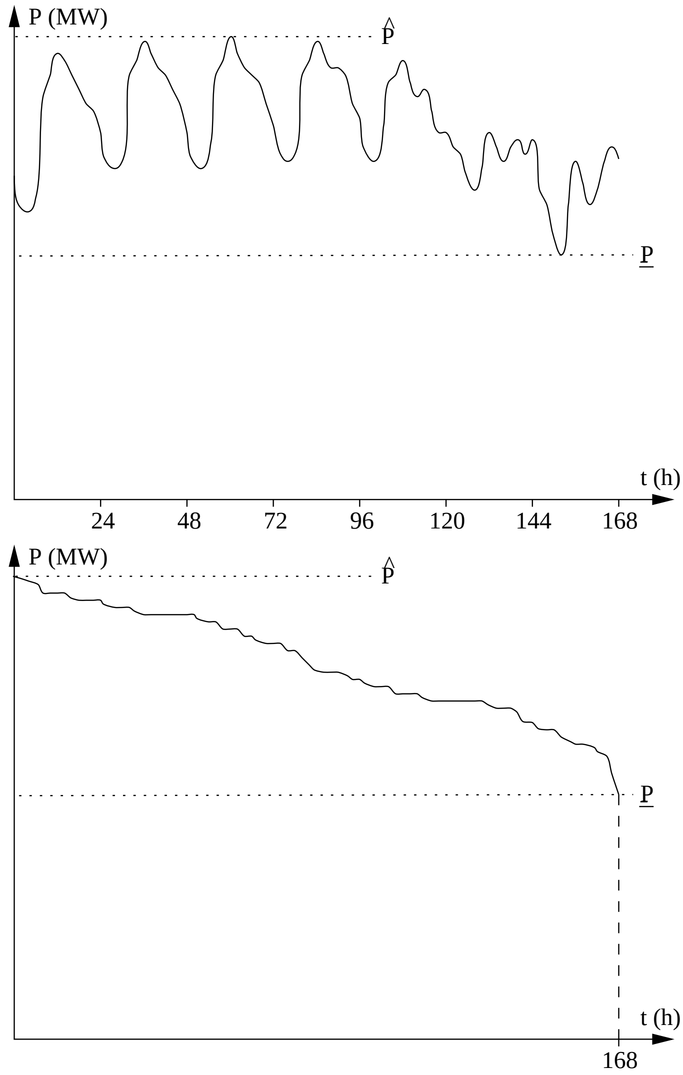

Figure 1: Load over time, above, and load-duration curve (LDC), below . (Data for a week — Monday to Sunday — in January).

⋆ *outage probability*: (q_j for the $j^{\underline{\text{th}}}$ unit) probability of a unit not being available when it is required to generate

⋆ *linear generation cost*: ($\widetilde{f_j}$ for the $j^{\underline{\text{th}}}$ unit) production cost in €/MWh

Other associated concepts are:

⋆ *merit order*: units are ordered according to their efficiency in generating electric power (€/MWh)

- all units will work at their maximum capacity since no unit should start to generate until the previous unit in the merit order is generating at its maximum capacity (because the price of the MWh it produces is lower!)

⋆ *loading order*: units will have load allocated to them in a given order

- loading order and merit order should coincide, but when there are other constraints to be satisfied, the most economical loading order may be different from the merit order.

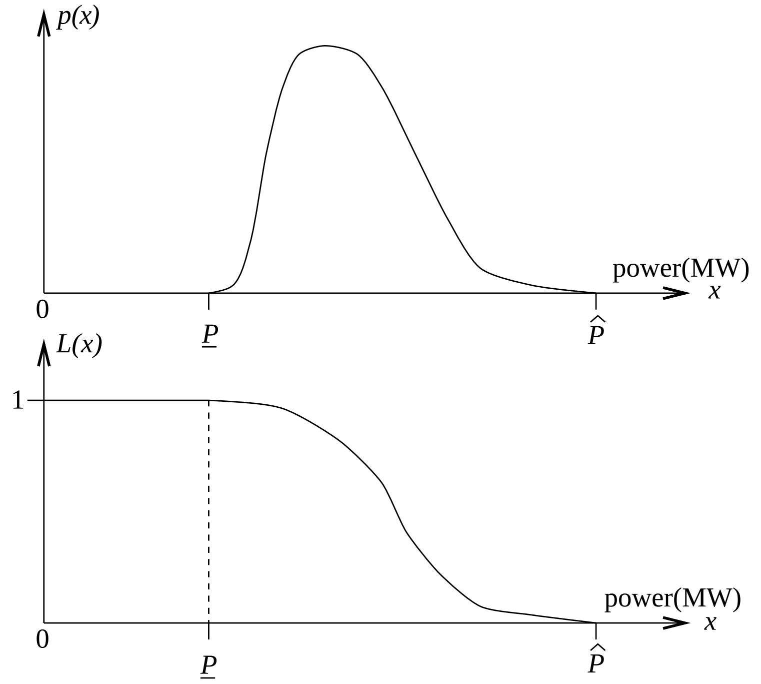

Figure 2: Probability density function of load $p(x)$ (above), and cumulative load distribution function $L_0(x)$ (below).

4 Matching the load-duration curve

Due to the outages of thermal units (whose probability is >0), the LDC does not coincide with the estimated production of thermal units. It is usual for the installed capacity to be higher than the peak load: $\sum_{j=1}^{N_u} C_j > \widehat{P}$, and it is normal to find that $\sum_{j=1}^{N_u} C_j \approx 1.4\widehat{P}$.

The *generation-duration curve* is the expected production of the thermal units over the time interval to which the LDC refers.

The energy generated by each unit is the slice of area under the generation-duration curve which corresponds to the capacity of the thermal unit.

The probability that there are time lapses within the time interval under consideration, where, due to outages, there is not enough generation capacity to cover the current load, is not null. Therefore, *external energy* (from other interconnected utilities) will have to be imported and paid for at a higher price than the most expensive unit in ownership.

The area under the LDC and the area under the generation-duration curve must coincide. (See in Fig. 3 the generation-duration curve corresponding to a given LDC.)

The peak power of the generation-duration curve is $\sum_{j=1}^{N_u} C_j + \widehat{P}$ and the area above power $\sum_{j=1}^{N_u} C_j$ is the external energy. In order to find the generation-duration curve from the LDC, the convolution method of Balériaux, Jamoulle & Linard de Guertechin [1] was implemented.

4.1 Convolution method of finding the generation-duration curve

The method calculates the production of each thermal unit, given a loading order. The load is modeled through its distribution (see Fig. 2):

$$L_0(x) = \begin{cases} 1, & \text{for } x \le \underline{P} \\ r \in [0,1], & \text{for } \underline{P} < x \le \widehat{P} \\ 0, & \text{for } x > \widehat{P}. \end{cases}$$

(Recall that $L_0(x)$ is the probability of requiring x MW, or more).

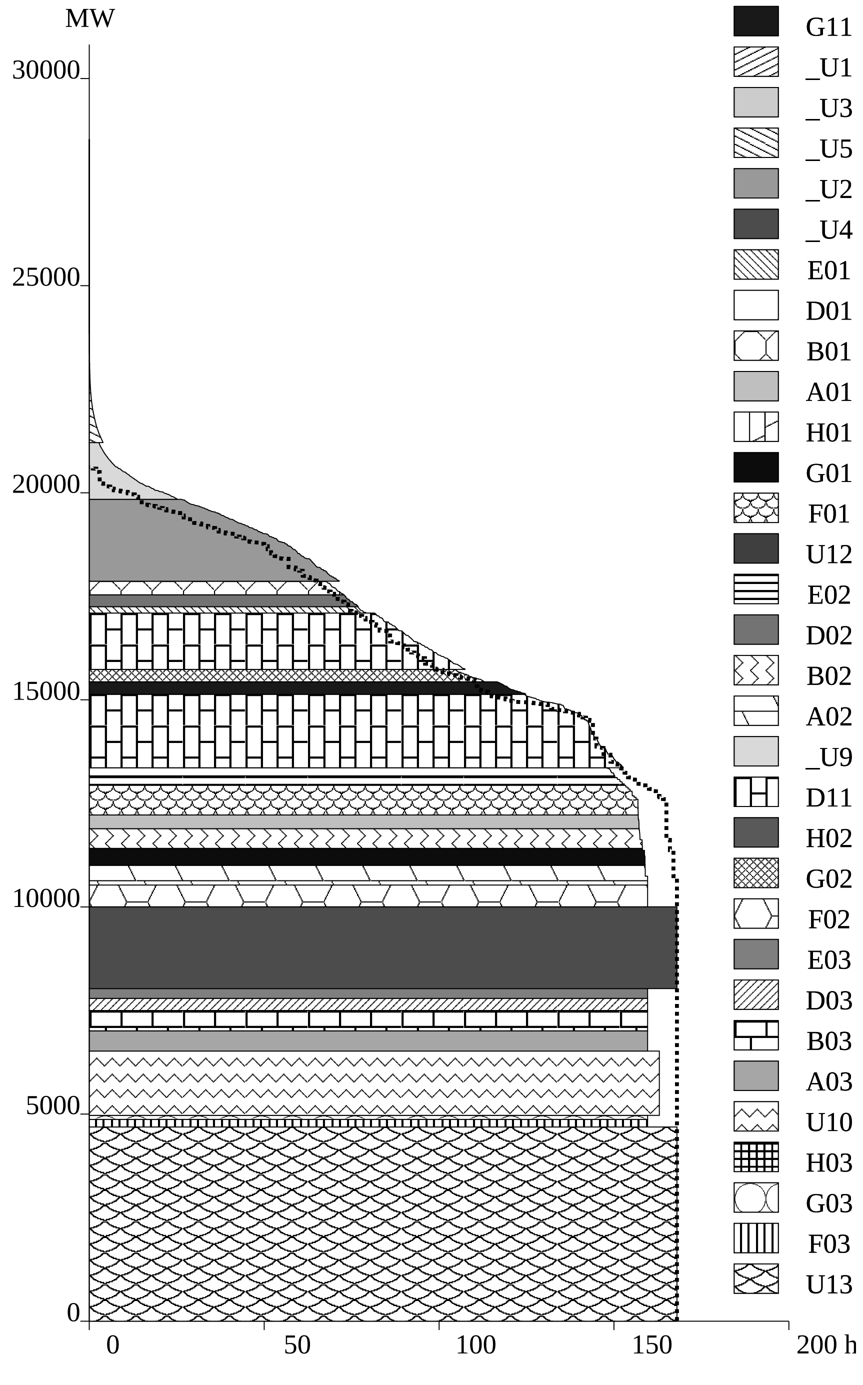

Figure 3: Generation-duration curve and load-duration curve (dotted line) for a weekly interval in a 32 unit problem.

4.2 Convolution of the $j^{\underline{\text{th}}}$ thermal unit

Let:

C_j	:	maximum power capacity in MW of unit j
q_j	:	outage probability of unit j
$1 - q_j$	:	in-service probability of unit j
U_j	:	set of unit indices $1, 2, \ldots, j$
$L_{U_{j-1}}(x)$	:	probability distribution of uncovered load after loading units $1, 2, ..., j-1$
$L_{U_j}(x)$	:	probability distribution of uncovered load after loading units $1, 2, ..., j-1, j$
x	:	load in MW

the convolution computes $L_{U_j}(x)$ from $L_{U_{j-1}}(x)$ as [1, 16]:

$$L_{U_j}(x) = q_j \, L_{U_{j-1}}(x) + (1 - q_j) \, L_{U_{j-1}}(x + C_j) \tag{1}$$

Recalling that $E{=}P{\cdot}T$, the energy generated by unit j is [1]:

$$E_j = (1 - q_j) \, T \int_0^{C_j} L_{U_{j-1}}(x) \, dx \; . \tag{2}$$

4.3 Unsupplied load after a set of thermal units is loaded

Let $L_0(x)$ be the cumulative probability distribution of the power load corresponding to the LDC, where x represents load. Once unit i (capacity C_i, and outage probability q_i) has been loaded, the cumulative power distribution of the unsupplied load will be

$$L_{\{i\}}(x) = q_i L_0(x) + (1 - q_i) L_0(x + C_i) \tag{3}$$

If we then load thermal unit j $(C_j,\ q_j)$ we are left with $L_{U_j}(x)$

$$L_{\{j,i\}}(x) = q_j L_{\{i\}}(x) + (1 - q_j) L_{\{i\}}(x + C_j)$$

which, taking into account (3), can be rewritten as

$$\begin{aligned} L_{\{j,i\}}(x) = \quad & q_i q_j L_0(x) + (1 - q_i) q_j L_0(x + C_i) + (1 - q_j) q_i L_0(x + C_j) \\ & + (1 - q_i)(1 - q_j) L_0(x + C_i + C_j) \end{aligned} \tag{4}$$

Note in (4) that, should we have loaded units i and j in the reverse order, $L_{\{i,j\}}(x)$ (having loaded unit i first and then unit j) would have been the same as $L_{\{j,i\}}j(x)$ (having loaded unit j first and unit i second).

It is not difficult to derive that, given a set of units whose indices 1,2, etc. are the elements of the set of indices Ω, the unsupplied load after loading all the units in Ω will have a cumulative probability distribution $L_\Omega(x)$

$$\begin{gathered} L_\Omega(x) = L_0(x) \prod_{m \in \Omega} q_m + \\ \sum_{U \subseteq \Omega} \Big(L_0(x + \sum_{i \in U} C_i)(1 - q_i) \prod_{i \in U} (1 - q_i) \prod_{i \in U} q_i\Big) \end{gathered}$$

We can thus say that the cumulative probability distribution $L_\Omega(x)$ of the unsupplied load is the same no matter the order in which the units in Ω have been loaded.

4.4 Computation of the unsupplied energy

The unsupplied energy $W(\Omega)$ is computed as:

$$W(\Omega) = \; T \int_0^{\widehat{P}} L_\Omega(x) \, dx \tag{5}$$

$L_\Omega(x)$ being the probability distribution of the unsupplied load after loading (in any order) all units from the set Ω.

The integration in (5) is to be carried out numerically.

4.5 Loading order and energy generation bounds of a given unit

Let Ω be a set of unit indices. It has been shown above that the distribution $L_\Omega(x)$ is independent of the order in which the units have been loaded. However, the contribution of each unit to matching the load will vary according to its position in the loading order, and the lower the order number the bigger the generation will be for the given unit. This is so because of (2), since $L_{U_{j+1}}(x) \le L_{U_j}(x)\ \forall x$.

A given unit of index $k \in \Omega$ will have its generation bounded by:

$$0 < (1-q_k)\, T \int_0^{C_k} L_{\Omega\backslash k}(x)\, dx = \underline{E}_k \le E_k \le \overline{E}_k = (1-q_k)\, T \int_0^{C_k} L_0(x)\, dx \qquad (6)$$

where $L_{\Omega\backslash k}(x)$ corresponds to the probability distribution of uncovered load after loading all units in Ω except that of index k.

$\underline{E}_k$ and $\overline{E}_k$ correspond to loading unit k the last and the first respectively.

5 Bloom & Gallant's model for matching the load-duration curve when there are non-load-matching constraints

Let the Bloom & Gallant formulation (for a single interval) [3] be given by:

$$\begin{aligned}
&\underset{E_j}{\text{minimize}} && \sum_{j=1}^{N_u+1} \widetilde{f}_j\, E_j && (7)\\
&\text{subject to} && \sum_{j\in U} E_j \le \widehat{E} - W(U) \quad \forall\, U \subset \Omega = \{1,\dots,N_u\} && (8)\\
& && A_{\ge}\, E_j \ge R_{\ge} && (9)\\
& && \sum_{j=1}^{N_u+1} E_j = \widehat{E} && (10)\\
& && E_j \ge 0 \qquad j = 1,\dots,N_u, N_u+1 && (11)
\end{aligned}$$

where:

$N_u{+}1$	is the index representing the external energy
$N_{\ge}$	is the total number of non-load-matching inequality constraints
$A_{\ge} \in \mathbb{R}^{N_{\ge}\times N_u}$	is the matrix of coefficients of non-load-matching inequality constraints
$R_{\ge}$	*rhs* of non-load-matching inequality constraints
U	subset of Ω
$W(U)$	unsupplied energy after loading all units $j \in U \subset \Omega$

The objective function (7) can be simplified using (10), which leads to:

$$\sum_{j=1}^{N_u} f_j\, E_j + \widetilde{f}_{N_u+1}\widehat{E} \qquad \text{where} \quad f_j = \widetilde{f}_j - \widetilde{f}_{N_u+1}$$

Given that $\widetilde{f}_{N_u+1}\widehat{E}$ is a constant, problem (7-11) can be recast as:

$$\begin{aligned}
&\underset{E_j}{\text{minimize}} && \sum_{j=1}^{N_u} f_j\, E_j \\
&\text{subject to} && \sum_{jinU} E_j \le \widehat{E} - W(U) \quad \forall\, U \subset \Omega = \{1,\dots,N_u\} && (12)\\
& && A_{\ge}\, E_j \ge R_{\ge} \\
& && E_j \ge 0 \qquad j = 1,\dots,N_u\,.
\end{aligned}$$

5.1 The case where no constraint (9) is active

Constraints (9) are the non-load-matching constraints. The Appendix of [8] contains a proof that the merit-order loading energies correspond to a minimum of the formulation (7–11) when there are no active constraints (9).

Assuming that units are ordered in order of merit, the active constraints at the minimizer of the set of inequalities (8) would be:

$$\begin{aligned} E_1 &= \quad \widehat{E} - W(1) \\ E_1 + E_2 &= \quad \widehat{E} - W(1,2) \\ E_1 + E_2 + E_3 &= \quad \widehat{E} - W(1,2,3) \\ &\cdots \\ E_1 + E_2 + E_3 + \cdots + E_{N_u} &= \quad \widehat{E} - W(1,2,\ldots,N_u) \end{aligned} \tag{13}$$

Subtracting the first from the second equality we get:

$$E_2 = W(1) - W(1,2) > 0$$

Subtracting the second from the third we get:

$$E_3 = W(1,2) - W(1,2,3) > 0\,.$$

Finally, subtracting the penultimate from the last we would have:

$$E_{N_u} = W(1,2,\ldots,N_u-1) - W(1,2,\ldots,N_u) > 0$$

It must be stressed that all E_i are $E_i{>}0$ because $W(1,2,\ldots,i-1){>}W(1,2,\ldots,i)$, and that however small $W(1,2,\ldots,i-1){-}W(1,2,\ldots,i)$ may be, it will never be zero, i.e., no nonnegativity bound (11) will be active. Moreover, the energies of all units will be within its bounds (6).

The energy generation E_j will coincide with that calculated in (2):

$$\begin{aligned} E_j &= W(1,...,j-1) - W(1,...,j-1,j) = T\int_0^{\widehat{P}} L_{U_{j-1}}(x)\,\mathrm{d}x - T\int_0^{\widehat{P}} L_{U_j}(x)\,\mathrm{d}x \\ &= T\int_0^{\widehat{P}} \left(L_{U_{j-1}}(x) - L_{U_j}(x)\right)\mathrm{d}x = T\int_0^{\widehat{P}} \left(L_{U_{j-1}}(x) - q_j L_{U_{j-1}}(x) - (1-q_j)L_{U_{j-1}}(x+C_j)\right)\mathrm{d}x \\ &= T(1-q_j)\int_0^{\widehat{P}} \left(L_{U_{j-1}}(x) - L_{U_{j-1}}(x+C_j)\right)\mathrm{d}x = T(1-q_j)\int_0^{C_j} L_{U_{j-1}}(x)\,\mathrm{d}x \end{aligned} \tag{14}$$

5.2 The case where a constraint (9) or nonnegativity bound (11) is active

In this case, at least one of the constraints in (9) or nonnegativity bound (11) will be active, which means that at least one of the active constraints in (13) will not be satisfied as an equality.

Let us assume that j, k, and l are three consecutive units in loading order (which may be different from the merit order), and that the k^{th} equation of system (13) is not active, while the first equations up to the $j^{\underline{\text{th}}}$ are, as are the equations from the $l^{\underline{\text{th}}}$ onwards. The values of the energies up to E_j can be obtained by subtracting the preceding equqation from the next, as done in 5.1.

By subtracting the active $j^{\underline{\text{th}}}$ equation from the $l^{\underline{\text{th}}}$ we get:

$$E_k + E_l = W(1,...,j) - W(1,...,j,k,l) \tag{15}$$

The actual value of E_k and E_l will be obtained as part of the solution of (7–11) and will satisfy (15) and the rest of the active constraints, including those of (9) and the nonnegativity bounds (11).

As noted in [3], as regards energies E_k and E_l, the solution can be viewed as a *splitting* of one unit by another.

5.3 The multi-interval Bloom and Gallant model

As power planning for a long time period cannot take into account changes over time of some parameters, the time period is subdivided into shorter *intervals* in which all parameters can be assumed to be constant. We will use superscript ^{i} to indicate that variables and parameters refer to the $i^{\underline{\text{th}}}$ interval.

Therefore some constraints refer only to variables of a single interval, while others may refer to variables in several intervals. E.g., constraints on the minimum consumption of gas may affect several or all the intervals, while emission limit constraints, or the constraint associated with the units composing a combined-cycle unit refer to each single interval.

The overhauling of thermal units must be taken into account. Therefore, there will be intervals where some units must remain idle. The set of available units in each interval may be different. Let Ω^i be the set of available units in the $i^{\underline{\text{th}}}$ interval, and let N_u^i be $N_u^i = |\Omega^i|$ (the cardinality of this set).

The Bloom and Gallant linear optimization model extended to N_i intervals, with inequality and equality non-load-matching constraints, can thus be expressed as:

$$\underset{Ej^i}{\text{minimize}} \quad \sum_{i=1}^{Ni} \sum_{j=1}^{N_u} f_j E_j^i \tag{16}$$

$$\text{subject to:} \quad \sum_{jinU} E_j^i \le \widehat{E}^i - W^i(U) \quad \forall U \subset \Omega^i \quad i = 1, \dots, N_i \tag{17}$$

$$A_{\geq}^i \, E^i \ge R_{\geq}^i \qquad i = 1, \dots, N_i \tag{18}$$

$$\sum_i A_{\geq}^{0i} \, E^i \ge R_{\geq}^0 \tag{19}$$

$$A_{=}^i \, E^i = R_{=}^i \qquad i = 1, \dots, N_i \tag{20}$$

$$\sum_i A_{=}^{0i} \, E^i = R_{=}^0 \tag{21}$$

$$E_j^i \ge \underline{0} \qquad j = 1, \dots, N_u, \quad i = 1, \dots, N_i \tag{22}$$

where:

$A_{\geq}^i \in \mathbb{R}^{N_{\geq}^i \times N_u}$ is the matrix of coefficients of inequalities that refer only to energies of $i^{\underline{\text{th}}}$ interval,

$A_{\geq}^{0i} \in \mathbb{R}^{N_{\geq}^0 \times N_u}$ is the matrix of coefficients of inequalities that refer to energies of more than one interval related to energies of $i^{\underline{\text{th}}}$ interval,

$R_{\geq}^i \in \mathbb{R}^{N_{\geq}^i}$ is the right-hand sides of inequalities that refer only to energies of $i^{\underline{\text{th}}}$ interval,

$R_{\geq}^0 \in \mathbb{R}^{N_{\geq}^0}$ is the right-hand sides of inequalities that refer to energies of more than one interval,

$A_{=}^i \in \mathbb{R}^{N_{=}^i \times N_u}$ is the matrix of coefficients of equalities that refer only to energies of $i^{\underline{\text{th}}}$ interval,

$A_{=}^{0i} \in \mathbb{R}^{N_{=}^0 \times N_u}$ is the matrix of coefficients of equalities that refer to energies of more than one interval related to energies of $i^{\underline{\text{th}}}$ interval,

$R_{=}^i \in \mathbb{R}^{N_{=}^i}$ is the right-hand sides of equalities that refer only to energies of $i^{\underline{\text{th}}}$ interval,

$R_{=}^0 \in \mathbb{R}^{N_{=}^0}$ is the right-hand sides of equalities that refer to energies of more than one interval.

The number of variables is now $\sum_i^{N_i} N_u^i$ and there are $\sum_i^{N_i} (2^{N_u^i} - 1)$ load-matching constraints plus $N_= = N_=^0 + \sum_i N_=^i$ non-load-matching equalities, and $N_\geq = N_\geq^0 + \sum_i N_\geq^i$ non-load-matching inequalities. Note that supraindices 0 indicate constraints which affect variables of more than one interval.

Should constraint sets (19) and (21), which are the multi-interval constraints, be empty, the problem would be separable into N_i subproblems, one for each interval. Otherwise a joint solution must be found.

5.4 Approximate model of long-term hydrogeneration

The long term model described is appropriate for thermal generation units but not for hydrogeneration, which requires additional variables to represent the variability of water storage in reservoirs and discharges necessary for the calculation of the hydroenergy generated.

A coarse model of hydrogeneration, which does not consider any of the reservoir dynamics, can be employed. All or a part of the reservoir systems of one or several basins are considered as a single pseudo-thermal unit H with cost $\widetilde{f}_H$=0, outage probability q_H=0 and capacity C_H (normally lower than the maximum installed hydropower capacity), with a constraint binding the intervals' hydrogenerations over the successive intervals so that they add up to a total expected hydrogeneration R_H^0 for the whole period:

$$\sum_{i}^{N_i} E_H^i = R_H^0 \,, \tag{23}$$

which is a constraint of the type (21).

6 Full model of long-term hydrogeneration

6.1 Hydrogeneration of a reservoir

The hydroelectric power production of reservoir m in the $i^{\underline{\text{th}}}$ interval depends on the amount of water d_m^i discharged over the interval and the water *head*, which is the difference between the water level in the reservoir and the water level in the reservoir outlet channel (see Fig. 4). A polynomial relates the volume of water stored v_m in reservoir m, and the water head s_m:

$$s_m = \mathrm{s}_{\mathrm{b}m} + \mathrm{s}_{\mathrm{l}m} v_m + \mathrm{s}_{\mathrm{q}m} {v_m}^2 + s_{cj} {v_m}^3$$

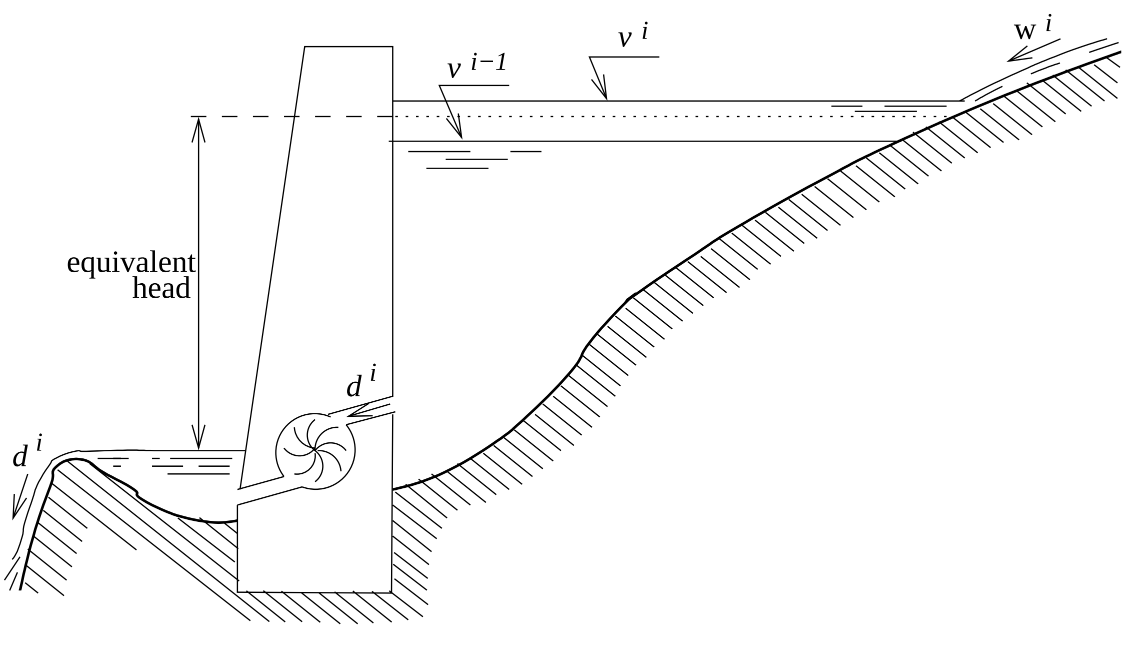

Figure 4: Cross-section of a reservoir showing water head and discharge.

where $\mathrm{s}_{\mathrm{b}m}$, $\mathrm{s}_{\mathrm{l}m}$, $\mathrm{s}_{\mathrm{q}m}$, and s_{cj} are the basic, linear, quadratic and cubic coefficients respectively, which are data to the problem.

The water head changes during the whole period considered, and also over the $i^{\underline{\text{th}}}$ interval. Let v_m^i represent the stored volume in reservoir m <u>at the end</u> of interval i, and v_m^{i-1} the stored volume at the end of the former interval, which is equivalent to the stored volume at the beginning of the $i^{\underline{\text{th}}}$ interval. The equivalent head $\widetilde{s}^i_m$ (the average of the interval i) will satisfy:

$$\widetilde{s}^i_m \times (v^i_m - v^{i-1}_m) = \int_{v^{i-1}_m}^{v^i_m} \left(\mathrm{s_{b}}_m + \mathrm{s_{l}}_m v_m + \mathrm{s_{q}}_m {v_m}^2 + s_{cj} {v_m}^3\right) \mathrm{d}v_m$$

Integrating, taking limits and simplifying we get:

$$\widetilde{s}^i_m = \mathrm{s_{b}}_m + \frac{\mathrm{s_{l}}_m}{2}(v^{i-1}_m + v^i_m) + \frac{\mathrm{s_{q}}_m}{3}(v^i_m - v^{i-1}_m)^2 + \mathrm{s_{q}}_m v^{i-1}_m v^i_m + \frac{\mathrm{s_{c}}_m}{4}({v^{i-1}_m}^2 + {v^i_m}^2)(v^{i-1}_m + v^i_m) \quad (24)$$

The difference in potential energy is $d^i_m \mathrm{g} \widetilde{s}^i_m$, the discharged water weight times the head. Given that water density is 1, the discharged water weight is $d^i_m \mathrm{g}$, g being the acceleration of gravity. Let ρ_m<1 be the efficiency of the turbine-alternator system and T^i the interval duration. Therefore, the average generated hydropower h^i_m over the $i^{\underline{\text{th}}}$ interval in reservoir m will be:

$$h^i_m = \frac{\rho_m \mathrm{g}}{T^i} d^i_m \widetilde{s}^i_m$$

Letting $\mathrm{k}_m = \rho_m \mathrm{g}$ and using (24) we have:

$$h^i_m = \frac{\mathrm{k}_m}{T^i} d^i_m \Big\{ \mathrm{s_{b}}_m + \tfrac{\mathrm{s_{l}}_m}{2}(v^{i-1}_m + v^i_m) + \tfrac{\mathrm{s_{q}}_m}{3}(v^i_m - v^{i-1}_m)^2 + \mathrm{s_{q}}_m v^{i-1}_m v^i_m + \tfrac{\mathrm{s_{c}}_m}{4}({v^{i-1}_m}^2 + {v^i_m}^2)(v^{i-1}_m + v^i_m) \Big\} \quad (25)$$

the term between braces being the equivalent head.

6.2 Temporary evolution of water in the reservoirs of a basin

The variables of the N_m reservoirs of one (or several) river basin(s) are related because the water outlet of a reservoir feeds the inflow of the downstream reservoir. A hydronetwork, such as that of Fig. 5, depicts the water flows in a river basin during an interval. The water balance in the node corresponding to reservoir m over the $i^{\underline{\text{th}}}$ interval would be:

$$v^{i-1}_m + \mathrm{w}^i_m + \sum_{j \in G_m} (d^i_j + p^i_j) = v^i_m + d^i_m + p^i_m \quad (26)$$

where w^i_m is the natural water inflow during the $i^{\underline{\text{th}}}$ interval, G_m is the set of indices of reservoirs up-stream of reservoir m (if any). Inflows w^i_m are a forecast for an immediate short interval (e.g., the current or next week), or a probability density function for a long, or future interval. p^i_m is the spillage or discharge not usable for generation.

The temporary evolution of water in reservoirs, over N_i successive intervals, can be modeled through a replicated hydronetwork [15] such as that in Fig. 6, where the initial volume in reservoir m during the $i^{\underline{\text{th}}}$ interval is the final volume of this reservoir in interval $i-1$. Initial and final water volumes in reservoirs, v^0_m and $\mathrm{v}^{N_i}_m$, are to be considered data for the problem. (Initial volumes are usually the current ones, and final volumes are the target values.)

All hydrovariables (stored volumes, discharges, and spillages) are flows in the replicated hydronetwork, and its node balances are expressed in equations such as (26). Note the existence in Fig. 6 of a sink node S (where the water outlet of one or more river basins flow), whose balance equation is redundant.

$$\mathrm{w_S} = -\left(\sum_{i=1}^{N_i} \sum_{m=1}^{N_m} \mathrm{w}^i_m + \sum_{m=1}^{N_m} \{\mathrm{v}^0_m - \mathrm{v}^{N_i}_m\} \right).$$

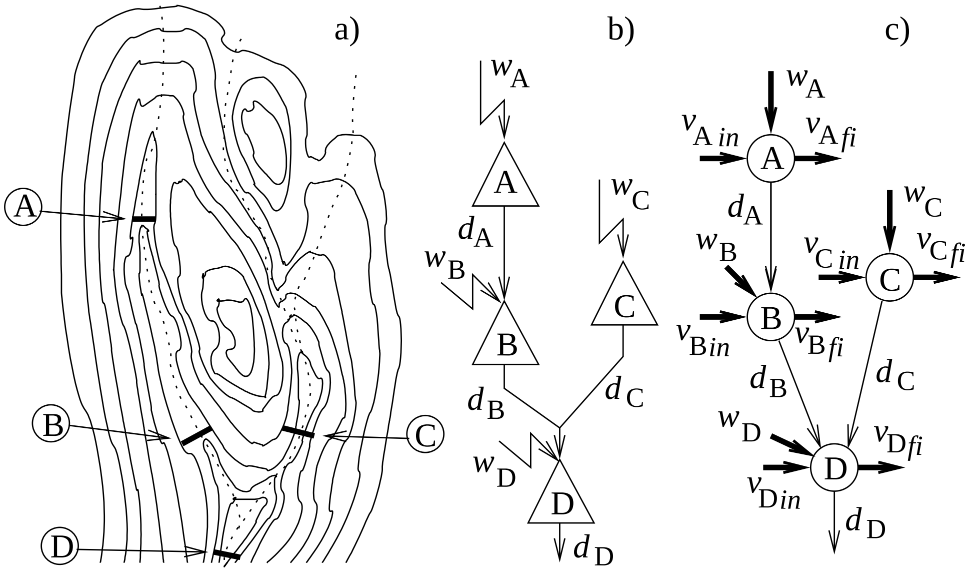

Figure 5: a) Reservoir system in a river basin. b) Hydro-connections of reservoirs. c) Reservoir system network in a time interval (showing water inflows).

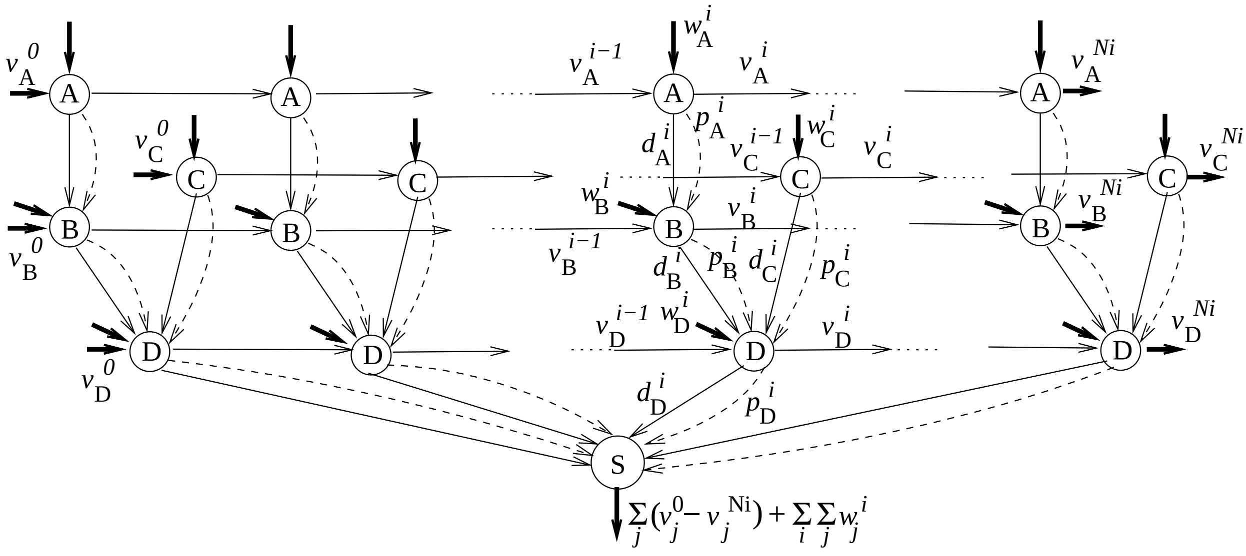

Figure 6: Replicated hydronetwork over N_i intervals. (The dashed arcs are spillages.)

6.3 Expected hydrogeneration and stochastic programming

Energies, including hydroenergies, that satisfy the load-matching constraints (17) are expected energies. Hydroenergies ($h_m^i T^i$) must also satisfy the hydropower equation (25), thus water volumes and discharges should also be expected values.

Stochastic programming [2] uses scenarios associated to probabilities, from which expected values can be calculated. In the case of hydrogeneration, water inflow scenarios ought to be employed [5], and, being a multi-stage problem (of N_i intervals), their number should be high enough in order to take into account the inflow variability. However, there are two principles of stochastic programming that do not appear to fit in well with long-term hydrogeneration optimization: the optimization for a particular scenario of long-term inflows may have no sense at all, because it could be infeasible for many other scenarios.

6.4 The multicommodity network flow model of long-term hydrogeneration as a two-stage stochastic programming problem

An alternative way to obtain long-term expected hydrogenerations is through a multicommodity network flow model, where, in a replicated hydronetwork such as that of Fig. 6, N_a commodities, corresponding to water inflows of different probability of occurrence, flow [7]; N_a is normally $2 \le N_a \le 6$. This leads to considering inflow probability density functions approximated by a block probability density function, as in Fig 7.

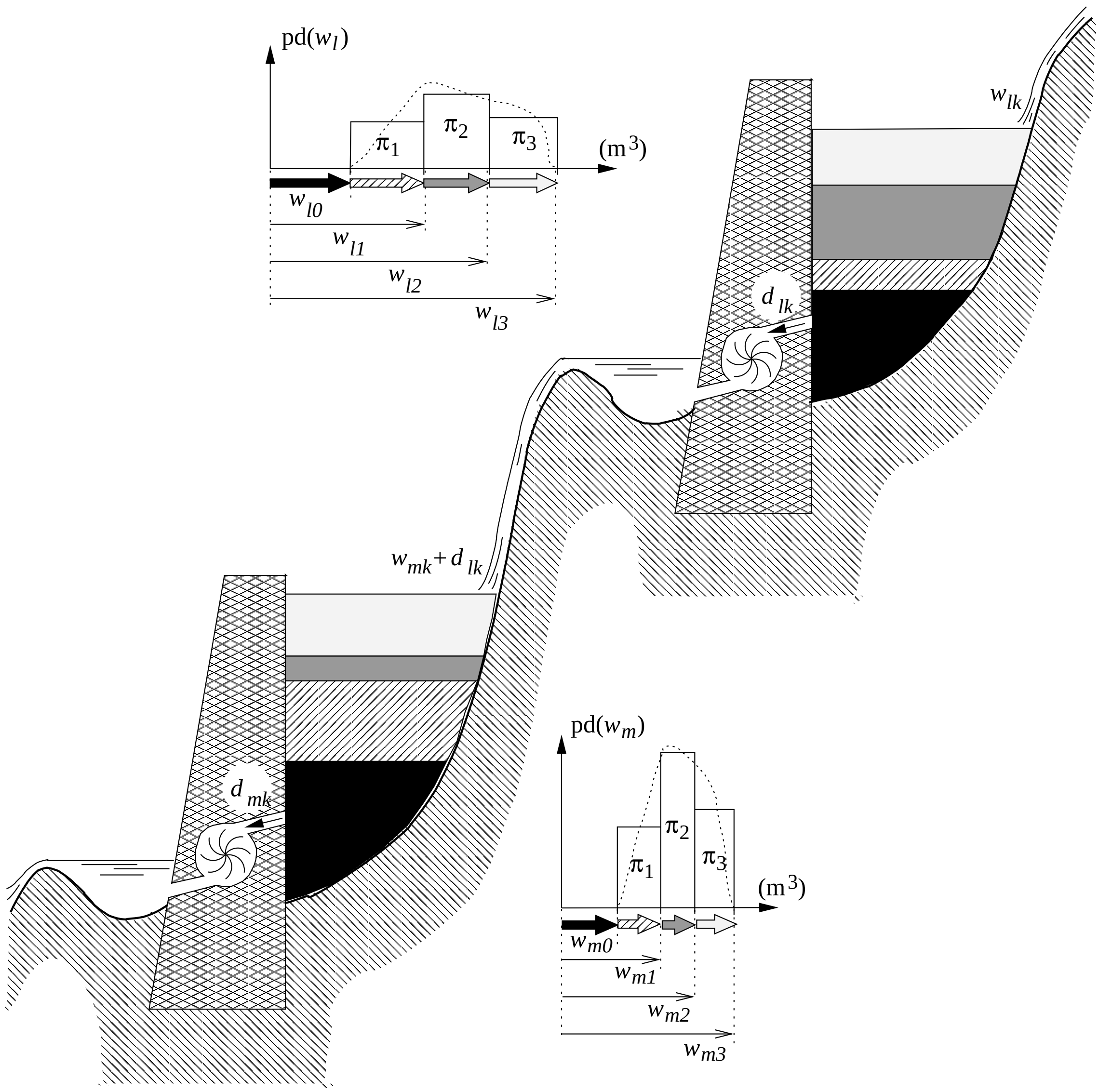

Figure 7: Sections of reservoirs with multicommodity inflows, volumes, and discharges, and block probability density functions of inflows for N_a=4.

This approach has some common points with scenario stochastic programming:

- The multicommodity procedure models hydrogeneration as a two-stage stochastic problem. The first stage is the first interval, which is assumed to have deterministic water inflows (forecasts for the current or next weekly interval). The second stage is the rest of the intervals, where only as many scenarios as commodities are assumed for inflows.
- Expected values of hydrogeneration are obtained from the discharges, volumes and probabilities associated to the multicommodity scenarios.

- The solution to the problem in both procedures has a form of a tree (with only as many branches as the number of commodities employed in the case of the multicommodity model).

The main differences of the multicommodity model from the scenario procedure are that:

⋆ Multicommodity scenarios are accumulative along all intervals, meaning that, in all intervals, a multicommodity scenario has higher inflows than the former multicommodity scenario. In stochastic programming, one scenario may have higher inflow in a particular interval than another scenario, but it may be the other way around in a different interval.

⋆ All multicommodity scenarios are optimized together in a single multicommodity network flow problem sharing the capacity constraints, whereas in stochastic programming each scenario is optimized separately.

The N_a commodities flowing in the network nodes are in fact $w^i_{m\,0}$, $(w^i_{m\,1}-w^i_{m\,0})$, $\ldots$, $(w^i_{m\,N_a-1}-w^i_{m\,N_a-2})$, $\forall\, m$, $\forall\, i$, with $w^1_{m\,0}=w^1_{m\,1}=\ldots=w^1_{m\,N_a-1}$, $\forall\, m$ in the first interval (current or next week).
The multicommodity expected solutions are feasible and are compatible with the long-term hydrogeneration planning practice of *anticipating* possible future large inflows by leaving room in reservoirs in order to avoid possible future spillages. (This is opposite to short-term planning, where all inflows are deterministic data and there is no need to *anticipate* possible future events.)

6.5 Multicommodity expected hydrogeneration and its linearization

Let N_a be the number of commodities considered and let π_1, π_2, $\ldots$, π_{N_a-1} be the probabilities associated with the block probability density function of a water inflow. An extra subindex k (k=$0,1,\ldots,N_a$$-$$1$) will be employed to distinguish to which commodity the inflows, volumes, discharges, and generations correspond. Hydrogeneration will now be:

$$h^i_{m\,k} = \frac{\mathrm{k}_m}{T^i} d^i_{m\,k} \Big\{ \mathrm{s}_{\mathrm{b}m} + \tfrac{\mathrm{s}_{\mathrm{l}m}}{2}(v^{i-1}_{m\,k} + v^i_{m\,k}) + \tfrac{\mathrm{s}_{\mathrm{q}m}}{3}(v^i_{m\,k} - v^{i-1}_{m\,k})^2 + \mathrm{s}_{\mathrm{q}m} v^{i-1}_{m\,k} v^i_{m\,k} + \tfrac{\mathrm{s}_{\mathrm{c}m}}{4}({v^{i-1}_{m\,k}}^2 + {v^i_{m\,k}}^2)(v^{i-1}_{m\,k} + v^i_{m\,k}) \Big\} \qquad k = 0,1,\ldots,N_a-1 \tag{27}$$

The expected hydropower generation $\widetilde{h}^i_m$ of reservoir m and total hydroenergy over the $i^{\underline{\text{th}}}$ interval is:

$$\widetilde{h}^i_m = \frac{\pi_1}{2} h^i_{m\,0} + \sum_{k=1}^{N_a-2} \frac{\pi_k + \pi_{k+1}}{2} h^i_{m\,k} + \frac{\pi_{N_a-1}}{2} h^i_{m\,N_a-1}\ , \quad m = 1,\ldots,N_m \qquad E^i_H = T^i \sum_m \widetilde{h}^i_m\,. \tag{28}$$

The (multicommodity) water balance equations (26) now become:

$$v^{i-1}_{m\,k} + \mathrm{w}^i_{m\,k} + \sum_{j\in G_m} (d^i_{j\,k} + p^i_{j\,k}) = v^i_{m\,k} + d^i_{m\,k} + p^i_{m\,k} \quad k = 0,1,\ldots,N_a-1\,, \tag{29}$$

and there will be the following bounds on the hydrovariables:

$$\underline{0} \le v^i_{m\,k} \le \overline{v}_m\,, \quad \underline{0} \le d^i_{m\,k} \le \overline{d}^i_m\,, \quad \underline{0} \le p^i_{m\,k} \le \overline{p}^i_m\,, \quad \forall\, m\,,\ \forall\, i\,,\ \forall\, k\,, \tag{30}$$

where the maximum volume $\overline{v}_m$ is independent of the interval, but the maximum discharge $\overline{d}^i_m$ or maximum spillage $\overline{p}^i_m$ are proportional to interval length T^i.
Note that if (27–30) are employed in lieu of an equality of type (23), we are using a full hydrogeneration model instead of an approximate one.
Equations (27) are fourth-order polynomials of hydrovariables, and may make the problem hard to solve, especially because there is also an exponential number of inequalities (17) to satisfy. A simplification of the problem could be to linearize equations (27) in the following way: the term in

braces of (27), which is the $m^{\underline{\text{th}}}$ reservoir head over the $i^{\underline{\text{th}}}$ interval, will be considered to be a constant $\widehat{s}^i_m$, as if the volumes $\widehat{v}^i_m$ it had been calculated with were known,

$$\widehat{s}^i_m = \mathrm{s}_{\mathrm{b}m} + \frac{\mathrm{s}_{\mathrm{l}m}}{2}(\widehat{v}^{i-1}_{m\,k} + \widehat{v}^i_{m\,k}) + \frac{\mathrm{s}_{\mathrm{q}m}}{3}(\widehat{v}^i_{m\,k} - \widehat{v}^{i-1}_{m\,k})^2 + \mathrm{s}_{\mathrm{q}m}\widehat{v}^{i-1}_{m\,k}\widehat{v}^i_{m\,k} + \frac{\mathrm{s}_{\mathrm{c}m}}{4}(\widehat{v}^{i-1\;2}_{m\,k} + \widehat{v}^{i\;\;2}_{m\,k})(\widehat{v}^{i-1}_{m\,k} + \widehat{v}^i_{m\,k})$$

and the succession of volumes employed $\widehat{v}^i_{m\,k}$, i=1,…,N_i could be a former solution, or a uniform variation from the initial $v^0_{m\,k}$ to the final volume $v^{N_i}_{m\,k}$. In this way, equations (27) become:

$$h^i_{m\,k} = \frac{\mathrm{k}_m \widehat{s}^i_m}{T^i} d^i_{m\,k} \qquad \forall m\ ,\ \forall i\ ,\ \forall k\ , \tag{31}$$

The linearized equations (31), together with (28–30), would be the linearized multicommodity hydrogeneration model to be appended to the optimization (16–22) instead of the constraint of set (21) which corresponds to (23).

7 Representation of the wind-power generation (and of other renewable energy sources) in long-term planning

Wind-Power generation, and other renewable energy sources such as photovoltaic generation, produce electric power when there is sufficient wind, or sunshine, in an appropriate direction. Their values can be approximately predicted in the short term given a weather forecast. In the long term, we can predict a generation duration curve, for wind power, or photovoltaic generation, for a given interval from historical records of wind-power or photovoltaic generations for the time interval considered. Fig. 8 shows the October 2002 metered wind-power generation in the Spanish power pool, sorted in descending power.

Assuming that October were one of the intervals in the long-term planning problem to solve, wind-power generation for this interval could be taken into account through the following procedure:

1. An expected wind-power generation duration curve should be predicted from available records of wind-power generation in the October months of various years. A low degree polynomial of $3^{\underline{\text{rd}}}$ or $4^{\underline{\text{th}}}$ order could be fitted, making sure it decreases monotonically within the interval duration. A maximum power capacity C^{Oct}_{WP} (normally well below the installed wind-power capacity) should be established. Note that the maximum metered wind-power generation in Fig. 8 is about 1740 MW, whereas the metered installed wind-power capacity is 2850 MW.

2. The expected wind-power generation duration curve should be split into a number of slices of wind power (as indicated in Fig. 9) and a wind power capacity should be associated to each slice. Three slices have been considered in Fig. 9, and $C^{Oct(1}_{WP}$+$C^{Oct(2}_{WP}$+ $C^{Oct(3}_{WP}$=C^{Oct}_{WP}.

3. Three pseudo units (as is the number of slices made) will be considered with capacities $C^{Oct(1}_{WP}$, $C^{Oct(2}_{WP}$ and $C^{Oct(3}_{WP}$, and with zero outage probability ($q^{Oct(k}_{WP}$=0), and zero cost ($\widehat{f}^{Oct(k}_{WP}$=0).

4. Let us assume that $a^{Oct(1}_{WP}$, $a^{Oct(2}_{WP}$ and $a^{Oct(3}_{WP}$, are the energies (areas in the wind-power generation duration curve) corresponding to each slice. The generations of each pseudo unit should be constrained to be exactly:

   $$E^{Oct(1}_{WP} = a^{Oct(1}_{WP} \qquad E^{Oct(2}_{WP} = a^{Oct(2}_{WP} \qquad E^{Oct(3}_{WP} = a^{Oct(3}_{WP}\ ,$$

   which are of type (20). The shape of the resulting generation duration curve for October will not coincide with that of the expected generation duration curve, but the areas of the sliced parts will coincide and they will be an acceptable approximation, and the more slices, the better the approximation.

This should be done for all intervals.

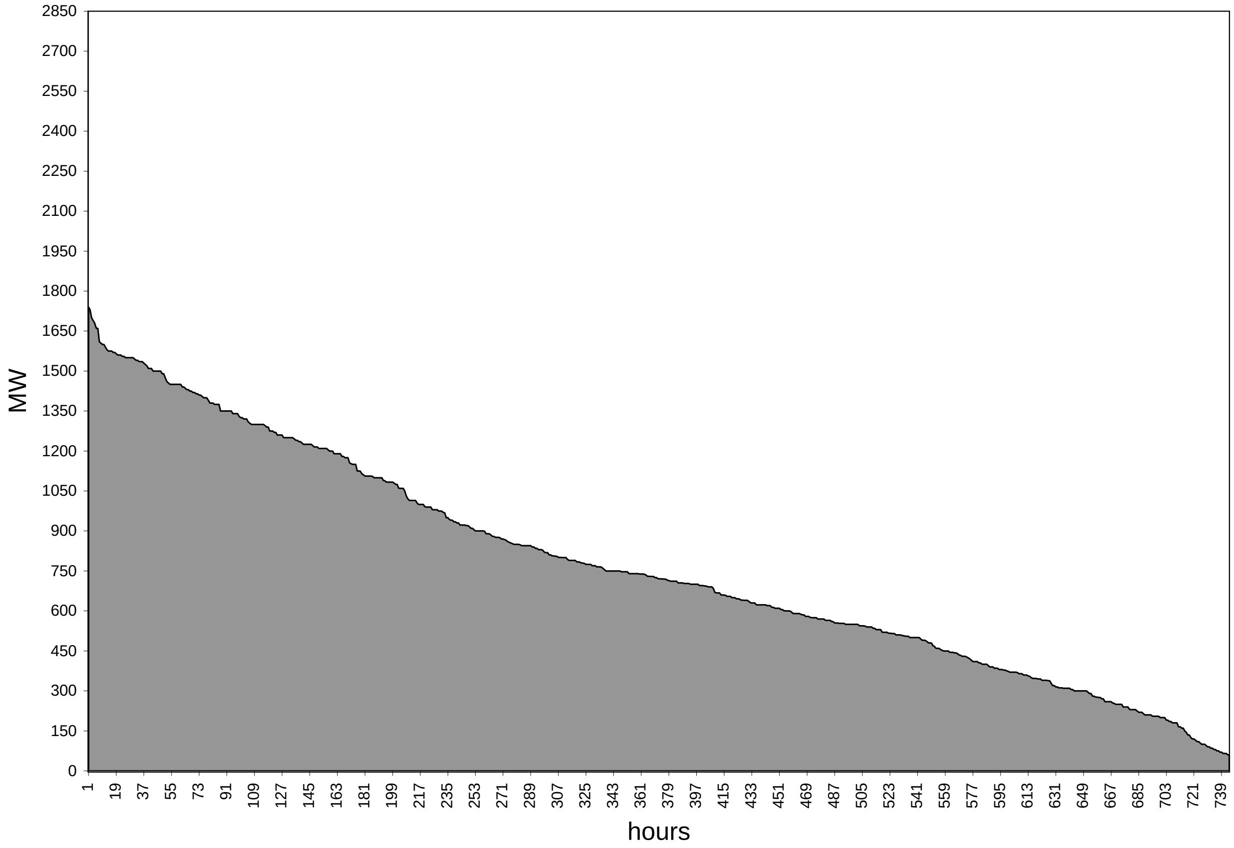

Figure 8: Metered wind-power generation, sorted in descending power, in the Spanish power pool in October 2002.

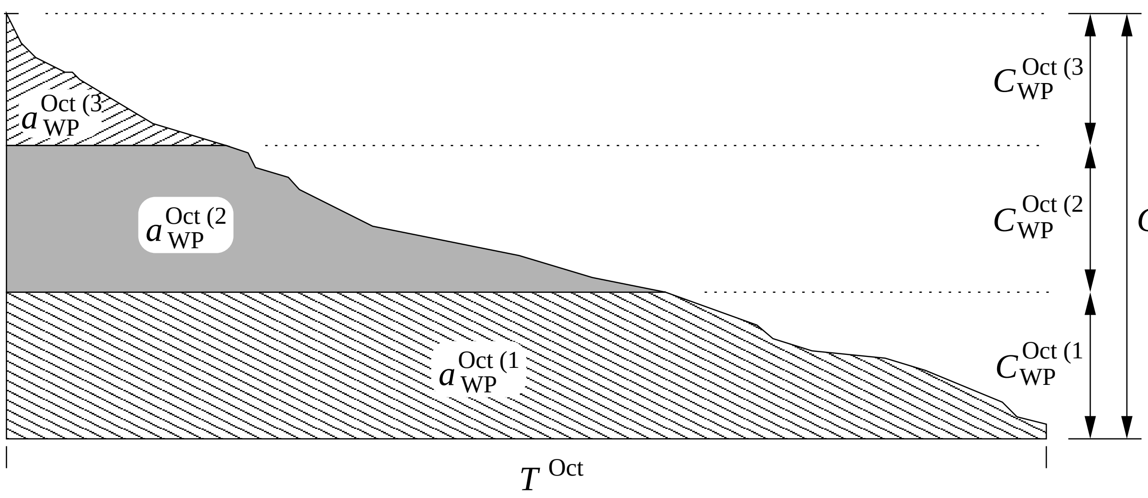

Figure 9: Splitting of wind-power generation during October in three pseudo-unit slices with capacities $C_{WP}^{Oct(1}$, $C_{WP}^{Oct(2}$ and $C_{WP}^{Oct(3}$, and generations $a_{WP}^{Oct(1}$, $a_{WP}^{Oct(2}$ and $a_{WP}^{Oct(3}$.

8 Long-term maximization of profit in a "competitive" market

In the classical electricity markets, utility companies have both generation and distribution of power. These companies have their own load to supply, corresponding to their clients plus other contracts, and try to minimize their generation cost. In "competitive" electricity markets, generation companies have no distribution, and therefore no load of their own. Generation companies must bid their generation to the market operator and a market price is determined for each hour by matching the demand with the generation of the lowest bids. Generation companies are no longer interested in generating at the lowest cost but in obtaining the maximum profit, which is the difference between market price and generation cost for all accepted generation bids. In long-term operation all accepted bids in a time interval (a week, or a month) must match the LDC of this interval.

There is no specific load to be matched by a specific generation company (SGC). The only known loads are the predicted LDC's for the whole market in each interval. As all generation companies pursue their maximum profit, it is natural to attempt to maximize the profit of all generation companies combined.

The SGC must thus solve the problem of the maximization of profit of all generation companies, taking into account the total market load. The SGC should introduce its own operation constraints (fuel and emission limits, contracts, etc.) and may also introduce a market-share constraint for its units in one or several intervals. (The Lagrange multiplier value of this constraint will tell whether the market share imposed, though feasible, is reasonable or not.) The long-term results will indicate how the SGC should program its units so that its profit be maximized while meeting all its operation constraints.

8.1 Long-term market price function of a given interval

From the records of past market-price and load series (see Fig. 10) it is possible to compute a market-price function for a given interval. This function is to be used with expected generations that match the LDC of the interval, so market prices should correspond in duration with the duration of loads, from peak to base load in the interval.

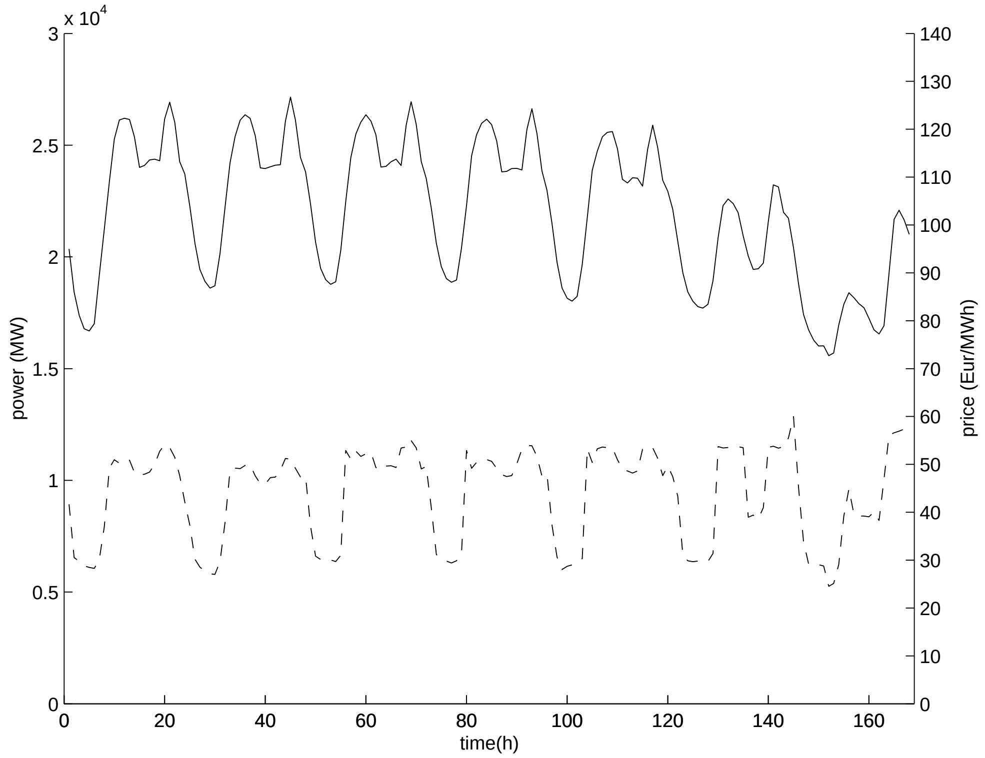

Figure 10: Hourly loads (continuous curve) and market prices (dashed) in a weekly interval.

Both the load and the market price series should be reordered in decreasing load order obtaining a LDC and a price-duration curve that corresponds to the loads in the LDC. The price-duration-curve obtained will be nonsmooth and may even be nondecreasing (see Fig. 11). However, fitting a straight line or a low order polynomial to it, a decreasing line or function will generally be obtained. Given the variability of the price-duration curve, it seems reasonable to fit a straight line to it. Let b^i and l^i be the basic and linear coefficient of such line for the $i^{\underline{\text{th}}}$ interval. (Predictions of b^i and l^i could be obtained taking into account both the series corresponding to the same interval in several successive years and that of successive intervals.)

8.2 Maximum profit objective function

In order to determine the maximum-profit objective function, a simplifying assumption is convenient regarding the shape of the unit contributions in the generation-duration curve (see Fig. 3). Instead of having some units (particulary those with the lowest loading order) with an irregular shape in its right side, it will be assumed that the contribution of all units will have a rectangular shape with height C_j (for unit j) and base length E^i_j/C_j as in Fig. 12.

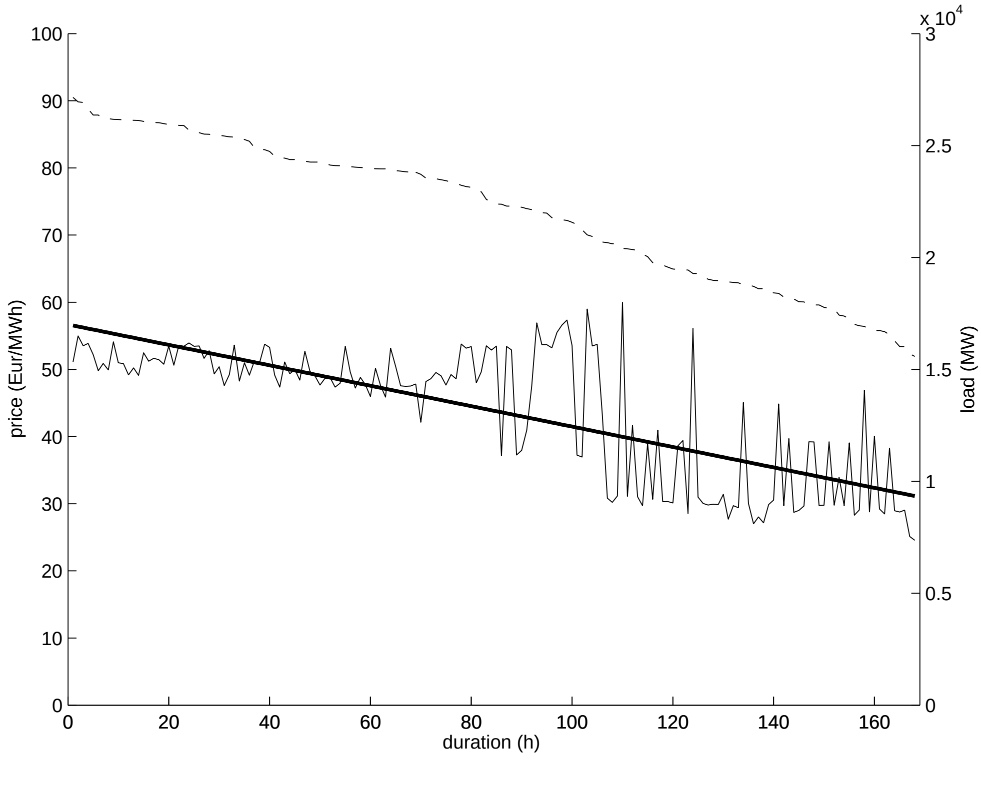

Figure 11: Market prices ordered by decreasing load power (thin continuous curve) in weekly interval, market-price linear function (thick line), and LDC (dashed).

The profit (price minus cost) of unit j in interval i will be:

$$\int_0^{E^i_j/C_j} C_j\{b^i + l^i t - \widetilde{f}_j\}\mathrm{d}t = (b^i - \widetilde{f}_j)E^i_j + \frac{l^i}{2C_j}{E^i_j}^2$$

and adding for all intervals and units, and taking into account the external energy, we get the profit function to be maximized:

$$\sum_i^{N_i}\left[\sum_j^{N_u}\left\{(b^i - \widetilde{f}_j)E^i_j + \frac{l^i}{2C_j}{E^i_j}^2\right\} - \widetilde{f}_{N_u+1}E^i_{N_u+1}\right]$$

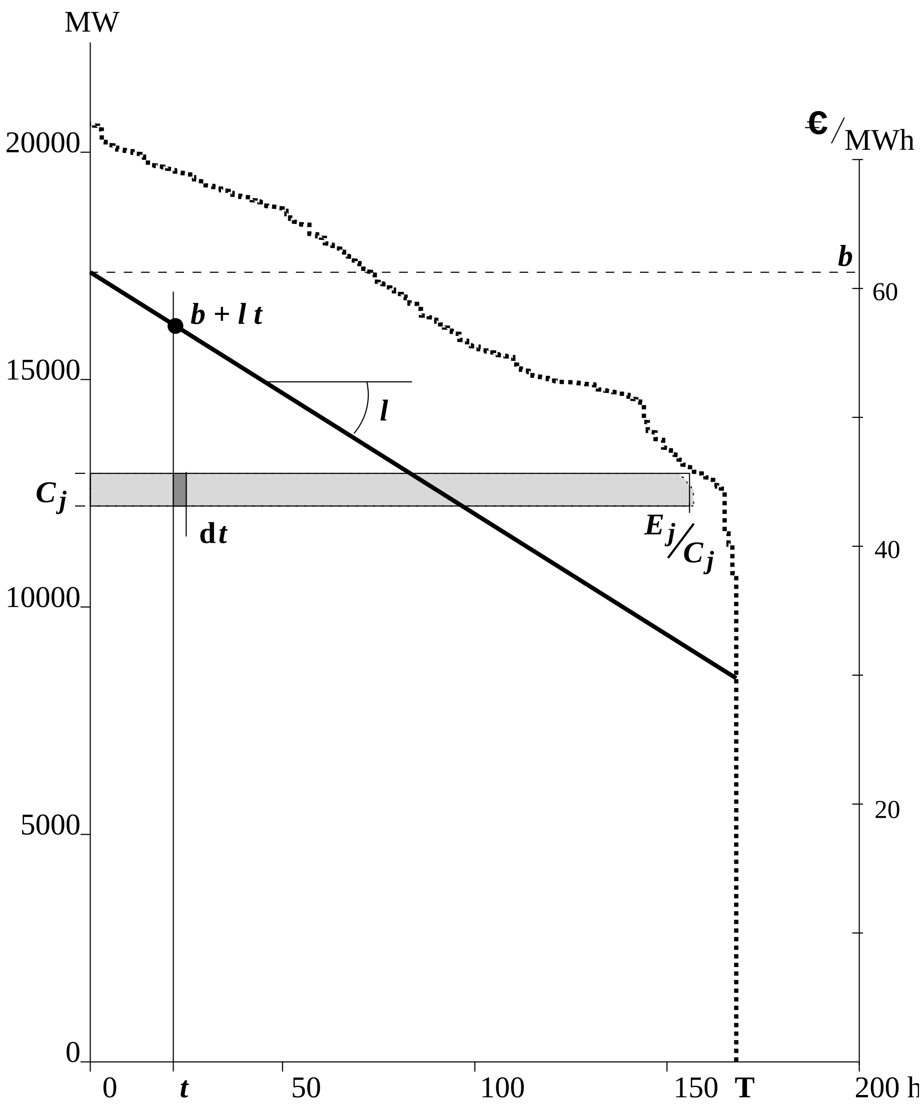

Figure 12: Long-term price function for a time interval and contribution of $j^{\underline{\text{th}}}$ unit.

which is quadratic in the generated energies. Using (10) we are led to the equivalent expression:

$$\sum_{i}^{N_i}\left[\sum_{j}^{N_u}\left\{\left(b^i - f_j\right)E_j^i + \frac{l^i}{2C_j}E_j^{i\;2}\right\} - \widetilde{f}_{N_u+1}\widehat{E}^i\right] \tag{32}$$

with $f_j{=}\widetilde{f}_j{-}\widetilde{f}_{N_u+1}$

Given that $f_{N_u+1}\widehat{E}^i$ is a constant, the problem to be solved is:

$$
\begin{array}{lll}
\underset{Ej^i}{\text{minimize}} & \sum_i^{N_i}\sum_j^{N_u}\left\{(f_j-b^i)E_j^i-\frac{l^i}{2C_j}E_j^{i\,2}\right\} & \\
\text{subject to:} & \sum_{jinU} E_j^i \le \widehat{E}^i - W^i(U) \quad \forall U\subset\Omega^i \quad i=1,\ldots,N_i & \\
 & A_{\ge}^i\,E^i \ge R_{\ge}^i \qquad\qquad i=1,\ldots,N_i & \\
 & \sum_i A_{\ge}^{0i}\,E^i \ge R_{\ge}^0 & (33)\\
 & A_{=}^i\,E^i = R_{=}^i \qquad\qquad i=1,\ldots,N_i & \\
 & \sum_i A_{=}^{0i}\,E^i = R_{=}^0 & \\
 & E_j^i \ge \underline{0} \qquad j=1,\ldots,N_u, \quad i=1,\ldots,N_i &
\end{array}
$$

It should be noted that, should all b^i and l^i be zero, the solution of the maximum profit problem (33) would be the same as that of the minimum cost problem (16-22). Otherwise, the cost of the maximum profit solution is higher than that of the minimum cost solution.

Given that $l^i{<}0$, the quadratic of the objective function of (33) is positive definite, thus problem (33) has a unique global minimizer . Moreover, the quadratic of the objective function is diagonal.

8.3 External nonbidding energy source

The decreasing market-price line induces more expensive units to generate only while market price is not lower than cost, while in the minimum-cost solution units generate as long as their price is not higher than the external energy (which is always the case).

There are two types of external energy: emergency external energy, which is the genuine (power-capacity unlimited) response of the external connected system to simultaneous failures in the units, and the external (nonbidding) units tied to the market system through contracts at a fixed price. This acts as an ordinary unit ν (with price f_ν, capacity C_ν, and outage probability q_ν) except that it does not participate in the bidding operation and therefore is not driven by profit but only by minimum cost. The complete objective function to consider is thus:

$$
\underset{Ej^i}{\text{minimize}} \quad \sum_i^{N_i}\left[f_\nu E_\nu^i + \sum_{j\neq\nu}^{N_u}\left\{(f_j-b^i)E_j^i-\frac{l^i}{2C_j}E_j^{i\,2}\right\}\right] \tag{34}
$$

9 Coding the load-matching constraints

The main difficulty of the direct solution of the Bloom and Gallant model is the exponential number of load-matching inequality constraints (17). These constraints are avoided in the application of the Dantzig-Wolfe column generation method [14, 11, 13], or are generated as they are required in the active set method [10]. In a direct solution by linear or quadratic programming all $N^i{\times}(2^{N_u}{-}1)$ constraints must be explicitly created.

Leaving aside the storage and processing time for these many load-matching inequality constraints, their creation has two parts: the linear coefficients, which is fast, and the *rhs*'s, which is very time consuming as it requires lots of calculation.

9.1 Creation of the linear coefficients

The linear coefficients of constraints (17) for each interval are ones for the units in all possible subsets ω of set Ω^i (where units with programmed overhauling in the $i^{\underline{\text{th}}}$ interval have been excluded). Assuming that $N_u^i{=}|\Omega^i|$, there will be $2^{N_u^i}{-}1$ constraints in the $i^{\underline{\text{th}}}$ interval.

A systematic way to create the ones of all these constraints is through the following procedure: the $m^{\underline{\text{th}}}$ constraint in the series from the $1^{\underline{\text{st}}}$ to the $(2^{N_u^i}-1)^{\underline{\text{th}}}$ will have a 1 or a 0 for each unit j according to the result of

$$(m\,\texttt{div}\,2^{j-1}) \bmod 2 \tag{35}$$

where `div` stands for the truncated quotient of the division of the left by the right operand (and mod is the *remainder* operator of the division of the left by the right operand)[2]. This amounts to expressing in base 2 all the numbers from 1 to $2^{N_u^i}-1$.

9.2 Calculation of the right-hand sides of the load-matching inequality constraints

For each interval i and for the units of each subset U of the set Ω^i, which can be determined using (35), we must first calculate $L_U^i(x)$ starting from $L_0^i(x)$ by successive convolution for all units j in U using (1), and then compute

$$\widehat{E}^i - W^i(U) = \widehat{E}^i - T^i \int_0^{\widehat{P}^i} L_U^i(x)\mathrm{d}x$$

using numerical integration.

This means a lot of arithmetic operations, and it has been found that a specific Fortran program that makes these calculations is far more efficient than letting these operations be performed in a *script* code in *AMPL*, where the solution to the optimization problem is executed after all *rhs*'s have been calculated.

10 Computational results

10.1 Test cases

Due to the limitations imposed by the number of load-matching constraints, the size of test problems must be small. However, an effort has been made to solve real problems. The loads of the Spanish electric power pool have been considered but generation units have had to be concentrated into a reduced number of psudo-units in order not to exceed 18, which is a practical limit, given that the calculation of its *rhs*'s requires over 19 hours of CPU time on a SPECfp2000 310 processor of a Hewlet Packard Netserver LC2000 U3, which is the computer employed.

The optimization considers one small specific generation company (SGC) in detail while the rest of the generation units are amalgamated in big pseudo-units so as to limit N_u. Hydrogeneration of the SGC considered has also been modeled in detail, while the rest of hydrogeneration is considered through the approximate model.

Problems of N_i=11 intervals have been considered, the first being a week in March, the second is the rest of March, the third is the month of April, from the forth to the seventh are the rest of the year in two-month intervals, and finally, intervals from the eighth to the eleventh are the next year in three-monthly intervals. This subdivisions allow us to impose the yearly domestic-coal incentive constraints to the coal-fired units, which is applicable in Spain.

Four cases with 13, 15, 17 and 18 units have been prepared. They are the same problem with disaggregation of some aggregated units. There are four subcases of each case, all of which have two hydro pseudo-units: one is the hydrogeneration of the SGC considered, and the other is the hydrogeneration of the rest of the Spanish power pool. Cases whose name starts with "ltp" have both hydrogenerations modeled with the approximate model, while cases starting with "lhtp" have the SGC hydrogeneration modeled by the linearized multicommodity full model, and the rest of hydrogeneration by the approximate model.

[2]The authors wish to gratefully acknowledge the suggestion of this well known procedure made to them by their colleague Mr. Jordi Laseras.

Table 1: Test cases for AMPL models of long-term electric power planning

case	N_i	N_u	$\sum_i N_u^i$	N_{Ha}	$N_{Ha=}$	$\sum_i N_=^i$	$N_=^0$	$\sum_i N_\geq^i$	$N_\geq^0$	$\sum_i(2^{N_u^i}-1)$	CPU (h) calc. *rhs*
ltp01a	11	13	140	–	–	0	2	0	2	79861	0.44
lhtp01a	11	13	140	384	275	0	2	0	2	79861	0.44
ltp01b	11	13	140	–	–	0	2	1	4	79861	0.44
lhtp01b	11	13	140	384	275	0	2	1	4	79861	0.44
ltp02a	11	15	162	–	–	0	2	33	3	319477	2.28
lhtp02a	11	15	162	384	275	0	2	33	3	319477	2.28
ltp02b	11	15	162	–	–	0	2	34	5	319477	2.28
lhtp02b	11	15	162	384	275	0	2	34	5	319477	2.28
ltp03a	11	17	183	–	–	0	2	54	5	1245173	9.52
lhtp03a	11	17	183	384	275	0	2	54	5	1245173	9.52
ltp03b	11	17	183	–	–	0	2	55	7	1245173	9.52
lhtp03b	11	17	183	384	275	0	2	55	7	1245173	9.52
ltp04a	11	18	193	–	–	0	2	64	6	2457589	19.27
lhtp04a	11	18	193	384	275	0	2	64	6	2457589	19.27
ltp04b	11	18	193	–	–	0	2	65	8	2457589	19.27
lhtp04b	11	18	193	384	275	0	2	65	8	2457589	19.27

Market-share constraints can be imposed. Cases whose name ends with "a" do not have any market-share constraint imposed. Cases ending with "b" have market-share constraints associated to the units of the SGC imposed and active.

As mentioned earlier, the purpose of these problems and computational tests is twofold:

- to test the models developed, described in this report, and to observe the influence of several parameters associated with the models, and
- to have reliable results (obtained with a reliable code for linear and quadratic programming: *Cplex 7.5*) for the problems posed with which to check alternative specialised algorithms to solve the same problems, specifically the Dantzig-Wolfe column generation algorithm, the active set algorithm, and other algorithms to be developed.

The characteristics of the test cases employed are summarized in Table 1. Column $\sum_i N_u^i$ is the number of variables and column $\sum_i(2^{N_u^i}-1)$ holds the number of load-matching inequality constraints, and the last column gives the CPU seconds required to compute their *rhs*'s.

$N_{ha}=N_h \times N_a$ is the total number of hydrovariables in the full hydromodel. N_h is $N_h=96$, which comes from a basin with $N_m=3$ reservoirs, each having one discharge and one spillage. The number of water commodities considered is $N_a=4$. The number of hydrobalance (29) and linearized hydrogeneration equality constraints (28,31) for all water commodities is $N_{Ha=}=275$.

All test cases have been solved with two different objective functions: the linear minimum cost (16) and the quadratic of maximum profit (34). The linear cost problems have been solved using the linear programming code in *Cplex 7.5* package [4], while for the quadratic profit problem the barrier separable QP solver [17] in *Cplex 7.5* package is employed, both through an *AMPL* [6] model and data files. Prior to the solution, the *rhs*'s of the load-matching inequality constraints (17) have been calculated using an separate program, whose required CPU time is reported in the last column of Table 1. The calculated *rhs*'s are a part of the *AMPL* data files used.

10.2 Solutions of long-term minimum cost planning and comparison with alternative solution procedures

The authors have developed alternative specialised algorithms, using the Dantzig-Wolfe column generation method [12, 13] and the active set method [9, 10], to solve the linear minimum cost problems not including the full model hydro variables. The solutions obtained with the Dantzig-Wolfe column generation, which is the most efficient of the two [13], are compared with those obtained through *AMPL* plus *Cplex 7.5* linear programming.

Table 2: Comparison of AMPL plus Cplex lp, and the Dantzig-Wolfe column generation method

	AMPL plus Cplex 7.5				Dantzig-Wolfe column gen.		
case	input (s)	lp iters.	lp (s)	$\sum_i \sum_j^{N_u+1} f_j E_j^i$ (€)	D.W. iters.	(s)	$\sum_i \sum_j^{N_u+1} f_j E_j^i$ (€)
ltp01a	1.3	781	1.3	4837512292	42	3.7	4837512061
ltp01b		2354	2.35	4854704625	41	3.6	4854767471
ltp02a	5.69	3285	11.0	3587429530	240	14.2	3587428413
ltp02b		7646	16.9	3622023526	385	23.0	3622186266
ltp03a	24.47	12622	56.8	3580260681	324	18.8	3580266252
ltp03b		23213	86.2	3624657306	873	28.7	3624656194
ltp04a	46.88	17447	115.1	3579624419	315	20.0	3579623303
ltp04b		42785	212.0	3624160513	910	29.8	3624160686

The second column of Table 2 has the input times required by the *AMPL* data files. These times are important because the data files, due to the *rhs*'s of the load-matching constraints, are very large, e.g., the data file for case ltp04a is over 100Mbyte.

It can be observed that the Dantzig-Wolfe column generation method is much more efficient than the solution of the entire problem by *AMPL* plus *Cplex 7.5* without considering the enormous time required to calculate the *rhs* terms of the load-matching constraints, whose number is exponential.

10.3 Solutions of long-term maximum profit planning and comparison with the minimum cost solution

$$\text{cost}: \sum_i \sum_j^{N_u+1} \widetilde{f}_j E_j^i \qquad \text{profit}: \sum_i^{N_i} \left[\sum_j^{N_u} \left\{ (b^i - \widetilde{f}_j) E_j^i + \frac{l^i}{2C_j} {E_j^i}^2 \right\} - \widetilde{f}_{N_u+1} E_{N_u+1}^i \right]$$

It is clear from the results that the maximization of profit with respect to the minimum cost solution brings about a greater increase in generation cost than an increase in profits.

10.4 Effect of market-share constraints

Three market-share constraints have been introduced in cases whose name ends with "b": one for the first interval, one for the intervals corresponding to the rest of the first year (intervals 2 to 7), and a third for the intervals of the second year (8 to 11). These three sets of successive intervals will be referred to with the supraindices ^{I}, ^{II} and ^{III} associated to the variables. The market-share constraints refer to the units of the SGC, and force their generation to add up to over a given percentage of the load in the corresponding intervals.

$$\sum_{i \in Ik} \sum_{j \in SGC} E_j^i \geq \mu^{Ik} \sum_{i \in Ik} \widehat{E}^i \qquad Ik: \; I, \; II, \; III, \tag{36}$$

which are of type (19), except set I (a single interval) which is of type (18).

Table 3: Minimum cost and maximum profit solutions with an approximate and linearized full hydro-model

	Minimum expected cost solution				Maximum expected profit solution			
case	lp iters.	lp (s)	expec. cost (€)	expec. profit (€)	b qp iters.	bar qp (s)	expec. profit (€)	expec. cost (€)
ltp01a	781	1.3	4837512292	9120093218	34	97.56	9552335013	5414756605
lhtp01a	5017	8.39	5851068559	8295313850	49	40.82	9395092981	5868003086
ltp01b	2354	2.35	4854704625	9097863493	46	55.09	9536489728	5427534028
lhtp01b	4342	8.11	5859268225	8382878646	51	39.64	9379988300	5877356849
ltp02a	3285	11.00	3587429530	10364887782	59	183.5	10986157177	3971181574
lhtp02a	9572	36.25	3567370363	10377588860	64	207.3	11005062681	3955156385
ltp02b	7646	16.92	3622023525	10291956272	56	176.9	10961049191	4018697239
lhtp02b	11514	44.31	3598955942	10334067931	60	186.3	10987917823	3992253653
ltp03a	12622	56.85	3580260681	10292023819	78	1020.3	11004938184	3978570531
lhtp03a	21659	154.2	3560201514	10399011268	78	1038.6	11023659735	3959557570
ltp03b	23213	86.16	3624657306	10287465543	75	977.7	10977720297	4025913005
lhtp03b	25669	187.0	3601589723	10324348019	77	1439.5	11004809653	3999334833
ltp04a	17220	113.7	3579624419	10306638029	87	4393.2	11006374461	3975489971
lhtp04a	34047	309.4	3559565252	10327962474	88	2282.7	11025091650	3959478586
ltp04b	42785	212.0	3624160513	10296858318	116	5787.0	10979064726	4025907694
lhtp04b	41866	426.2	3600704161	10318778719	98	2479.9	11006729490	3998309774

The criterion employed to fix a market-share μ^{Ik} for the units in the set SGC is based on the Lagrange multiplier values of the market-share constraints λ^{Ik}_{m-s} and the expected profit rate in the power pool r^{Ik}: total profit over total load. The Lagrange multipliers λ^{Ik}_{m-s} express the rate of change in pool profit due to a market-share increase by the SGC. The reaction of competitor generating companies to a market-share increase by the SGC would be proportional to the resulting $\lambda^{Ik}_{m-s}/r^{Ik}$. Therefore, attainable market-shares are those that produce a small enough value $\lambda^{Ik}_{m-s}/r^{Ik}$. In the cases reported in Table 4 the market-shares μ^{Ik} of the SGC have been pushed up until the ratio $\lambda^{Ik}_{m-s}/r^{Ik}$ was close to but did not exceed $\frac{1}{3}$.

Table 4: Effect of market share constraints on the profit of the SGC

case	μ^{I} %	λ^{I}_{m-s}	r^{I}	μ^{II} %	λ^{II}_{m-s}	r^{II}	μ^{III} %	λ^{III}_{m-s}	r^{III}	total profit (€)	SGC profit (€)
lhtp01a	3.79	0.0		3.11	0.0		3.03	0.0		9395092981	276449929
lhtp01b	5.0	7.66	25.7	4.3	7.27	27.1	4.2	6.34	24.6	9379988300	296324589
lhtp02a	2.48	0.0		2.23	0.0		2.75	0.0		11005062681	215212171
lhtp02b	3.6	7.11	29.4	3.3	8.35	31.2	3.5	8.65	29.2	10987917823	252856724
lhtp03a	2.70	0.0		2.54	0.0		2.38	0.0		11023659736	251555580
lhtp03b	4.0	8.27	29.4	3.7	8.94	31.2	3.9	8.87	29.2	11004809653	291410819
lhtp04a	2.70	0.0		2.55	0.0		2.77	0.0		11025091650	256426270
lhtp04b	5.2	7.32	29.3	3.7	8.69	31.2	3.9	8.75	29.3	11006729490	297019059

There are also cases whose name ends with “a” in Table 4. These cases are the same as those ending in “b” but without the market-share constraints. They are thus equivalent to having imposed a nonactive market share, lower than the share the SGC gets in the solution. It should be noted that the market-share constraints imposed slightly decrease the overall profit, but they noticeably increase the SGC profit.

11 Conclusions

- The long-term hydrothermal planning of the electricity generation problem has been presented and an extension of the Bloom and Gallant model has been put forward in order to solve it.
- An alternative procedure to account for the stochasticity of long-term hydrogeneration, based on a multicommodity network flow model, has been described and a comparative analysis with classical scenario based stochastic programming has been presented.
- A procedure to model long-term wind-power generation (which is also applicable to other types of renewable energy sources) in long-term planning has been described.
- A new way of formulating the long-term profit maximization of generating companies in a competitive market has been described.
- Implementation details of the solution with *AMPL* of the minimum cost and the maximum profit long-term planning problems have been given.
- The computational experience with the *AMPL* plus *Cplex 7.5* linear programming and barrier quadratic programming has been reported. This includes:
  - The calculation of the *rhs*'s of the load-matching constraints for the data files required by *AMPL*, which is extremely time-consuming, and which is fairly time-consuming to be read in the solution process. This lengthy calculation, requiring extremely long files to store the results, makes this procedure impractical to use for real cases (where the number of units to consider may be well above one hundred).
  - The solution of the minimum cost and the maximum profit long-term problems.
  - The alternative use of an approximated model and a linearized full model of long term hydrogeneration.
  - The comparison of the linear *AMPL* plus *Cplex 7.5* minimum cost solution, using the approximate hydrogeneration representation, with that obtained using a code implementing the Dantzig-Wolfe column-generation algorithm, which turns out to be much faster than *Cplex 7.5* linear programming for the entire problem.
  - The analysis of the effect of market-share constraints for a SGC in the maximum profit solution.
- Further work by the authors on this subject is underway. Its main objectives are:
  - To prepare realistic test cases of larger sizes with the data available of the Spanish electric power pool.
  - To develop a quadratic version of the Dantzig-Wolfe column generation algorithm to solve long-term maximum profit problems.
  - To adapt both the linear and quadratic programming Dantzig-Wolfe codes to include the full hydrogeneration model.

12 Acknowledgements

The authors wish to thank prof. J. Gondzio of Univ. of Edinburgh, Edinburgh G.B. for his careful reading of the first version of this report and many helpful and constructive suggestions.

13 Glossary of symbols

^{0}	(supraindex) refers to constraints on energies with nonzero terms corresponding to more than one single interval
$A^{i}_{\geq}$	matrix of coefficients of inequalities that refer only to energies of $i^{\underline{\text{th}}}$ interval
$A^{0i}_{\geq}$	matrix of coefficients of inequalities that refer to energies of more than one interval related to energies of $i^{\underline{\text{th}}}$ interval
$A^{i}_{=}$	matrix of coefficients of equalities that refer only to energies of $i^{\underline{\text{th}}}$ interval
$A^{0i}_{=}$	matrix of coefficients of equalities that refer to energies of more than one interval related to energies of $i^{\underline{\text{th}}}$ interval
$a^{i(k}_{WP}$	area (energy) of $k^{\underline{\text{th}}}$ slice of wind power generation in $i^{\underline{\text{th}}}$ interval
b^{i}	basic term of energy-market price line of $i^{\underline{\text{th}}}$ interval
C_j	power capacity in MW of $j^{\underline{\text{th}}}$ generating unit
d^{i}_{m}	water discharge from the $m^{\underline{\text{th}}}$ reservoir over the $i^{\underline{\text{th}}}$ interval
E^{i}	total energy in MWh of LDC of $i^{\underline{\text{th}}}$ interval
E^{i}_{j}	energy generated in MWh by unit j over the $i^{\underline{\text{th}}}$ interval
$\widehat{E}^{i}$	total energy in MWh requested over the $i^{\underline{\text{th}}}$ interval
$\overline{E}_j$	upper bound of energy production for unit j
$\underline{E}_j$	lower bound of energy production for unit j.
f_j	$=\widetilde{f}_j-\widetilde{f}_{N_u+1}$
$\widetilde{f}_j$	generation cost (linear) in €/MWh of $j^{\underline{\text{th}}}$ unit
$\widetilde{f}_{N_u+1}$	generation cost (linear) in €/MWh of external emergency (power-unlimited) energy
G_m	set of indices of reservoirs up-stream of reservoir m
h^{i}_{m}	(average) hydrogeneration power (MW) at the $m^{\underline{\text{th}}}$ reservoir over the $i^{\underline{\text{th}}}$ interval
$\widetilde{h}^{i}_{m}$	expected hydrogeneration power (MW) at the $m^{\underline{\text{th}}}$ reservoir in $i^{\underline{\text{th}}}$ interval
i	(supraindex) indication of $i^{\underline{\text{th}}}$ interval
Ik	$k^{\underline{\text{th}}}$ set of successive intervals
j	(subindex) indication of $j^{\underline{\text{th}}}$ generating unit
l^{i}	linear term of energy-market price line of $i^{\underline{\text{th}}}$ interval
$L^{i}_{0}(x)$	cumulative probability of electric load to be supplied over the $i^{\underline{\text{th}}}$ interval
$L^{i}_{U_i}(x)$	cumulative probability of electric load still to be supplied over the $i^{\underline{\text{th}}}$ interval after loading (in any order) units $1, 2, ..., j$
LDC	load duration curve
m	(subindex) indication of $m^{\underline{\text{th}}}$ reservoir
N_a	number of water inflow commodities considered (with different probability of occurrence)
$N_{=}$,$N_{\geq}$	total numbers of equality and inequality constraints
N_h,N_i	numbers of hydrovariables and of intervals
N_m,N_u	numbers of reservoirss and of generating units
N^{i}_{u}	number of available thermal units in $i^{\underline{\text{th}}}$ interval
$\widehat{P}^{i}$	Peak power load of $i^{\underline{\text{th}}}$ interval (MW)
p^{i}_{m}	water spillage (of no use for generation) in reservoir m over the $i^{\underline{\text{th}}}$ interval
$p(x)$	probability density function of x
q_j	outage probability of unit j
r^{Ik}	expected profit rate (profit over load) in intervals in set Ik
$R^{0}_{=}$	right-hand sides of equalities that refer to energies of more than one interval
$R^{0}_{\geq}$	right-hand sides of inequalities that refer to energies of more than one interval
$R^{i}_{=}$	right-hand sides of equalities that refer only to energies of $i^{\underline{\text{th}}}$ interval
$R^{i}_{\geq}$	right-hand sides of inequalities that refer only to energies of $i^{\underline{\text{th}}}$ interval
S^{0}	surpluses of inequalities that refer to energies of more than one interval
S^{i}	surpluses of inequalities that refer only to energies of $i^{\underline{\text{th}}}$ interval
s_m	water head (m) in reservoir m
$\widetilde{s}^{i}_{m}$	equivalent head in $i^{\underline{\text{th}}}$ interval (head of center of gravity of water layer delimited

	by volumes v_m^{i-1} and v_m^i)
s_{bm},s_{lm},s_{qm},s_{cm}	basic, linear, quadratic, and cubic coefficients of polynomial that approximates water head in terms of stored volume in reservoir m
SGC	specific generation company in a competitive market
t	time in interval (h)
T^i	total duration of $i^{\underline{\text{th}}}$ interval (h)
U	subset of the set of indices of generating units Ω
U_j	subset of set Ω containing the indices $1, 2, \ldots, j$
v_m^i	stored water volume in the $m^{\underline{\text{th}}}$ reservoir at the end of the $i^{\underline{\text{th}}}$ interval
$W(\Omega)$	unsupplied load after loading (in any order) units $j \in \Omega$
w_m^i	natural water inflow in the $m^{\underline{\text{th}}}$ reservoir over the $i^{\underline{\text{th}}}$ interval
WP	(subindex) indication of wind power generation
x	electric power load to be supplied (MW)
λ_{m-s}^{Ik}	Lagrange multiplier of market-sharey constraint of intervals in set Ik
μ^{Ik}	per unit market share of SGC over intervals in set Ik
π_k	probability associated with the $k^{\underline{\text{th}}}$ block probability density function of a multicommodity water inflow, k=1, ..., N_a-1
ρ_m^i	generation efficiency and unit conversion coefficient of $m^{\underline{\text{th}}}$ reservoir
Ω	set of indices of generating units.

References

[1] H. Balériaux, E. Jamoulle, and F. Linard de Gertechin. Simulation de l'explotation d'un parc de machines thermiques de production d'électricité couplé à des stations de pompage. *Revue E de la Soc. Royale Belge des Electriciens*, 7:225–245, 1967.

[2] J.R. Birge and F. Louveaux. *Introduction to stochastic programming.* Springer-Verlag, New York, USA, 1997.

[3] J.A. Bloom and L. Gallant. Modeling dispatch constraints in production cost simulations based on the equivalent load method. *IEEE Transactions on Power Systems*, 9(2):598–611, 1994.

[4] CPLEX Optimization Inc. *Using the CPLEX Callable Library. User's Manual, version 4.0*, 1985.

[5] L.F. Escudero, J. Salmerón, I. Paradinas, and M. Sánchez. SEGEM: a simulation approach for electric generation management. *IEEE Transactions on Power Systems*, 13(3):738–748, 1998.

[6] R. Fourer, D.M. Gay, and B.W. Kernighan. *AMPL a modeling language for mathematical programming.* Boyd and Fraser Publishing Company, Danvers, MA 01293, USA, 1993.

[7] N. Nabona. Multicommodity network flow model for long-term hydro-generation optimization. *IEEE Transactions on Power Systems*, 8(2):395–404, 1993.

[8] N. Nabona and M.A. Díez. Long-term thermal multi-interval thermal generation scheduling using Bloom and Gallant's linear model. Technical Report DR9803, Dept. Estadística i Inv. Operativa, Univ. Politèc. de Catalunya, 08028 Barcelona, 1998.

[9] N. Nabona and C. Gil. The MIBG code of multi–interval long–term thermal generation scheduling using Bloom and Gallant's model. Technical Report DR9901, Dept. Estadística i Inv. Operativa, Univ. Politèc. de Catalunya, 08028 Barcelona, 1999.

[10] N. Nabona, C. Gil, and J. Albrecht. Long-term thermal power planning at VEW ENERGIE using a multi-interval Bloom and Gallant method. *IEEE Transactions on Power Systems*, 16(1):69–77, 2001.

[11] A. Pagès. Modelització amb AMPL de la coordinació tèrmica a llarg termini de la generació elèctrica segons la formulació de Bloom i Gallant. Final-year project, Facultat de Matemàtiques i Estadística, Univ. Politècnica de Catalunya, 08028 Barcelona, 2002.

[12] A. Pagès. Modelització amb AMPL de la coordinació tèrmica a llarg termini de la generació elèctrica segons la formulació de Bloom i Gallant. Supervised assignment, Facultat de Matemàtiques i Estadística, Univ. Politècnica de Catalunya, 08028 Barcelona, 2002.

[13] A. Pagès and N. Nabona. Long-term electric power planning using the Bloom and Gallant's linear model. A comparison of solution techniques. Technical Report DR2002/25, Dept. Estadística i Inv. Operativa, Univ. Politèc. de Catalunya, 08028 Barcelona, 2002.

[14] J. Pérez-Ruiz and A.J. Conejo. Multi-period probabilistic production cost model including dispatch constraints. *IEEE Transactions on Power Systems*, 15(2):502–507, 2000.

[15] R.E. Rosenthal. A nonlinear network flow algorithm for maximization of benefits in a hydroelectric power system. *Operations Research*, 29(4):763–784, 1981.

[16] A.J. Wood and B.F. Wollenberg. *Power generation, operation and control.* John Wiley & Sons, New York, USA, 1984.

[17] S.J. Wright. *Primal-dual interior-point methods.* SIAM, Philadelphia, PA 19104-2688, USA, 1997.