Mean-Risk Optimization of Electricity Portfolios

W. Römisch

Humboldt-University Berlin
Institute of Mathematics
10099 Berlin, Germany

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

(A. Eichhorn, H. Heitsch)

Operations Research 2008, Universität Augsburg, 2.9.-5.9.08

DFG Research Center MATHEON
Mathematics for key technologies

Bundesministerium
für Bildung
und Forschung

Introduction

- Electricity portfolio optimization models often contain uncertain parameters (e.g., electricity spot prices, electrical load, wind speed, inflows to reservoirs) for which statistical data is available.

- The uncertain parameters may be represented approximatively by a finite number of scenarios and their probabilities.

- Scenarios become tree-structured if they appear in a process of recursive observations and decisions,

- Advantages of such stochastic programming models:
  - Decisions are robust with respect to random perturbations,
  - The risk of decisions can be modeled properly and minimized,
  - Simulation studies show the better performance.

Mathematical Model

Let $\{\xi_t\}_{t=1}^T$ be a discrete-time stochastic data process defined on some probability space $(\Omega, \mathcal{F}, I\!\!P)$ 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:

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

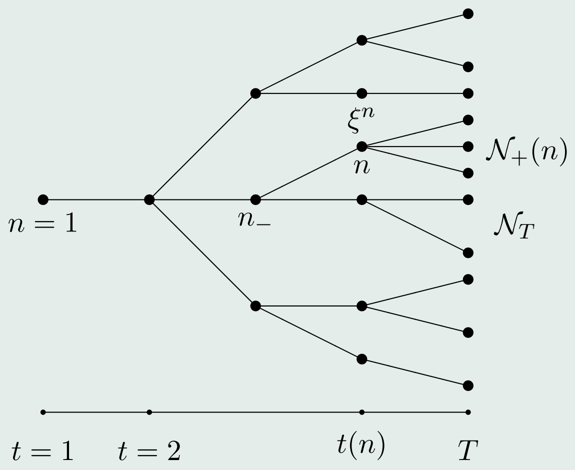
where the sets X_t, $t = 1, \ldots, T$, are closed subsets of $I\!\!R^{m_t}$ (containing linear and (possibly) integrality constraints), the vectors $b_t(\cdot)$ and $h_t(\cdot)$ depend affine linearly on ξ_t.

To have the model well defined as optimization problem in infinite dimensions one may assume, for example, that for some $p \geq 1$

$$x_t \in L_p(\Omega, \mathcal{F}_t, I\!\!P; I\!\!R^{m_t}) \quad (t = 1, \ldots, T).$$

Data process approximation by scenario trees

The process $\{\xi_t\}_{t=1}^{T}$ is approximated by a process whose scenarios are tree-structured with nodes from a finite set $\mathcal{N} \subset \mathbb{N}$.



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:

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

How to solve that (mixed-integer) linear program ?

- Standard software (e.g., CPLEX, X-PRESS).

- Decomposition methods for (very) large scale programs.

- Implementation for general (”irregular”) scenario trees.

Mathematical challenges

- Decomposition methods of the resulting large scale (mixed-integer) linear programming models,
- Generation of scenarios from statistical models (e.g., simulation from time series models, resampling techniques),
- Generation of scenario trees out of given scenarios and their eventual reduction,
- Risk modeling and minimization.

Decomposition methods

Dual decomposition approaches:

(i) Scenario decomposition by Lagrangian relaxation of nonanticipativity constraints,
(ii) nodal decomposition by Lagrangian relaxation of dynamic constraints,
(iii) geographical decomposition by Lagrangian relaxation of (decision) coupling constraints (for block separable models).

The dual can be solved by bundle subgradient methods followed by branch-and-bound techniques or Lagrangian heuristics (in case of small duality gaps).

Result: (Dentcheva-Rö 04)

Nodal decomposition leads to larger duality gaps than scenario as well as geographical decomposition. The relation of the size of duality gaps of the latter two decomposition schemes depends on the structure of the stochastic programs.

Geographical decomposition

In electricity optimization the tree representation of multistage stochastic programs often has block separable structure

$$\min \left\{ \sum_{n \in \mathcal{N}} \pi^n \sum_{i=1}^{k} \langle b^i_{t(n)}(\xi^n), x^n_i \rangle \;\middle|\; \begin{array}{l} x^n_i \in X^i_{t(n)} \\ \sum_{i=1}^{k} B^i_{t(n)}(\xi^n) x^n_i \geq g_{t(n)}(\xi^n) \\ A^i_{t(n),0} x^n_i + A^i_{t(n),1} x^{n-}_i = h^i_{t(n)}(\xi^n) \\ i = 1, \ldots, k, n \in \mathcal{N} \end{array} \right\}$$

Lagrange relaxation of coupling constraints: $L(x, \lambda) =$

$$\sum_{n \in \mathcal{N}} \pi^n \left(\sum_{i=1}^{k} \langle b^i_{t(n)}(\xi^n), x^n_i \rangle + \left\langle \lambda^n, (g_{t(n)}(\xi^n) - \sum_{i=1}^{k} B^i_{t(n)}(\xi^n) x^n_i \right\rangle \right)$$

The dual problem

$$\max_{\lambda \geq 0} \inf_{x} L(x, \lambda)$$

decomposes into k geographical subproblems and is solved by bundle subgradient methods. For nonconvex models the duality gap is typically small allowing for Lagrangian heuristics.

Generation of scenario trees

Some recent approaches:

(1) Bound-based approximation methods: Kuhn 05, Casey-Sen 05.

(2) Monte Carlo-based schemes: Shapiro 03, 06.

(3) Quadrature formulae: Pennanen 05, 06 (Quasi Monte Carlo), Chen-Mehrotra 07 (sparse grids).

(4) Moment-matching principle: Wallace et al. 01, 03.

(5) Stability-based approximations: Pflug 01, Hochreiter-Pflug 07, Mirkov-Pflug 07, Heitsch-Rö 05.

Stability of linear multistage models

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-Rö-Strugarek 06)

Let (A1) – (A3) be satisfied and X_1 be bounded.

There exist constants $L > 0$ and $\delta > 0$ such that the optimal value function $v(\cdot)$ satisfies the calmness condition

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

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

Here, $d_{\mathrm{f},p}(\xi, \tilde{\xi})$ denotes the filtration distance of ξ and $\tilde{\xi}$ defined by

$$d_{\mathrm{f},p}(\xi, \tilde{\xi}) := \sup_{\|x\|_p \leq 1} \sum_{t=2}^{T-1} \|\mathbb{E}(x_t|\mathcal{F}_t(\xi)) - \mathbb{E}(x_t|\mathcal{F}_t(\tilde{\xi}))\|_{r'}$$

where r and r' satisfy $\frac{1}{r} + \frac{1}{r'} = 1$ and $p \geq 1$.

Constructing scenario trees:

Let ξ be the original stochastic process and $\hat{\xi}$ be an approximate stochastic process having (many) scenarios (paths) $\xi^i = (\xi_1^i, \ldots, \xi_T^i)$ with probabilities π^i, $i = 1, \ldots, N$, which coincide at $t = 1$, i.e., $\xi_1^1 = \ldots = \xi_1^N =: \xi_1^*$.

Idea: Recursive scenario reduction and bundling on $[1, t]$, $t = 2, \ldots, T$ such that the scenario tree process $\xi_{\rm tr}$ satisfies

$$\|\hat{\xi} - \xi_{\rm tr}\|_r + d_{{\rm f},r'}(\hat{\xi}, \xi_{\rm tr}) \le \varepsilon$$

(Forward and backward tree construction algorithms (Heitsch-Rö 05)).

(The algorithms are currently under implementation in GAMS-SCENRED.)

(Scenario reduction algorithms for mixed-integer models are available and allow for extensions.)

Reduction of scenario trees:

Let $\xi_{\rm tr}$ be a scenario tree process and determine a scenario tree process $\xi_{\rm red}$ having less nodes than $\xi_{\rm tr}$ such that

$$\|\xi_{\rm tr} - \xi_{\rm red}\|_r + d_{{\rm f},\infty}(\xi_{\rm tr}, \xi_{\rm red}) \le \varepsilon$$

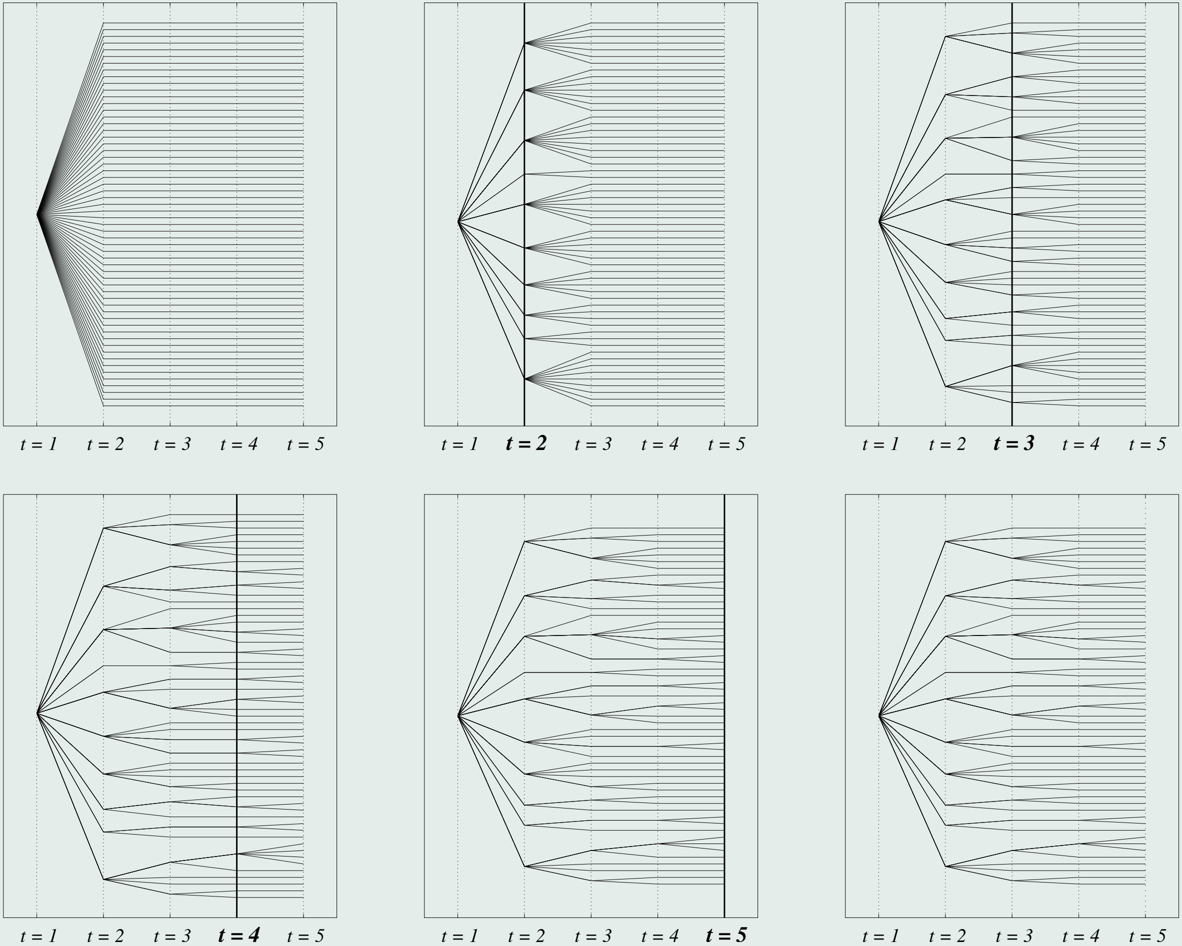


Illustration of the forward tree construction for an example including T=5 time periods starting with a scenario fan containing N=58 scenarios

<Start Animation>

Risk Functionals

A risk functional ρ assigns a real number to any (real) random variable Y (possibly satisfying certain moment conditions). Recently, it was suggested that ρ should satisfy the following axioms for all random variables $Y, \tilde{Y}$, $r \in \mathbb{R}$, $\lambda \in [0,1]$:

(A1) $\rho(Y + r) = \rho(Y) - r$ (translation-antivariance),

(A2) $\rho(\lambda Y + (1-\lambda)\tilde{Y}) \leq \lambda\rho(Y) + (1-\lambda)\rho(\tilde{Y})$ (convexity),

(A3) $Y \leq \tilde{Y}$ implies $\rho(Y) \geq \rho(\tilde{Y})$ (monotonicity).

A risk functional ρ is called coherent if it is, in addition, positively homogeneous, i.e., $\rho(\lambda Y) = \lambda\rho(Y)$ for all $\lambda \geq 0$ and random variables Y.
Given a risk functional ρ, the mapping $\mathcal{D} = \mathbb{E} + \rho$ is also called deviation risk functional.

References: Artzner-Delbaen-Eber-Heath 99, Föllmer-Schied 02, Fritelli-Rosazza Gianin 02

Examples:

(a) Average Value-at-Risk $\mathbb{A}\mathsf{V@R}_\alpha$:

$$\begin{aligned}\mathbb{A}\mathsf{V@R}_\alpha(Y) &:= \frac{1}{\alpha}\int_0^\alpha \mathbb{V}\mathsf{@R}_u(Y)(u)du \\ &= \inf\left\{x + \frac{1}{\alpha}\mathbb{E}([Y+x]^-) : x \in \mathbb{R}\right\} \\ &= \sup\left\{-\mathbb{E}(YZ) : \mathbb{E}(Z) = 1, 0 \le Z \le \frac{1}{\alpha}\right\}\end{aligned}$$

where $\alpha \in (0,1]$, $\mathbb{V}\mathsf{@R}_\alpha := \inf\{y \in \mathbb{R} : \mathbb{P}(Y \le y) \ge \alpha\}$ is the Value-at-Risk and $[a]^- := -\min\{0, a\}$.

Reference: Rockafellar-Uryasev 02

(b) Lower semi standard deviation corrected expectation:

$$\rho(Y) := -\mathbb{E}(Y) + (\mathbb{E}([Y - \mathbb{E}(Y)]^-)^2)^{\frac{1}{2}}$$

Reference: Markowitz 52

Multi-Period Risk Functionals

Let $\xi = (\xi_1, \ldots, \xi_T)$ be some input random vector. We assume that all random vectors $Y = (Y_1, \ldots, Y_T)$ have the property that Y_t only depends on $(\xi_1, \ldots, \xi_t)$.
A functional ρ that assigns to each such random vector $Y = (Y_1, \ldots, Y_T)$ a real number is called a multi-period risk functional if it satisfies the following conditions for all random vectors $Y = (Y_1, \ldots, Y_T)$ and $\tilde{Y} = (\tilde{Y}_1, \ldots, \tilde{Y}_T)$:

(A1) $\rho(Y_1 + W_1, \ldots, Y_T + W_T) = -\sum_{t=1}^{T} \mathbb{E}(W_t) + \rho(Y_1, \ldots, Y_T)$ for all W belonging to some convex subset of random vectors $\mathcal{W}$ (possibly depending on ξ) ($\mathcal{W}$-translation-antivariance),

(A2) ρ is convex (convexity),

(A3) $Y_t \leq \tilde{Y}_t$, for all t, implies $\rho(Y_1, \ldots, Y_T) \geq \rho(\tilde{Y}_1, \ldots, \tilde{Y}_T)$ (monotonicity).

The set $\mathcal{W}$ is related to the set of available financial instruments for hedging the risk.

References: Artzner-Delbaen-Eber-Heath-Ku 07, Fritelli-Scandolo 06, Pflug-Rö 07

Example: (for the set $\mathcal{W}$)

(a) $\mathcal{W} = \{(x, 0, \ldots, 0) \in I\!R^T : x \in I\!R\} = I\!R \times \{0\}^{T-1}$

(Artzner-Delbaen-Eber-Heath-Ku 07).

(b) $\mathcal{W} = I\!R^T$.

(c) $\mathcal{W} = \{W = (W_1, \ldots, W_T) : \sum_{t=1}^{T} W_t$ is deterministic$\}$.

(Fritelli-Scandolo 06).

(d) $\mathcal{W} = \{W = (W_1, \ldots, W_T) : W_t$ depends only on $(\xi_1, \ldots, \xi_{t-1})\}$

(Pflug-Ruszczynski 04).

Polyhedral risk functionals:

Multi-period risk functionals are called polyhedral if they preserve linearity structures (stability and decomposition properties) of stochastic programming models (when inserted into them) although such functionals are nonlinear by nature. They may be represented by (classical) linear stochastic programs.

Reference (for polyhedral risk functionals): Eichhorn-Rö 05.

Examples:

(a) Expectation of accumulated incomes $\sum_{\tau=1}^{t} Y_\tau$ at risk measuring time steps t_j, $j = 1, \dots, J$, with $t_J = T$:

$$\rho_0(Y_{t_1}, \dots, Y_{t_J}) := \sum_{j=1}^{J} \mathbb{E}\left(\sum_{t=1}^{t_j} Y_t\right)$$

(b) Sum of Average Value-at-Risk's at risk measuring time steps:

$$\rho_1(Y_{t_1}, \dots, Y_{t_J}) := \frac{1}{J}\sum_{j=1}^{J} \mathbb{A}\mathsf{V@R}_\alpha\left(\sum_{t=1}^{t_j} Y_t\right)$$

(c) Average Value-at-Risk of the average at risk measuring time steps:

$$\rho_4(Y_{t_1}, \dots, Y_{t_J}) := \mathbb{A}\mathsf{V@R}_\alpha\left(\frac{1}{J}\sum_{j=1}^{J} \sum_{t=1}^{t_j} Y_t\right)$$

(d) Average Value-at-Risk of the minimum at risk measuring time steps:

$$\rho_6(Y_{t_1}, \dots, Y_{t_J}) := \mathbb{A}\mathsf{V@R}_\alpha\left(\min_{j=1,\dots,J} \sum_{t=1}^{t_j} Y_t\right)$$

All examples are polyhedral risk functionals and satisfy $\mathbb{R} \times \{0\}^{T-1}$-translation-antivariance.

Stochastic programming problem with risk objective:

$$\min_{x} \left\{ \rho(Y_1, \ldots, Y_T) \;\middle|\; \begin{array}{l} Y_t = \langle b_t(\xi_t), x_t \rangle, \\ x_t = x_t(\xi_1, \ldots, \xi_t) \in X_t, \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h_t(\xi_t) \\ (t = 1, \ldots, T) \end{array} \right\}$$

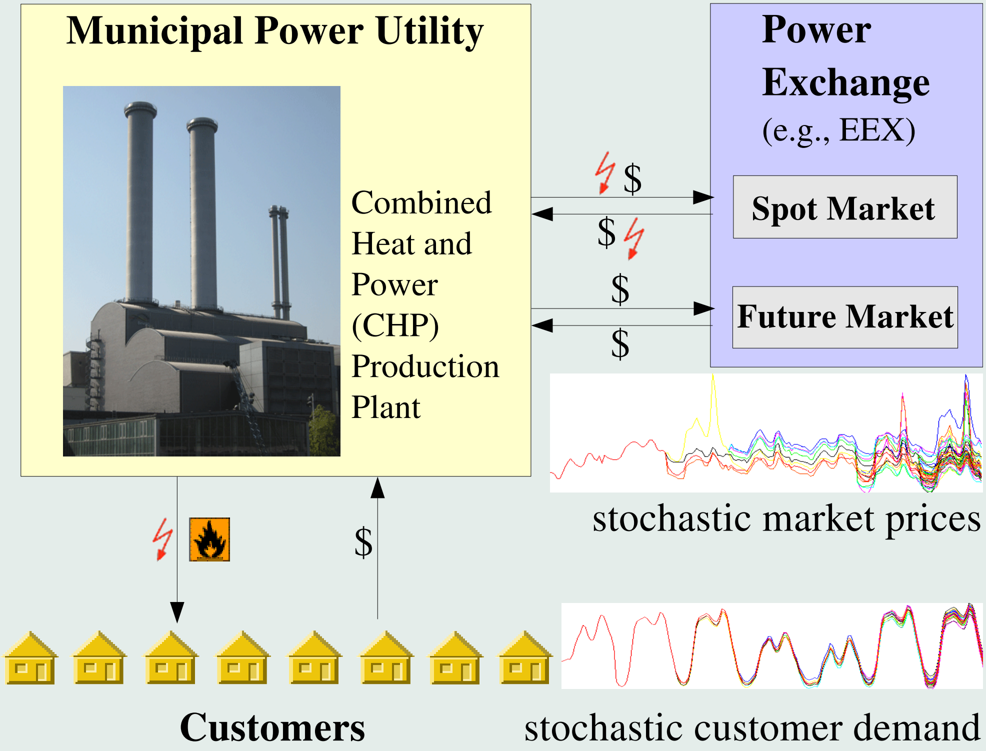
Polyhedral risk functional (evaluated at risk measuring time steps):

$$\rho(Y) = \inf \left\{ \mathbb{E}\left(\sum_{j=0}^{J} \langle c_j, v_j \rangle \right) \;\middle|\; \begin{array}{l} v_j = v_j(\xi_1, \ldots, \xi_{t_j}) \in V_j, \\ \sum_{k=0}^{j} B_{j,k} v_{j-k} = r_j \\ (j = 0, \ldots, J), \\ \sum_{k=0}^{j} \langle a_{j,k}, v_{j-k} \rangle = \sum_{t=1}^{t_j} Y_t \\ (j = 1, \ldots, J) \end{array} \right\}$$

Equivalent linear stochastic programming model:

$$\min_{(v,x)} \left\{ \mathbb{E}\left(\sum_{j=0}^{J} \langle c_j, v_j \rangle \right) \;\middle|\; \begin{array}{l} x_t = x_t(\xi_1, \ldots, \xi_t) \in X_t, \\ v_j = v_j(\xi_1, \ldots, \xi_{t_j}) \in V_j, \\ \sum_{s=0}^{t-1} A_{t,s}(\xi_t) x_{t-s} = h_t(\xi_t), \\ \sum_{k=0}^{j} B_{j,k} v_{j-k} = r_j, \\ \sum_{k=0}^{j} \langle a_{j,k}, v_{j-k} \rangle = \sum_{t=1}^{t_j} \langle b_t(\xi_t), x_t \rangle \end{array} \right\}$$

Mean-Risk Electricity Portfolio Management



We consider the electricity portfolio management of a German municipal electric power company. Its portfolio consists of the following positions:

- power production (based on company-owned thermal units),
- bilateral contracts,
- (physical) (day-ahead) spot market trading (e.g., European Energy Exchange (EEX)) and
- (financial) trading of futures.

The time horizon is discretized into hourly intervals. The underlying stochasticity consists in a multivariate stochastic load and price process that is approximately represented by a finite number of scenarios. The objective is to maximize the total expected revenue and to minimize the risk. The portfolio management model is a large scale (mixed-integer) multi-stage stochastic program.

Electricity portfolio management: statistical models and scenario trees

For the stochastic input data of the optimization model (here yearly electricity and heat demand, and electricity spot prices), a 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,

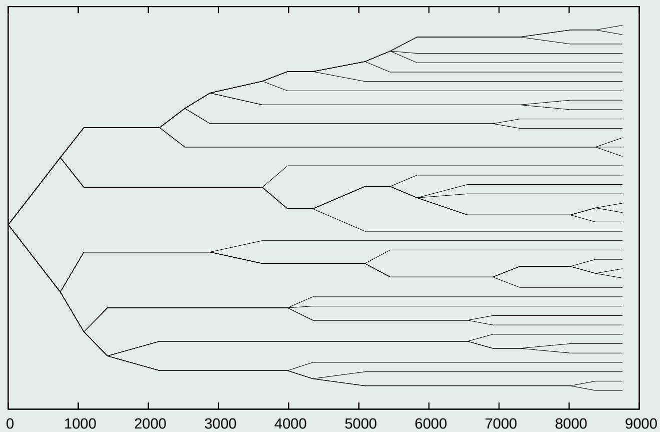
- generation of scenario trees (Heitsch-Rö 05).

Electricity portfolio management: Results

Test runs were performed on real-life data of a German municipal power company leading to a linear program containing $T = 365 \cdot 24 = 8760$ time steps, a scenario tree with 40 demand-price scenarios (see below) with about 150.000 nodes. The objective function is of the form

$$\text{Minimize} \quad \gamma\rho(Y) - (1-\gamma)\mathbb{E}\left(\sum\nolimits_{t=1}^{T} Y_t\right)$$

with a (multiperiod) risk functional ρ with risk aversion parameter $\gamma \in [0,1]$ ($\gamma = 0$ corresponds to no risk).



Home Page
Title Page
Contents
Page 22 of 31
Go Back
Full Screen
Close
Quit

Single-period and multi-period risk functionals are computed for the accumulated income at $t = T$ and at the risk time steps t_j, $j = 1, \ldots, J = 52$, respectively. The latter correspond to 11 pm at the last trading day of each week.

It turns out that the numerical results for the expected maximal revenue and minimal risk

$$\mathbb{E}\left(\sum_{t=1}^{T} Y_t^{\gamma *}\right) \qquad \text{and} \qquad \rho(Y_{t_1}^{\gamma *}, \ldots, Y_{t_J}^{\gamma *})$$

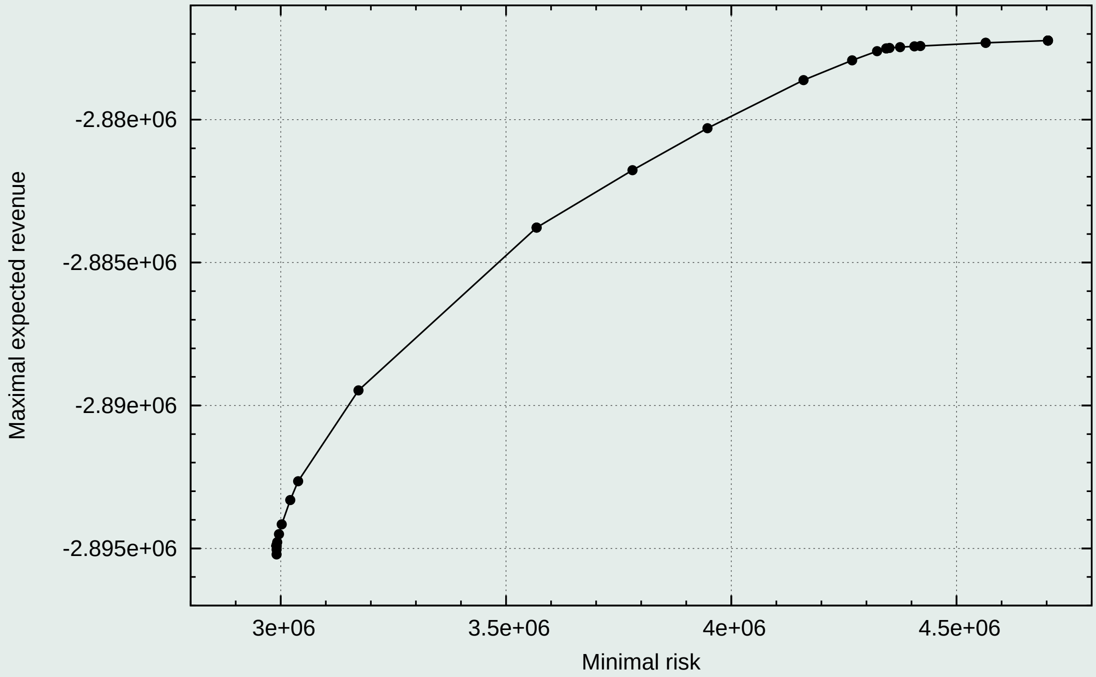
with the optimal income process $Y^{\gamma *}$ are **identical** for $\gamma \in [0.15, 0.95]$ and all risk functionals used in the test runs.

The efficient frontier

$$\gamma \mapsto \left(\rho(Y_{t_1}^{\gamma *}, \ldots, Y_{t_J}^{\gamma *}), \mathbb{E}\left(\sum_{t=1}^{T} Y_t^{\gamma *}\right)\right)$$

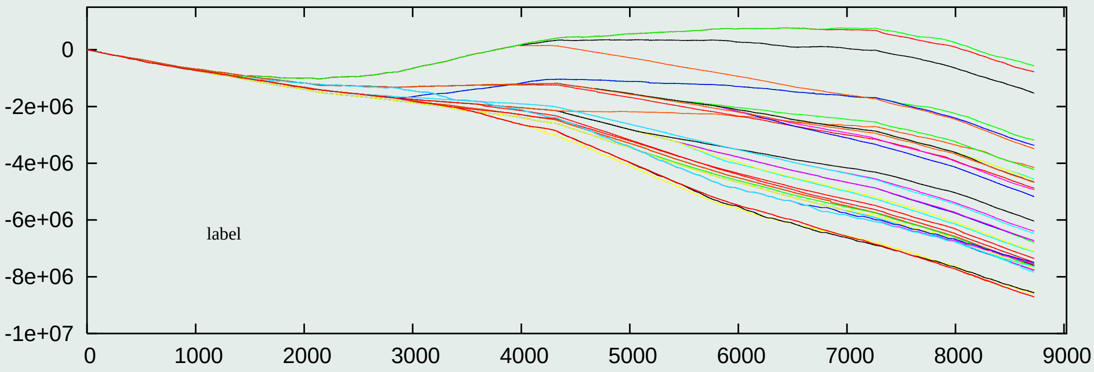
is concave for $\gamma \in [0, 1]$.

Risk aversion costs less than 1% of the expected overall revenue.

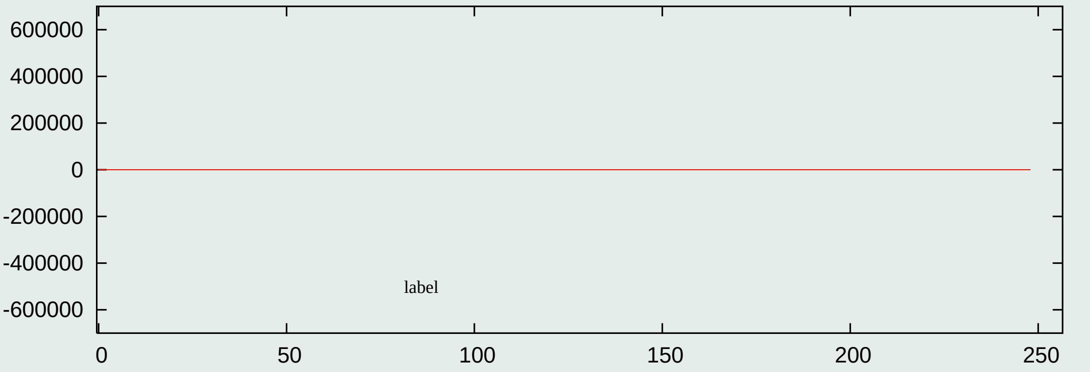


Efficient frontier

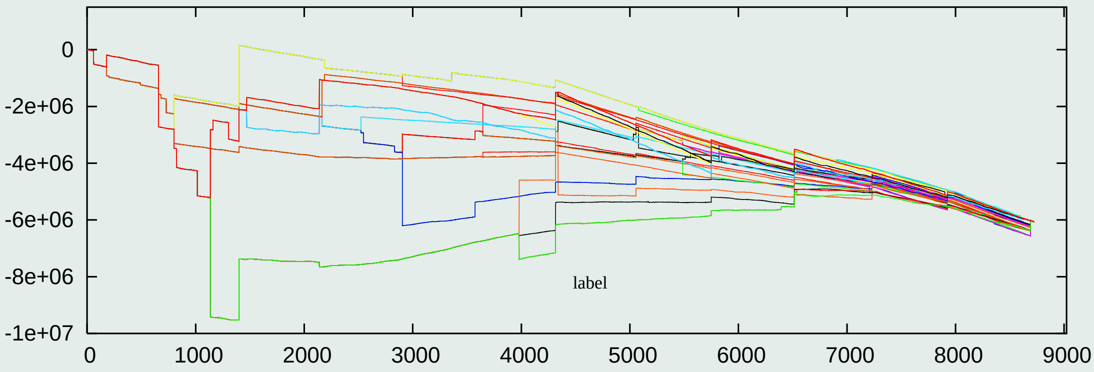
The LP is solved by CPLEX 9.1 in about 1 h running time on a 2 GHz Linux PC with 1 GB RAM.



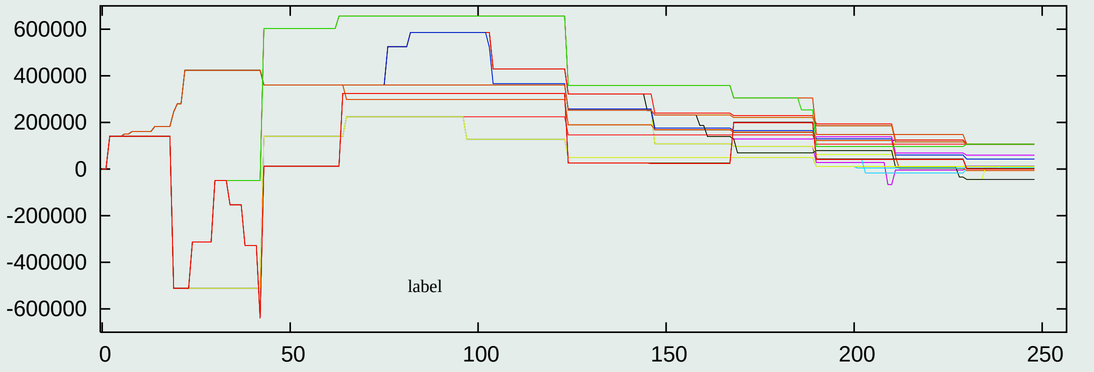
Overall revenue scenarios for $\gamma = 0$



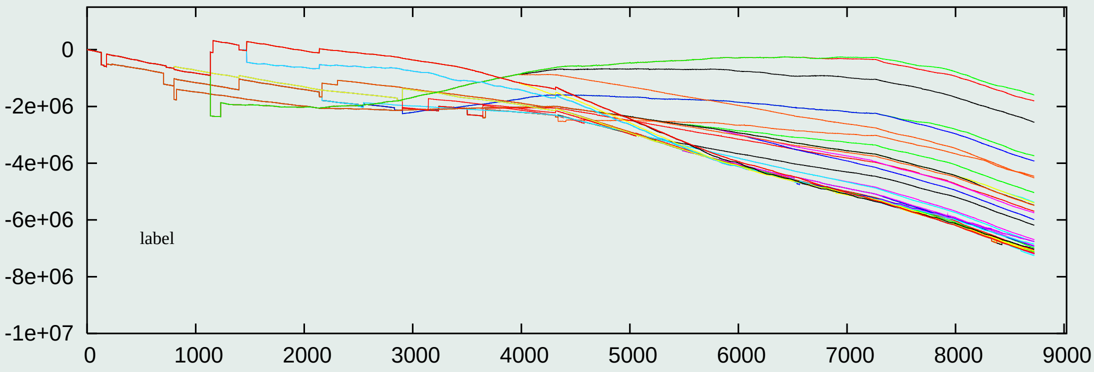
Future trading for $\gamma = 0.9$



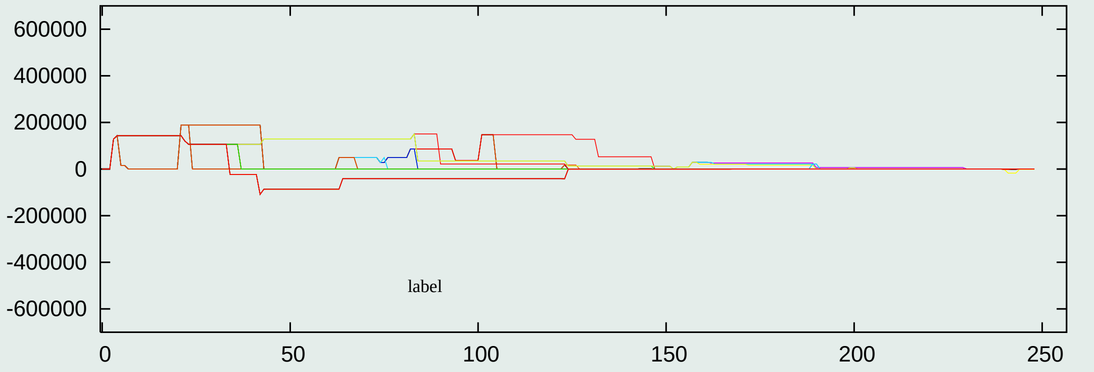
Overall revenue scenarios with $\mathbb{A}\text{V@R}_{0.05}$ and $\gamma = 0.9$



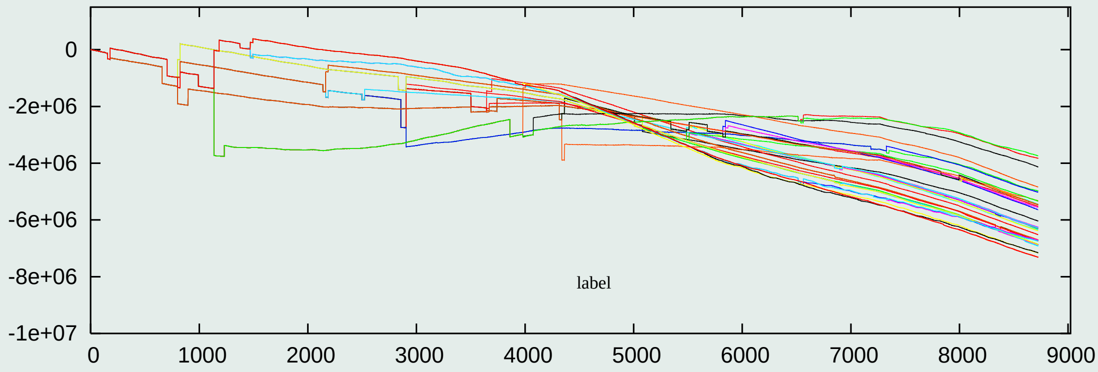
Future trading with $\mathbb{A}\text{V@R}_{0.05}$ and $\gamma = 0.9$



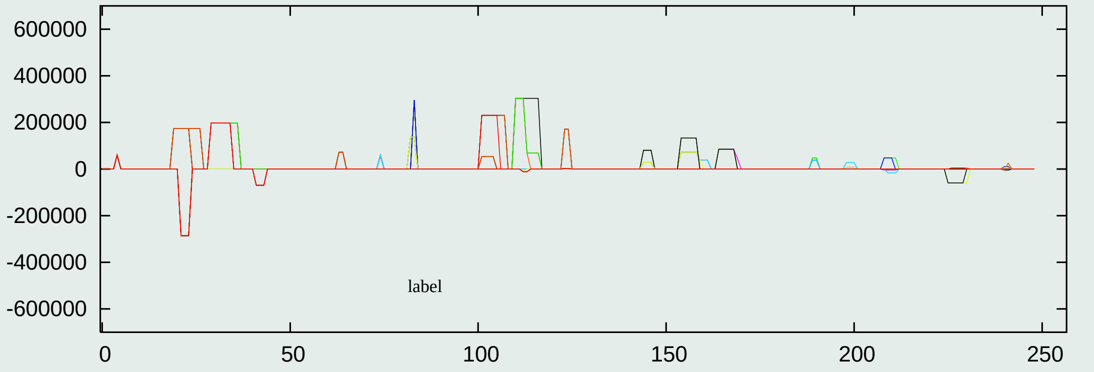
Overall revenue scenarios with ρ_1 and $\gamma = 0.9$



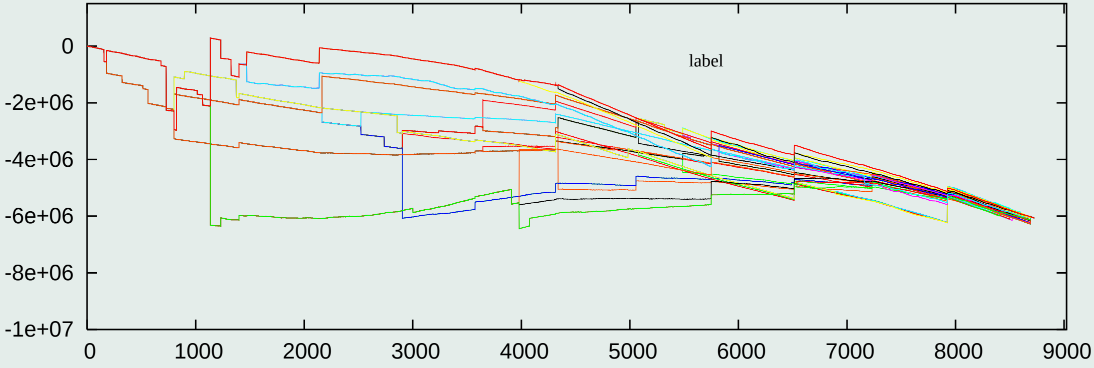
Future trading for ρ_1 and $\gamma = 0.9$



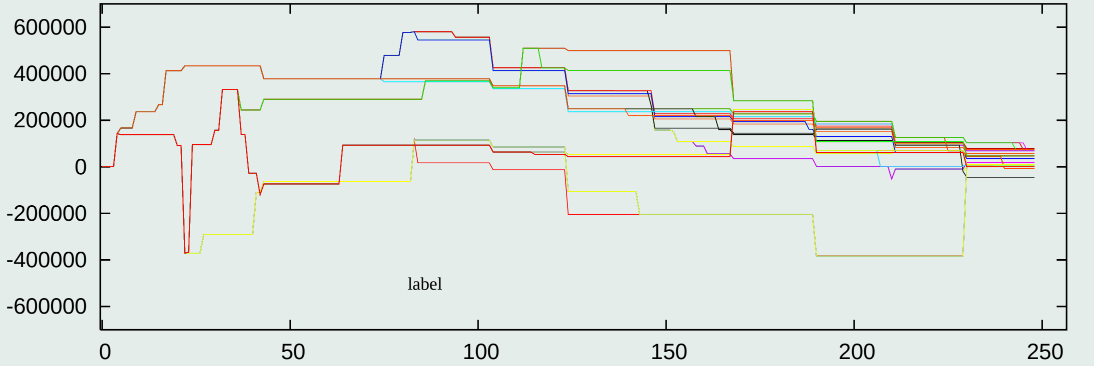
Overall revenue scenarios with ρ_4 and $\gamma = 0.9$



Future trading with ρ_4 and $\gamma = 0.9$



Overall revenue scenarios with ρ_6 and $\gamma = 0.9$



Future trading with ρ_6 and $\gamma = 0.9$

Thank you

References

Dentcheva, D., Römisch, W.: Duality gaps in nonconvex stochastic optimization, *Mathematical Programming* 101 (2004), 515–535.

Eichhorn, A., Römisch, W.: Polyhedral risk measures in stochastic programming, *SIAM Journal on Optimization* 16 (2005), 69–95.

Eichhorn, A., Römisch, W.: Dynamic risk management in electricity portfolio optimization via polyhedral risk functionals, IEEE PES General Meeting 2008 (Pittsburgh, July 2008).

Gröwe-Kuska, N., Römisch, W.: Stochastic unit commitment in hydro-thermal power production planning, *Applications of Stochastic Programming* (S.W. Wallace, W.T. Ziemba eds.), MPS-SIAM Series in Optimization, 2005, 633–653.

Heitsch, H., Römisch, W.: A note on scenario reduction for two-stage stochastic programs, *Operations Research Letters* 35 (2007), 731–736.

Heitsch, H., Römisch, W.: Scenario tree modeling for multistage stochastic programs, *Mathematical Programming* (to appear).

Heitsch, H., Römisch, W.: Scenario tree reduction for multistage stochastic programs, Preprint, DFG Research Center MATHEON Berlin, 2008.

Heitsch, H., Römisch, W., Strugarek, C.: Stability of multistage stochastic programs, *SIAM Journal Optimization* 17 (2006), 511–525.

Henrion, R., Küchler, C., Römisch, W.: Discrepancy distances and scenario reduction in two-stage stochastic mixed-integer programming, *Journal Industr. Managem. Optimization* 4 (2008), 363–384.

Pflug, G. Ch.; Römisch, W.: *Modeling, Measuring and Managing Risk*, World Scientific, Singapore, 2007.

Römisch, W., Vigerske, S.: Quantitative stability of fully random mixed-integer two-stage stochastic programs, *Optimization Letters* 2 (2008), 377–388.

Römisch, W.; Wets, R. J-B: Stability of ε-approximate solutions to convex stochastic programs, *SIAM Journal on Optimization* 18 (2007), 961–979.