Scenario reduction and scenario tree construction for power management problems

N. Gröwe-Kuska, H. Heitsch and W. Römisch

Humboldt-University Berlin

Institute of Mathematics

IEEE Bologna POWER TECH 2003

Bologna, June 23-26, 2003

1. Electricity portfolio management

We consider a power utility owning a hydro-thermal generation system and producing and trading electric power during some time horizon $[1, T]$ (e.g., weekly, monthly, yearly).

Objective: Maximization of (expected) revenue

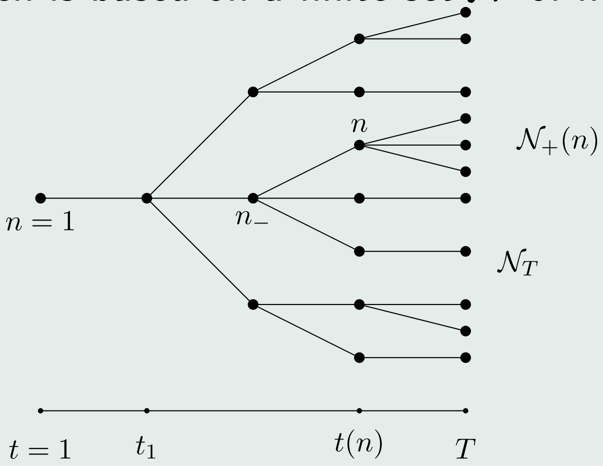
Decisions: Mixed-integer (large scale)

System and trading constraints: Capacity, reservoir, operational, load, reserve constraints

Stochastic data processes: Electrical load, fuel and electricity prices, inflows

1.1. Data process approximation by scenario trees

The data process $\xi = \{\xi_t\}_{t=1}^{T}$ is approximated by a process forming a scenario tree which is based on a finite set $\mathcal{N}$ of nodes.



Scenario tree with $t_1 = 2$, $T = 5$, $|\mathcal{N}| = 23$ and 11 leaves

The root node $n = 1$ stands for period $t = 1$. Every other node n has a unique predecessor n_- and a set $\mathcal{N}_+(n)$ of successors. Let $\text{path}(n)$ be the set $\{1, \ldots, n_-, n\}$ of nodes from the root to node n, $t(n) := |\text{path}(n)|$ and $\mathcal{N}_T := \{n \in \mathcal{N} : \mathcal{N}_+(n) = \emptyset\}$ the set of leaves. A scenario corresponds to $\text{path}(n)$ for some $n \in \mathcal{N}_T$. With the given scenario probabilities $\{\pi_n\}_{n \in \mathcal{N}_T}$, we define recursively node probabilities $\pi_n := \sum_{n_+ \in \mathcal{N}_+(n)} \pi_{n_+}$, $n \in \mathcal{N}$.

1.2. Stochastic power management model

Stochastic process: $\{\xi_t = (d_t, r_t, \gamma_t, \alpha_t, \beta_t, \zeta_t)\}_{t=1}^{T}$
(electrical load, spinning reserve, inflows, (fuel or electricity) prices)
given as a (multivariate) scenario tree
Mixed-integer programming problem:

$$\min \sum_{n\in\mathcal{N}} \pi_n \sum_{i=1}^{I} [C_i^n(p_i^n, u_i^n) + S_i^n(u_i)] \quad \text{s.t.}$$

$$p_{it(n)}^{\min} u_i^n \le p_i^n \le p_{it(n)}^{\max} u_i^n, \quad u_i^n \in \{0,1\}, \quad n \in \mathcal{N},\ i = 1{:}\, I,$$

$$u_i^{n-\tau} - u_i^{n-(\tau+1)} \le u_i^n, \quad \tau = 1{:}\, \bar{\tau}_i - 1,\ n \in \mathcal{N},\ i = 1{:}\, I,$$

$$u_i^{n-(\tau+1)} - u_i^{n-\tau} \le 1 - u_i^n, \quad \tau = 1{:}\, \underline{\tau}_i - 1,\ n \in \mathcal{N},\ i = 1{:}\, I,$$

$$0 \le v_j^n \le v_{jt(n)}^{\max},\ 0 \le w_j^n \le w_{jt(n)}^{\max},\ 0 \le l_j^n \le l_{jt(n)}^{\max},\ n \in \mathcal{N},\ j = 1{:}\, J,$$

$$l_j^n = l_j^{n-} - v_j^n + \eta_j w_j^n + \gamma_j^n, \quad n \in \mathcal{N},\ j = 1{:}\, J,$$

$$l_j^0 = l_j^{\text{in}}, \quad l_j^n = l_j^{\text{end}}, \quad n \in \mathcal{N}_T,\ j = 1{:}\, J,$$

$$\sum_{i=1}^{I} p_i^n + \sum_{j=1}^{J} (v_j^n - w_j^n) \ge d^n, \quad n \in \mathcal{N},$$

$$\sum_{i=1}^{I} (u_i^n p_{it(n)}^{\max} - p_i^n) \ge r^n, \quad n \in \mathcal{N}.$$

C_i^n are fuel or trading costs and S_i^n start-up costs of unit i at node $n \in \mathcal{N}$:

$$C_i^n(p_i^n, u_i^n) := \max_{l=1,\ldots,\bar{l}} \{\, \alpha_{il}^n p_i^n + \beta_{il}^n u_i^n \,\}$$

$$S_i^n(u_i) := \max_{\tau=0,\ldots,\tau_i^c} \zeta_{i\tau}^n (u_i^n - \sum_{\kappa=1}^{\tau} u_i^{n-\kappa})$$

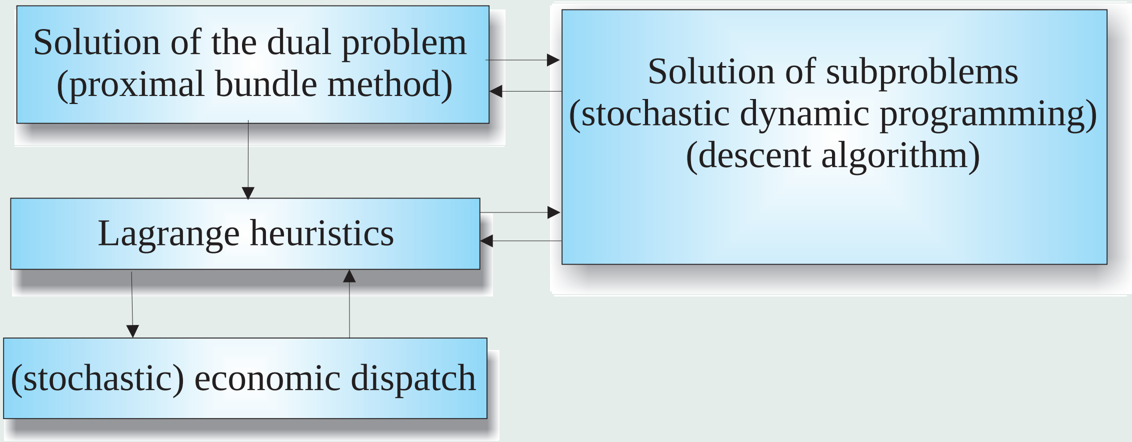
1.3. Solving the stochastic power management model

$\lvert\mathcal{N}_T\rvert$	$\lvert\mathcal{N}\rvert$	variables		constraints	nonzeros
		binary	continuous		
1	168	4200	6652	13441	19657
20	1176	29400	45864	94100	137612
50	2478	61950	96642	198290	289976
100	4200	105000	163800	336100	491500

Dimension of the model for $T = 168$, $I = 25$ and $J = 7$

$\Rightarrow$Primal approaches seem to be hopeless in general!

$\Rightarrow$Lagrangian relaxation of coupling constraints



2. Generation of scenario trees

(i) Development of a stochastic model for the data process ξ (parametric [e.g. time series model], nonparametric [e.g. resampling]) and generation of simulation scenarios;

(ii) Construction of a scenario tree out of the stochastic model or of the simulation scenarios;

(iii) optional scenario tree reduction.

Approaches for (ii):

(1) Barycentric scenario trees (conditional expectations w.r.t. a decomposition of the support into simplices);

(2) Fitting of trees with prescribed structure to given moments;

(3) Conditional sampling by integration quadratures;

(4) Clustering methods for bundling scenarios;

(5) Scenario tree construction based on optimal approximations w.r.t. certain probability metrics.

3. Distances of probability distributions

Let P denote the probability distribution of the stochastic data process $\{\xi_t\}_{t=1}^T$, where ξ_t has dimension r, i.e., P has support $\Xi \subseteq \mathbb{R}^{rT} = \mathbb{R}^s$.

The Kantorovich functional or transportation metric takes the form

$$\mu_c(P,Q) := \inf\{\int_{\Xi\times\Xi} c(\xi,\tilde{\xi})\eta(d\xi,d\tilde{\xi}) : \pi_1\eta = P,\ \pi_2\eta = Q\},$$

where $c : \Xi \times \Xi \to \mathbb{R}$ is a certain cost function.

Example: $c(\xi,\tilde{\xi}) := \max\{1, \|\xi\|^{p-1}, \|\tilde{\xi}\|^{p-1}\}\|\xi - \tilde{\xi}\| \quad (p \geq 1)$

Approach:

Select a probability metric a function $c : \Xi \times \Xi \to \mathbb{R}$ such that the underlying stochastic optimization model is stable w.r.t. μ_c.

Given P and a tolerance $\varepsilon > 0$, determine a scenario tree such that its probability distribution P_{tr} has the property

$$\mu_c(P, P_{tr}) \leq \varepsilon\,.$$

Distances of discrete distributions

P: scenarios ξ_i with probabilities p_i, $i = 1, \ldots, N$,
Q: scenarios $\tilde{\xi}_j$ with probabilities q_j, $j = 1, \ldots, M$.

Then

$$\mu_c(P,Q) = \sup\{\sum_{i=1}^{N} p_i u_i + \sum_{j=1}^{M} q_j v_j : u_i + v_j \le c(\xi_i, \tilde{\xi}_j) \ \forall i,j\}$$

$$= \inf\{\sum_{i,j} \eta_{ij} c(\xi_i, \tilde{\xi}_j) : \eta_{ij} \ge 0, \sum_j \eta_{ij} = p_i, \sum_i \eta_{ij} = q_j\}$$

(*optimal value of linear transportation problems*)

(a) Distances of distributions can be computed by solving specific linear programs.

(b) The principle of optimal scenario generation can be formulated as a best approximation problem with respect to μ_c. However, it is nonconvex and difficult to solve.

(c) The best approximation problem simplifies considerably if the scenarios are taken from a specified finite set.

4. Scenario Reduction

We consider discrete distributions P with scenarios ξ_i and probabilities p_i, $i = 1, \ldots, N$, and Q having a subset of scenarios ξ_j, $j \in J \subset \{1, \ldots, N\}$, of P, but different probabilities q_j, $j \in J$.

Optimal reduction of a given scenario set J:
The best approximation of P with respect to μ_c by such a distribution Q exists and is denoted by $\bar{Q}$. It has the distance

$$D_J = \mu_c(P, Q) = \sum_{i \in J} p_i \min_{j \notin J} c(\xi_i, \xi_j)$$

and the probabilities $\bar{q}_j = p_j + \sum_{i \in J_j} p_i$, $\forall j \notin J$, where $J_j := \{i \in J : j = j(i)\}$ and $j(i) \in \arg\min_{j \notin J} c(\xi_i, \xi_j)$, $\forall i \in J$, i.e., the optimal redistribution consists in adding the deleted scenario weight to that of some of the closest scenarios.

However, finding the optimal scenario set with a fixed number n of scenarios is a combinatorial optimization problem.

5. Fast reduction heuristics

Algorithm 1: (Simultaneous backward reduction)

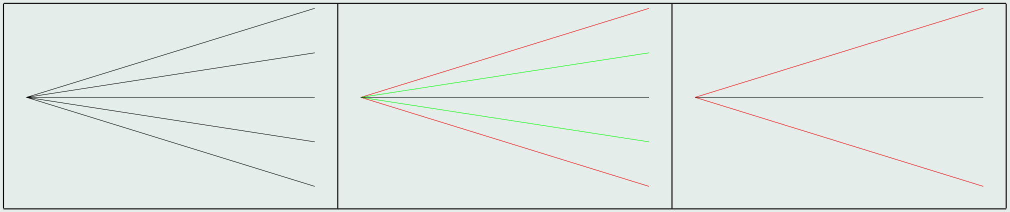
Step [0]: Sorting of $\{c(\xi_j, \xi_k) : \forall j\}, \forall k,$

$J^{[0]} := \emptyset\,.$

Step [i]: $l_i \in \arg\min_{l \notin J^{[i-1]}} \sum_{k \in J^{[i-1]} \cup \{l\}} p_k \min_{j \notin J^{[i-1]} \cup \{l\}} c(\xi_k, \xi_j).$

$J^{[i]} := J^{[i-1]} \cup \{l_i\}\,.$

Step [N-n+1]: Optimal redistribution.



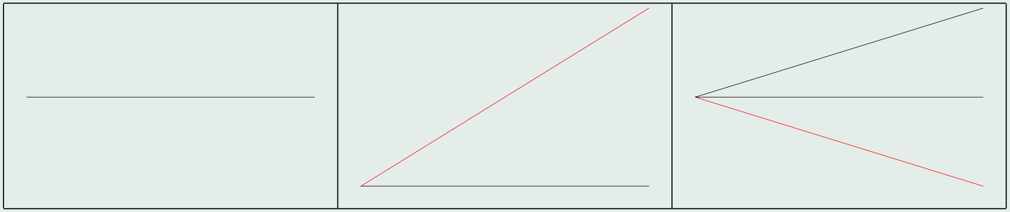
Home Page
Title Page
Contents
Page 10 of 20
Go Back
Full Screen
Close
Quit

Algorithm 2: (Fast forward selection)

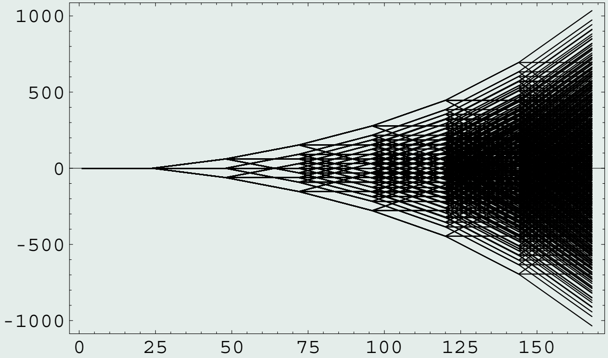
Step [0]: Compute $c(\xi_k, \xi_u)$, $k, u = 1, \ldots, N$,
$$J^{[0]} := \{1, \ldots, N\}.$$

Step [i]:
$$u_i \in \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \setminus \{u\}} p_k \min_{j \notin J^{[i-1]} \setminus \{u\}} c(\xi_k, \xi_j),$$
$$J^{[i]} := J^{[i-1]} \setminus \{u_i\}.$$

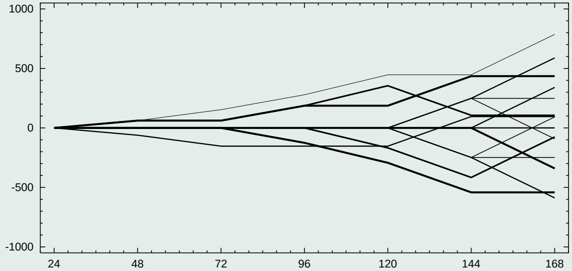
Step [n+1]: Optimal redistribution.



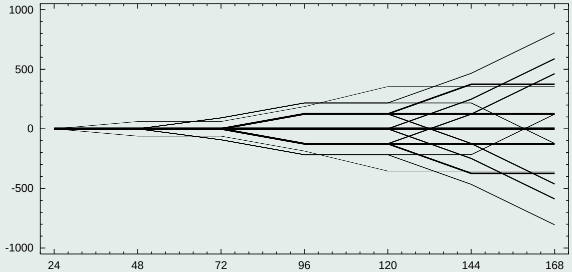
Original load scenario tree



Reduced load scenario tree / backward

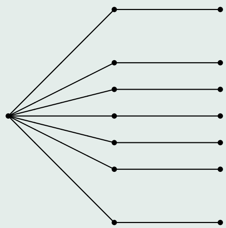


Reduced load scenario tree / forward



6. Constructing scenario trees from data scenarios

Let a fan of data scenarios $\xi^i = (\xi_1^i, \ldots, \xi_T^i)$ with probabilities π^i, $i = 1, \ldots, N$, be given, i.e., all scenarios coincide at the starting point $t = 1$, i.e., $\xi_1^1 = \ldots = \xi_1^N =: \xi_1^*$. Hence, it has the form



$t = 1$ may be regarded as the root node of the scenario tree consisting of N scenarios (leaves).

Now, P is the (discrete) probability distribution of ξ. Let c be adapted to the underlying stochastic program containing P.

We describe an **algorithm** that produces, for each $\varepsilon > 0$, a scenario tree with distribution P_ε, root node ξ_1^*, less nodes than P and

$$\mu_c(P, P_\varepsilon) < \varepsilon.$$

Recursive reduction algorithm:

Let $\varepsilon_t > 0$, $t = 1, \ldots, T$, be given such that $\sum_{t=1}^{T} \varepsilon_t \leq \varepsilon$, set $t := T$, $I_{T+1} := \{1, \ldots, N\}$, $\pi_{T+1}^i := \pi^i$ and $P_{T+1} := P$.

For $\mathbf{t = T, \ldots, 2}$:

Step t: Determine an index set $I_t \subseteq I_{t+1}$ such that

$$\mu_{c_t}(P_t, P_{t+1}) < \varepsilon_t \,,$$

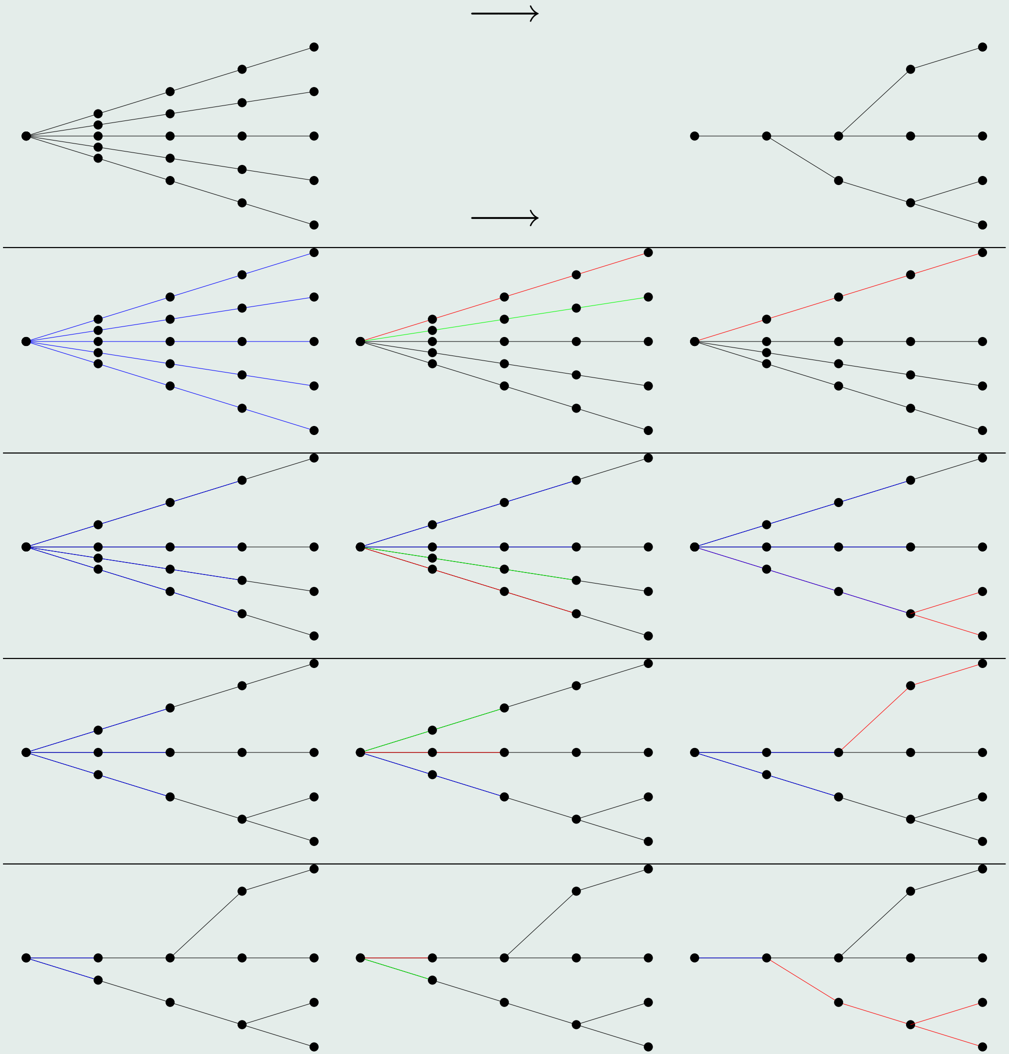
where $\{\xi^i\}_{i \in I_t}$ is the support of P_t and c_t is defined by $c_t(\xi, \tilde{\xi}) := c((\xi_1, \ldots, \xi_t, 0, \ldots, 0), (\tilde{\xi}_1, \ldots, \tilde{\xi}_t, 0, \ldots, 0))$;

(scenario reduction w.r.t. the time horizon $[1, t]$)

Step 1: Determine a probability measure P_ε such that its marginal distributions $P_\varepsilon \Pi_t^{-1}$ are $\delta_{\xi_1^*}$ for $t = 1$ and

$$P_\varepsilon \Pi_t^{-1} = \sum_{i \in I_t} \pi_t^i \delta_{\xi_t^i} \quad \text{and} \quad \pi_t^i := \pi_{t+1}^i + \sum_{j \in J_{t,i}} \pi_{t+1}^j \,,$$

where $J_{t,i} := \{j \in I_{t+1} \setminus I_t : i_t(j) = i\}$, $i_t(j) \in \arg\min_{i \in I_t} c_t(\xi^j, \xi^i)\}$ are the index sets according to the redistribution rule.



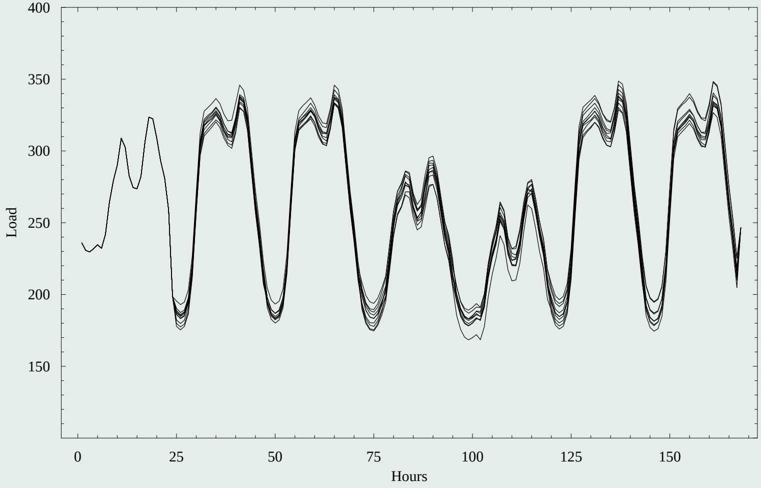
Blue: compute c-distances of scenarios; delete the green scenario & add its weight to the red one

Application:

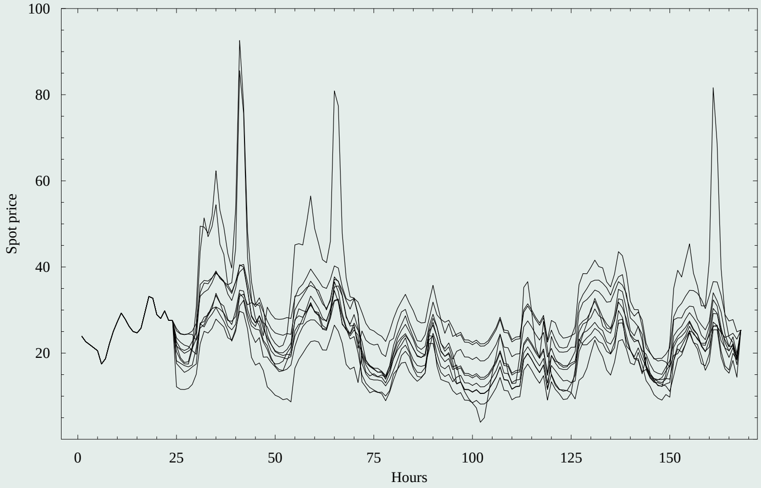
ξ is the bivariate weekly data process having the components
a) electrical load,
b) hourly electricity spot prices (at EEX).
Data scenarios are obtained from a stochastic model calibrated to the historical load data of a (small) German power utility and historical price data of the European Energy Exchange (EEX) at Leipzig.
We choose $N = 50$, $T = 7$, $\varepsilon = 0.05$, $\varepsilon_t = \frac{\varepsilon}{T}$, and arrive at a tree with 4608 nodes (instead of 8400 nodes of the original fan).

t	hours	$\lvert I_t \rvert$
1	$1 \cdots 24$	1
2	$25 \cdots 48$	12
3	$49 \cdots 72$	23
4	$73 \cdots 96$	31
5	$97 \cdots 120$	37
6	$121 \cdots 144$	42
7	$145 \cdots 168$	46

Scenario tree for the electrical load



Scenario tree for hourly spot prices



7. GAMS/SCENRED

- GAMS/SCENRED introduced to GAMS Distribution 20.6 (May 2002)

- SCENRED is a collection of C++ routines for the optimal reduction of scenarios or scenario trees

- GAMS/SCENRED provides the link from GAMS programs to the scenario reduction algorithms. The reduced problems can then be solved by a deterministic optimization algorithm provided by GAMS.

- SCENRED contains three reduction algorithms:
  - FAST BACKWARD method
  - Mix of FAST BACKWARD/FORWARD methods
  - Mix of FAST BACKWARD/BACKWARD methods

  Automatic selection (best expected performance w.r.t. running time)

Details: www.scenred.de, www.scenred.com

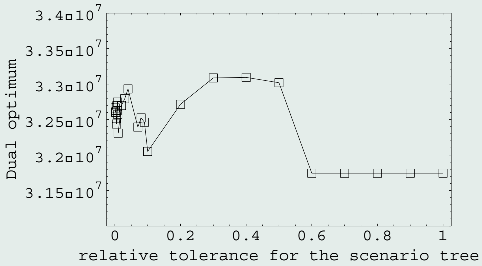
8. Numerical tests

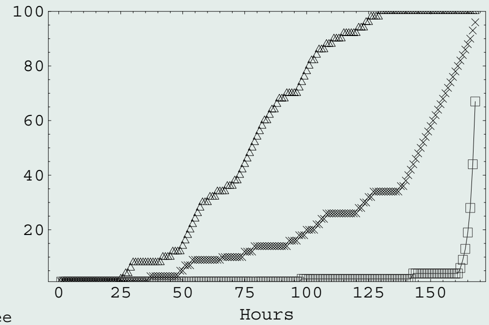
We tested the link between the Lagrangian relaxation and the scenario tree construction algorithms.

- Portfolio management problem for 25 thermal generation units and 7 pumped-storage hydro units
- Time horizon: 1 week; Discretization: 1 hour
- Initial fan of 100 load scenarios simulated from a statistical model for the load process (combines a time series model for the daily mean load with regression models for the intra-day behaviour)

Dimension and solution time for the dual

$\varepsilon_{\mathrm{rel}}$	$\lvert\mathcal{N}_T\rvert$	$\lvert\mathcal{N}\rvert$	Variables		Nonzeros	time[s]
			binary	continuous		
0.6	1	168	4200	7728	44695	7.83
0.1	67	515	12875	23690	137459	17.09
0.05	81	901	22525	41446	240233	37.82
0.01	94	2660	66500	122360	708218	150.14
0.005	96	3811	95275	175306	1014398	291.65
0.001	100	9247	231175	425362	2460402	1176.38





Dual optimum and number of scenario bundles $|I_t|$ $(t = 1, \ldots, T)$ for scenario trees with relative tolerance $\varepsilon_{\mathrm{rel}} = 0.001$ ($\triangle$), 0.005 ($\times$), 0.01 ($\square$)