Optimal Power Dispatch via Multistage Stochastic Programming *

M.P. Nowak[1] and W. Römisch[1]

Abstract. The short-term cost-optimal dispatch of electric power in a generation system under uncertain electricity demand is considered. The system comprises thermal and pumped-storage hydro units. An operation model is developed which represents a multistage mixed-integer stochastic program and a conceptual solution method using Lagrangian relaxation is sketched. For fixed start-up and shut-down decisions an efficient algorithm for solving the multistage stochastic program is described and numerical results are reported.

1 Introduction

Mathematical models for cost-optimal power scheduling in hydro-thermal systems often combine several difficulties such as a large number of mixed-integer variables, nonlinearities, and uncertainty of problem data. Typical examples for the latter are uncertain prices in electricity trading, the future electric power demand, and future inflows into reservoirs of hydro plants. Incorporating the uncertainties directly into an optimization model leads to stochastic programming problems. In the context of power scheduling such models are developed e.g. in [4], [5], [9], [11].

In the present paper we consider a short-term optimization model for the dispatch of electric power in a hydro-thermal generation system over a certain time horizon in the presence of uncertain demand. The generation system comprises (coal-fired and gas-burning) thermal and pumped-storage hydro units (without inflows) which is typical for the eastern part of Germany. Short- and long-term energy contracts are regarded (and modelled) as (particular) thermal units. The operation of such a generation system is very complex, because it creates a link between a decision in a given time interval and the future consequences of this decision. Even for optimal on-line power scheduling future costs created by actual decisions have to be taken into account. Since a longer time horizon (e.g. one

[1]Humboldt University Berlin, Institute of Mathematics, 10099 Berlin, Germany

*This research is supported by the Schwerpunktprogramm “Echtzeit-Optimierung großer Systeme” of the Deutsche Forschungsgemeinschaft (DFG)

week) is often needed due to the pumping cycle of the hydro storage plants, the stochastic nature of the demand cannot be ignored. The optimization model thus represents a multistage stochastic program containing mixed-integer (stochastic) decisions which reflect the on/off schedules and production levels of the generating units for all time intervals of the horizon. The increase of scenarios in the stochastic programming model corresponds to a decrease of information on the power demand.

The stochastic model will be developed and discussed in some detail in section 2 (for more information we refer to [3]). In section 3 we sketch a conceptual decomposition method by applying Lagrangian relaxation to the loosely coupled multistage stochastic program, and in section 4 an efficient algorithm for solving the stochastic program for fixed on/off decisions is described and numerical results are reported.

2 Stochastic Model

The mathematical model represents a mixed-integer multistage stochastic program with linear constraints. Let T denote the number of (hourly or shorter) time intervals in the optimization horizon and $\{\mathbf{d}^t \;:\; t = 1, \ldots, T\}$ the stochastic demand process (on some probability space $(\Omega, \mathcal{A}, P)$) reflecting the stochasticity of the electric power demand. It is assumed that the information on the demand is complete for $t = 1$ and that it decreases with increasing t. This is modelled by a filtration of σ-fields

$$\mathcal{A}_1 = \{\emptyset, \Omega\} \subseteq \mathcal{A}_2 \subseteq \ldots \subseteq \mathcal{A}_t \subseteq \ldots \subseteq \mathcal{A}_T \subseteq \mathcal{A},$$

where $\mathcal{A}_t$ is the σ-field generated by the random vector $(\mathbf{d}^1, \ldots, \mathbf{d}^t)$. Let I and J denote the number of thermal and pumped-storage hydro units in the system, respectively. According to the stochasticity of the demand process the decisions for all thermal and hydro units

$$\begin{array}{ll} \{(\mathbf{u}_i^t, \mathbf{p}_i^t) \;:\; t = 1, \ldots, T\} & (i = 1, \ldots, I) \\ \{(\mathbf{s}_j^t, \mathbf{w}_j^t) \;:\; t = 1, \ldots, T\} & (j = 1, \ldots, J) \end{array}$$

are also stochastic processes being adapted to the filtration of σ-fields. The latter condition means that the decisions at time t only depend on the demand vector $(\mathbf{d}^1, \ldots, \mathbf{d}^t)$ (nonanticipativity). Here, $\mathbf{u}_i^t \in \{0, 1\}$ and $\mathbf{p}_i^t$ denote the on/off decision and the production level for the thermal unit i and time interval t, respectively, and $\mathbf{s}_j^t$, $\mathbf{w}_j^t$ are the generation and pumping levels, for the pumped-storage plant j during time interval t, respectively. Further, let $\mathbf{l}_j^t$ denote the water level (in terms of electrical energy) in the upper reservoir of plant j at the end of interval t.

The objective function is given by the expected value of the total fuel and start-up costs of the thermal units

$$\mathbb{E}\left[\sum_{i=1}^{I}\sum_{t=1}^{T} FC_i(\mathbf{p}_i^t, \mathbf{u}_i^t) + SC_i(\mathbf{u}_i(t))\right] \tag{2.1}$$

where $\mathbb{E}$ denotes the expectation, FC_i the fuel cost function and SC_i the start-up costs for the operation of the i-th thermal unit. It is assumed that the functions FC_i are monotonically increasing and piecewise linear convex with respect to $\mathbf{p}_i^t$ and that $SC_i(\mathbf{u}_i(t))$ is determined by $(\mathbf{u}_i^t, \ldots, \mathbf{u}_i^{t_{s_i}})$ where $t - t_{s_i}$ is the preceding down-time of the unit i (see e.g. [12] for typical start-up cost functions). All the (stochastic) variables mentioned above have finite lower and upper bounds reflecting unit capacity limits and reservoir capacities of the generation system:

$$\begin{array}{c} p_{it}^{min}\mathbf{u}_i^t \leq \mathbf{p}_i^t \leq p_{it}^{max}\mathbf{u}_i^t, \; i = 1, \ldots, I, \; t = 1, \ldots, T \\ 0 \leq \mathbf{s}_j^t \leq s_{jt}^{max}, 0 \leq \mathbf{w}_j^t \leq w_{jt}^{max}, 0 \leq \mathbf{l}_j^t \leq l_{jt}^{max}, j = 1, \ldots, J, \; t = 1, \ldots, T \end{array} \tag{2.2}$$

The constants p_{it}^{min}, p_{it}^{max}, s_{jt}^{max}, w_{jt}^{max} and l_{jt}^{max} denote the minimal/maximal outputs and maximal water levels in the upper reservoir, respectively. During the whole time horizon reservoir constraints have to be maintained for all pumped storage plants. These are modelled by the equations:

$$\begin{array}{ll} \mathbf{l}_j^t = \mathbf{l}_j^{t-1} - \mathbf{s}_j^t + \eta_j \mathbf{w}_j^t, & t = 1, \ldots, T, \; j = 1, \ldots, J \\ \mathbf{l}_j^0 = l_j^{in}, \; \mathbf{l}_j^T = l_j^{end}, & j = 1, \ldots, J. \end{array} \tag{2.3}$$

Here, l_j^{in} and l_j^{end} denote the initial and terminal water level in the upper reservoir, respectively, and η_j is the efficiency of the j-th pumped-storage plant. Moreover, there are minimum down times τ_i and possible must-on/off constraints for each thermal unit i. Minimum down times are imposed to prevent the thermal stress and high maintenance costs due to excessive unit cycling. They are described by the inequalities:

$$\mathbf{u}_i^{t-1} - \mathbf{u}_i^t \leq 1 - \mathbf{u}_i^\tau, \; \tau = t+1, .., \min\{t + \tau_i - 1, T\}, \; i = 1, .., I, \; t = 2, .., T. \tag{2.4}$$

Load coverage for each time interval t of the horizon is described by the equations:

$$\sum_{i=1}^{I} \mathbf{p}_i^t + \sum_{j=1}^{J} \left(\mathbf{s}_j^t - \mathbf{w}_j^t \right) = \mathbf{d}^t, \; t = 1, \ldots, T. \tag{2.5}$$

In order to compensate sudden load increases or unforeseen events on-line, some spinning reserve level $\mathbf{r}^t$ for the thermal units is required leading to the constraints:

$$\sum_{i=1}^{I} \left(p_{it}^{max}\mathbf{u}_i^t - \mathbf{p}_i^t \right) \geq \mathbf{r}^t, \; t = 1, \ldots, T. \tag{2.6}$$

Altogether, (2.1)-(2.6) represents a multistage stochastic program with $2(I+J)T$ stochastic decision variables. For large power generation systems like that of VEAG Vereinigte Energiewerke AG in the eastern part of Germany (with $I = 25$, $J = 7$ and $T = 168$ which corresponds to an hourly discretization of one week) these models involve an enormous number of stochastic decisions.

Numerical approaches for solving (2.1)-(2.6) are mostly based on designing discretization schemes (scenario trees) for the probability distribution of the random demand vector $(\mathbf{d}^1, \ldots, \mathbf{d}^T)$ and lead to large-scale mixed-integer programs with a huge number of variables. In general, such problems are too large from the viewpoint of even the latest solution techniques for multistage stochastic programs with discrete distributions. However, it is possible to make use of the fact that the problem (2.1)-(2.6) is loosely coupled via the constraints (2.5) and (2.6) with respect to the operation of different units.

3 Lagrangian relaxation

For deterministic models of the form (2.1)-(2.6) (i.e. for $\mathcal{A}_T = \{\emptyset, \Omega\}$) the authors of [10] came to the conclusion that for solving realistic problems a clear consensus is presently tending toward the Lagrangian relaxation approach over other methodologies. The approach is based on introducing a partial Lagrangian for constraints linking the operation of different units and on solving the nondifferentiable concave dual maximization problem by modern nonsmooth optimization methods. This approach is also suggested for certain multistage stochastic models in [11] by relying on the same arguments as in the deterministic case, namely, that the dual problems decompose into single unit subproblems and the duality gap becomes small under certain circumstances (see [1], [3]).

In order to describe the Lagrangian relaxation approach for the model (2.1)-(2.6), let $\{\boldsymbol{\lambda}^t : t = 1, \ldots, T\}$ and $\{\boldsymbol{\mu}^t : t = 1, \ldots, T\}$ be stochastic processes in $L^1(\Omega, \mathcal{A}, P; I\!R^T)$ adapted to the filtration and consider the (partial) Lagrangian:

$$L(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w}; \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbb{E} \sum_{t=1}^T \Big\{ \sum_{i=1}^I \{FC_i(\mathbf{p}_i^t, \mathbf{u}_i^t) + SC_i(\mathbf{u}_i(t))\}$$
$$+\boldsymbol{\lambda}^t \left(\mathbf{d}^t - \sum_{i=1}^I \mathbf{p}_i^t - \sum_{j=1}^J (\mathbf{s}_j^t - \mathbf{w}_j^t)\right) + \boldsymbol{\mu}^t \left(\mathbf{r}^t - \sum_{i=1}^I (\mathbf{u}_i^t p_{it}^{max} - \mathbf{p}_i^t)\right) \Big\}. \quad (3.1)$$

The dual problem then reads

$$\max \{D(\boldsymbol{\lambda}, \boldsymbol{\mu}) \; : \; \boldsymbol{\mu} \geq 0, \boldsymbol{\lambda}\}, \quad (3.2)$$

where $D(\boldsymbol{\lambda}, \boldsymbol{\mu})$ denotes the infimum of $L(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w}; \boldsymbol{\lambda}, \boldsymbol{\mu})$ subject to $(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w})$ satisfying the constraints (2.2)-(2.4). Since (2.2)-(2.4) represent exclusively single unit constraints, the Lagrangian dual function can be written as:

$$D(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \sum_{i=1}^I D_i(\boldsymbol{\lambda}, \boldsymbol{\mu}) + \sum_{j=1}^J \hat{D}_j(\boldsymbol{\lambda}) + \mathbb{E} \sum_{t=1}^T \left[\boldsymbol{\lambda}^t \mathbf{d}^t + \boldsymbol{\mu}^t \mathbf{r}^t\right].$$

where the functions D_i and $\hat{D}_j$ are defined by

$$D_i(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min_{(\mathbf{u}_i, \mathbf{p}_i)} \mathbb{E} \sum_{t=1}^T \left[FC_i(\mathbf{p}_i^t, \mathbf{u}_i^t) + SC_i(\mathbf{u}_i(t)) - (\boldsymbol{\lambda}^t - \boldsymbol{\mu}^t)\mathbf{p}_i^t - \boldsymbol{\mu}^t \mathbf{u}_i^t p_{it}^{max}\right] \quad (3.3)$$

$$\hat{D}_j(\boldsymbol{\lambda}) = \min_{(\mathbf{s}_j,\mathbf{w}_j)} \mathbb{E} \sum_{t=1}^{T} \left[-\boldsymbol{\lambda}^t \left(\mathbf{s}_j^t - \mathbf{w}_j^t \right) \right] \tag{3.4}$$

and the corresponding minimization is carried out subject to the constraints (2.2), (2.4) and (2.2), (2.3), respectively.

To solve the dual problem, an iterative bundle-type method ([6], [7]) is used for updating the Lagrange multipliers $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ and maximizing the concave function D, respectively. For given $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ the value of $D(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is computed by solving the single thermal unit subproblem (3.3) by dynamic programming (as described in [11]) and the single hydro subproblem (3.4) by a fast descent method developed in [8]. After solving the dual problem a heuristic approach is applied to obtain primal decisions $(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w})$ which are feasible for (2.5) and (2.6). This approach consists in a modification of the search for reserve-feasible solutions in [12] for the case of hydro-thermal systems.

4 Economic Dispatch

Having the binary variables $\mathbf{u}_i^t$ fixed, one has to solve the minimization problem with respect to $\mathbf{p}_i^t$, $\mathbf{s}_j^t$ and $\mathbf{w}_j^t$, i.e. the economic dispatch problem. It means that the objective function

$$\mathbb{E} \sum_{t=1}^{T} \sum_{i=1}^{I} FC_i(\mathbf{p}_i^t, \mathbf{u}_i^t) \tag{4.1}$$

has to be minimized subject to the constraints (2.2),(2.3),(2.5) and (2.6).

Taking the right-hand side of $\sum_{i=1}^{I} \mathbf{p}_i^t = \mathbf{d}^t - \sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t)$ as a parameter $\mathbf{v}^t$ the problem for one time period and one realization of $\mathbf{v}^t$ reads:

$$\text{minimize} \sum_{i=1}^{I} FC_i(\mathbf{p}_i^t, \mathbf{u}_i^t) \text{ s.t. } \sum_{i=1}^{I} \mathbf{p}_i^t = \mathbf{v}^t,\ \mathbf{u}_i^t p_{it}^{min} \leq \mathbf{p}_i^t \leq \mathbf{u}_i^t p_{it}^{max},\ i = 1, .., I \tag{4.2}$$

Since $FC_i(\mathbf{p}_i^t, \mathbf{u}_i^t)$ are piecewise linear with respect to $\mathbf{p}_i^t$, the optimal value function $\phi^t(\mathbf{v}^t)$ of (4.2) is piecewise linear, too. Sorting the segments (see figure 1) of the cost functions of all thermal units, the computation of $\phi^t(\mathbf{v}^t)$ consists of a look up in a list. Then the problem (4.1) consists in minimizing

$$\mathbb{E} \sum_{t=1}^{T} \phi^t \left(\mathbf{d}^t - \sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t) \right) \tag{4.3}$$

subject to (2.2), (2.3).

This problem can be solved by a modification of the algorithm described in [8]. The crucial point in this descent algorithm consists in selecting a direction from a prescribed subset of descent directions. Since the objective function is not linear as in [8], the algorithm has to regard the kinks of the piecewise linear function. At

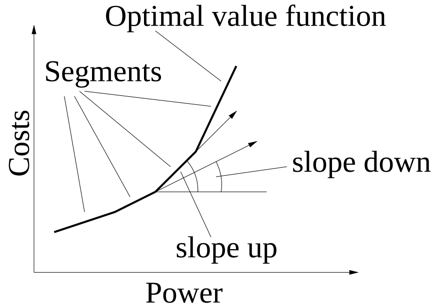

Figure 1: overall cost function

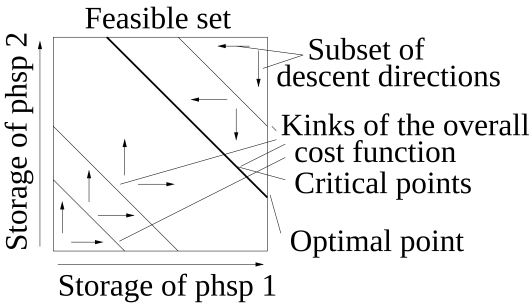

Figure 2: piecewise linear function

these kinks two different slopes (up and down in figure 1) have to be considered. Since the kinks of the objective function (see figure 2) do not coincide with the sufficiently large subset of directions of [8], one has to avoid critical points.

Under the additional assumption $l_i^{in} = l_i^{end}$, $i = 1, \ldots, I$ the point $\mathbf{s}_i^t \equiv \mathbf{w}_i^t \equiv 0$ is feasible. With this point as a starting point and choosing the direction of steepest descent the algorithm converges to an optimal solution.

5 Computational results

The algorithm described above is implemented in *C++*. The performance of the corresponding code ECDISP has been compared with CPLEX 4.0 [2] on several examples. Here we report computation times for both codes on an example including 25 thermal power units, 8 pumped hydro storage plants, 192 stages and 1 scenario. This corresponds to a linear program with 14200 columns, 17856 rows, 46256 nonzero elements of the matrix. The computation time of ECDISP is 50.95 seconds. For CPLEX we display the computation times (in seconds) for different methods and pricing strategies.

CPLEX function	Pricing strategy primal/dual					
	-1	0	1	2	3	4
Simplex/primal	1232.47	1188.4	1918.15	2664.14	2440.7	1696.9
Simplex/dual		1086.18	946.24	1103.48	1466.54	1083.8
baropt	94.78					
hybbaropt/primal	114.71	114.32	114.36	486.55	114.45	114.35
hybbaropt/dual		115.08	114.69	693.03	1424.86	114.84
hybnetopt/primal	957.66	910.39	1298.03	2252.83	1960.93	1162.68
hybnetopt/dual		1393.82	1253.76	1412.06	1833.96	1392.3

All these computations are done on a SPARCstation IPX (4/50) with 64 MB Main Memory and 40 MHz CPU-frequency.

Further comparisons with several numbers of scenarios are performed, too. At this time the code ECDISP is compared with the *baropt* function of CPLEX only.

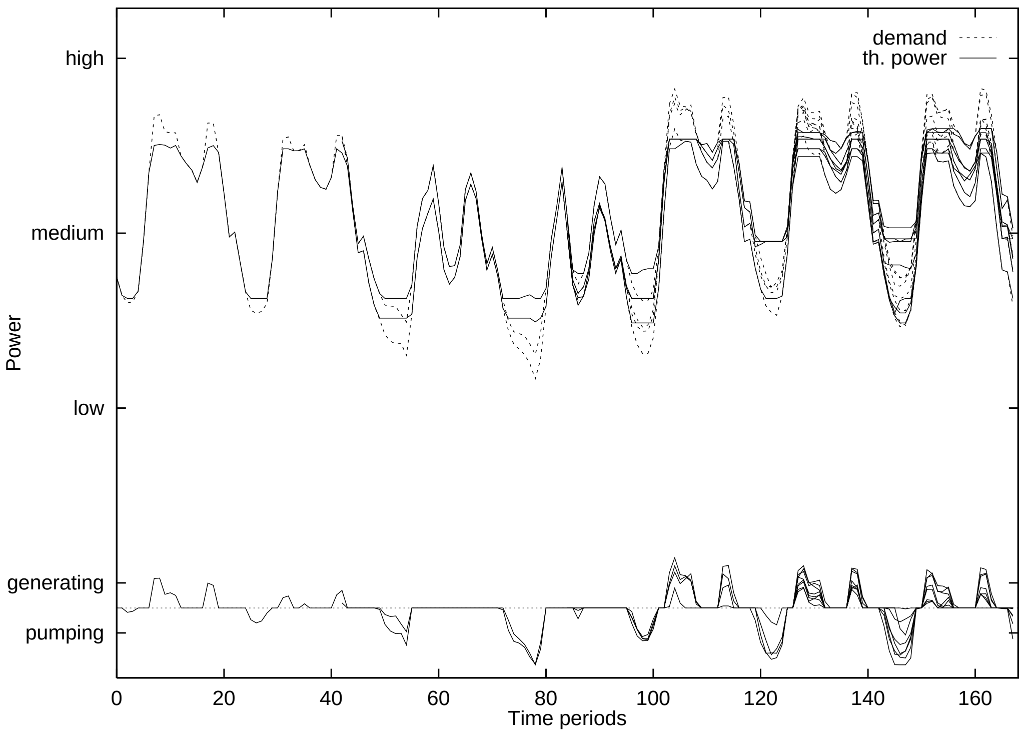

Figure 3: Stochastic solution

The number of nodes is the number of nodes in the scenario tree. Here, advantage denotes the advantage of using ECDISP versus CPLEX.

scenarios	nodes	columns	rows	nonzeros	ECDISP	CPLEX	advantage
2	252	18632	23436	60708	11.02	64.35	5.84
6	504	37248	46872	121408	37.17	228.76	6.15
10	723	53422	67239	174155	61.59	289.79	4.71
14	966	71372	89838	232686	116.81	534.78	4.58
18	1064	78592	98952	256272	103.33	504.41	4.88
22	1260	93064	117180	303476	128.18	794.93	6.20

The amount of memory needed by CPLEX exceeds the memory of the workstation if examples are computed with more than 22 scenarios, but ECDISP can even handle problems with more than 500 scenarios.

For testing the algorithm on a stochastic power dispatch model a scenario tree for approximating the stochastic demand process is generated as follows:

$d^t_{sampled} = d^t_{given} + \alpha(d^{t-1}_{sampled} - d^{t-1}_{given}) + \beta\varepsilon$

where d^t_{given} is the given load, $d^t_{sampled}$ a sample of the demand tree for the time interval t, ε is a standard normal (i.e. $N(0,1)$) random variable, and α, β are sampling parameters. A numerical example for the demand scenario and the corresponding stochastic schedule is given in figure 3.

Acknowledgement

We wish to thank G. Scheibner, G. Schwarzbach and J. Thomas (VEAG Vereinigte Energiewerke AG) for the fruitful and productive cooperation.

References

[1] D.P. Bertsekas, G.S. Lauer, N.R. Sandell, and T.A. Posbergh. Optimal short-term scheduling of large-scale power systems. *IEEE Transactions on Automatic Control*, 28(1983)1–11.

[2] CPLEX Optimization, Inc. *Using the CPLEX Callable Library*, 1989–1995.

[3] D. Dentcheva and W. Römisch. Optimal power generation under uncertainty via stochastic programming. Preprint Nr. 96-35, Humboldt-Universität Berlin, Institut für Mathematik, 1996.

[4] N. Gröwe, W. Römisch, and R. Schultz. A simple recourse model for power dispatch under uncertain demand. *Annals of Operations Research*, 59(1995)135–164.

[5] J. Jacobs, G. Freeman, J. Grygier, D. Morton, G. Schultz, K. Staschus, and J. Stedinger. SOCRATES: A system for scheduling hydroelectric generation under uncertainty. *Annals of Operations Research*, 59(1995)99–133.

[6] K.C. Kiwiel. Proximity control in bundle methods for convex nondifferentiable minimization. *Mathematical Programming*, 46(1990)105–122.

[7] C. Lemarechal. Lagrangian decomposition and nonsmooth optimization: bundle algorithm, prox iteration, augmented lagrangian. In: F. Gianessi (Ed.) *Nonsmooth Optimization Methods and Applications*, Gordon and Breach 1992,201–216.

[8] M.P. Nowak. A fast descent method for the hydro storage subproblem in power generation. Working Paper 96–109, International Institute for Applied Systems Analysis, September 1996.

[9] M.V.F Pereira and L.M.V.G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1991)359–375.

[10] G.B. Sheble and G.N. Fahd. Unit commitment literature synopsis. *IEEE Transactions on Power Systems*, 9(1994)128–135.

[11] S. Takriti, J.R. Birge, and E. Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(1996)1397–1506.

[12] F. Zhuang and F.D. Galiana. Towards a more rigorous and practical unit commitment by lagrangian relaxation. *IEEE Transactions on Power Systems*, 3(1988)763–773.