The short-term electricity production management problem at EDF

Grace Hechme-Doukopoulos, Sandrine Brignol-Charousset, Jérôme Malick, Claude Lemaréchal

Grace Hechme-Doukopoulos, Sandrine Brignol-Charousset, Jérôme Malick, Claude Lemaréchal

The short-term electricity production management problem at EDF

1 Introduction

Production management aims at meeting the demand of customers at minimum cost. Since electricity cannot be stored, an electricity producer mainly faces the permanent challenge of matching generation and demand to avoid physical failures of the production system. This article presents how EDF (French Electricity Board) solves this optimization problem: we sketch some issues and the resolution scheme, and we highlight parts of the numerical optimization process of the short term management.

From long- to short-term management

EDF manages a mix of generation units, composed of nearly 60 nuclear plants, 100 classical thermal plants (coal, fuel, gas, combined cycles), and 500 hydraulic plants dispatched in 50 valleys (a hydro-valley is a set of rivers, grouped in such a way that two different valleys are geographically independent). The electricity generation management at EDF consists in determining a strategy (illustrated by Figure 1) which manages the chain of decisions.

The overall decision-making problem is highly complex and way too difficult to be solved globally. Starting from long-term decisions to the short-term ones, a time horizon decomposition amenable to dynamic programming is thus used. Each horizon computes Bellman values for stocks (reservoirs, nuclear plants, ...) to be used by the shorter horizon.

On the longer-term horizons, we take into account uncertainties by using statistical forecast models within a stochastic optimization framework; technical constraints are drastically simplified. By contrast, a very fine vision of the generation mix is necessary for the short term. Real technical constraints make the use of stochastics prohibitive; rather, deterministic counterparts are solved.

The main decisions at each time-horizon are the following ones:

- On the long term (five to twenty years): to design the generation mix: planning investments, choosing the right kind of plants, forecasting polluting emissions, ...
- On the mid term (one to five years): to define the planning for nuclear outages (refueling and overhaul), to calculate management strategies for the main reservoirs (hydraulic reservoirs, demand side management, polluting emission, fuel stocks, ...), to buy fossil fuel and to evaluate the failure risks and associated hedging decisions.

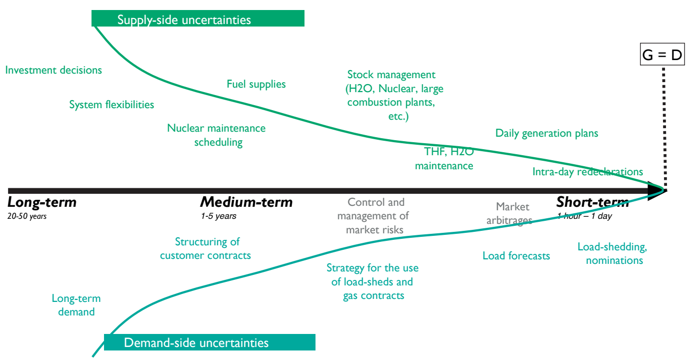

Figure 1. *The EDF generation management decision chain*

- On the short term (a few days to a few hours): to compute production schedules that satisfy the technical constraints and the demand/production equilibrium, and provide marginal costs of the system.

Aspects of short-term management

We focus here on the short-term management problem, commonly called unit-commitment. The decision-making problem is expressed as an optimization problem whose main characteristics are the following:

- *large size* (10^6 variables, 10^6 constraints): all production plants are modeled with a large number of technical constraints and on a 48 hours horizon discretized in half-hourly time steps;
- *nonconvex and noncontinuous* nature: some production costs have discontinuities and the production variables are discrete;
- strict *computational* limits, due to a very tense operational process: the latest data collection phase ends at 12:30 and the feasible schedules have to be sent to the transport system operator at 16:30. Moreover post-optimization treatments involving human expertise are required; altogether, about 15 minutes are left for solving the optimization problem.

This article is mainly devoted to modeling aspects of the problem: we describe the technical characteristics of the production units in some details, and we outline the general solution methodology. For more details on the latter, as well as numerical results, the reader is referred to [4].

The structure of this paper is then as follows. Section 2 presents technical aspects of the production units (thermal and hydraulic), and their mathematical modeling. Section 3 sketches the methods that are used in practice to tackle the production optimization problem. Current issues and research directions are discussed in Section 4.

2 Description of the short-term management problem

2.1 *Overall optimization problem*

The goal of generation management is to compute technically feasible production schedules with a good supply-demand balance at a minimal operating cost.

At a high level, this optimization problem is written as follows. Let T be the discretized time horizon and $D \in \mathbb{R}^T$ the total predicted demand. A production unit v has a production $P_v \in \mathbb{R}^T$, an operating cost c_v, and a technical functioning domain $\mathcal{P}_v \subset \mathbb{R}^T$. The unit-commitment problem is

$$\min \sum_v c_v(P_v), \quad \text{s.t.} \quad P_v \in \mathcal{P}_v, \quad \sum_v P_v = D. \tag{1}$$

The production units are thus subject to two types of constraints: the global linking demand constraint, and local structural constraints. These local constraints are detailed in the next two sections. Then section 3 explains how to deal with all of these constraints when solving problem (1).

2.2 *Constraints and cost for thermal units*

Technical aspects

Thermal units consist of both classical (coal, oil, gas) and nuclear units since their operating domains are similar. We give a simplified description in the following.

When a thermal unit is turned on, the production level must remain between a minimum and a maximum value, which can vary in time. For instance, the maximum level is equal to zero during refueling periods. All levels between the maximum and minimum cannot be used, and the production levels are discrete. Moreover, production variations must follow several rules, namely: a minimal duration between consecutive level changes, upper and lower bound values, and variation prohibition for the rest of the time period after a level decrease.

Switching on or off a thermal unit is not instantaneous. Specific start-up and shut-down curves must be followed, as well as minimum durations of shut-down. There are also some daily constraints: the number of start-ups, shut-downs, and production level variations are limited in a day.

The production cost of a thermal plant is composed by two parts: a fuel cost depending on the generation level, plus some start-up costs depending on the duration of the previous shut-down.

Mathematical formulation

Denote by P_v^t the production level of unit v at time step $t \in T$. Let T_v^{on} [T_v^{off}] be the set of time steps when the unit is online (that is with $P_v^t > 0$) [offline (with $P_v^t = 0$)]. The constraints on the production levels can be expressed as follows:

- Static constraints: N_v being the number of discrete generation levels for v, and $0 < \underline{P}_v^t = P_{v,1}$ and $\overline{P}_v^t = P_{v,N_v}$ the minimal and maximal generation levels, the static constraints are just

$$\forall t \in T_v^{\text{on}}, \quad P_v^t \in \{P_{v,1}, \dots, P_{v,n}, \dots, P_{v,N_v}\},$$

  where $P_{v,1} < \dots < P_{v,n} < \dots < P_{v,N_v}$.
- Dynamic constraints: the three constraints on the production variations are: the minimal duration between level variations,

$$\text{if} \quad P_v^{t+1} \neq P_v^t \quad \text{then} \quad \forall \tau \in [t+2; t+d_v^{\min}], \quad P_v^\tau = P_v^{t+1},$$

  the variation prohibition after a decrease

$$\text{if} \quad P_v^{t+1} - P_v^t < 0 \quad \text{then} \quad \forall \tau \in [t+2; T], \quad P_v^\tau = P_v^{t+1},$$

  and the bound constraints:

$$\underline{\Delta}_v \le \frac{P_v^{t+1} - P_v^t}{d^t} \le \overline{\Delta}_v,$$

  $\underline{\Delta}_v$ and $\overline{\Delta}_v$ being the minimum and maximum values, and d^t the duration of time step t.
- Start-up or shut-down curves constraints: A start-up [shut-down] curve is a set of different generation levels at each time-step of the start-up [shut-down] period:

$$\{0 \le P_v^{\text{start},i} \le \underline{P}_v, \quad i \in [1, d_v^{\text{start}}]\},$$

  and

$$\{0 \le P_v^{\text{stop},i} \le \underline{P}_v, \quad i \in [1, d_v^{\text{stop}}]\},$$

  where d_v^{start} [d_v^{stop}] is the duration of start-up [shut-down]. Note that the particular curve to follow depends on the duration of the previous offline/online period. The constraints are then described as follows:
  - A plant has to follow a minimum duration for any offline period: if $P_v^t \neq 0$ and $P_v^{t+1} = 0$ then:

$$\forall \tau \in [t+1; t+d_v^{\text{stop}}], \quad P_v^\tau = 0.$$

  - A plant has to follow the adequate starting-up curve when going online: if $P_v^t = 0$ and $P_v^{t+1} \neq 0$ then:

$$\forall \tau \in [t+1; t+d_v^{\text{start}}], \quad P_v^\tau = P_v^{\text{start},\tau-t}.$$

  - A plant has to follow the adequate shut-down curve when going offline: if $P_v^t \neq 0$ and $P_v^{t+1} = 0$ then:

$$\forall \tau \in [t+1; t+d_v^{\text{stop}}], \quad P_v^\tau = P_v^{\text{stop},\tau-t}.$$

- The aforementioned daily constraints turn out to make the problem intractable. They are therefore incorporated as penalties p_{daily} in the objective function.

In view of the previous discussions, the total generation cost has the form:

$$c_v(P_v) = \sum_{t \in T_v^{on}} \{c_{v,t}^{\text{fixed}} + c_{v,t}^{\text{prop}} P_v^t + p_{\text{daily}}(P_v^t, P_v^{t-1})\} + c_v^{\text{start}}, \quad (2)$$

where c_v^{start} is a start-up cost depending on the previous offline period.

2.3 Constraints and cost for hydraulic units

Technical aspects

A hydro-plant consists of a set of turbines that discharge water from its upstream reservoir into its downstream one. The reverse is also possible for some plants equipped with pumping units: pumping up water at low demand hours allows one to re-use the water at higher demand ones. Unlike thermal or nuclear units, the production of a hydro-plant is not computed individually. It is rather optimized in a more global entity, a hydro-valley, that depicts the interaction between a set of hydro-plants and the reservoirs connecting them.

The power delivered by a hydro-plant can take only a finite number of values (designated in what follows as discrete production points). These values correspond to the power produced by its turbines that are switched on successively. Since the time period is short, the considered turbines' rates are fixed, just because the water level in the upstream reservoir is considered as constant.

The production of a hydro-valley is subject to a set of constraints that deal with technical functioning aspects, aimed at preventing a fast degradation of the units or simply at following some external regulations. Of course, in addition to the flows induced by pumping or turbining, a reservoir is subject to outer water inputs due to rain, snow or spillage. Hence, through the time period, the volume of a reservoir is governed by an equilibrium flow constraint that rules these factors.

As for the power plants, their production variations are subject to upper and lower bound constraints. A minimal delay of one hour is also imposed between two production variations of opposite nature. Furthermore, when two reservoirs are connected with both turbining and pumping plants, simultaneous pumping and turbining is forbidden, and a minimal halt of thirty minutes before switching from pumping into turbining (and vice versa) is imposed.

Mathematical formulation

Given a hydro-valley v, we denote by U its set of production plants and by R its reservoirs. Each plant $u \in U$ is described by a set of turbining/pumping units $G(u)$. Each unit is characterized by its flow capacity $\overline{F}_{u,g}$ and its power rate $\rho_{u,g}$; units are ranked according to decreasing ρ's. At time step t, the state of unit g is given by a binary variable $e_{u,g}^t \in \{0,1\}$ so that we have

$$F_u^t = \sum_{g \in G(u)} e_{u,g}^t \overline{F}_{u,g}, \quad \text{(the flow capacity of } u\text{)}$$

$$P_u^t = \sum_{g \in G(u)} e_{u,g}^t \rho_{u,g} \overline{F}_{u,g} \quad \text{(the power of } u\text{).}$$

For all u, the binary variables must follow a sequence constraint:

$$\forall t, \quad \forall g \in G(u), \qquad e_{u,g+1}^t \le e_{u,g}^t,$$

to have P_u^t equal to a discrete production point at each time step t. Note that it is then sufficient to apply the pumping/turbining technical constraints for $g = 1$ only.

Denote by u_T and u_P the turbining and pumping plants respectively. The constraints

$$\forall t, \quad e^t_{u_T,1} + e^t_{u_P,1} \leq 1$$

prohibit simultaneous pumping and turbining, while the minimal halt delay before a flow mode switch can be imposed by:

$$\forall t \in [1, T-1], \quad e^t_{u_P,1} + e^{t+1}_{u_T,1} \leq 1, \quad e^{t+1}_{u_P,1} + e^t_{u_T,1} \leq 1.$$

The production variation constraints are expressed as:

$$\forall t \in [2, T-1], \quad \forall g \in G(u), \quad -1 \leq e^t_{u,g} - e^{t-1}_{u,g} - e^{t+1}_{u,g} \leq 0,$$

for the minimal delay between production variations, and

$$\forall t \in [1, T-1], \quad \underline{\delta_u} \leq \sum_{g \in G(u)} (e^{t+1}_{u,g} - e^t_{u,g}) \overline{F}_{u,g} \leq \overline{\delta_u}$$

for the bound constraints, $\underline{\delta_u}$ and $\overline{\delta_u}$ being the lower and upper bounds respectively.

Denote by V^t_r the volume of reservoir $r \in R$ at time step t. The flow constraint has the following form:

$$V^t_r = V^{t-1}_r + \sum_{u \in N_\uparrow(r)} F^{t-d(u,r)}_u - \sum_{u \in N_\downarrow(r)} F^{t+d(r,u)}_u + I^t_r,$$

where $N_\uparrow(r)$ $[N_\downarrow(r)]$ is the set of hydro-plants up [down] reservoir r, I^t_r is the outer water input, and $d(u,r)$ is the travel time of water between reservoir r and plant u and vice versa. Note finally that the volume is also subject to bound constraints (resulting from the hydraulicity, environment, or regulations due to the recreational use of the reservoir).

The production cost of a hydro-valley is the global water loss through the time horizon:

$$c_v(P_v) = \sum_{r \in R} \omega_r (V^0_r - V^T_r).$$

The value ω_r of water of reservoir r is estimated with marginal indicators resulting from mid-term models, giving the future gain if the water is not discharged (remember Figure 1).

3 Optimization methods

3.1 Solving the overall problem via decomposition

In our overall optimization problem (1), v indices the thermal plants of Section 2.1 and the hydro-valleys of Section 2.2. Then we see that each $\mathcal{P}_v$ depends on no other $P_{v'}$, and that the balance constraint $\sum_v P_v = D$ is the only link between the "local agents" P_v. The problem is thus clearly decomposable and, as already seen in [1], Lagrangian relaxation is an attractive approach. Thus, for given dual variable $\lambda \in \mathbb{R}^T$, (1) is replaced by the *decomposed* problem

$$\begin{aligned} \theta(\lambda) \quad &:= \min_{P_v \in \mathcal{P}_v} \sum_v c_v(P_v) + \lambda \cdot \Big(D - \sum_v P_v\Big) \\ &= \lambda \cdot D + \sum_v \min_{P_v \in \mathcal{P}_v} (c_v(P_v) - \lambda \cdot P_v) \end{aligned} \tag{3}$$

and the issue becomes that of finding an adequate λ, so as to reproduce a solution of (1). To this end, duality theory (see, e.g., [6, 8, 5, 9]) tells us that:

(i) If an optimal solution $P(\lambda)$ of (3) is feasible in (1), then it is also optimal in (1).
(ii) To achieve this, λ must maximize the *dual function* θ of (3).
(iii) This function is concave, $D - \sum_v P_v(\lambda) \in \mathbb{R}^T$ being a subgradient (of $-\theta$ at λ).
(iv) However, the converse in (ii) is false: finding a $\hat{\lambda}$ maximizing θ does not necessarily yield a primal optimal $P(\hat{\lambda})$. Usually, $\sum_v P_v(\lambda) \neq D$ for any $\lambda \in \mathbb{R}^T$, even for $\lambda = \hat{\lambda}$.
(v) Nevertheless, the *dual problem* (ii) does provide a certain primal point $\hat{P} = \{\hat{P}_v\}_v$ which solves a certain *convexified* form of (1); the $\hat{P}_v$'s need not lie in $\mathcal{P}_v$ but $\sum_v \hat{P}_v = D$.
(vi) Besides, a dual optimum $\hat{\lambda}$ gives the marginal cost of the linking constraints $\sum_v P_v = D$, associated with the convexification alluded to in (v).

Accordingly, a two-phase strategy is adopted.

- *Phase I – maximizing* $\boldsymbol{\theta(\lambda)}$. Property (iii) makes (ii) possible and common approaches use the popular subgradient algorithm. Instead, the present operational software uses the bundle algorithm of [10], which in turn uses a quadratic solver written by K. C. Kiwiel. A first advantage is *robustness*: a reliable dual optimum $\hat{\lambda}$ is computed which, according to (vi), provides useful information on the marginal prices of the demand D in (1). Besides, a primal point $\hat{P}$ as described in (v) is also obtained. We will see in the forthcoming sections that hydraulic valleys cannot always be optimized exactly; this results in a noisy θ, which is handled by the technique of [7].
- *Phase II – producing schedules*. Because of (iv), solving the dual problem as above can only be viewed as a first step toward solving (1). It is therefore followed by a second phase, aimed at computing schedules that do lie in $\mathcal{P}_v$, while realizing a good compromise between minimizing the cost and satisfying the balance equation. The method currently used is based on *augmented Lagrangian* [3] where a quadratic stabilization term is added to the dual function. Since the quadratic term destroys the decomposability property, a "partial linearization" as in [2] is applied. Altogether the local problems of (3) are replaced by

$$\min_{P_v \in \mathcal{P}_v} (c_v(P_v) - \lambda \cdot P_v + r|P_v - Q_v|^2) \tag{4}$$

where the "stability center" Q_v is just the previous iterate P^{k-1}_v (thus an initialization P^0 is required).

In order to assess schedules satisfying all technical constraints of Section 2, while not matching the linking constraints, a "total cost"

$$C(P) := \sum_v c_v(P_v) + \pi\Big(D - \sum_v P_v\Big) \tag{5}$$

is introduced, where π penalizes the balance mismatch. Convergence of the model is measured as the gap between the value $\theta(\hat{\lambda})$ of the dual function computed in Phase I and the total cost of the primal solution obtained in Phase II (from weak duality, this gap gives a bound for the optimal cost C).

3.2 Solving the sub-problems

Thermal units

A standard MIP formulation would be quite complex and require considerably high computational time to obtain a satisfactory solution. A specific dynamic programming approach has been developed to solve the thermal sub-problems, which can be outlined as follows.

A four-dimensional state $\{S_1, S_2, S_3, S_4\}$ is defined: S_1 represents the online/offline state ($S_1 = 1$ if the plant is online, $S_1 = 0$ if not), $S_2 = P_v$ is the discrete production level, S_3 represents the sign of the last production variation ($S_3 = 1$ for an increase, 0 otherwise), $S_2 = d_v$ represents the duration of the current online/offline period. At each time step t, the set of authorized states is calculated taking into account all the production constraints (limits on the production levels, halts ...). Authorized transitions between states at different timesteps are then computed taking into account the timing constraints (minimum durations, start-up curves ...).

Costs associated to each transition are then calculated: a transition to an offline state costs 0; a transition to a generation state is associated to the cost

$$c_{v,t}^{\text{fixed}} + c_{v,t}^{\text{prop}} P_v^t + p_{\text{daily}}(P_v^t, P_v^{t-1}) - \lambda_t P_v^t$$

in the Lagrangian (2) (3); a transition to a start-up curve is associated to the sum of the generation cost and the start-up cost.

Dynamic programming is then applied, calculating backward the Bellman value of each state. The graph is a set of nodes and a set of transitions between nodes. Call Tr_i the set of transitions going to node i; then

$$Vb_i^v = \min_{\mathrm{tr}\in\mathrm{Tr}_i} [C(\mathrm{tr}) + Vb_{\mathrm{tr}^{-1}(i)}^v]$$

is the Bellman value of node i, where $\mathrm{tr}^{-1}(i)$ is the node connected to node i by transition tr, $C(\mathrm{tr})$ is the cost associated to transition tr.

Hydro-valleys
The hydraulic sub-problems, modeled as MIP's described in Section 2.2, cannot be solved to optimality in a reasonable computing time. This is the reason why three versions of the model are at stake:

(a) the actual implementation currently in operation uses a continuous relaxation; combinatorial constraints are taken into account only in Phase II with the help of various heuristics;
(b) a forthcoming version will keep the continuous relaxation in Phase I only, while Phase II will solve inaccurately the actual model 2.2;
(c) a third version is planned, where inaccurate solutions of the actual model will be computed in both phases.

Effort is currently focused on version (b); Section 4 will explain why version (c) has been postponed so far.

3.3 Numerical illustration

Let us illustrate the behaviour of the forthcoming version b) above. The code is distributed and runs on a cluster of 32 processors. It performs around 500 total iterations for each phase. An iteration takes about 0.1 seconds for Phase I, and 1.5 seconds for Phase II (remember that Phase I solves "easy" hydraulic subproblems). The total computational time of one run is therefore around 900 seconds.

On average over one year, Phase I optimizes θ within 0.1 % and produces a convexified $\hat{P}$ satisfying the balance constraint within 0.1 % as well. Phase II produces a schedule $\bar{P}$, whose total cost $C(\bar{P})$ is optimal within 1.3 %.

Two figures below give an idea of the progression of the algorithm, for both phases. The evolution of the dual function (3) through Phase I is depicted in Figure 2. As for Phase II, Figure 3 displays two curves: the balance mismatch $|\sum P_v - D|$ (averaged over the time period $[0, T]$), and the total cost of (5). Both figures use a logarithmic scale; for industrial privacy, all values have been normalized as follows:

$$\text{for} \quad n = 1, \ldots, N, \quad a_n \to \hat{a}_n = \frac{a_n - \underline{a}}{a_1 - \underline{a}}, \quad \text{with } \underline{a} = \min(a_n).$$

Good convergence behavior is sometimes difficult to achieve (especially in Phase II). For instance, we can remark in Figure 3 that the best solution was found at iteration 414 ($\hat{a}_{414} = 0$), and the method carried on for almost 100 additional iterations without being able to improve it.

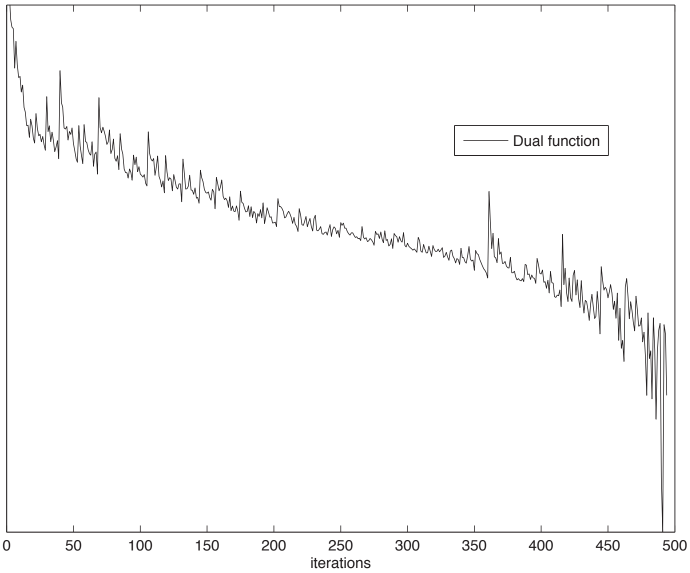

Figure 2. *Evolution of the (normalized) dual function along the iterations*

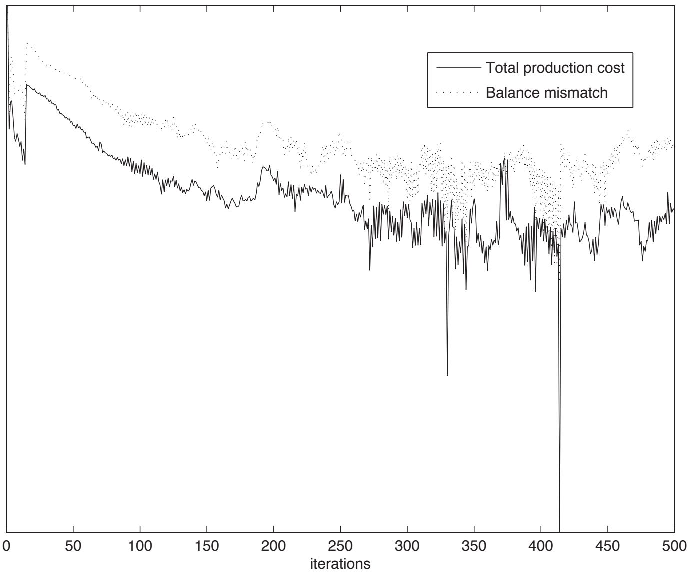

Figure 3. *Evolution of the (normalized) balance mismatch and total production cost*

4 Perspectives

Including all possible real-life constraints of an electricity production management problem is pure dream (remember the human postprocessing mentioned in Section 1). The model is therefore regularly improved, in order to better reflect reality and to achieve better numerical performances.

For instance, Phase I of the actual implementation (a) "sees" a less constrained hydraulic model; implementing version (c) should be desirable. However, it turns out that inserting the true model in Phase I results in chaotic $\hat{\lambda}$. This undesired behaviour might be due to stiff hydraulic constraints (which impacts the balance constraints), or to inaccurate computations of the dual function (3). Further analysis is needed to fix this question; an important question because Phase I not only initializes Phase II but also provides marginal indicators about costs of the demand ($\hat{\lambda}$ mentioned in (vi) of Section 3.1).

Another point concerns Phase II, which uses a local search method strongly dependent on the initial point. Moreover the algorithm does not have guaranteed convergence properties on this nonconvex problem. Alternative methods are worth investigating to improve this phase.

Grace Hechme-Doukopoulos, EDR R&D, 1 avenue du Général de Gaulle, 92141 Clamart, France. grace.doukopoulos@edf.fr
Sandrine Brignol-Charousset, EDR R&D, 1 avenue du Général de Gaulle, 92141 Clamart, France. sandrine.charousset@edf.fr
Jérôme Malick, CNRS, Lab. J. Kunztmann, INRIA, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier, France. jerome.malick@inria.fr
Claude Lemaréchal, INRIA, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier, France. claude.lemarechal@edf.fr

References

[1] D.P. Bertsekas, G.S. Lauer, N.R. Sandell, and T.A. Posberg. Optimal short-term scheduling of large-scale power systems. *IEEE Transactions on Automatic Control*, AC-28:1–11, 1983.
[2] G. Cohen. Auxiliary problem principle and decomposition of optimization problems. *Journal of Optimization Theory and Applications*, 32:277–305, 1980.
[3] G. Cohen and D. L. Zhu. Decomposition coordination methods in large scale optimization problems. the nondifferentiable case and the use of augmented Lagrangians. In J.B. Cruz, editor, *Advances in Large Scale Systems*, 1, pages 203–266. JAI Press Inc., Greenwich, Connecticut, 1984.
[4] L. Dubost, R. Gonzalez, and C. Lemaréchal. A primal-proximal heuristic applied to the french unit-commitment problem. *Mathematical Programming*, 104(1):129–152, 2005.
[5] A.M. Geoffrion. Duality in nonlinear programming: a simplified applications-oriented development. *SIAM Review*, 13(1):1–37, 1971.
[6] R.C. Grinold. Lagrangian subgradients. *Management Science*, 17(3):185–188, 1970.
[7] K.C. Kiwiel. A proximal bundle method with approximate subgradient linearizations. *SIAM Journal on Optimization*, 16(4):1007–1023, 2006.
[8] L. Lasdon. *Optimization Theory for Large Systems*. Macmillan Series in Operations Research, 1970.
[9] C. Lemaréchal. Lagrangian relaxation. In M. Jünger and D. Naddef, editors, *Computational Combinatorial Optimization*, pages 112–156. Springer Verlag, 2001.
[10] C. Lemaréchal and C. Sagastizábal. Variable metric bundle methods: from conceptual to implementable forms. *Mathematical Programming*, 76(3):393–410, 1997.

Discussion Column

Antonio Frangioni

Unit Commitment problems: A tale in Lagrangian optimization

The paper provides an account of the current status of a multi-luster collaboration between academia and industry about the application of Lagrangian techniques for the solution of Unit Commitment (UC) problems in electrical power production.

UC problems have played an important role, perhaps as significant as that of multicommodity flows, to popularize solution techniques based on Lagrangian relaxation both in the mathematical programming community and within practitioners (in particular, in this case in the electrical engineering community). This is due to a number of factors:

- UC is indeed a large-scale problem with both nonlinear and discrete components. As such, it has been until recently firmly out of reach of solution techniques based on general-purpose mixed-integer solvers, and specialized approaches have been a necessity. However, single-unit scheduling problems, both for thermal and hydro units, are relatively easy to solve with appropriate approaches (typically dynamic programming and flow/linear programming techniques).
- As a consequence, UC, at least in some of its "easiest" versions, is incredibly well-suited for the technique. Lower bounds obtained by Lagrangian relaxation can have a ludicrously small inherent gap, as low as a small fraction of a percentage point, especially for the largest hydro-thermal instances that used to be the norm in several relevant application environments (and still are in that described in the Scientific Contribution). Each Lagrangian iteration is quite fast, which coupled with a good method to update the multipliers produces these terrific bounds quickly enough.
- UC has really tight operational constraints, making the development of methods capable of quickly and reliably producing good-quality solution of utmost importance.
- UC is, or at least used to be, the almost perfect example of an "easy sell" for advanced applied research. The problem is huge in terms of costs involved and has to be solved daily, thus small savings rapidly add up to incredibly vast sums. Optimizing the schedules of production units has no noticeable negative effect, and therefore no meaningful opposition by any of the parties involved. The problem used to be of concern of a single monopolistic producer – often state-owned – with almost unlimited financial resources.

This is not to say that UC is an easy problem. Indeed, successful application of Lagrangian techniques to UC (as well as to other difficult combinatorial problems) requires several nontrivial steps which have motivated relevant theoretical contributions, to which the authors of the Scientific Contribution are by no means unconnected:

- algorithmic recovery [7] of the continuous solutions of the "primal counterpart" of the Lagrangian Dual [8];
- fast converging bundle methods using techniques like disaggregation in the master problem and preconditioning [2];
- general yet efficient approaches for recovering primal feasible solutions out of a Lagrangian dual [3], comprised an entire new class of Lagrangian-based heuristics [6].

All this illustrates a very nice instance of a central credence (hope? wishful thinking?) in the mathematical programming community, and in applied mathematics in general: important practical problems motivate relevant theoretical developments, sophisticated mathematical theory is necessary to solve crucial practical applications.

Research on UC problems is by no means over, due to several factors:

- Most countries in the world have been transitioning from electrical systems based on monopolistic producers to those based on free market, where competition between producers and consumers (regulated by a central authority, typically taking care of the electrical network) is supposed to increase the overall efficiency. In this setting, the UC problem takes different forms for different actors, sometimes complicating matters considerably (most of the system is in the hand of other decision makers, whose behavior is unknown), sometimes simplifying them somewhat (each decision maker needs to model only his units, which are a subset of the whole system, and the effects of the constraints on the transmission network is somewhat less pronounced).
- It is possible (perhaps desirable) that the future will bring very relevant changes in the characteristics of the generating units. Other than several hundreds of large and relatively reliable plants, many thousands of smaller and less reliable units based on renewables – or even much more with fuel cells cars in the envisaged hydrogen economy – may have to be taken into account. Distribution grids may also undergo substantial updates to the so-called "smart grids". All this may clearly have very profound impacts on the UC models to be solved.