PRIMAL AND DUAL METHODS FOR UNIT COMMITMENT IN A HYDRO-THERMAL POWER SYSTEM

R. Gollmer[1], A. Möller[2], M.P. Nowak[3], W. Römisch[3], R. Schultz[1]

[1] Gerhard-Mercator-Universität Duisburg, FB Mathematik, Lotharstr. 65, D-47048 Duisburg, Germany
[2] Weierstraß-Institut für Angewandte Analysis und Stochastik, Mohrenstr. 39, D-10117 Berlin, Germany
[3] Humboldt Universität Berlin, Institut für Mathematik, Unter den Linden 6, D-10099 Berlin, Germany

Abstract: The unit commitment problem in a power generation system comprising thermal and pumped-storage hydro units is addressed. A large-scale mixed-integer optimization model for unit commitment in a real power system is developed and solved by primal and dual approaches. Both solution methods employ state-of-the-art algorithms and software. Results of test runs are reported.

Keywords: Unit commitment, mixed-integer linear programming, polyhedral combinatorics, Lagrangian relaxation, bundle methods

1 INTRODUCTION

Unit commitment in power operation planning aims at the cost optimal scheduling of on/off decisions and output levels for generating units. The power mix of the generation system has an essential impact on the design of mathematical models and algorithms for solving unit commitment problems. In the present paper, the interaction of a fair number of big coal fired blocks with several pumped storage plants of differing efficiencies provides the main challenge. This reflects the energy situation encountered at the German utility VEAG Vereinigte Energiewerke AG Berlin. Employing modern tools from mathematical optimization we demonstrate how to solve unit commitment problems for the VEAG system ranging over time horizons of up to 6 months with hourly discretizations. The presentation starts with the mathematical model followed by primal and dual solution approaches both accompanied by reports on some characteristic test runs.

This research was supported by a grant of the German Federal Ministry of Education, Science, Research and Technology (BMBF).

2 MODEL

In our model, T denotes the number of subintervals of the optimization horizon, I is the number of thermal and J the number of pumped storage hydro units. The variable $\boldsymbol{u}_i^t \in \{0,1\},\ i = 1,\dots,I; t = 1,\dots,T$ indicates whether the thermal unit i is in operation at time t. Variables $\boldsymbol{p}_i^t, \boldsymbol{s}_j^t, \boldsymbol{w}_j^t, i = 1,\dots,I; j = 1,\dots,J;\ t = 1,\dots,T$ reflect the output levels for the thermal units, the hydro units in generation mode and the hydro units in pumping mode, respectively. Moreover, we have variables $\boldsymbol{l}_j^t$ reflecting the fill (in energy) of the upper dam of the hydro unit j at the end of the time interval t, $j = 1,\dots,J;\ t = 1,\dots,T$.

The objective function to be minimized reads

$$\sum_{t=1}^{T}\sum_{i=1}^{I} C_i(\boldsymbol{p}_i^t, \boldsymbol{u}_i^t) + \sum_{t=1}^{T}\sum_{i=1}^{I} S_i^t(\boldsymbol{u}_i). \qquad (1)$$

Here, C_i denotes the fuel costs for unit i which often are a convex function of power output. We will consider linear and piecewise linear versions of C_i. The start-up costs $S_i^t(\boldsymbol{u}_i)$ of the i-th unit depend on its preceding down time.

When formulating the constraints we place accent on linear expressions although there are elegant alternatives using nonlinearities as well. Sticking to linearity at this place is motivated by the far more powerful collection of the then available mathematical tools (LP-based branch-and-bound and Lagrangian relaxation, polyhedral combinatorics).

Bounds for the power output of units and the fill of the upper dams read

$$\begin{array}{ll} p_{it}^{min}\boldsymbol{u}_i^t \leq \boldsymbol{p}_i^t \leq p_{it}^{max}\boldsymbol{u}_i^t, & \\ 0 \leq \boldsymbol{s}_j^t \leq s_{jt}^{max}, & i=1,\dots,I; \\ 0 \leq \boldsymbol{w}_j^t \leq w_{jt}^{max}, & j=1,\dots,J; \\ 0 \leq \boldsymbol{l}_j^t \leq l_j^{max}, & t=1,\dots,T. \end{array} \quad (2)$$

Here, $p_{it}^{min}, p_{it}^{max}, s_{jt}^{max}, w_{jt}^{max}$ denote minimal and maximal outputs, respectively, and l_j^{max} is the maximal fill of the upper dam.

The equilibrium between total generation and electrical load is covered by the following equations

$$\sum_{i=1}^{I} \boldsymbol{p}_i^t + \sum_{j=1}^{J} (\boldsymbol{s}_j^t - \boldsymbol{w}_j^t) \geq D^t, \quad t=1,\dots,T, \quad (3)$$

where D^t denotes the electrical load at time t. Moreover, at each time, a spinning reserve R_t has to be ensured which is modeled by

$$\sum_{i=1}^{I} (\boldsymbol{u}_i^t p_{it}^{max} - \boldsymbol{p}_i^t) \geq R^t, \quad t=1,\dots,T. \quad (4)$$

For the whole time horizon, the following balances in the pumped storage plants have to be maintained:

$$\begin{array}{lll} \boldsymbol{l}_j^t = \boldsymbol{l}_j^{t-1} - (\boldsymbol{s}_j^t - \eta_j \boldsymbol{w}_j^t), & & j=1,\dots,J; \\ \boldsymbol{l}_j^0 = l_j^{in}, & \boldsymbol{l}_j^T = l_j^{end}, & t=1,\dots,T. \end{array} \quad (5)$$

Here, l_j^{in}, l_j^{end} are the initial and final contents (in energy) of the upper dams, η_j denote the pumping efficiencies.

Constraints avoiding simultaneous generation and pumping in the hydro plants are dispensable since it can be shown that such a deficiency can not occur in optimal points.

Finally, we have minimum down times τ_i for the thermal units. These are modeled via

$$\begin{array}{ll} \boldsymbol{u}_i^{t-1} - \boldsymbol{u}_i^t \leq 1 - \boldsymbol{u}_i^l, & i=1,\dots,I; \\ & t=2,\dots,T-1; \\ & l=t+1,\dots,t+\tau_i \end{array} \quad (6)$$

where the constraints for the time intervals $t > T - \tau_i + 1$ have to be modified accordingly.

3 PRIMAL METHODS

LP-based branch-and-bound is among the earliest mathematical approaches to unit commitment, cf. [11]. It is based on formulating, possibly after exploiting proper equivalences, the unit commitment problem as a mixed-integer linear program that quickly becomes large-scale. Fixing integer (often Boolean) variables leads to subdivisions of the feasible region (branching) on which the objective function is bounded below by solving LP-relaxations and bounded above by trying to find feasible points, e.g., by rounding heuristics. Members of the subdivisions are cancelled (pruned) if their lower bounds are above the least upper bound found so far or if they correspond to empty subregions. The difference between the least upper bound and the minimum lower bound in the current subdivision provides an indication on how accurately optimality has been reached in the course of the algorithmic scheme. The ratio of the difference and the minimum lower bound is usually referred to as the certificate of optimality reached so far. The problem being large-scale, a zero certificate is rather utopic, and certificates in the lower per cents or per mills are usually accepted as sufficiently good.

The efficiency of LP-based branch-and-bound essentially depends on the way the subdivision is organized, on the speed for solving the LP-relaxations and on the quality of the heuristics for finding feasible points. Moreover, all this has to be embedded into a powerful implementation. The early branch-and-bound approaches to unit commitment suffered from the comparatively poor mathematical methodology and software technology available at that time. Meanwhile, this has changed drastically, both with respect to mathematical algorithms and software implementations. Let alone advances in hardware. General purpose codes like the CPLEX Callable Library [1] combine latest LP-methodology with a variety of options for arranging the subdivision and setting up the heuristics for the upper bounds. In fact, the CPLEX Callable Library forms the algorithmic backbone of our primal approach to unit commitment.

To make LP-based branch-and-bound work for the above model the costs in (1) have to be expressed by linear terms, possibly involving integer variables. As to the fuel costs C_i, this is possible for the linear and the piecewise linear situations. For the start-up costs S_i which depend exponentially on the preceding downtime we used approximations via step functions. The numbers of linearity regions for

model dimensions	variant with groups of aggregated units and fixed start-up costs			variant with individual units and a 3-step function for start-up costs		
	1 week	1 month	6 months	1 week	1 month	6 months
integer variables	2112	8184	56472	5420	20832	130320
real variables	9781	37867	217608	15210	65442	383033
constraints	8053	31237	204576	22902	83364	594619
nonzeroes	31448	121877	760110	196803	749430	6363009

Table 1: Model dimensions for both model variants

CPU-time and accuracy	variant with groups of aggregated units and fixed start-up costs			variant with individual units and a 3-step function for start-up costs		
	1 week	1 month	6 months	1 week	1 month	6 months
CPU-time / min	0:58.9	7:40.9	234:02.9	7:44.3	161:32.9	out of
accuracy bound / %	0.086	0.073	0.133	0.391	0.389	memory

Table 2: Computing times on a HP 9000 (770/J180) and accuracy bounds of the primal method

C_i and steps for S_i proved critical for the size of the model and hence for memory requirements and run times. Therefore, proper selections based on the concrete VEAG data were made at this place. The concrete data situation was also exploited for guiding the subdivision (branching start-up variables according to the load profile), for improving model properties (introducing integer instead of Boolean variables for units with identical design which leads to subdivisions with smaller cardinality) and for designing a fast heuristic to come up with a first feasible solution. Moreover, some first experiments with cutting planes based on polyhedral combinatorics [10] were made to tighten the lower bounds.

Test runs based on real-life data were performed on an HP 9000 (770/J180). Time horizons considered are 1 week, 1 month, and 6 months, with an hourly discretization. The generation system included 34 thermal and 7 hydro units. The generating costs were approximated by a linear function. Two different approximations to the start-up costs were made, leading to two different variants of the model. In the first variant the start-up costs were considered to be constant. This allows for the reformulation as a general mixed integer problem by aggregating the groups of technically identical generating units. In the second variant a step function with three steps per unit was used for the start-up costs. This prohibits aggregation of units and leads to much bigger models with the implication of an increase in computing time. The tighter bounds on the accuracy of the solution value in the first model variant are due to the fact, that the gap between the LP-relaxation and the feasible set is smaller in this case. The first table gives an impression of the dimensions of the problems to be solved.

The solution process, which in case of the branch-and-bound approach at least theoretically could be continued until a satisfactory accuracy bound is achieved, was finished, when a solution with a bound of less than 1% was reached. For all the models considered this was achieved with the first feasible solution, determined by a problem-specific rounding heuristic, which normally is followed by a branch-and-bound procedure for the full problem. The a priori knowledge of a good feasible solution in this step allows to cut off a lot of branches in the tree developed by the branch-and-bound procedure and thus leads to savings in memory requirements and computing time. Computing times and accuracy bounds are displayed in Table 2.

In general, the primal approach via LP-based branch-and-bound allows ample model enrichment as long as this is expressible in mixed-integer linear terms. In particular, further internal coupling of the model caused by the introduction of additional constraints is not critical. This has been exploited when extending the above model towards more sophisticated re-

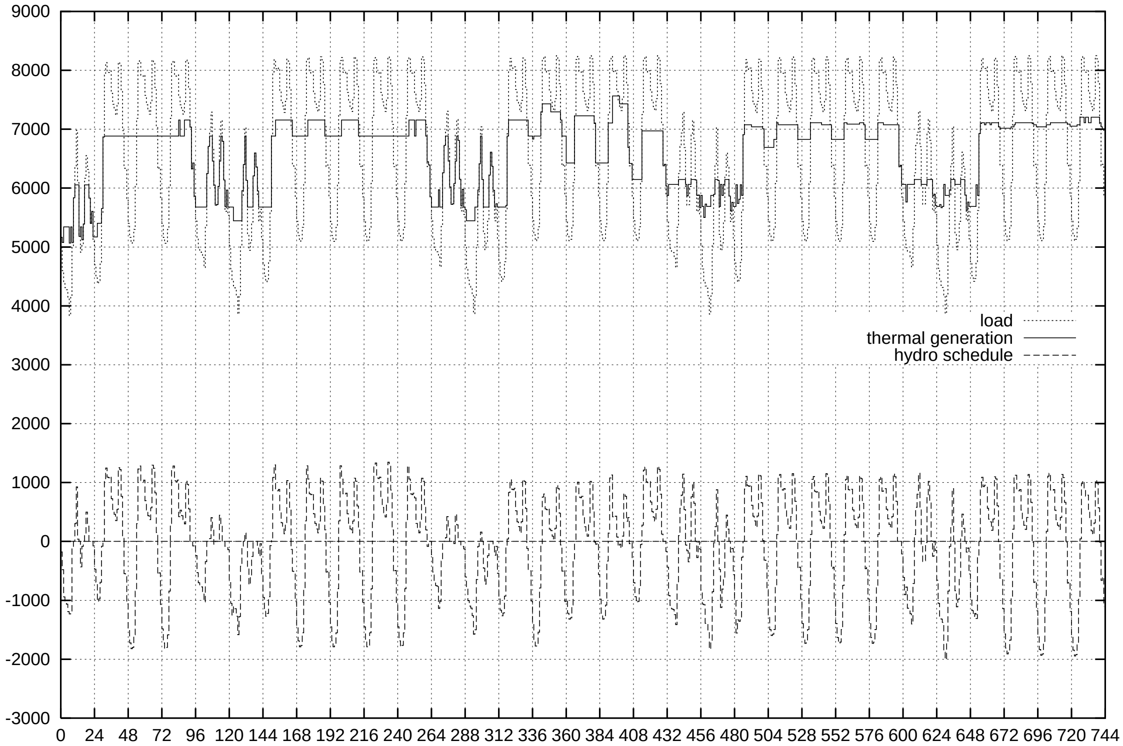

Figure 1: Solution of the primal method for 1 month

serve policies involving hydro units [5] or towards staggered fuel prices [4]. On the other hand, always the full model has to be handled which may become prohibitive even if advanced methodology is used for solving the LP-relaxations. This paves the way for decomposition which will be discussed next.

4 DUAL METHODS

Dual methods which are referred to as Lagrangian relaxation have become a very popular approach in unit commitment (cf. [11]). Recently, three aspects made Lagrangian relaxation attractive and applicable to large-scale unit commitment problems: the progress of bundle methods for solving the nondifferentiable Lagrangian dual, the fact that the relative duality gap is usually small and the progress in developing fast Lagrangian heuristics for finding good primal feasible solutions. We refer to [3], [8] and [12] for a discussion of these aspects in case of thermal generation systems.

For the hydro-thermal model in Section 2, the Lagrangian relaxation approach is based on associating Lagrange multipliers with the loading and reserve constraints, where we use the modified reserve constraint

$$\sum_{i=1}^{I} \boldsymbol{u}_i^t p_{it}^{max} + \sum_{j=1}^{J} (\boldsymbol{s}_j^t - \boldsymbol{w}_j^t) \geq D^t + R^t, \qquad (7)$$

$(t = 1, \ldots, T)$ instead of (4).

The dual problem reads

$$\max_{(\boldsymbol{\lambda},\boldsymbol{\mu}) \in I\!R_+^T \times I\!R_+^T} d(\boldsymbol{\lambda}, \boldsymbol{\mu}), \qquad (8)$$

where $\boldsymbol{\lambda}, \boldsymbol{\mu}$ are the Lagrange multipliers and the dual function d is defined by taking the infimum of the Lagrangian with respect to the vector ($\mathbf{p}$, $\mathbf{u}$, $\mathbf{s}$, $\mathbf{w}$) and the constraints (2), (5) and (6). d exhibits the separable form

$$\begin{aligned} d(\boldsymbol{\lambda}, \boldsymbol{\mu}) := \sum_{i=1}^{I} d_i(\boldsymbol{\lambda}, \boldsymbol{\mu}) + \sum_{j=1}^{J} \tilde{d}_j(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ + \sum_{t=1}^{T} [\boldsymbol{\lambda}^t D^t + \boldsymbol{\mu}^t (D^t + R^t)] \end{aligned} \qquad (9)$$

where the functions d_j and $\tilde{d}_j$ denote the optimal values of certain single-unit thermal and hydro subproblems, respectively.

$$d_j(\boldsymbol{\lambda},\boldsymbol{\mu}) := \min_{\boldsymbol{u}_i} \Big\{ \sum_{t=1}^{T} [S_i(\boldsymbol{u}_i(t)) - \boldsymbol{\mu}^t \boldsymbol{u}_i^t p_{it}^{max} + \min_{\boldsymbol{p}_i^t} \{C_i(\boldsymbol{p}_i^t, \boldsymbol{u}_i^t) - \boldsymbol{\lambda}^t \boldsymbol{p}_i^t\}] : \boldsymbol{u}_i^t, \boldsymbol{p}_i^t \text{ satisfy (2) and (6)} \Big\}$$

$$\tilde{d}_j(\boldsymbol{\lambda},\boldsymbol{\mu}) := \min_{\boldsymbol{s}_j, \boldsymbol{w}_j} \Big\{ \sum_{t=1}^{T} (\boldsymbol{\lambda}^t + \boldsymbol{\mu}^t)(\boldsymbol{w}_j^t - \boldsymbol{s}_j^t) : \boldsymbol{s}_j^t, \boldsymbol{w}_j^t \text{ satisfy (2) and (5)} \Big\}$$

The inner minimization of the thermal subproblems w.r.t. $\boldsymbol{p}_i^t$ is carried out explicitly while the outer minimization w.r.t. $\boldsymbol{u}_i$ is performed by dynamic programming. For solving the hydro subproblems a fast descent algorithm developed in [9] is used. Since the dual function d is concave and subgradient information of d is available, powerful bundle-type algorithms [6] may be used for the solution of the Lagrangian dual (8). After having solved the dual we obtain a lower bound for the minimal costs of the model in Section 2 and together with the optimal multipliers $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ we have solutions of the thermal and hydro subproblems. But, in general these solutions lead to a violation of the load and reserve constraints (3) and (4). Hence, a low-cost (primal) feasible solution has to be determined by a Lagrangian heuristics.

Altogether, the Lagrangian relaxation algorithm consists of the following steps:
Step 1: Initialize the multipliers $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$.
Step 2: Solve the dual problem (8) by a bundle method.
Step 3: Determine a primal feasible solution by a Lagrangian heuristics.
In both Step 2 and Step 3 thermal and hydro subproblems are solved repeatedly.

To initialize the multiplier $\boldsymbol{\lambda}$ a priority list approach is used. In each time interval thermal units are switched on in ascending order of the relative costs at their generation maximum until the sum of the maximum generation levels is greater than or equal to the demand or until all units are online. The relative costs of the last unit switched on are then used to initialize $\boldsymbol{\lambda}^t$. The multipliers $\boldsymbol{\mu}^t$ are initialized by zero in all time intervals.

The software package NOA 3.0 [7] is used for solving the dual problem. The underlying proximal bundle algorithm [6] maximizes a polyhedral approximation of the dual function by subgradients obtained from previous iterations. This approximation is extended by a quadratic term to keep the new iterates for $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ in a trust region of the current iterates.

Two different Lagrangian heuristics have been developed and implemented.
The first Lagrangian heuristics (LH1) consists of three steps and begins with reducing the value $D^t + R^t + \sum_{j=1}^{J}[\boldsymbol{w}_j^t - \boldsymbol{s}_j^t]$ by modifying the schedule of the hydro plants if the reserve constraint (7) is violated at time t and the value of this sum is the largest in a certain set of considered intervals. This procedure may result in new violations of the reserve contraint in intervals with small values of the sum above. In a second step the hydro variables are kept fixed and we search for binary variables $\boldsymbol{u}_i^t$ which satisfy the condition $\sum_{i=1}^{I} \boldsymbol{u}_i^t p_{it}^{max} \geq D^t + R^t + \sum_{j=1}^{J}[\boldsymbol{w}_j^t - \boldsymbol{s}_j^t]$ by the method described in [13]. Its main idea consists in determining the interval where the preceding condition is violated most and in computing the necessary increase of μ^t to switch on just as many thermal units as needed to satisfy this condition. This procedure is repeated until the reserve constraint (7) is satisfied in all intervals. After having fixed the binary variables $\boldsymbol{u}_i^t$, the economic dispatch problem is solved by CPLEX [1] in the last step.

The second Lagrangian heuristics (LH2) exploits the structure of the dual problem (8). The idea is to look at all primal solutions $(\mathbf{p}, \mathbf{u}, \mathbf{s}, \mathbf{w})$ that correspond to (nearly) optimal multipliers $(\boldsymbol{\lambda},\boldsymbol{\mu})$. For convex models each primal solution $(\mathbf{p}, \mathbf{u}, \mathbf{s}, \mathbf{w})$ that is feasible and corresponds to optimal multipliers, is also optimal. Of course, this is no longer valid in the mixed-integer situation. Instead we determine a set of primal solutions that correspond to slightly perturbed optimal multipliers. This is done by recording first the set of possible binary decision which are obtained from the thermal subproblems by dynamic programming. It turned out in our test runs that only a small percentage of binary variables changes their values. Fixing the remaining binary decisions leads to a drastic reduction of dimensionality. Then a sequence of binary decisions $\mathbf{u}$ is constructed that decreases componentwise. In each step a time period t is selected where the available reserve

NOA 3.0 optimality tolerance: 10^{-4}		optimization horizon					
		1 week		1 month		6 months	
production costs	start-up costs	time/ min	bound of gap/ %	time/ min	bound of gap/ %	time/ min	bound of gap/ %
linear	constant	0:17	1.10	2:36	0.93	60:15	0.84
linear	time dependent	0:20	1.13	3:04	0.98	63:04	0.73
piecew. lin.	constant	0:28	1.09	5:33	0.86	110:23	0.79
piecew. lin.	time dependent	0:30	1.07	5:28	0.96	119:02	0.69

Table 3: CPU-time in minutes on HP 9000 (770/J180) and upper bound of the duality gap of the dual method (with LH1)

$\sum_{i=1}^{I}(\boldsymbol{u}_i^t p_{it}^{max} - \boldsymbol{p}_i^t) - R^t$ is large and the multipliers are used to determine in which previous and subsequent time periods some unit can be switched off. For each element of the sequence an economic dispatch problem is solved by a modification of the descent method from [9]. The element of the sequence having the least optimal value provides a reasonable good solution of the problem (1)-(6).

The numerical results in Table 3, Table 4 and Figure 2 are based on the same data as for the primal method and were obtained on the same hardware. Compared to Section 3 exponential approximations for start-up costs are used leading to more than 50 different steps. Moreover, test examples with piecewise linear fuel costs were run. Compared with the primal approach, the dual method is faster but yields wider accuracy bounds. A more detailed modelling of down time dependence in start-up costs is less time critical in the dual approach since the time for solving the thermal subproblems grows only linearly with the number of steps used for approximating start-up costs. The Lagrangian heuristics LH2 yields smaller accuracy bounds than LH1. Piecewise linear fuel costs lead to quite substantial increases of computing time when using LH1 and CPLEX for the economic dispatch. This effect does not occur in LH2, since LH2 employs a specific descent algorithm for economic dispatch.

ACKNOWLEDGEMENTS

We are grateful to K.C. Kiwiel (Polish Academy of Sciences, Warsaw) for the permission to use the NOA 3.0 package and to G. Schwarzbach and J. Thomas (VEAG Vereinigte Energiewerke AG Berlin) for their outstanding collaboration.

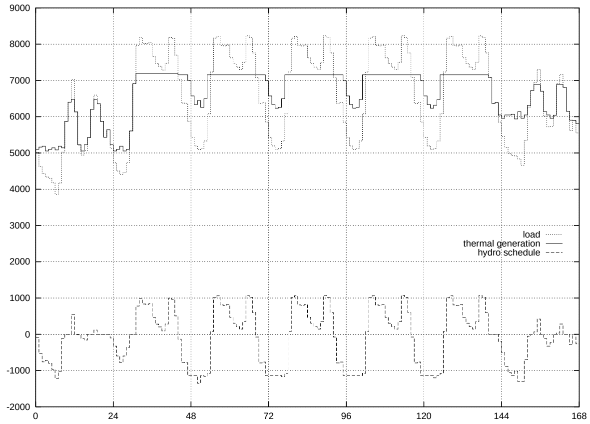

Figure 2: Solution of the dual method (with LH1) for 1 week

REFERENCES

[1] Using the CPLEX Callable Library, CPLEX Optimization, Inc. 1994.

[2] D. Dentcheva, R. Gollmer, A. Möller, W. Römisch and R. Schultz: Solving the unit commitment problem in power generation by primal and dual methods, in: *Progress in Industrial Mathematics at ECMI 96* (M. Brøns, M.P. Bendsøe, M.P. Sørensen Eds.), Teubner, Stuttgart, 1997, 332 - 339.

[3] S. Feltenmark, K.C. Kiwiel and P.O. Lindberg: Solving unit commitment problems in power production planning, in: *Operations Research Proceedings 1996* (U. Zimmermann et al. Eds.), Springer-Verlag, Berlin 1997, 236-241.

[4] M. Fischer: Gestaffelte Brennstoffpreise in der Kraftwerkseinsatzoptimierung, Diplomarbeit, Technische Universität Berlin, 1997.

		optimization horizon					
		1 week			1 month		
production costs	start-up costs	# economic dispatch runs	bound of gap/ %	time/ min	# economic dispatch runs	bound of gap/ %	time/ min
linear	constant	2	0.44	0:19			
linear	time dep.	2	0.28	0:18			
piecew. lin.	constant	2	0.20	0:21	7	0.30	6:24
piecew. lin.	time dep.	2	0.29	0:22	8	0.42	5:26

Table 4: Number of economic dispatch problems solved, upper bound for the duality gap and CPU-time on HP 9000 (J280) of the dual method (with LH2)

[5] R. Gollmer, A. Möller, W. Römisch, R. Schultz, G. Schwarzbach und J. Thomas: Optimale Blockauswahl bei der Kraftwerkseinsatzplanung der VEAG, in: *Optimierung in der Energieversorgung II*, VDI-Berichte 1352, 1997, 71 - 85.

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

[7] K.C. Kiwiel: User's Guide for NOA 2.0/3.0: A Fortran Package for Convex Nondifferentiable Optimization. Polish Academy of Science, System Research Institute, Warsaw, 1993/1994.

[8] C. Lemaréchal, C. Sagastizábal, F. Pellegrino and A. Renaud: Bundle methods applied to the unit commitment problem, in: *System modelling and optimization* (J. Doležal and J. Fidler Eds.), Proceedings of the 17th IFIP TC7 Conference, Chapman&Hall, London 1996, 395 - 402.

[9] M.P. Nowak: A fast descent method for the hydro storage subproblem in power generation, Working Paper WP-96-109, IIASA, Laxenburg (Austria), 1996.

[10] T. van Roy and L.A. Wolsey: Valid inequalities for mixed 0-1 programs, *Discrete Applied Mathematics* 14(1986), 199 - 213.

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

[12] S. Takriti: The unit commitment problem, in: *Operations Research in Industry* (T.A. Ciriani, S. Gliozzi and E.L. Johnson Eds.), Macmillan Press, 1999 (to appear).

[13] 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.