Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

CLAUS C. CARØE RÜDIGER SCHULTZ

A Two-Stage Stochastic Program for Unit Commitment Under Uncertainty in a Hydro-Thermal Power System

Preprint SC 98-11 (February 1998)

A Two-Stage Stochastic Program for Unit Commitment Under Uncertainty in a Hydro-Thermal Power System

Claus C. Carøe
University of Copenhagen, Institute of Mathematics,
Universitetsparken 5, DK-2100 Copenhagen, Denmark.
E-mail: `caroe@math.ku.dk`.

Rüdiger Schultz
Universität Leipzig, Mathematisches Institut
Augustusplatz 10–11, D-04109 Leipzig, Germany.
E-mail: `schultz@mathematik.uni-leipzig.de`.

February 1998

Abstract

We develop a two-stage stochastic programming model with integer first-stage and mixed-integer recourse for solving the unit commitment problem in power generation in the presence of uncertainty of load profiles. The solution methodology rests on a novel scenario decomposition method for stochastic integer programming. This method combines Lagrangian relaxation of non-anticipativity constraints with branch-and-bound. It can be seen as a decomposition algorithm for large-scale mixed-integer linear programs with block-angular structure. With realistic data from a German utility we validate our model and carry out test runs. Sizes of these problems go up to 20.000 integer and 150.000 continuous variables together with up to 180.000 constraints.

Subject classifications: Programming, stochastic: Scenario Decomposition of mixed-integer programs. Natural resources, energy: Unit commitment under uncertainty.

Unit commitment aims at finding a fuel cost optimal scheduling of start-up/shut-down decisions and operation levels for power generation units over some time horizon. This is a central task in reliable and efficient operation of power systems. Solution strategies for the unit commitment problem are influenced by the power mix of the generation system. In the present paper we consider a hydro-thermal system as it is met with the German power company VEAG Vereinigte Energiewerke AG Berlin. This system comprises conventional coal and gas fired thermal units as well as pumped-storage plants. The latter imply substantial coupling over time of scheduling decisions which, compared with purely thermal systems, further complicates the problem.

Mathematically, unit commitment is handled as a large-scale mixed-integer optimization problem. Advances in mathematical methodology, software engineering and hardware have led to practicable solutions of more and more complex unit commitment problems. However, a characteristic feature of power system operation is that decisions have to be taken with incomplete knowledge of future economical and technological data. Due to the inherent complexity of unit commitment, in the past this feature mostly had to be faded out in mathematical models. The main challenge is to combine traditional unit commitment and incomplete information into a model that is accessible to computation. The machinery of stochastic programming offers a vehicle to handle uncertainty of problem data in optimization models. Particularly well developed methodology exists for linear models with continuous variables. Recent progress in the field concerns incorporation of integer variables which makes stochastic programming attractive for unit commitment under uncertainty. First contributions along this line were made in Takriti, Birge, and Long (1994, 1996), Dentcheva and Römisch (1998) and Takriti, Krasenbrink, and Wu (1997).

In the present paper we elaborate a mixed-integer two-stage stochastic programming model for unit commitment accounting for uncertainty of electrical load (Section 1). The stochastic program comprises both traditional features of unit commitment and constraints on the availability of information. With incomplete information scheduling of power systems involves additional costs caused by readjustments after the unknown data are reveiled. The stochastic program includes future readjustment costs into the optimization of decisions that have to be taken prior to reveiling uncertainty. From the modeling viewpoint this benefit has to be paid for by an enormous enlargement of the traditional deterministic unit commitment model. Roughly speaking, the number of variables and constraints has to be multiplied by the number of scenarios for the uncertain load profile. In addition, information constraints induce coupling between variables corresponding to different scenarios. Altogether, the stochastic program is intractable by algorithms and software developed for deterministic unit commitment.

The critical issue that prevents application of existing stochastic programming methodology for our model is the presence of integer variables. In Section 2 we present a dual decomposition method that breaks down the stochastic program into single-scenario unit-commitment-like subproblems. Integer variables lead to duality gaps that are overcome by a branch-and-bound procedure on top of the dual decomposition. The basics of this algorithm were already developed in Carøe and Schultz (1997). In the present paper we demonstrate successful application of the method to an important planning problem with realistic size.

Computational experience reported in Section 3 shows that our implementation is able to handle mixed-integer stochastic programs with up to 20.000 integer and 150.000 continuous variables on an advanced workstation. The design of our algorithm allows to benefit from achievements in non-smooth optimization, cf. Kiwiel (1990, 1994), and from recent methods for deterministic unit commitment reported in Dentcheva et al. (1997).

1 Model

Due to uncertain electrical load the planning of electricity production faces a major problem when determining the schedule of generating units. Our aim is to develop a model that takes this uncertainty into account and yields robust decisions. Before formalizing the problem let us explain in more detail some operational characteristics of our power system that are important for the model. Starting up a coal fired block involves some time delay before the block becomes available for electricity generation. Therefore, switching decisions for these units have to be taken well in advance and cannot be employed as short-term corrective actions. Gas turbines, however, have a sufficiently small delay in this respect such that switching is feasible for short-term corrections.

The output levels of thermal units in operation are adjustable at speeds depending on the technological design. Here we assume these speed limits to be sufficiently wide which, of course, narrows the scope of the model in its present form but is removable at the cost of increased model complexity. Due to their adjustability, decisions on the thermal units' outputs are suitable for short-term corrective actions.

The pumped storage hydro units at VEAG-plants operate with a fixed amount of water without additional in- or outflows. Our model will cover precisely this situation, but other types of hydro units can easily be modelled. Water in pumped storage plants can either be sent to turbines (electricity generation) or pumps (electricity consumption). This involves an efficiency less than one and has to match both with input/output bounds on storage dams and water balances over time. An important characteristic of pumped storage plants is their quick availability which makes them appropriate for short-term corrections.

Our stochastic programming model basically works as follows. For a finite optimization horizon with a suitable division into subintervals we attach a probability distribution to the electrical load. Decisions are grouped into long-term policies that have to be taken before observation of load values and short-term corrections. Grouping of decisions follows the operational characteristics above and we refer to the groups as first- and second-stage decisions, respectively. Start-up decisions for coal fired units will become first-stage, whereas start-up decisions for gas turbines and output levels for all units will be the second-stage decisions. Finally, optimization aims at minimizing the sum of direct costs for first-stage decisions and the expected value of the costs induced by the first- together with the second-stage decisions.

Let $t = 1, \ldots, T$ denote the subintervals (e.g., hours) of the optimization horizon. Suppose that there are $i = 1, \ldots, I$ coal fired thermal units, $k = 1, \ldots, K$ gas turbines, and $j = 1, \ldots, J$ pumped storage plants. The stochastic behavior of electrical load is represented by a random variable d on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $\mathbb{R}^T$. We assume that d follows a discrete distribution with finite support and use the symbol $d^\omega, \omega \in \Omega$ to denote its realizations. Following a usual convention, these realizations will be called scenarios.

By $u_{it} \in \{0,1\}$, $i = 1, \ldots, I$ and $t = 1, \ldots, T$ we denote the first-stage variables indicating whether the coal fired unit i is on or off at time t. Correspondingly, $u_{kt}^\omega \in \{0,1\}$, $k = 1, \ldots, K$, $t = 1, \ldots, T$, $\omega \in \Omega$ denotes the on/off decision for gas turbine k in time interval t under scenario d^ω. Along with the on/off decisions we have output levels $p_{it}^\omega, p_{kt}^\omega$ of the mentioned units where the superscript ω indicates linkage to the corresponding scenario. The output limitations of thermal units then read as follows

$$p_i^{\min} u_{it} \;\le\; p_{it}^\omega \;\le\; p_i^{\max} u_{it},\; i = 1, \ldots, I, \quad t = 1, \ldots, T,\; \omega \in \Omega \tag{1}$$

$$p_k^{\min} u_{kt}^\omega \;\le p_{kt}^\omega \;\le p_k^{\max} u_{kt}^\omega,\; k = 1, \ldots, K,\; t = 1, \ldots, T,\; \omega \in \Omega. \tag{2}$$

Here, $p_i^{\min}$, $p_k^{\min}$ and $p_i^{\max}$, $p_k^{\max}$ are the minimal and maximal outputs of the respective units. By s_{jt}^ω and w_{jt}^ω, $j = 1, \ldots, J$, $t = 1, \ldots, T$, $\omega \in \Omega$ we denote the levels of generation and pumping in pumped storage plant j at time t under scenario d^ω. With upper bounds $s_j^{\max}, w_j^{\max}$ we have the following box constraints on these variables

$$0 \;\le\; s_{jt}^\omega \;\le\; s_j^{\max},\; j = 1, \ldots, J,\;\; t = 1, \ldots, T,\; \omega \in \Omega \tag{3}$$

$$0 \;\le w_{jt}^\omega \;\le w_j^{\max},\; j = 1, \ldots, J,\;\; t = 1, \ldots, T,\; \omega \in \Omega \tag{4}$$

It is convenient to introduce variables $l_{jt}^{\omega},\ j=1,\ldots,J,\ t=1,\ldots,T,\ \omega\in\Omega$ for the fills (in energy) of the upper dams. Water balances in the pumped storage plants can then be expressed as follows

$$\begin{aligned} 0 \le l_{jt}^{\omega} &\le l_j^{\max}, & t&=1,\ldots,T, & (5)\\ l_{jt}^{\omega} &= l_{jt-1}^{\omega} - s_{jt}^{\omega} + \eta_j w_{jt}^{\omega}, & t&=2,\ldots,T, & (6)\\ l_{j1}^{\omega} &= l_j^{\text{ini}} - s_{j1}^{\omega} + \eta_j w_{j1}^{\omega}, & & & (7)\\ l_{jT}^{\omega} &= l_j^{\text{end}} & & & (8) \end{aligned}$$

for $j=1,\ldots,J,\ \omega\in\Omega$. The constants $\eta_j,\ 0<\eta_j<1$, are the pumping efficiencies, and $l_j^{\max}, l_j^{\text{ini}}, l_j^{\text{end}}$ denote the maximal, initial and final fills, respectively. Inequalities (5) state that the fills must not exceed certain bounds and equations (6) - (8) display the fill dynamics with (8) avoiding empty dams at the end of the time horizon by prescribing proper end conditions. The equilibrium between total generation and electrical load reads

$$\sum_{i=1}^{I} p_{it}^{\omega} + \sum_{k=1}^{K} p_{kt}^{\omega} + \sum_{j=1}^{J} (s_{jt}^{\omega} - w_{jt}^{\omega}) \;=\; d_t^{\omega}, \quad t=1,\ldots,T,\ \omega\in\Omega. \tag{9}$$

Fuel costs can be divided into start-up costs and operation costs for the power units. There are no direct fuel costs for pumped storage plants, although the latter have indirect impact on fuel costs by the pumping energy needed to establish the necessary levels in the upper dams. Start-up costs for thermal units depend on the preceding down time of the block. Here, we will neglect this dependence and assume constant costs a_i, a_k for start-ups in coal fired blocks and gas turbines, respectively. Again this narrows the scope of our model, but can be removed on the cost of higher model complexity. Total start-up costs for the coal fired units then compute as

$$\sum_{t=2}^{T} \sum_{i=1}^{I} a_i \max\{u_{it} - u_{it-1}, 0\}. \tag{10}$$

Denoting by $\mathbb{E}_{\omega}$ expectation with respect to P, the expected value of total start-up costs for the gas turbines reads

$$\mathbb{E}_{\omega}\Big[\sum_{t=2}^{T} \sum_{k=1}^{K} a_k \max\{u_{kt}^{\omega} - u_{kt-1}^{\omega}, 0\}\Big]. \tag{11}$$

For reasons of clarity we prefer the (nonlinear) maximum term in (10) and (11). For computations, of course, these are transformed in the usual way into linear terms by introducing further variables and including additional linear constraints.

Fuel costs of coal and gas fired thermal units in operation are assumed to be affinely linear with coefficients c_i, c_i^0 and c_k, c_k^0, respectively. Although we found that this fairly well reflects the situation met at VEAG refinements are again possible. Piecewise linear costs can, in case of convexity, be handled as maximum terms. In the non-convex case, they can be expressed using additional Boolean variables. In our model, total expected operation costs compute as

$$\mathbb{E}_{\omega}\Big[\sum_{t=1}^{T} \big(\sum_{i=1}^{I} u_{it}(c_i p_{it}^{\omega} + c_i^0) \;+\; \sum_{k=1}^{K} u_{kt}^{\omega}(c_k p_{kt}^{\omega} + c_k^0)\big)\Big].$$

Here, nonlinearities can be removed using (1), (2), and we obtain

$$\sum_{t=1}^{T}\sum_{i=1}^{I} c_i^0 u_{it} \;+\; \mathbb{E}_\omega\Big[\sum_{t=1}^{T}\sum_{i=1}^{I} c_i p_{it}^\omega + \sum_{t=1}^{T}\sum_{k=1}^{K} c_k p_{kt}^\omega + \sum_{t=1}^{T}\sum_{k=1}^{K} c_k^0 u_{kt}^\omega\Big]. \tag{12}$$

Altogether, we have the following optimization problem:

$$\begin{aligned}
\text{Minimize}\quad & \sum_{t=1}^{T}\sum_{i=1}^{I} c_i^0 u_{it} \;+\; \sum_{t=2}^{T}\sum_{i=1}^{I} a_i \max\{u_{it}-u_{it-1},0\} \\
& +\; \mathbb{E}_\omega\Big[\sum_{t=1}^{T}\sum_{i=1}^{I} c_i p_{it}^\omega + \sum_{t=1}^{T}\sum_{k=1}^{K} c_k p_{kt}^\omega + \sum_{t=1}^{T}\sum_{k=1}^{K} c_k^0 u_{kt}^\omega \\
& \qquad\qquad + \sum_{t=2}^{T}\sum_{k=1}^{K} a_k \max\{u_{kt}^\omega - u_{kt-1}^\omega, 0\}\Big] \qquad (13) \\
\text{subject to}\quad & (1)-(9) \\
& \text{and} \\
& u_{it}, u_{kt}^\omega \in \{0,1\},\; i=1,\dots,I,\; k=1,\dots,K,\; t=1,\dots,T,\; \omega\in\Omega.
\end{aligned}$$

This model can be classified as a two-stage stochastic program with integer requirements in both stages. Behind two-stage stochastic programs (also called stochastic programs with recourse) there is a process of alternating decisions and observations. First, decisions have to be taken that do not anticipate future data of the model. At the time these data are known, second-stage (or recourse) decisions follow which exploit the additional information and depend on the variables that were fixed in the first stage. Costs of the decision process are given as a sum of deterministic costs depending on the first-stage decisions and random costs depending on the first-stage decisions, the outcome of the random data and the second-stage decisions. This leads to the natural optimization criterion of minimizing the sum of the deterministic costs and the expectation of the random costs.

In terms of our practical model this leads to an optimization of start-up decisions under uncertainty. Due to the operational characteristics of coal fired thermal units, start-ups indeed have to be fixed without the precise knowledge of the load profile for the optimization horizon. The stochastic program (13) incorporates the probabilistic information and selects a start-up schedule that hedges against uncertainty and is cost optimal among all such hedging policies. The main point is that uncertainty is dealt with using scenarios for possible load profiles. Power companies maintain extensive records of past load data that allow extraction of realistic load profiles and a probability distribution as needed for our model. A desired property of the hedging policy is the ability to handle unexpected or unusual demands. Load scenarios should therefore consider factors like generator failures, high demand increases or decreases, perturbations in the timing of peak periods, etc. In practice the problem of accurately forecasting future load is done by schedulers for a time horizon of usually one week (i.e., 168 hours). Scenarios should possibly take into account past experience of unforeseen load changes. On the other hand, as the computing times grow rapidly with the size of (13), only a limited number of scenarios can be considered and demands should be sufficiently diverse to all contribute to the hedging policy. Here we distinguish between modeling generator failures and inaccurate load forecasts, but the two cases can of course be treated at the same time.

Generator failures occur due to unscheduled maintenance, technical problems etc. A common practice in power system operation is to impose reserve constraints in the unit commitment problem, e.g., a spinning reserve that requires excess capacity at any time point to be greater than some specified percentage of the peak load. In our model these constraints will be handled by scenario modelling. Following Takriti, Birge, and Long (1996), generator failures are modelled by creating scenarios with demand increases equal to the expected capacity loss over some time period, and with appropriate probabilities. For the power system of VEAG, the largest coal fired unit has a capacity of about 750 MW per hour. A scenario with a demand increase of 750 MW over some time period means that this capacity has to be obtained from the remaining units.

All the output levels of generating units are second-stage decisions in our model. Scenario modelling then entails that scenarios should have different timing of the on/off decisions. Thus we consider not only high demand increases and decreases but also earlier or delayed peak demand in the mornings and afternoons as well as longer peak periods. Moreover the coupling of hydro and thermal power units should be taken into account. Pumped storage plants are usually employed for storing energy by pumping at night and generating during peak periods. Higher demand in the night periods means that possibly more coal fired units should be used in peak periods, since pumped storage plants will then not be available.

If the set Ω of possible realizations consists of just one element, i.e., if the electrical load is perfectly known in advance, then (13) reduces to a deterministic unit commitment problem as developed in Dentcheva et al. (1997). For this class of problems there is a rich literature on solution methods and their implementations in practice, cf. the annotated bibliography of Sheblé and Fahd (1994). Here, Lagrangian relaxation of unit coupling constraints, e.g., equilibria between total generation and load, has become quite popular and successful, see e.g., Aoki et al. (1989) or Bertsekas et al. (1983). The work reported in Dentcheva et al. (1997) employs both Lagrangian relaxation and LP-based branch-and-bound. Subsequently, the LP-based branch-and-bound part of that method will be used as a subroutine in our algorithm for problem (13).

The first application of stochastic programming methodology in unit commitment was reported in Takriti, Birge, and Long (1994, 1996). In contrast to the two-stage scheme of alternating decisions and observations in the present paper, Takriti, Birge, and Long (1994, 1996) adopt a multi-stage point of view. Decisions are made separately for each subinterval of the optimization horizon alternating with observation of the random load values. As a counterpart to (13) a multi-stage stochastic program with integer requirements in each time stage arises. This aims at unified control of both planning (start-up scheduling) and operation (generation allocation). The present paper places accent rather on robust planning. Our hedging schedules concern start-ups only but are implementable for the whole time horizon. Hedging policies in Takriti, Birge, and Long, (1994, 1996) concern both start-ups and generation allocation but are implementable for the first time interval only. Their solution method employs progressive hedging (Rockafellar and Wets 1991), a scenario decomposition method whose convergence is granted in case of continuous variables only. Nevertheless, the authors empirically observed proper convergence for their class of mixed-integer problems. Progressive hedging splits the multi-stage stochastic program into single-scenario problems that are very close to unit commitment. In Takriti, Birge, and Long, (1994, 1996), these are tackled by Lagrangian relaxation of load balances (c.f. constraints (9) above).

Our algorithm for solving (13) differs from that in Takriti, Birge, and Long (1994, 1996) in two respects. First, in theory, finite convergence can be established for our decomposition method. Second, instead of Lagrangian relaxation based dual techniques we employ for the subproblems the primal method that was successfully applied to deterministic unit commitment problems in Dentcheva et al. (1997).

Another two-stage stochastic program for unit commitment was recently proposed in Dentcheva and Römisch (1998). The main difference to (13) is that both start-up and output variables for all generation units are included into the first stage and that second-stage decisions are adjustments of the first-stage ones. The authors present a conceptual algorithm based on Lagrangian relaxation of load balances. This splits the problem into single-unit two-stage stochastic programs of which the thermal-unit subproblems are mixed-integer ones.

2 Scenario Decomposition

The preceding section showed that the mathematical model of our problem can be written as

$$\min\{cx + Q(x) \ : \ Ax \leq b, \ x \in X\} \tag{14}$$

where

$$Q(x) = \mathbb{E}_\omega \phi(h^\omega - Tx)$$

and

$$\phi(s) = \min\{qy \ : \ Wy \leq s, \ y \in Y\}.$$

Here x, y refer to the first- and second-stage variables and X, Y denote restrictions requiring some or all of the variables to be binary. Accordingly, the data vectors and matrices are derived from (13), with the mentioned remodelling of the maximum expressions in the objective.

Since h^ω follows a discrete distribution, problem (13) is equivalent to the following large-scale linear mixed-integer program where $h^\nu, \pi^\nu, \nu = 1, \dots, r$ denote mass points and probabilities, respectively :

$$\begin{aligned} \min\Big\{cx + \sum_{\nu=1}^{r} \pi^\nu q y^\nu \ : Ax \leq b, \ x \in X, \\ Tx + Wy^\nu \leq h^\nu, \ y^\nu \in Y, \ \nu = 1, \dots, r\Big\}. \end{aligned} \tag{15}$$

The constraints matrix of (15) has the block-angular structure depicted in Figure 1. Due to the sheer size of (15) for realistic problems, state-of-the-art mixed-integer LP (MILP) codes inevitably fail when being applied to the above problem directly; see Schultz, Stougie, and Van der Vlerk (1998) for an example with the same structure but in a different context. It is therefore crucial to look for appropriate decomposition strategies. For linear stochastic programs with continuous variables various such strategies have been developed in the past. It has become quite common to distinguish primal decomposition which splits the problem according to time stages and dual decomposition which splits according to scenarios. For both approaches, however, convergence is based on intrinsic properties of linear and convex programs with continuous variables, respectively.

In primal decomposition methods (Birge 1985, Higle and Sen 1991, Ruszczyński 1986, Van Slyke and Wets 1969), first-stage decisions are iterated in a convex non-smooth master

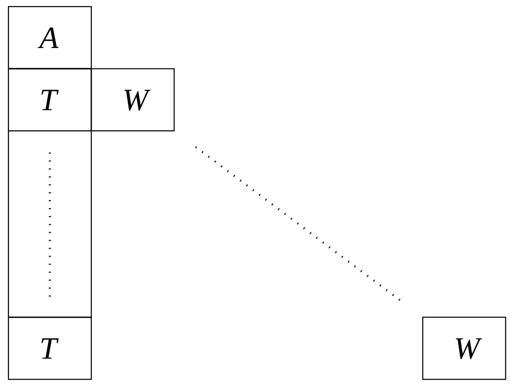

Figure 1: Constraints matrix structure of (15).

problem by suitable subgradient methods in an outer loop. In the inner loop, the second-stage problem is solved for various right-hand sides. Convexity of the master is inherited from the convexity of the value function in linear programming. In dual decomposition, (Mulvey and Ruszczyński 1995, Rockafellar and Wets 1991), a convex non-smooth function of Lagrange multipliers is minimized in an outer loop. Here, convexity is granted by fairly general reasons that would also apply with integer variables in (15). In the inner loop, subproblems differing only in their right-hand sides are to be solved. Linear (or convex) programming duality is the driving force behind this procedure that is mainly applied in the multi-stage setting.

When following the idea of primal decomposition in the presence of integer variables one faces discontinuity of the master in the outer loop. This is caused by the fact that the value function of an MILP is merely lower semicontinuous in general. Computations have to overcome the difficulty of lower semicontinuous minimization for which no efficient methods exist up to now. In Carøe and Tind (1998) this is analyzed in more detail. In the inner loop, MILPs arise which differ in their right-hand sides only. Application of Gröbner bases methods from computational algebra has led to first computational techniques that exploit this similarity in case of pure-integer second-stage problems, see Schultz, Stougie, and Van der Vlerk (1998).

With integer variables, dual decomposition runs into trouble due to duality gaps that typically arise in integer optimization. In Løkketangen and Woodruff (1996) and Takriti, Birge, and Long (1994, 1996), Lagrange multipliers are iterated along the lines of the progressive hedging algorithm in Rockafellar and Wets (1991) whose convergence proof needs continuous variables in the original problem. Despite this lack of theoretical underpinning the computational results in Løkketangen and Woodruff (1996) and Takriti, Birge, and Long (1994, 1996), indicate that for practical problems acceptable solutions can be found this way. A branch-and-bound method for stochastic integer programs that utilizes stochastic bounding procedures was derived in Ruszczyński, Ermoliev, and Norkin (1994). In Carøe and Schultz (1997) a dual decomposition method was developed that combines Lagrangian relaxation of non-anticipativity constraints with branch-and-bound. We will apply this method to the model from Section 1 and describe the main features in the remainder of the present section.

The idea of scenario decomposition is well known from stochastic programming with continuous variables where it is mainly used in the multi-stage case. For stochastic integer programs scenario decomposition is advantageous already in the two-stage case. The idea is to let $x^1, \ldots, x^r$ be copies of the first-stage variable x and rewrite (15) as

$$\begin{aligned} \min \Big\{ \sum_{\nu=1}^{r} \pi^\nu (cx^\nu + qy^\nu) \ : & \ Ax^\nu \leq b, \ x^\nu \in X, \\ & Tx^\nu + Wy^\nu \leq h^\nu, \ y^\nu \in Y, \ \nu = 1, \ldots, r, \\ & x^1 = \ldots = x^r \Big\}. \end{aligned} \tag{16}$$

The equations $x^1 = \ldots = x^r$ express independence of first-stage decisions on the realizations of h^ω and are called non-anticipativity constraints. Of course, there are several ways to express this property. To be flexible in this respect and for notational convenience we assume that non-anticipativity is represented by the constraint $\sum_{\nu=1}^{r} H^\nu x^\nu = 0$ where $H = (H^1, \ldots, H^r)$ is a suitable matrix. The block structure of the constraints matrix of formulation (16) can be seen in Figure 2 . Separability of (16) can be achieved when removing the non-anticipativity

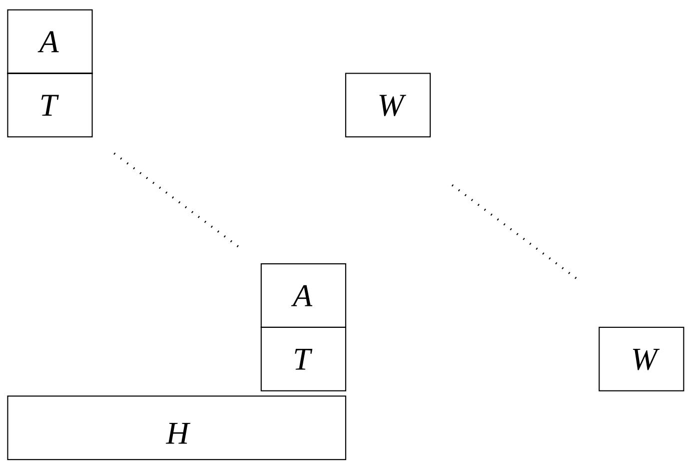

Figure 2: Constraints matrix of the scenario formulation (16).

conditions from the constraints. This leads to considering the following Lagrangian relaxation of (16):

$$\begin{aligned} D(\lambda) = \min \Big\{ \sum_{\nu=1}^{r} L^\nu (x^\nu, y^\nu, \lambda) \ : & \ Ax^\nu \leq b, \ x^\nu \in X, \\ & Tx^\nu + Wy^\nu \leq h^\nu, \, y^\nu \in Y, \, \nu = 1, \ldots, r \Big\}, \end{aligned} \tag{17}$$

where

$$L^\nu (x^\nu, y^\nu, \lambda) = \pi^\nu (cx^\nu + qy^\nu) + \lambda (H^\nu x^\nu) \ \text{ for } \ \nu = 1, \ldots, r.$$

The problem $\max_\lambda D(\lambda)$ is called the Lagrangian dual of (16). From the theory of integer linear programming it is well known (cf. Nemhauser and Wolsey 1988) that the optimal value of the Lagrangian dual is a lower bound to the optimal value of (16) which is strict in general but greater than or equal on the lower bound given by the LP relaxation of (16). If for some

λ the corresponding solution (x^ν, y^ν), $\nu = 1, \ldots, r$ of the Lagrangian relaxation (17) fulfills $x^1 = \ldots = x^r$ then it is optimal for (16).

The Lagrangian dual is a non-smooth concave maximization problem which can be solved by subgradient or bundle methods. At each iteration point a function value and a subgradient of D are needed. These are obtained by solving (17) whose major advantage is that it splits into separate subproblems for each scenario and which are identical to conventional unit commitment problems except for changes in the cost coefficients. Therefore, existing methodology, for instance that developed in Dentcheva et al. (1997), can be employed. The Lagrangian dual is thus tractable but in general only provides lower bounds for (16) unless the solution of (17) fulfills $x^1 = \ldots = x^r$. In Carøe and Schultz (1997) this led to designing a branch-and-bound algorithm where bounds are computed via the Lagrangian dual and branching is directed to establishing the identity $x^1 = \ldots = x^r$.

In the following $\mathcal{P}$ denotes the list of current problems and $z_{\mathrm{LD}} = z_{\mathrm{LD}}(P)$ is a lower bound associated with problem $P \in \mathcal{P}$. The outline of the algorithm from Carøe and Schultz (1997) is as follows:

Step 1 *Initialization:* Set $\bar{z} = +\infty$ and let $\mathcal{P}$ consist of problem (16).

Step 2 *Termination:* If $\mathcal{P} = \emptyset$ then the solution $\hat{x}$ that yielded $\bar{z} = c\hat{x} + Q(\hat{x})$ is optimal.

Step 3 *Node selection:* Select and delete a problem P from $\mathcal{P}$ and solve its Lagrangian dual. If the optimal value $z_{\mathrm{LD}}(P)$ hereof equals $+\infty$ (infeasibility of a subproblem) then go to Step 2.

Step 4 *Bounding:* If $z_{\mathrm{LD}}(P) \geq \bar{z}$ go to Step 2 (this step can be carried out as soon as the value of the Lagrangian dual rises above $\bar{z}$).

- **(i)** The scenario solutions x^ν, $\nu = 1, \ldots, r$, are identical: If $cx\nu + Q(x^\nu) < \bar{z}$ then let $\bar{z} = cx\nu + Q(x\nu)$ and delete from $\mathcal{P}$ all problems P' with $z_{\mathrm{LD}}(P') \geq \bar{z}$. Go to Step 2.
- **(ii)** The scenario solutions x^ν, $\nu = 1, \ldots, r$ differ: Compute the average $\bar{x} = \sum_{\nu=1}^r \pi^\nu x^\nu$ and round it by some heuristic to obtain $\bar{x}^{\mathrm{R}}$. If $c\bar{x}^{\mathrm{R}} + Q(\bar{x}^{\mathrm{R}}) < \bar{z}$ then let $\bar{z} = c\bar{x}^{\mathrm{R}} + Q(\bar{x}^{\mathrm{R}})$ and delete from $\mathcal{P}$ all problems P' with $z_{\mathrm{LD}}(P') \geq \bar{z}$. Go to Step 5.

Step 5 *Branching:* Select a component $x_{(m)}$ of x and add two new problems to $\mathcal{P}$ obtained from P by setting $x_{(m)} = 0$ and $x_{(m)} = 1$.

Since there are only finitely many feasible first-stage solutions in (13) the above algorithm in theory terminates after a finite number of steps.

Compared with LP-based branch-and-bound for general integer programs we relax non-anticipativity instead of integrality requirements. Thus feasibility is obtained when scenario solutions are identical. With a non-zero duality gap, however, our Lagrangian dual will lead to primal solutions that are not feasible such that we included Step 4(ii) where a heuristic procedure is used to quickly come up with feasible points. Its simplicity results from the fact that our Lagrangian relaxation concerns a class of constraints that are fairly simply structured. Further implementational details about node selection and branching rules we will report in the computation section.

3 Computational Experience

The presented algorithm was coded in Fortran using NOA 3.0 by Kiwiel (1994, 1990) and the CPLEX 4.0 Callable Library (1995) and run on realistic problem instances. The power system of VEAG comprises 17 coal fired blocks, 8 gas fired units and 7 pumped storage plants. For a typical weekly unit commitment problem involving 168 time periods, each of the scenario subproblems has approximately 14.000 constraints and 16.000 variables, of which 4200 are binary. A characteristic feature of the power system met at VEAG is that several of the coal-fired thermal blocks are identical. This allows us to reduce the size of the scenario subproblems by aggregation of these units. The start-up/shut-down decisions for these units are then represented by one integer variable, namely the number of units which are turned on. In the following we present results for both formulations.

We generated 16 scenarios for a generator failure example and an inaccurate load forecast example, respectively, and build 3 instances with 4, 10 and 16 scenarios for each of the examples. The sizes of the corresponding deterministic equivalent problems are given in Table I.

Table I: Problem sizes of deterministic equivalents.

Formulation	Scenarios	Constraints	Variables	Integers	Multipliers
Binary	4	47159	47327	7560	11424
	10	113639	109775	14616	28560
	16	180119	172223	21672	45696
Integer	4	32049	37257	5880	4704
	10	78369	89625	12936	11760
	16	124689	141993	19992	18816

We solved the Lagrangian dual problems using the proximal bundle method NOA 3.0 of K. C. Kiwiel and described in Kiwiel (1990, 1994). In contrast to standard subgradient procedures, where careful tuning of parameters and stepsize rules are often necessary for each test problem, the bundle method proved very efficient and robust for our application. To save computation time, the Lagrangian dual is solved at the root node only, and the obtained multipliers are then used for solving the Lagrangian relaxation of non-root nodes. A crucial step when applying Lagrangian relaxation techniques is the ability to generate feasible solutions. Due to the simple structure of the non-anticipativity constraints, candidates for feasible solutions can easily be found in our case. We used the average $\bar{x}$ and rounded all components to the nearest integer. For all our test runs good feasible solutions were obtained this way already in the root node of the branching algorithm.

The main workload of our algorithm consists of solving the scenario subproblems and it is therefore crucial to solve them as fast as possible. To this end, we store and update optimal bases for the LP-relaxations of all scenario subproblems every 5 iterations of the NOA code. During NOA-iterations we employ primal Simplex for solving LP-relaxations since primal feasibility is maintained and multipliers affect cost coefficients. For non-root nodes dual feasibility is maintained since variable bounds are being fixed, thus dual Simplex is employed during the branching procedure for solving LP-relaxations.

As observed in Dentcheva et al. (1997) the LP-relaxations seem to be quite tight and good feasible solutions within 0.5% of optimality can be obtained fast using the built-in rounding heuristic in CPLEX. The key observation is that runtimes can be significantly reduced by *relaxing first-stage variables to continuous variables.* Although for given multipliers the Lagrangian relaxation becomes weaker, it allows us (for a given timelimit) to perform a significantly larger number of NOA steps, thus providing a better overall bound on the optimal value. Also notice that scenario solutions do not necessarily have to be integer in order for the rounding heuristic to produce feasible solutions.

Special attention has been paid to the branching and node selection criteria of our branch-and-bound algorithm. For problems with only binary integer variables we take the average $\bar{x}$ and branch on the component for which the fractional part is closest to 0.5. Intuitively, this corresponds to the component for which the non-anticipativity condition is violated most. For the formulation with general integer variables we calculate the sum of squared deviations, $\sum_{\nu=1}^{r} \pi^{\nu}(x^{\nu} - \bar{x})^2$, and choose the component for which this dispersion measure is largest.

In the bounding step of our algorithm we check whether nodes in the branching tree can be fathomed. However, since the subproblems are not solved exactly, $z_{\rm LD}$ will not necessarily be a lower bound for the optimal value of the initial problem if it is computed on the basis of *approximate* solutions to the subproblems. Instead of approximate solutions we use the subproblem (lower) bounds returned by CPLEX when computing $z_{\rm LD}$. This guarantees that the $z_{\rm LD}$ are indeed lower bounds, however, not necessarily the best possible, so a larger branching tree may result. As node selection strategy we use a best-bound search to quickly improve the lower bound obtained from the Lagrangian dual of the root node problem.

Following the lines in Takriti, Birge, and Long (1996) we tested our algorithm at a generation shortage example and at an example with scenarios reflecting different load profiles, in particular where load curves are out of phase. The results were achieved on a Digital Alpha Personal Workstation with 500 MHz processor. Although in principle the algorithm is finite and eventually will come up with an optimal solution, it is of greater importance to produce provably good feasible solutions within reasonable time. Hence we report the results after 10 minutes of CPU-time. However, due to the tradeoff between the number of NOA-iterations, the quality of lower bounds and time needed to generate bounds and feasible solutions, the maximum number of bundle iterations in NOA is the smaller the higher the number of scenarios.

3.1 Generator failure example

The 3 instances of the generator failure example were generated as described in Section 1. All instances have a basic scenario with no generator failures whereas the rest of the scenarios have demand increases corresponding to the capacity loss when generators fall out. For the power system operated by VEAG, the coal fired units have maximum capacities between 400 and 750 MW per hour and the maximum load increase in some scenario is 1150 MW. In all our runs the multipliers were initially set to zero, e.g., the problem is decomposed into independent unit commitment problems with identical cost coefficients and load curves corresponding to the scenarios. To examine the effect of using Lagrangian duality as the basis of the bounding procedure in the branch-and-bound algorithm we also ran the algorithm with multipliers fixed to zero. Of course, NOA is superflouos in this case, but branching may be necessary to obtain a feasible solution. The column "Without NOA" reports the duality gap estimate obtained when the first feasible solution is found. The numbers in this column underline the fact that finding good multipliers is important for the success of the algorithm and may be essential for finding good feasible solutions. In Figure 3 the load profiles for an

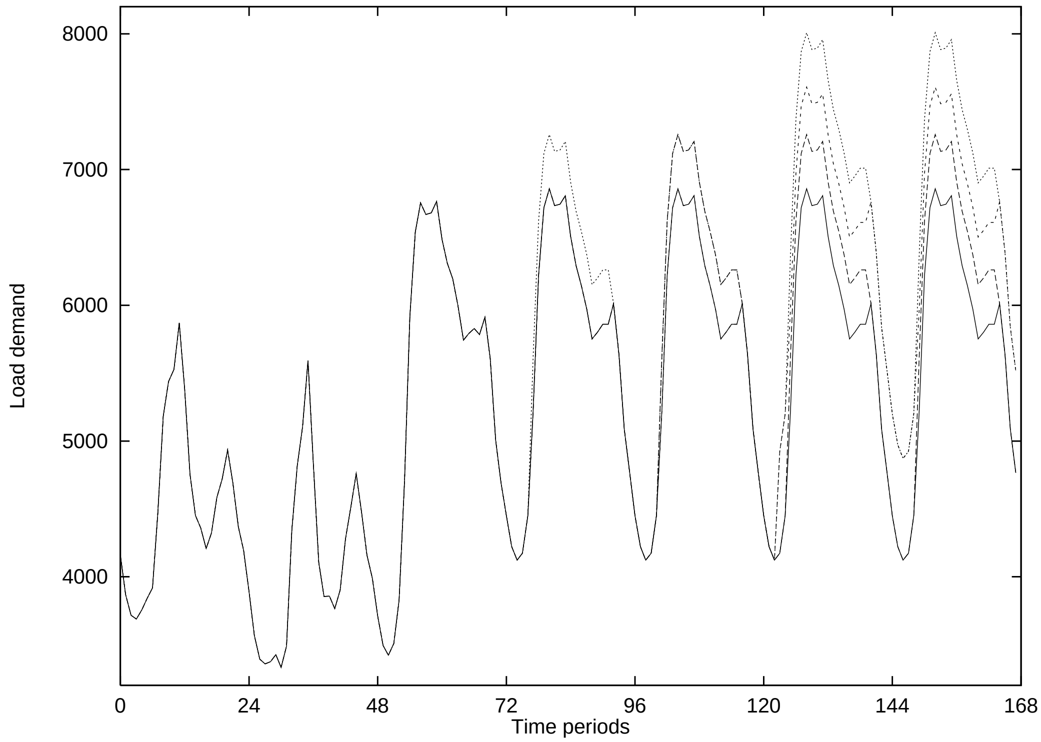

Figure 3: Load profiles for generator failure example with 4 scenarios.

instance with 4 scenarios are depicted. The computational results are summarized in Table II. First feasible solutions were found in 15 to 400 seconds of CPU-time, however, for several

Table II: Computational results for generator failure instances.

Formulation	Scenarios	NOA Steps	Best solution	Lower Bound	Gap	Without NOA
Binary	4	30	3.6417	3.6134	0.8%	3.2%
	10	10	3.6329	3.6050	0.8%	11.1%
	16	5	3.7869	3.6852	2.8%	9.7%
Integer	4	100	3.6249	3.6206	0.1%	3.2%
	10	40	3.6306	3.6251	0.2%	1.7%
	16	25	3.7208	3.7098	0.3%	2.2%

problems feasible solutions could only be obtained after several branching steps. In all cases, keeping first-stage variables as integers, turned out to give feasible solutions in the root node.

3.2 Uncertain load profile example

To test our algorithm on a problem where the scenarios have different load profile patterns we created 16 scenarios with different timing of peak periods and peak levels for the last 5 days of the time horizon. We assume the scenarios in all 3 instances have probabilities according to a uniform distribution. The load profiles for a 4 scenario instance are depicted in Figure 4.

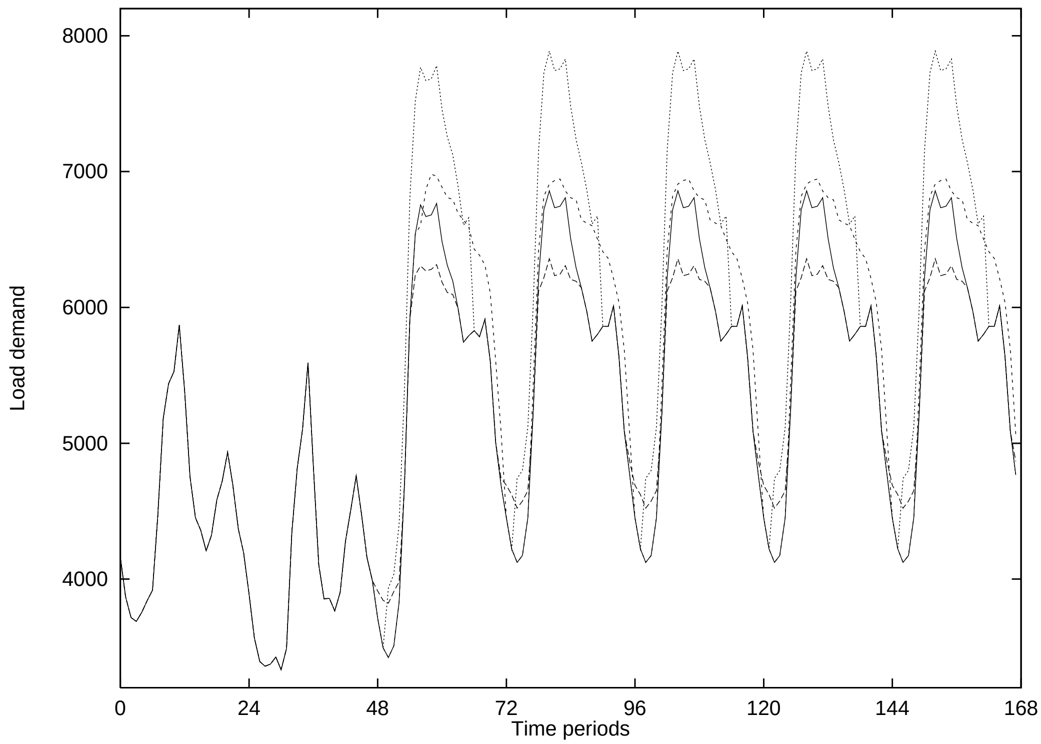

Figure 4: Load profiles for inaccurate load forecast example with 4 scenarios

The computational results for the 3 instances are given in Table III. From Tables II and III it

Table III: Computational results for inaccurate load forecast instances.

Formulation	Scenarios	NOA Steps	Best solution	Lower Bound	Gap	Without NOA
Binary	4	30	3.6598	3.6411	0.5%	1.4%
	10	10	3.6955	3.5781	3.3%	8.3%
	16	5	3.6225	3.5276	2.7%	10.1%
Integer	4	100	3.6579	3.6527	0.1%	4.1%
	10	40	3.6195	3.6080	0.3%	3.1%
	16	25	3.5698	3.5556	0.4%	2.5%

can be seen that the integer formulation is superior to the binary formulation. This is partly due to the fact that the subproblem LP-relaxations in the integer formulation are smaller and easier to solve than those of the binary formulation. However, also the performance of NOA is considerably better for the integer formulation. We conjecture this is so because the integer formulation, with first-stage variables relaxed, give rise to stronger Lagrangian relaxations than the binary formulation, but can also be credited to bad choice of initial multipliers. This said, better feasible solutions as well as tighter lower bounds can easily be obtained for the binary formulation at the expense of more computing time.

It should be noticed that the so-called Value of the Stochastic Solution (VSS), that is, the difference between the expectation of the expected value solution and the stochastic programming solution, in relative terms is quite small. However, due to the size of the costs, the stochastic programming solution represents significant savings over the deterministic model. For confidence reasons, costs displayed in the tables do not allow conclusions to the precise real figures. For a rough impression, however, it might be interesting to note that weekly fuel costs are in the range of several millions of German Marks.

There are several ways of improving the algorithm, for instance to employ several, different, rounding heuristics to come up with further candidates for feasible solutions. Moreover, the bound on the duality gap can be improved by solving the scenario subproblems with greater accuracy, at the expense of more time consuming iterations.

4 Conclusions

We have elaborated an optimization model for unit commitment under uncertainty of electrical load using modelling techniques from two-stage stochastic programming. Due to the substantial share of integer variables in both stages, this model is intractable by the majority of algorithms in stochastic programming, which hinge upon continuity of variables. As an MILP, the model is large-scale with block-angular constraints matrix. The sheer size of the model prevents application of state-of-the-art MILP solvers.

Using the scenario decomposition method from Carøe and Schultz (1997) it is possible to break down the large-scale MILP into smaller problems that are very close to conventional deterministic unit commitment problems, and for whose solution existing methods like the one in Dentcheva et al. (1997) can be employed. Our computational experience indicates that the method is able to produce feasible solutions for realistic problem instances with optimality certificates in the range of 1% within 10 minutes of computing time on a modern workstation.

Potential for improvement lies in speeding up subproblem solving and elaborating implementational issues such as the optimal tradeoff between the number of subgradient iterations and the objective function increase in the Lagrangian dual or alternative heuristics for node selection and generation of feasible points in the branch-and-bound procedure on top of the scenario decomposition.

Acknowledgements

We are grateful to Krzysztof C. Kiwiel (Polish Academy of Sciences, Warsaw) for permission to use the NOA 3.0 package and to Andrzej Ruszczyński (Rutgers University, New Brunswick) for helpful discussions on the model and algorithmic approach. VEAG Vereinigte Energiewerke AG Berlin is acknowledged for providing us with the data used in our implementation.

This research was supported by the Schwerpunktprogramm "Echtzeit - Optimierung großer Systeme" of the Deutsche Forschungsgemeinschaft and a Konrad Zuse Fellowship for the second author. Part of the work was carried out while the first author was visiting the Konrad-Zuse-Zentrum für Informationstechnik Berlin.

References

Aoki, K., M. Itoh, T. Satoh, K. Nara, and M. Kanezashi. 1989. Optimal long-term unit commitment in large scale systems including fuel constrained thermal and pumped-storage hydro. *IEEE Transactions on Power Systems* **4**, 1065–1073.

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

Birge, J. R. 1985. Decomposition and partitioning methods for multi-stage stochastic programs. *Operations Research* **33**, 989–1007.

Carøe, C. C. and R. Schultz. 1997. Dual decomposition in stochastic integer programming. Preprint SC 96-46, Konrad-Zuse-Zentrum für Informationstechnik Berlin. Submitted in revised form to *Operations Research Letters*.

Carøe, C. C. and J. Tind. 1995. L-shaped decomposition of two-stage stochastic programs with integer recourse. To appear in *Mathematical Programming*.

CPLEX Optimization, Inc. 1995. *Using the CPLEX Callable Library*. CPLEX Optimization, Inc.

Dentcheva, D., R. Gollmer, A. Möller, W. Römisch, and R. Schultz. 1997. Solving the unit commitment problem in power generation by primal and dual methods. M. Brøns, M. P. Bendsøe and M. P. Sørensen (eds.), *Progress in Industrial Mathematics at ECMI 96*, Stuttgart: Teubner, 324-331.

Dentcheva, D. and W. Römisch. 1998. Optimal power generation under uncertainty via stochastic programming. Preprint 96–35, Institut für Mathematik, Humboldt-Universität Berlin. To appear in K. Marti and P. Kall (Eds.), *Stochastic Programming: Numerical Techniques and Engineering Applications*, Lecture Notes in Economics and Mathematical Systems **458**, Berlin: Springer-Verlag.

Higle, J. L. and S. Sen. (1991). Stochastic decomposition: an algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research* **16**, 650–669.

Kiwiel, K. C. 1990. Proximity control in bundle methods for convex nondifferentiable optimization. *Mathematical Programming* **46**, 105–122.

Kiwiel, K. C. 1994 *User's Guide for NOA 2.0/3.0: A Fortran Package for Convex Nondifferentiable Optimization*. Newelska 6, 01-447 Warsaw, Poland: Systems Research Institute, Polish Academy of Sciences.

Løkketangen, A. and D. L. Woodruff. 1996. Progressive hedging and tabu search applied to mixed integer (0,1) multi-stage stochastic programming. *Journal of Heuristics* **2**, 111–128.

Mulvey, J. M. and A. Ruszczyński. 1995. A new scenario decomposition method for large-scale stochastic optimization. *Operations Research* **43**, 477–490.

Nemhauser, G. L. and L. A. Wolsey. 1988. *Integer and Combinatorial Optimization*. New York: Wiley-Interscience.

Rockafellar, R. T. and R. J.-B. Wets. 1991. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research* **16**, 119–147.

Ruszczyński, A. 1986. A regularized decomposition method for minimizing a sum of polyhedral functions. *Mathematical Programming* **35**, 309–333.

RUSZCZYŃSKI, A., Y. M. ERMOLIEV, AND V. I. NORKIN. 1994. On optimal allocation of indivisibles under uncertainty. Working-Paper WP-94-021, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria. To appear in *Operations Research*.

SCHULTZ, R., L. STOUGIE, AND M. H. VAN DER VLERK. 1998. Solving stochastic programs with integer recourse by enumeration: a framework using gröbner basis reductions. To appear in *Mathematical Programming*.

SHEBLÉ, G. B. AND G. N. FAHD. 1994. Unit commitment literature synopsis. *IEEE Transactions on Power Systems* **9**, 128–135.

TAKRITI, S., J. R. BIRGE, AND E. LONG. 1994. Intelligent unified control of unit commitment and generation allocation. Technical Report 94-26 (revised 1995), Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor.

TAKRITI, S., J. R. BIRGE, AND E. LONG. 1996. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems* **11**, 1497–1508.

TAKRITI, S., B. KRASENBRINK, AND L. S.-Y. WU. 1997. Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. IBM Research Report RC 21066, Yorktown Heights, New York.

VAN SLYKE, R. M. AND R. J.-B. WETS. 1969. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics* **17**, 638–663.