Polyhedral risk measures in electricity portfolio optimization

Andreas Eichhorn[*] [1], **Werner Römisch**[1], and **Isabel Wegner**[1]

[1] Humboldt-University Berlin, Department of Mathematics, 10099 Berlin, Germany

We compare different multiperiod risk measures taken from the class of polyhedral risk measures with respect to the effect they show when used in the objective of a stochastic program. For this purpose, simulation results of a stochastic programming model for optimizing the electricity portfolio of a German municipal power utility are presented and analyzed. This model aims to minimize risk and expected overall cost simultaneously.

1 Introduction

The risk of high losses of uncertain outcomes is quantified with so-called risk measures, i.e., mappings from some space of random variables (or processes) to the real line that have certain properties (cf. [1, 5, 7]). In particular, in case the risk of a value process over a finite number of time periods has to be considered, multiperiod risk measures are needed (cf. [2, 6]).

Risk measures and stochastic programs fit together in a natural way: they both rest on stochastic models and it is an expedient goal to minimize risk (cf. [7, 8]). To be more precise: the aim is to find a reasonable tradeoff between low risk and low expected cost (since minimal risk does not come for free). The choice of the risk measure is a crucial factor, because it determines in which way extreme events are avoided. The usage of a risk measure that is not appropriate for a certain situation is likely to lead to higher expected cost, i.e., to a non-optimal solution.

However, stochastic programs incorporating risk measures are usually harder to solve. An unfavorable choice of the risk measure may easily result in a problem that is no longer solvable in practice, especially if integer variables are incorporated. Therefore, one has to restrict the choice of the risk measure to those with favorable properties for the structure of the respective stochastic program. To this end, the class of polyhedral risk measures was introduced in [3] for which theses favorable properties are guaranteed. Instances of this class were suggested for the multiperiod case.

In this paper, such risk measures and their effect in stochastic programs will be compared in a simulative study of a real world application model. We use the electricity portfolio optimization model presented in [4] which is a multistage stochastic programming model set up for a municipal power utility to optimize power production and electricity trading under uncertainty over a period of one year. The objective is to minimize the expected overall cost and a multiperiod risk measure simultaneously.

2 Optimization model

The model used for the analysis is set up for a power utility with limited power production capacities. It is based on hourly discretization, i.e. time is considered in terms of time steps $t = 1, ..., T$ where $T = 365 \cdot 24$. The objective is to satisfy an uncertain time-dependent electricity demand in an optimal manner by utilizing available power production facilities as well as several types of contracts with larger power companies, electricity spot market at the European Energy Exchange (EEX), and certain energy derivative products from EEX that can be used to hedge risk (so-called *futures*). On a high level of abstraction the model reads

$$\min \left\{ \gamma\rho(z_{t_1}, ..., z_{t_k}) + (1-\gamma)\mathbb{E}\left[-z_T\right] : z_t = \textstyle\sum_{\tau=1}^{t} b_\tau(\xi_\tau)\cdot x_\tau \,,\ x \in \mathcal{X}(\xi) \right\} \tag{1}$$

with $1 \le t_1 \le ... \le t_k = T$. The objective is a weighted sum of a k-period polyhedral risk measure ρ and the expectation of the final value with some fixed weighting parameter $\gamma \in [0, 1]$. Thereby, the uncertainty is represented by a multivariate data process $\xi = (\xi_1, ..., \xi_T)$ containing electricity demand, spot and future prices. The vectors $x = (x_1, ..., x_T)$ denote the decisions of the model at each time step that have to satisfy several restrictions symbolized by the set $\mathcal{X}(\xi)$ consisting of polyhedral, integrality, and non-anticipativity constraints. The book values at each time step, i.e., the accumulated revenues, are represented by the variables $z_1, ..., z_T$. The vectors b_τ are cost coefficients depending on the random data. For solving the model, the data process ξ is approximated by a finite scenario tree.

It is assumed that, within certain bounds, any amounts of electricity can be traded or produced, hence, these decisions can be modelled with continuous variables. The only integer variables in the model are the decisions whether a certain contract is to make or not. See [4] for further details.

[*] Corresponding author: e-mail: eichhorn@math.hu-berlin.de, Phone: +49/(0)30/2093-5498, Fax: +49/(0)30/2093-2232

3 Quantifying risk of value processes

We consider a finite number T of time periods, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a filtration $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq ... \subseteq \mathcal{F}_T$ of σ-fields, e.g., $\mathcal{F}_t = \sigma(\xi_1, ..., \xi_t)$ with some random process ξ. Suppose the (uncertain) value process is represented by random variables $z_1, z_2, ..., z_T$ with $z_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ ($p \geq 1$) for which large outcomes are preferred to lower ones. In [2] it is claimed that *from an economic point of view* multiperiod risk measures ρ should at least satisfy the following conditions:
(i) If $z_t \leq \tilde{z}_t$ a.s., $t = 1, ..., T$, then $\rho(z_1, ..., z_T) \geq \rho(\tilde{z}_1, ..., \tilde{z}_T)$ *(inverse monotonicity)*
(ii) for each $r \in \mathbb{R}$ we have $\rho(z_1 + r, ..., z_T + r) = \rho(z) - r$ *(translation equivariance)*
(iii) $\rho(\mu z_1 + (1-\mu)\tilde{z}_1, ..., \mu z_T + (1-\mu)\tilde{z}_T) \leq \mu\rho(z_1, ..., z_T) + (1-\mu)\rho(\tilde{z}_1, ..., \tilde{z}_1)$ for $\mu \in [0, 1]$ *(convexity)*
(iv) for $\mu \geq 0$ we have $\rho(\mu z_1, ..., \mu z_T) = \mu\rho(z_1, ..., z_T)$ *(positive homogeneity)*.
If so, the functional ρ is called a *multiperiod coherent risk measure*. If a functional ρ is continuous from below, then it is a multiperiod coherent risk measure iff there exists a convex set $\mathcal{P}_\rho \subseteq \mathcal{D}_T := \{f \in \times_{t=1}^T L_1(\Omega, \mathcal{F}_t, \mathbb{P}) : f_t \geq 0,\ \sum \mathbb{E}[f_t] = 1\}$ such that $\rho(z_1, ..., z_T) = \sup\{-\sum_{t=1}^T \mathbb{E}[f_t z_t] : f \in \mathcal{P}_\rho\}$ (cf. [2, 5, 3]).

Multiperiod *polyhedral risk measures* were defined in [3] as optimal values of certain simple multistage stochastic programs:

$$\rho(z_1, ..., z_T) = \inf \left\{ \mathbb{E}\left[\sum_{t=1}^T \langle c_t, y_t \rangle\right] \middle| \begin{array}{l} y_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{k_t}),\ y_t(\omega) \in Y_t, \\ \sum_{\tau=0}^{t-1} \langle w_{t,\tau}, y_{t-\tau}(\omega) \rangle = z_t(\omega) \end{array} \quad (t = 1, \ldots, T) \right\} \tag{2}$$

with some $k_t \in \mathbb{N}$, $c_t \in \mathbb{R}^{k_t}$, $t = 1, \ldots, T$, $w_{t\tau} \in \mathbb{R}^{k_{t-\tau}}$, $t = 1, \ldots, T$, $\tau = 0, ..., t-1$, and polyhedral cones $Y_t \subseteq \mathbb{R}^{k_t}$, $t = 1, \ldots, T$. It is shown in [3] that risk measures of this form have favorable properties for stochastic programs with respect to stability and algorithmic structures. If complete recourse and dual feasibility is imposed (standard assumptions in stochastic programming guaranteeing finiteness), it is shown that then ρ is Lipschitz continuous and allows the dual representation

$$\begin{aligned} &\rho(z_1, ..., z_T) = \sup \left\{ -\mathbb{E}\left[\sum_{t=1}^T \lambda_t z_t\right] : \lambda \in \Lambda_\rho \right\} \\ &\Lambda_\rho := \left\{ \lambda \in \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) : c_t + \sum_{\nu=t}^T \mathbb{E}[\lambda_\nu | \mathcal{F}_t]\, w_{\nu,\nu-t} \in -Y_t^* \right\} \quad \text{with} \quad \tfrac{1}{p} + \tfrac{1}{p'} = 1 \end{aligned} \tag{3}$$

Hence, in this case ρ is a multiperiod coherent risk measure if $\Lambda_\rho \subseteq \mathcal{D}_T$. All these assumptions are satisfied for the examples suggested in [3]. For our analysis, we selected the following instances that are, hence, multiperiod coherent risk measures:

No.	primal representation (2)	dual multipliers according to (3)
ρ_1	$\inf \left\{ \sum\limits_{t=2}^T \frac{1}{T-1} \left(y_1^{(t)} + \frac{1}{\alpha}\mathbb{E}\left[y_t^{(2)}\right] \right) \middle\vert \begin{array}{l} y_1 \in \mathbb{R}^T \text{ constant},\ y_1^{(1)} = z_1 \\ y_t \in \mathbb{R}_+ \times \mathbb{R}_+\ \mathcal{F}_t\text{-measurable} \\ y_t^{(1)} - y_t^{(2)} = z_t + y_1^{(t)}\ \ (t = 2, ..., T) \end{array} \right\}$	$0 \leq \lambda_t \leq \frac{1}{\alpha(T-1)}$, $\mathbb{E}[\lambda_2] = ... = \mathbb{E}[\lambda_T] = \frac{1}{T-1}$
ρ_2	$\inf \left\{ y_1^{(1)} + \sum\limits_{t=2}^T \frac{1}{\alpha(T-1)}\mathbb{E}\left[y_t^{(2)}\right] \middle\vert \begin{array}{l} y_1 \in \mathbb{R} \times \mathbb{R} \text{ constant},\ y_1^{(2)} = z_1 \\ y_t \in \mathbb{R}_+ \times \mathbb{R}_+\ \mathcal{F}_t\text{-measurable} \\ y_t^{(1)} - y_t^{(2)} = z_t + y_1^{(1)}\ \ (t = 2, ..., T) \end{array} \right\}$	$0 \leq \lambda_t \leq \frac{1}{\alpha(T-1)}$ $(t = 2, ..., T)$, $\sum_{t=1}^T \mathbb{E}[\lambda_t] = 1$
ρ_3	$\inf \left\{ y_1^{(1)} + \sum\limits_{t=2}^T \frac{1}{\alpha(T-1)}\mathbb{E}\left[y_t^{(2)}\right] \middle\vert \begin{array}{l} y_1 \in \mathbb{R} \times \mathbb{R} \text{ constant},\ y_1^{(2)} = z_1 \\ y_t \in \mathbb{R}_+ \times \mathbb{R}_+\ \mathcal{F}_t\text{-meas.}\ (t = 2, ..., T) \\ y_2^{(1)} - y_2^{(2)} = z_2 + y_1^{(1)} \\ y_t^{(1)} - y_t^{(2)} = z_t + y_1^{(1)} + y_{t-1}^{(2)}\ \ (t = 3, ..., T) \end{array} \right\}$	$\lambda_t + \mathbb{E}[\lambda_{t+1} \vert \mathcal{F}_t] \leq \frac{1}{\alpha(T-1)}$, $0 \leq \lambda_t,\ (t = 2, ..., T-1)$, $0 \leq \lambda_T \leq \frac{1}{\alpha(T-1)}$, $\sum_{t=1}^T \mathbb{E}[\lambda_t] = 1$
ρ_4	$\inf \left\{ \frac{1}{T-1} \left(y_1^{(1)} + \sum\limits_{t=2}^T \frac{1}{\alpha}\mathbb{E}\left[y_t^{(2)}\right] \right) \middle\vert \begin{array}{l} y_1 \in \mathbb{R} \times \mathbb{R} \text{ constant},\ y_1^{(2)} = z_1 \\ y_t \in \mathbb{R} \times \mathbb{R}_+\ \mathcal{F}_t\text{-meas.}\ (t = 2, ..., T-1) \\ y_T \in \mathbb{R}_+ \times \mathbb{R}_+ \\ y_t^{(1)} - y_t^{(2)} = z_t + y_{t-1}^{(1)}\ \ (t = 2, ..., T) \end{array} \right\}$	$0 \leq \lambda_t \leq \beta_t\, (t = 2, ..., T)$, $\lambda_t = \mathbb{E}[\lambda_{t+1} \vert \mathcal{F}_t]$ $(t = 2, ..., T-1)$ $\mathbb{E}[\lambda_2] = ... = \mathbb{E}[\lambda_T] = \frac{1}{T-1}$

All these examples can be considered as multiperiod extensions of the one-period Conditional-Value-at-Risk $CVaR_\alpha(z) = \inf_{r \in \mathbb{R}} \left\{ r + \frac{1}{\alpha}\mathbb{E}[(z+r)^-] \right\}$ with $\alpha \in (0, 1)$ small, e.g. $\alpha = 0.05$ (cf. [7]). The primal representation (2) is suitable for being incorporated in the objective (1) since the two nested minimization problems can be reduced to one. However, to understand how a respective risk measure works, it is more suggestive to regard the dual representation according to (3). Note that the maximization there aims to choose λ big where z is small in compliance with the respective restrictions. Hence, $\rho(z)$ can be understood as a kind of (negative) worst case weighted expectation of z.

4 Simulation Results

Now we are ready to present simulation results of the model. Thereby, the data process ξ is approximated by a scenario tree with 8760 timesteps and 21 scenarios (cf. [4]).

Most important for the power utility is, of course, the book value at each time step, especially if one focuses on liquidity. Figure 1, 2 and 3 show plots of these value processes from optimal portfolios according to (1) for each risk measure, respectively ($\gamma = 0.25$, $\alpha = 0.05$). The treelike curve structure in each figure corresponds, of course, to the input scenario tree.

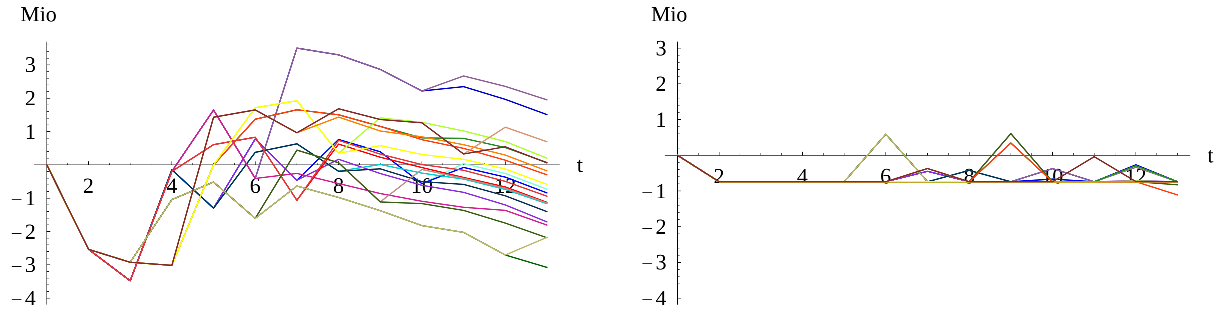


Fig. 1 $\mathbb{E}[-z_T]$ $\quad$ $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot \rho_2(z_{t_1}, ..., z_T)$

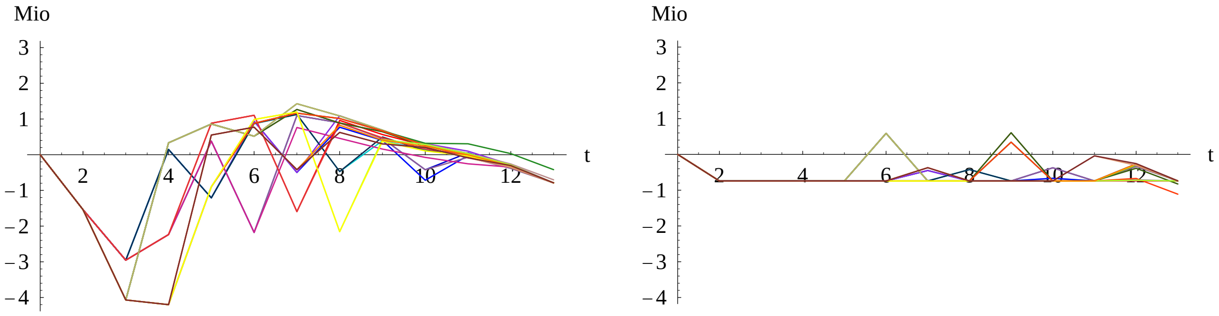


Fig. 2 $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot CVaR_{0.05}(z_T)$ $\quad$ $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot \rho_3(z_{t_1}, ..., z_T)$

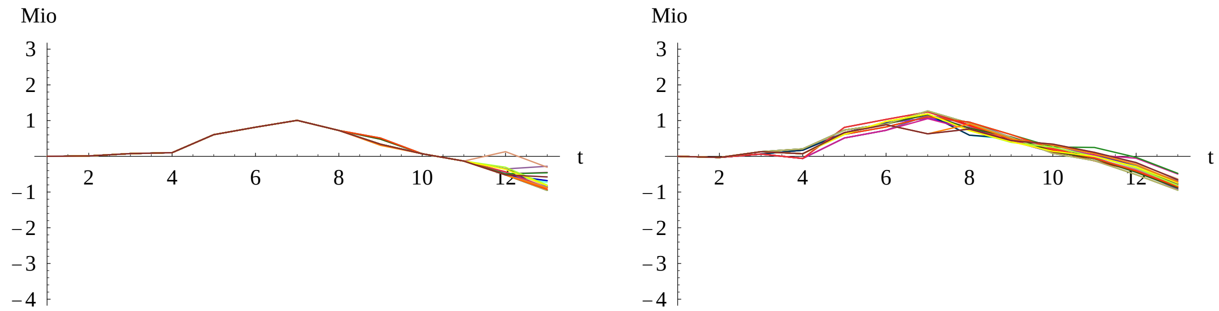


Fig. 3 $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot \rho_1(z_{t_1}, ..., z_T)$ $\quad$ $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot \rho_4(z_{t_1}, ..., z_T)$

Optimizing without risk or with $CVaR$ applied to the value at the last time step only, leads to high spreading and to very low intermediate values for a considerably high number of scenarios. The usage of a multiperiod risk measure that takes the intermediate time steps into account corrects both, spreading and negativity of values. The way this is achieved, however, differs among the risk measures. Obviously, they can be divided into two groups, ρ_2 and ρ_3, on the one hand, and ρ_1 and ρ_4, on the other hand.

The effect of ρ_1 and ρ_4 is that, roughly speaking, the values of the scenarios run closer together. A difference is that for ρ_1 (sum of one-period $CVaR$s) the runs are pushed closer together at the beginning for the price of a higher spread at the end, i.e., at the time when the portfolio value tends to be low. Hence, the effect of pushing the value runs together turns out to be more uniform when ρ_4 is used. This is, of course, advantageous since it is the spread at low levels that should be avoided.

This spread at the end is even smaller if ρ_2 or ρ_3 are used. In this case one can make out a level that is attempted not to be underrun whereas upward deviation is not avoided. For ρ_2, this level corresponds to the number r in a reformulation of ρ_2 which reads $\rho_2(z_1, ..., z_T) = \inf_{r \in \mathbb{R}} \left\{ r + \sum_{t=2}^{T} \frac{1}{\alpha(T-1)} \mathbb{E}[(z_t + r)^-] \right\}$ (cf. [3]). Thereby, another sort of uniformity is achieved which seems to be very desirable from the point of view of liquidity. However, this uniformity is achieved by a higher amount of future trading (see below).

For other values of γ and α the results are qualitatively the same and quantitatively similar in the majority of cases. Significant differences can be observed in the case that α is enlarged for ρ_2 or ρ_3: Figure 4 demonstrates that larger α causes that the level described above is higher for the price of spread and decrease at the end. Although the latter is unavoidable since expected cost cannot be lower than in the purely expectation based model, the behavior for $\alpha = 0.2$ seems to be less desirable.

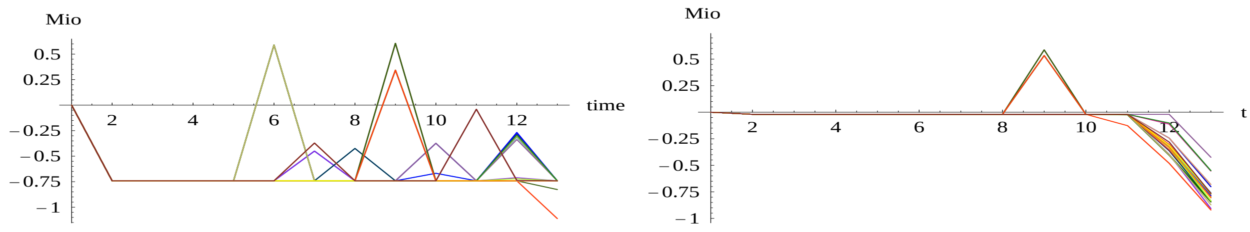


Fig. 4 $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot \rho_2(z_{t_1}, ..., z_T), \;\; \alpha = 0.05$ $\qquad$ $0.75 \cdot \mathbb{E}[-z_T] + 0.25 \cdot \rho_2(z_{t_1}, ..., z_T), \;\; \alpha = 0.2$

Enlarging γ slightly does not change the results significantly, but for $\gamma \gtrapprox 0.5$ (depending on the risk measure) there occurs a switch to the effect that the contract described in Section 2 is closed (cf. [4]).

Future trading activity differs among the risk measures in amount and in time. It is relatively low for ρ_4 whereas for ρ_1 it is high at the end of the time horizon and for ρ_2 and ρ_3 it is very high in the midway (cf. Figure 5).

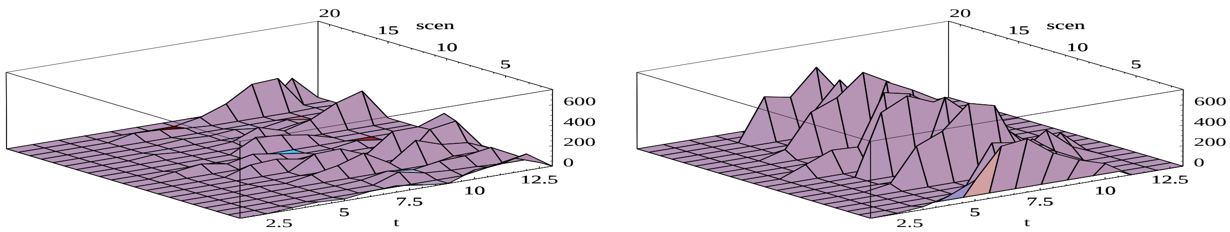
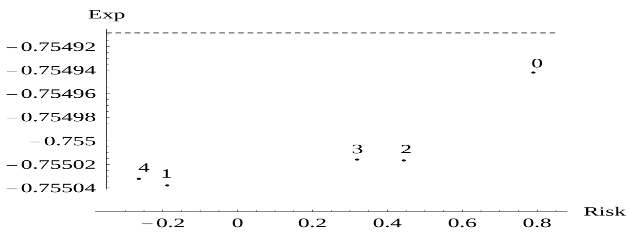


Fig. 5 Future trading for ρ_1 $\qquad$ Future trading for ρ_2

As mentioned above in the introduction, lower risk does not come for free, i.e., incorporating risk measures in (1) leads to higher expected overall cost. In the figure below the risk values and the values of the expected revenues are shown for each risk measure. Thereby, index $1, ..., 4$ denotes values of the optimal portfolios for ρ_1,...,ρ_4, respectively, and index 0 addresses the $CVaR$ applied to the last time. The horizontal line shows the maximal expected revenue without consideration of risk.



Hence, ρ_2 and ρ_3 produce fewer extra cost compared to ρ_1 and ρ_4 but, of course, more than $CVaR$ applied to the final value only. The relatively small differences of the optimal expected revenues are due to the fact that the model considers fair prices for the futures, and no transaction cost are taken into account.

Acknowledgements This research was supported by the DFG Research Center *Mathematics for Key Technologies* in Berlin and by the BMBF under grant 03-RLM5B3. The authors wish to thank Holger Heitsch (Humboldt-University Berlin) for useful hints and discussion and Gorden Spangardt (Fraunhofer-Institute UMSICHT Oberhausen) for his contribution to the optimization model.

References

[1] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk, Mathematical Finance 9 (1999), 203-228.
[2] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Ku, H.: Coherent multiperiod risk adjusted values and Bellman's principle, manuscript (2004), downloadable from <`www.math.ethz.ch/~delbaen`>.
[3] Eichhorn, A., Römisch, W.: Polyhedral risk measures in stochastic programming, Preprint 2004-05, Department of Mathematics, Humboldt-University Berlin (2004), submitted to SIAM Journal on Optimization.
[4] Eichhorn, A., Gröwe-Kuska, N., Liebscher, A., Römisch, W., Spangardt, G., Wegner, I.: Mean risk optimization of electricity portfolios, Proceedings in Applied Mathematics and Mechanics (PAMM) (2004).
[5] Föllmer, H., Schied, A.: Stochastic Finance: An Introduction in Discrete Time, Walter de Gruyter, Berlin (2002).
[6] Pflug, G., Ruszczyński, A.: Risk measures for income streams, in: Risk Measures for the 21st Century (Szegö, G. ed.), Wiley (2004).
[7] Rockafellar, R. T., Uryasev, S.: Conditional value-at-risk for general loss distributions, Journal of Banking & Finance 26 (2002), 1443-1471.
[8] Ruszczyński, A., Shapiro, A.: Optimization of convex risk functions, Stochastic Programming E-Print Series 2004-08, downloadable from <`www.speps.info`>.