Mean-Risk Optimization Models for Electricity Portfolio Management

A. Eichhorn, H. Heitsch, W. Römisch, I. Wegner-Specht

Humboldt-University Berlin
Institute of Mathematics
10099 Berlin, Germany

http://www.math.hu-berlin.de/~romisch

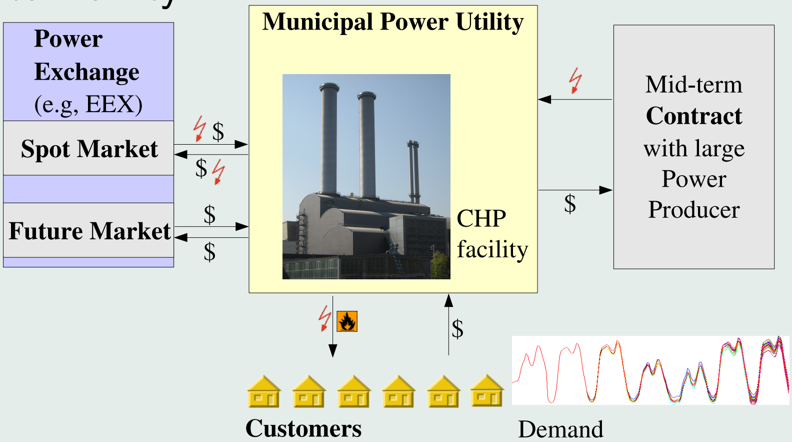
Energy Workshop, University Vienna, March 17, 2006

WWTF – Wiener Wissenschafts-, Forschungs- und Technologiefonds

DFG Research Center MATHEON
Mathematics for key technologies

1. Introduction

We consider a typical German municipal power utility, which has to serve an electricity demand and a heat demand of customers in a city and its vicinity.



Scheme of the optimization model components

The power utility owns a *combined heat and power* (CHP) production facility that can serve the heat demand completely and the electricity demand partly. Further electricity can be obtained by purchasing volumes for each hour at the (day-ahead) spot market of the European Energy Exchange (EEX), and by signing a supply contract for a medium term horizon with a larger power producer.

Portfolio:

- own power production,
- (mid-term) supply contract (no, fix, flexible),
- (day-ahead) spot market (EEX),
- electricity futures.

Objective:

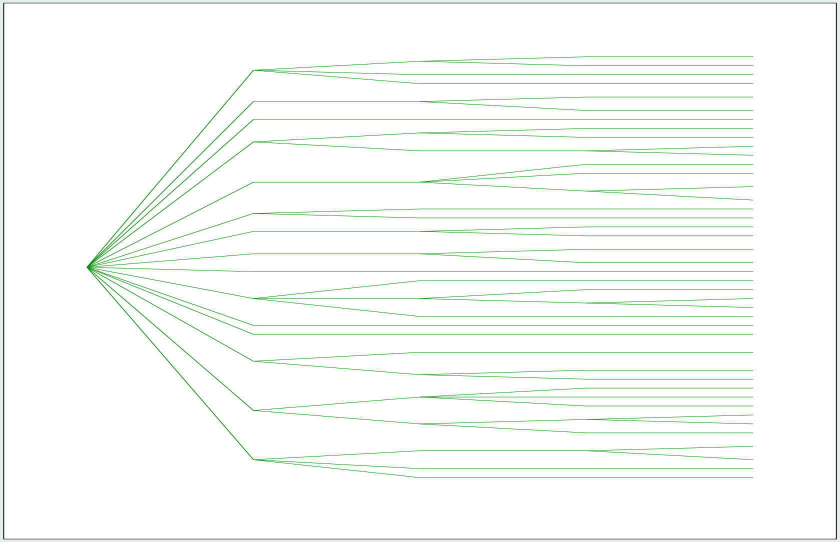
Maximize mean overall revenue (and minimize risk)

Time horizon:

One year with hourly time discretization

Uncertainty:
electricity demand, heat demand, spot market prices, (future prices) over time (represented by a variety of scenarios with their probabilities)

Modeling requirement:
recursive observation and decision (represented by branchings of scenarios at prescribed time points)



Tree of scenarios for the future

2. Mathematical Model

Let $\{\xi_t\}_{t=1}^T$ be a discrete-time stochastic data process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with ξ_1 deterministic. The stochastic decision x_t at period t is assumed to be measurable with respect to $\mathcal{F}_t := \sigma(\xi_1, \ldots, \xi_t)$ (nonanticipativity).

Multistage stochastic optimization model:

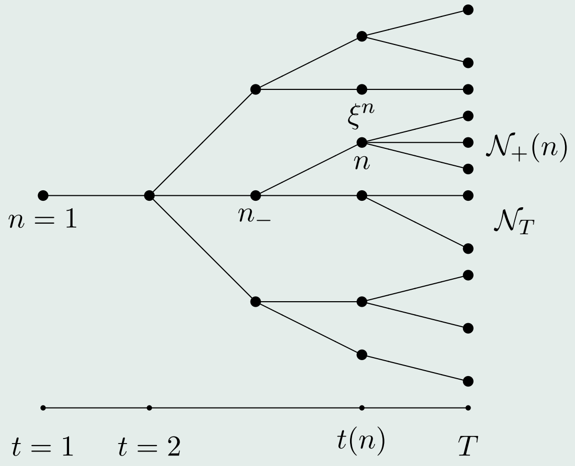
$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \langle b_t(\xi_t), x_t \rangle\right] \middle| \begin{array}{l} x_t \in X_t, t = 1, \ldots, T, A_{1,0}x_1 = h_1(\xi_1), \\ x_t \text{ is } \mathcal{F}_t - \text{measurable}, t = 1, \ldots, T, \\ A_{t,0}x_t + A_{t,1}x_{t-1} = h_t(\xi_t), t = 2, ., T \end{array} \right\}$$

where the sets X_t, $t = 1, \ldots, T$, are polyhedral cones, the vectors $b_t(\cdot)$ and $h_t(\cdot)$ depend affine linearly on ξ_t.

If the process $\{\xi_t\}_{t=1}^T$ has a finite number of scenarios, they exhibit a scenario tree structure.

3. Data process approximation by scenario trees

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



Scenario tree with $T = 5$, $N = 22$ and 11 leaves

$n = 1$ root node, n_- unique predecessor of node n, $\mathrm{path}(n) = \{1, \ldots, n_-, n\}$, $t(n) := |\mathrm{path}(n)|$, $\mathcal{N}_+(n)$ set of successors to n, $\mathcal{N}_T := \{n \in \mathcal{N} : \mathcal{N}_+(n) = \emptyset\}$ set of leaves, $\mathrm{path}(n)$, $n \in \mathcal{N}_T$, scenario with (given) probability π^n, $\pi^n := \sum_{\nu \in \mathcal{N}_+(n)} \pi^\nu$ probability of node n, ξ^n realization of $\xi_{t(n)}$.

Tree representation of the optimization model

$$\min\left\{\sum_{n\in\mathcal{N}}\pi^n\langle b_{t(n)}(\xi^n),x^n\rangle\,\middle|\,\begin{array}{l}x^n\in X_{t(n)},n\in\mathcal{N},A_{1,0}x^1=h_1(\xi^1)\\ A_{t(n),0}x^n+A_{t(n),1}x^{n-}=h_{t(n)}(\xi^n),n\in\mathcal{N}\end{array}\right\}$$

How to solve the optimization model ?

- Standard software (e.g., CPLEX)

- Decomposition methods for (very) large scale models
  (Ruszczynski/Shapiro (Eds.): Stochastic Programming, Handbook, 2003)

Open questions:

- How to model and incorporate risk ?

- How to generate (multivariate) scenario trees ?

4. Generation of scenario trees

(i) Development of a statistical model for the stochastic 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 statistical model or of the simulation scenarios.

Approaches for (ii):

(1) Bound-based approximation methods,

(Frauendorfer 96, Edirisinghe 99, Casey/Sen 05).

(2) Monte Carlo-based schemes (inside or outside decomposition methods) (e.g. Shapiro 03, 06, Higle/Rayco/Sen 01, Chiralaksanakul/Morton 04).

(3) the use of Quasi Monte Carlo integration quadratures

(Pennanen 05, 06).

(4) EVPI-based sampling schemes (inside decomposition schemes)

(Dempster/Thompson 99).

(5) Moment-matching principle (Høyland/Wallace 01, Høyland/Kaut/Wallace 03).

(6) (Nearly) best approximations based on probability metrics

(Pflug 01, Hochreiter/Pflug 02, Gröwe-Kuska/Heitsch/Römisch 01, Heitsch/Römisch 05).

Survey: Dupačová/Consigli/Wallace 00

5. Stability of multistage models

Stability studies the behavior of the optimization problem if the input ξ is approximated.

We assume that the stochastic input process $\xi = \{\xi_t\}_{t=1}^T$ belongs to the linear space $\times_{t=1}^T L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ for some $r > 1$. The optimization model is regarded in the space $\times_{t=1}^T L_{r'}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{m_t})$ for some $r' > 1$, where both spaces are endowed with the norms

$$\|\xi\|_r := \Big(\sum_{t=1}^T \mathbb{E}[|\xi_t|^r]\Big)^{\frac{1}{r}}$$

$$\|x\|_{r'} := \Big(\sum_{t=1}^T \mathbb{E}[|x_t|^{r'}]\Big)^{\frac{1}{r'}}$$

respectively. Here, $|\cdot|$ denotes some norm on the relevant Euclidean spaces and r' is defined by

$$r' := \begin{cases} \frac{r}{r-1} & , \text{ if costs are random ,} \\ r & , \text{ if only right-hand sides are random.} \end{cases}$$

Let $v(\xi)$ denote the optimal value of the stochastic optimization model.

Assumptions:

(A1) $\mathbb{E}[|\xi|^r] < \infty$,

(A2) The optimization model has relatively complete recourse,

(A3 The objective function is level-bounded locally uniformly at ξ.

Theorem: (Heitsch/Römisch/Strugarek 05)

Let (A1), (A2) and (A3) be satisfied.

There exist constants $L > 0$ and $\delta > 0$ such that the estimate

$$|v(\xi) - v(\tilde{\xi})| \leq L(\|\xi - \tilde{\xi}\|_r + D_{\mathrm{f}}(\xi, \tilde{\xi}))$$

holds for all $\tilde{\xi} \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$ with $\|\tilde{\xi} - \xi\|_r \leq \delta$.

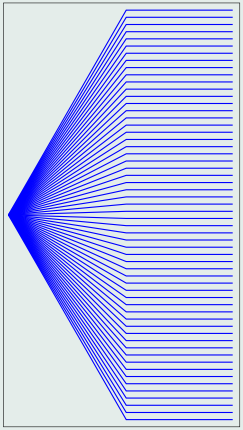
Here, D_{f} is the filtration distance

$$D_{\mathrm{f}}(\xi, \tilde{\xi}) = \inf \sum_{t=2}^{T-1} \max\{\|x_t - \mathbb{E}[x_t|\mathcal{F}_t(\tilde{\xi})]\|_{r'}, \|\tilde{x}_t - \mathbb{E}[\tilde{x}_t|\mathcal{F}_t(\xi)]\|_{r'}\},$$

where x and $\tilde{x}$ are solutions of the stochastic optimization problem.

6. Constructing scenario trees

Let ξ be the original stochastic process and ξ^{f} a stochastic process consisting of a fan of (multivariate) scenarios (paths) $\xi^i = (\xi_1^i, \ldots, \xi_T^i)$ with probabilities π^i, $i = 1, \ldots, N$, i.e., all scenarios coincide at $t = 1$, i.e., $\xi_1^1 = \ldots = \xi_1^N =: \xi_1^*$.



The fan may be regarded as a scenario tree with $1+N(T-1)$ nodes. We assume that $r \in \mathbb{N}$ is adapted to the underlying stochastic program with input process $\{\xi_t\}_{t=1}^T$ by stability arguments.

Idea: Recursive scenario reduction and bundling on $[1, t]$, $t = 2, \ldots, T$.

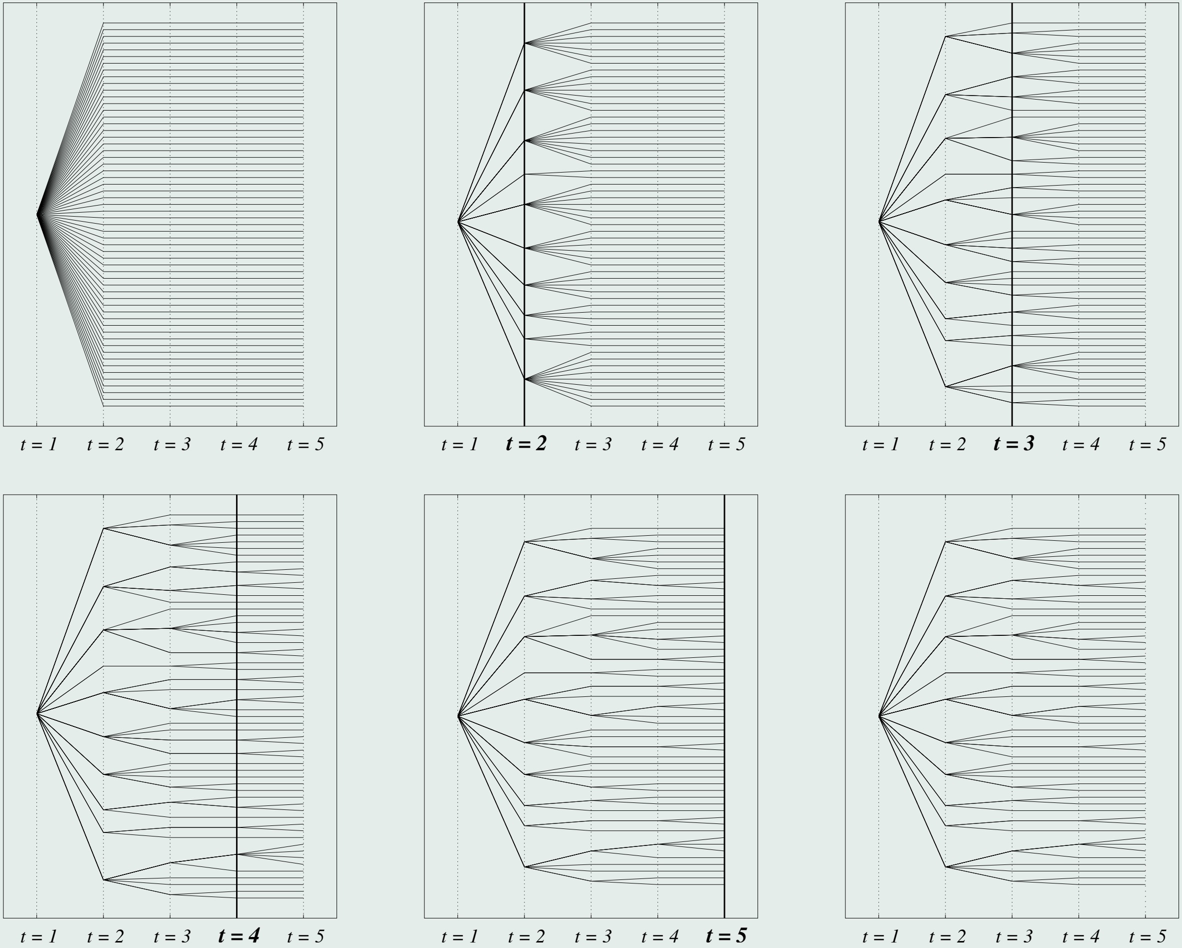
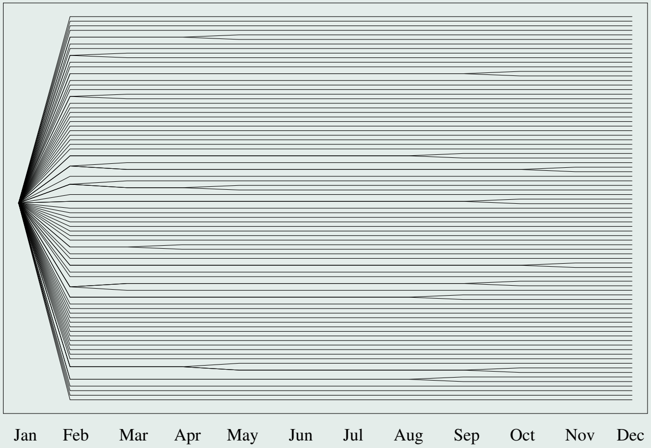
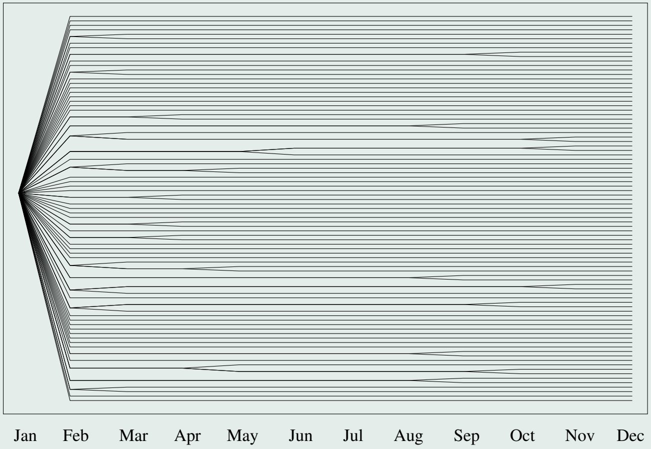


Illustration of the forward tree construction for an example including T=5 time periods starting with a scenario fan containing N=58 scenarios

<Start Animation>

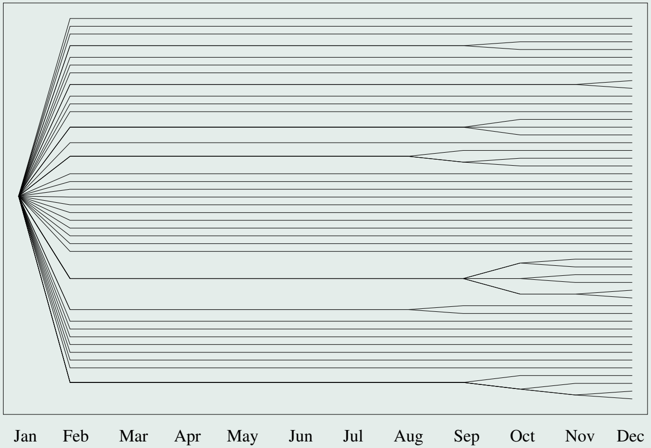


a) Forward tree construction with filtration level 0.35

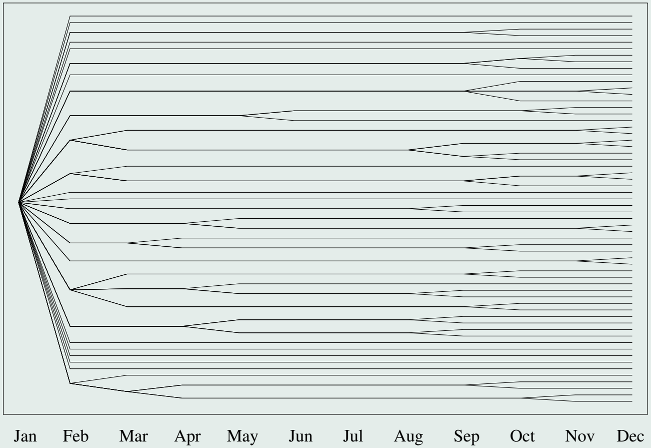


b) Forward tree construction with filtration level 0.45

Yearly demand-price scenario trees with relative tolerance 0.25



a) Modified forward tree construction with filtration level 0.6



b) Modified forward tree construction with filtration level 0.7

Yearly demand-price scenario trees with relative reduction level 0.5 (Heitsch/Römisch 05)

7. Risk functionals

Let z be a real random variable on some probability space $(\Omega, \mathcal{F}, I\!P)$. Assume that z is the revenue depending on a decision in some stochastic optimization model. The traditional objective of such models consists in maximizing the expected revenue, i.e.,

$$\max \; I\!E[z].$$

However, the revenue z of some or many decisions might have fat tails, in particular, to the left. Looking only at the expectation of z hides any tail information.

Examples of risk functionals:

Upper semivariance:

$$sV_+(z) := I\!E[[I\!E[z] - z]_+^2] = I\!E[\max\{I\!E[z] - z, 0\}^2]$$

Value-at-Risk:

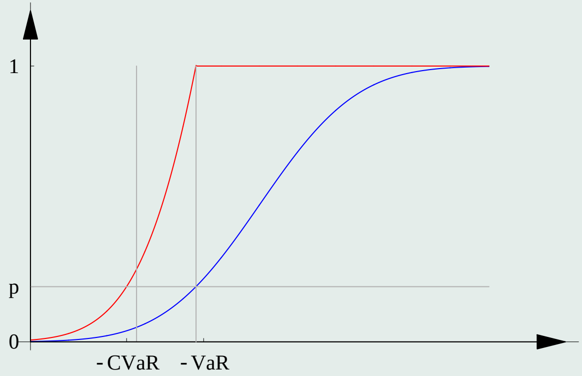
$$VaR_p(z) := -\inf\{r \in I\!R : I\!P(z \le r) \ge p\} \quad (p \in (0,1))$$

Conditional Value-at-risk:

$$CVaR_p(z) := \text{mean of the tail distribution function } F_p$$

where $F_p(t) := \begin{cases} 1 & t \geq -VaR_p(z), \\ \frac{F(t)}{p} & t < -VaR_p(z) \end{cases}$ and

$F(t) := I\!P(\{z \leq t\})$ is the distribution function of z.



$VaR_p(z)$ and $CVaR_p(z)$ for a continuously distributed z

Axiomatic characterization of risk:

Let $\mathcal{Z}$ denote a linear space of real random variables on some probability space $(\Omega, \mathcal{F}, I\!P)$. We assume that $\mathcal{Z}$ contains the constants. A functional $\rho : \mathcal{Z} \to I\!R$ is called a risk measure if it satisfies the following two conditions for all z, $\tilde{z} \in \mathcal{Z}$:

(i) If $z \leq \tilde{z}$, then $\rho(z) \geq \rho(\tilde{z})$ (monotonicity).

(ii) For each $r \in I\!R$ we have $\rho(z + r) = \rho(z) - r$ (equivariance).

A risk measure ρ is called convex if it satisfies the condition

$$\rho(\lambda z + (1 - \lambda)\tilde{z}) \leq \lambda \rho(z) + (1 - \lambda)\rho(\tilde{z})$$

for all z, $\tilde{z} \in \mathcal{Z}$ and $\lambda \in [0, 1]$.
A convex risk measure is called coherent if it is positively homogeneous, i.e., $\rho(\lambda z) = \lambda \rho(z)$ for all $\lambda \geq 0$ and $z \in \mathcal{Z}$.

References: Artzner/Delbaen/Eber/Heath 99, Föllmer/Schied 02

8. Multiperiod polyhedral risk functionals

When (real) random variables $z_1, ..., z_T$ with $z_t \in L_p(\Omega, \mathcal{F}_t, I\!P)$, $1 \le p \le +\infty$, are considered that evolve over time and unveil the available information with the passing of time, it may become necessary to use multiperiod risk functionals. We consider the filtration of σ-fields $\mathcal{F}_t = \sigma\{z_1, \ldots, z_T\}$, $t = 1, ..., T$, i.e., $\mathcal{F}_t \subseteq \mathcal{F}_{t+1} \subseteq \mathcal{F}$, and $\mathcal{F}_1 = \{\emptyset, \Omega\}$, i.e. that z_1 is always deterministic.

Definition: (Artzner et al. 01, 02)
A functional ρ on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!P)$ is called multiperiod risk functional if
(i) If $z_t \le \tilde{z}_t$ a.s., $t = 1, ..., T$, then $\rho(z_1, ..., z_T) \ge \rho(\tilde{z}_1, ..., \tilde{z}_T)$ *(monotonicity)*,
(ii) For each $r \in I\!R$ we have $\rho(z_1 + r, ..., z_T + r) = \rho(z) - r$ *(equivariance)*,
are satisfied. It is called a multiperiod coherent risk measure, if ρ is convex and positively homogeneous on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!P)$ in addition.

It is a natural idea to introduce risk functionals as optimal values of certain multistage stochastic programs.

Definition: A multiperiod risk functional ρ on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!P)$ is called polyhedral if there are $k_t \in I\!N$, $c_t \in I\!R^{k_t}$, $t = 1, \ldots, T$, $w_{t\tau} \in I\!R^{k_{t-\tau}}$, $t = 1, \ldots, T$, $\tau = 0, ..., t-1$, and polyhedral cones $Y_t \subset I\!R^{k_t}$, $t = 1, \ldots, T$, such that

$$\rho(z) = \inf \left\{ I\!E \left[\sum_{t=1}^{T} \langle c_t, y_t \rangle \right] \middle| \begin{array}{c} y_t \in L_p(\Omega, \mathcal{F}_t, I\!P; I\!R^{k_t}),\ y_t \in Y_t \\ \sum_{\tau=0}^{t-1} \langle w_{t,\tau}, y_{t-\tau} \rangle = z_t,\ \ t = 1, \ldots, T \end{array} \right\}.$$

Remark: A convex combination of (negative) expectation and of a multiperiod polyhedral risk functional is again a multiperiod polyhedral risk functional.

Polyhedral risk functionals preserve linearity and decomposition structures of optimization models.

(Eichhorn/Römisch 05)

Theory: Characterization of coherent multiperiod polyhedral risk functionals based on duality results.

Example: (Multiperiod extensions of CVaR)
A first idea is to incorporate the Conditional-Value-at-Risk at all time periods and to consider the weighted sum

$$\rho_1(z) := \sum_{t=2}^{T} \gamma_t CVaR_{\alpha_t}(z_t) = \sum_{t=2}^{T} \gamma_t \inf_{r \in I\!R} \left\{ r + \frac{1}{\alpha_t} I\!E\left[(r+z)^-\right] \right\}$$

with some weights $\gamma_t \geq 0$, $\sum_{t=1}^{T} \gamma_t = 1$, and some confidence levels $\alpha_2, \alpha_3, \ldots, \alpha_T \in (0,1)$. Here, $a^- = \max\{0, -a\}$.
Then ρ is a coherent multiperiod polyhedral risk functional.
By interchanging sum and minimization one arrives at the variant

$$\rho_2(z) = \inf_{r \in I\!R} \left\{ r + \sum_{t=2}^{T} \beta_t I\!E\left[(z_t + r)^-\right] \right\}.$$

It is also coherent and polyhedral.

9. Electricity portfolio management: statistical models and scenario trees

For the stochastic input data of the optimization model here (electricity demand, heat demand, and electricity spot prices), a very heterogeneous statistical model is employed. It is adapted to historical data in the following way:

- cluster classification for the intra-day (demand and price) profiles

- 3-dimensional time series model for the daily average values (deterministic trend functions, a trivariate ARMA model for the (stationary) residual time series)

- simulation of an arbitrary number of three dimensional sample paths (scenarios) by sampling the white noise processes for the ARMA model and by adding on the trend functions and matched intra-day profiles from the clusters afterwards.

- generation of scenario trees by the forward tree construction technique

10. Electricity portfolio management: Results

Test runs were performed on real-life data of the utility DREWAG Stadtwerke Dresden GmbH leading to three different linear programs (supply contract: no, flexible, fix) containing $T = 365 \cdot 24 = 8760$ time steps, a scenario tree with 40 demand-price scenarios and about $N = 150.000$ nodes. The objective function is of the form

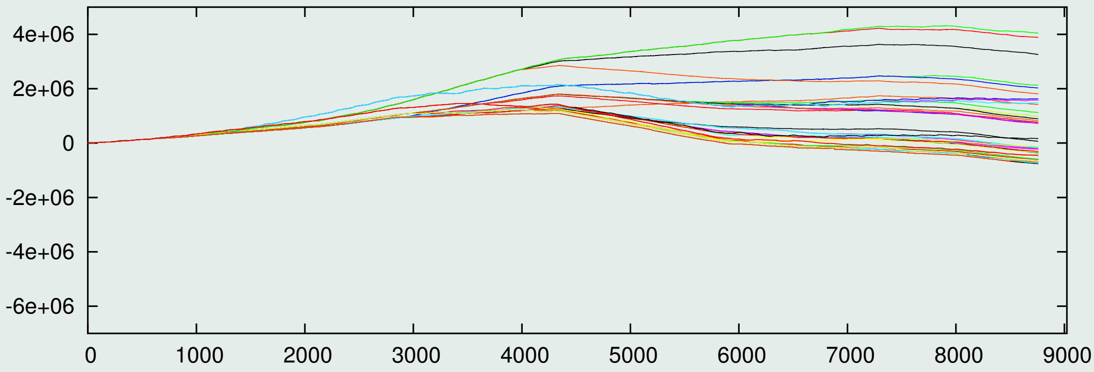
$$\text{Minimize} \quad \gamma\rho(z) - (1-\gamma)\mathbb{E}[z_T]$$

with a (multiperiod) risk functional ρ with coefficient $\gamma \in [0,1]$ ($\gamma = 0$ corresponds to no risk). $\mathbb{E}[z_T]$ denotes the overall expected revenue.

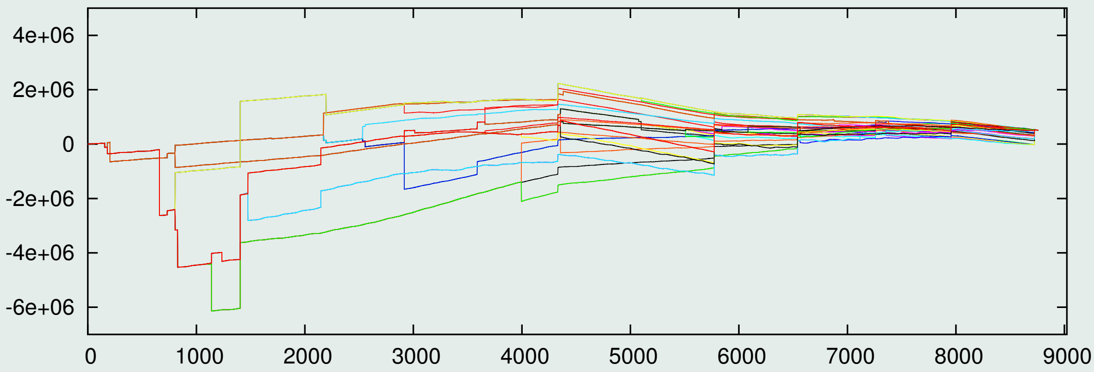
Result:

No contract is the best alternative in terms of expected revenue and, surprisingly, in terms of risk!

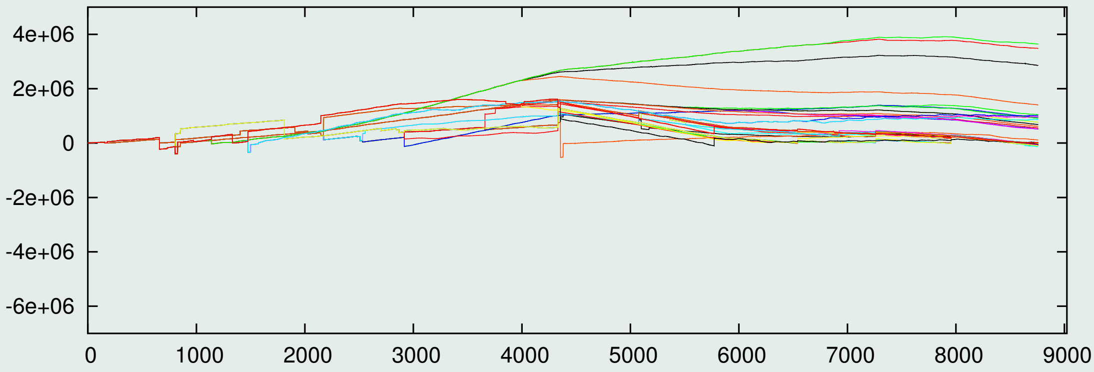
The model is implemented and solved with ILOG CPLEX 9.1 on a 2 GHz Linux PC with 1 GB memory.



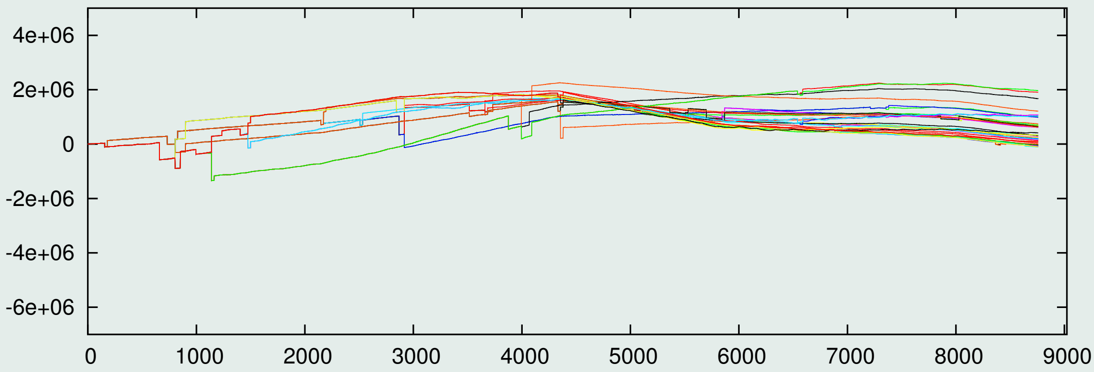
Total revenue and $\gamma = 0$



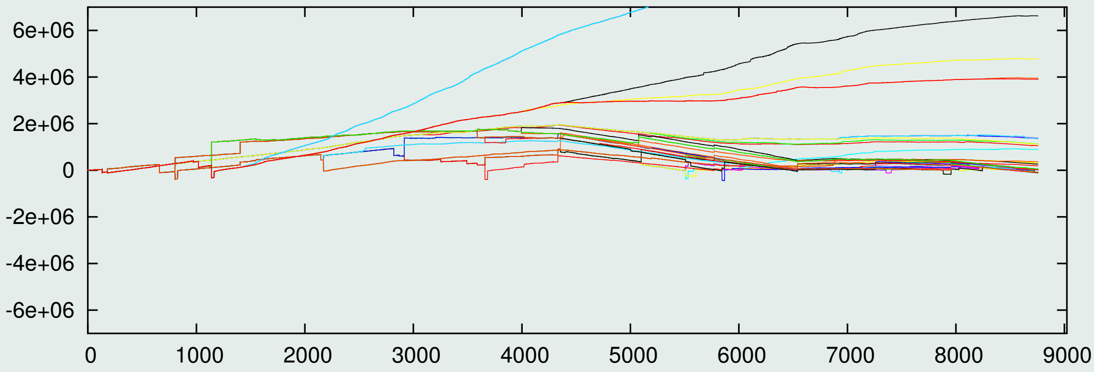
Total revenue with $CVaR_{0.05}$ and $\gamma = 0.9$



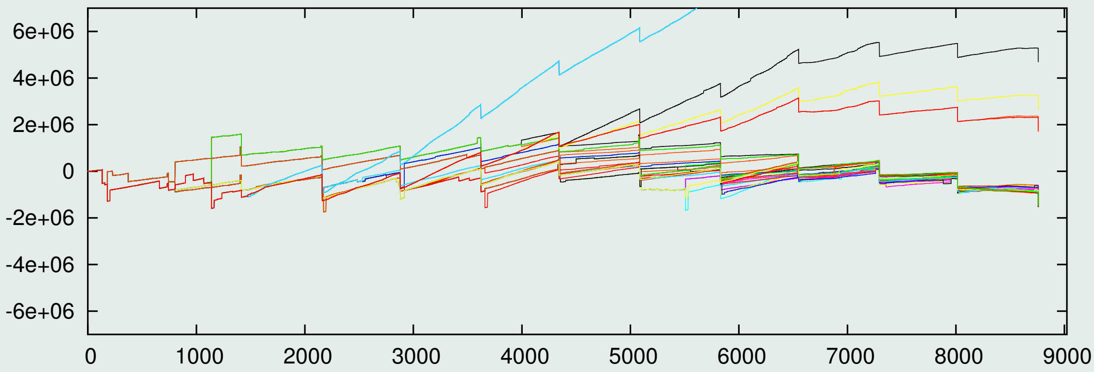
Total revenue with ρ_2 and $\gamma = 0.9$



Total revenue with ρ_4 and $\gamma = 0.9$



Total revenue with ρ_2, $\gamma = 0.9$ and the fix contract



Total revenue with ρ_2, $\gamma = 0.9$ and the flexible contract