Jitka Dupačová and scenario reduction

W. Römisch

Humboldt-University Berlin
Institute of Mathematics

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

Session in honor of Jitka Dupačová

ICSP 2016, Buzios (Brazil), June 29, 2016

Introduction

Most approaches for solving stochastic programs of the form

$$\min \left\{ \int_{\Xi} f_0(x,\xi) P(d\xi) : x \in X \right\}$$

with a probability measure P on $\Xi \subset \mathbb{R}^d$ and a (normal) integrand f_0, require to replace P by a some discrete probability measure or, equivalently, to replace the integral by some quadrature formula with nonnegative weights

$$\int_{\Xi} f_0(x,\xi) P(d\xi) \approx \sum_{i=1}^{n} p_i f_0(x,\xi_i),$$

where $p_i = P(\{\xi_i\})$, $\sum_{i=1}^{n} p_i = 1$, are the probabilities and $\xi_i \in \Xi$, $i = 1, \ldots, n$, the scenarios. This leads to the scenario-based stochastic program

$$\min \left\{ \sum_{i=1}^{n} p_i f_0(x,\xi_i) : x \in X \right\}$$

Since f_0 is often expensive to compute, the number n should be as small as possible.

Working Paper

Scenario Based Stochastic Programs: Strategies for Deleting Scenarios

Jitka Dupačová

WP-95-014
February 1995

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

With $v(P)$ and $S(P)$ denoting the optimal value and solution set of the stochastic program, respectively, the following estimates are known

$$
\begin{aligned}
|v(P) - v(Q)| &\leq \sup_{x \in X} \left| \int_{\Xi} f_0(x,\xi)(P-Q)(d\xi) \right| \\
\emptyset \neq S(Q) &\subseteq S(P) + \Psi_P^{-1}\left(\sup_{x \in X} \left| \int_{\Xi} f_0(x,\xi)(P-Q)(d\xi) \right| \right),
\end{aligned}
$$

where X is assumed to be compact, Q is a probability distribution approximating P and Ψ_P is the growth function of the objective near the solution set, i.e.,

$$
\Psi_P(t) := \inf \left\{ \int_{\Xi} f_0(x,\xi) P(d\xi) - v(P) : x \in X, d(x, S(P)) \geq t \right\}.
$$

Hence, the distance $d_{\mathcal{F}}$ with $\mathcal{F} := \{f_0(x,\cdot) : x \in X\}$

$$
d_{\mathcal{F}}(P,Q) := \sup_{f \in \mathcal{F}} \left| \int_{\Xi} f(\xi)(P-Q)(d\xi) \right|
$$

becomes important when approximating P.

For given $n \in \mathbb{N}$ the best possible choice of elements $\xi_i \in \Xi$ (scenarios) and probabilities p_i, $i = 1, \ldots, n$, is obtained by minimizing

$$\sup_{x \in X} \left| \int_{\Xi} f_0(x, \xi) P(d\xi) - \sum_{i=1}^{n} p_i f_0(x, \xi_i) \right|,$$

i.e., by solving the best approximation problem

$$\min_{Q \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, Q)$$

where $\mathcal{P}_n(\Xi) := \{Q : Q$ is a discrete probability measure with n scenarios$\}$.

It may be reformulated as a semi-infinite program. and is known as optimal quantization of P with respect to the function class $\mathcal{F}$. Such optimal quantization problems of probability measures are often extremely difficult to solve.

Idea: Enlarging the class $\mathcal{F}$!

Aim of the talk:

Solving the best approximation problem for discrete probability measures P having many scenarios and for function classes $\mathcal{F}$, which are relevant for two-stage stochastic programs (optimal scenario reduction).

Linear two-stage stochastic programs

$$\min\left\{\langle c,x\rangle+\int_{\Xi}\Phi(q(\xi),h(\xi)-T(\xi)x)P(d\xi):x\in X\right\},$$

where $c\in\mathbb{R}^m$, Ξ and X are polyhedral subsets of $\mathbb{R}^d$ and $\mathbb{R}^m$, respectively, P is a probability measure on Ξ and the $s\times m$-matrix $T(\cdot)$, the vectors $q(\cdot)\in\mathbb{R}^{\overline{m}}$ and $h(\cdot)\in\mathbb{R}^s$ are affine functions of ξ.

Furthermore, Φ and D denote the infimum function of the linear second-stage program and its dual feasibility set, respectively, i.e.,

$$\begin{aligned}\Phi(u,t) &:= \inf\{\langle u,y\rangle : Wy=t, y\in Y\} \quad ((u,t)\in\mathbb{R}^{\overline{m}}\times\mathbb{R}^s)\\ D &:= \{u\in\mathbb{R}^{\overline{m}}:\{z\in\mathbb{R}^s: W^\top z-u\in Y^*\}\neq\emptyset\},\end{aligned}$$

where W is the $s\times\overline{m}$ recourse matrix, $W^\top$ the transposed of W and Y^* the polar cone to the polyhedral cone Y in $\mathbb{R}^{\overline{m}}$.

Theorem: (Walkup-Wets 69)
The function $\Phi(\cdot,\cdot)$ is finite and continuous on the polyhedral set $D \times W(Y)$. Furthermore, the function $\Phi(u,\cdot)$ is piecewise linear convex on the polyhedral set $W(Y)$ for fixed $u \in D$, and $\Phi(\cdot,t)$ is piecewise linear concave on D for fixed $t \in W(Y)$.

Assumptions:

(A1) *relatively complete recourse*: for any $(\xi,x) \in \Xi \times X$, $h(\xi) - T(\xi)x \in W(Y)$;

(A2) *dual feasibility*: $q(\xi) \in D$ holds for all $\xi \in \Xi$.

(A3) *existence of second moments:* $\int_\Xi \|\xi\|^2 P(d\xi) < +\infty$.

Note that (A1) is satisfied if $W(Y) = \mathbb{R}^s$ (complete recourse). In general, (A1) and (A2) impose a condition on the support Ξ of P. (A1) and (A2) imply that $\Phi(q(\cdot), h(\cdot) - T(\cdot)x)$ is a finite linear-quadratic function on Ξ.

Extensions to certain random recourse models, i.e., to $W(\xi)$, exist.

Idea: Extend the class $\mathcal{F}$ such that it covers all two-stage models.

Fortet-Mourier metrics: (as canonical distances for two-stage models)

$$\zeta_r(P,Q) := \sup \left| \int_\Xi f(\xi)(P-Q)(d\xi) : f \in \mathcal{F}_r(\Xi) \right|,$$

where $r \geq 1$ ($r \in \{1,2\}$ if $W(\xi) \equiv W$)

$$\mathcal{F}_r(\Xi) := \{ f : \Xi \mapsto \mathbb{R} : f(\xi) - f(\tilde{\xi}) \leq c_r(\xi, \tilde{\xi}), \forall \xi, \tilde{\xi} \in \Xi \},$$

$$c_r(\xi,\tilde{\xi}) := \max\{1, \|\xi\|^{r-1}, \|\tilde{\xi}\|^{r-1}\} \|\xi - \tilde{\xi}\| \quad (\xi, \tilde{\xi} \in \Xi).$$

Proposition: (Rachev-Rüschendorf 98)
If Ξ is bounded, ζ_r may be reformulated as transportation problem

$$\zeta_r(P,Q) = \inf \left\{ \int_{\Xi \times \Xi} \hat{c}_r(\xi,\tilde{\xi}) \eta(d\xi, d\tilde{\xi}) : \pi_1 \eta = P, \pi_2 \eta = Q \right\},$$

where $\hat{c}_r$ is a metric (reduced cost) with $\hat{c}_r \leq c_r$ and given by

$$\hat{c}_r(\xi,\tilde{\xi}) := \inf \left\{ \sum_{i=1}^{n-1} c_r(\xi_{l_i}, \xi_{l_{i+1}}) : n \in \mathbb{N}, \xi_{l_i} \in \Xi, \xi_{l_1} = \xi, \xi_{l_n} = \tilde{\xi} \right\}.$$

Let P and Q be two discrete distributions with finite support, where ξ_i are the scenarios with probabilities p_i, $i = 1, \ldots, N$, of P and $\tilde{\xi}_j$ the scenarios and q_j, $j = 1, \ldots, n$, the probabilities of Q. Let Ξ denote the union of both scenario sets. Then

$$\begin{aligned}
\zeta_r(P, Q) &= \inf\left\{ \int_{\Xi\times\Xi} \hat{c}_r(\xi, \tilde{\xi})\eta(d\xi, d\tilde{\xi}) : \pi_1\eta = P, \pi_2\eta = Q \right\} \\
&= \inf\Big\{ \sum_{i=1}^{N}\sum_{j=1}^{n} \eta_{ij}\hat{c}_r(\xi_i, \tilde{\xi}_j) : \sum_{j=1}^{n} \eta_{ij} = p_i, \sum_{i=1}^{N} \eta_{ij} = q_j, \eta_{ij} \geq 0, \\
&\qquad\qquad i = 1, \ldots, N,\ j = 1, \ldots, n \Big\} \\
&= \sup\Big\{ \sum_{i=1}^{N} p_i u_i - \sum_{j=1}^{n} q_j v_j : p_i - q_j \leq \hat{c}_r(\xi_i, \tilde{\xi}_j), i = 1, \ldots, N, \\
&\qquad\qquad j = 1, \ldots, n \Big\}
\end{aligned}$$

These two formulas represent the primal and dual representations of $\zeta_r(P, Q)$ and at the same time primal and dual linear programs.

Math. Program., Ser. A 95: 493–511 (2003)

Digital Object Identifier (DOI) 10.1007/s10107-002-0331-0

J. Dupačová · N. Gröwe-Kuska · W. Römisch

Scenario reduction in stochastic programming

An approach using probability metrics

Received: July 2000 / Accepted: May 2002
Published online: February 14, 2003 –

Abstract. Given a convex stochastic programming problem with a discrete initial probability distribution, the problem of optimal scenario reduction is stated as follows: Determine a scenario subset of prescribed cardinality and a probability measure based on this set that is the closest to the initial distribution in terms of a natural (or canonical) probability metric. Arguments from stability analysis indicate that Fortet-Mourier type probability metrics may serve as such canonical metrics. Efficient algorithms are developed that determine optimal reduced measures approximately. Numerical experience is reported for reductions of electrical load scenario trees for power management under uncertainty. For instance, it turns out that after 50% reduction of the scenario tree the optimal reduced tree still has about 90% relative accuracy.

Key words. stochastic programming – quantitative stability – Fortet-Mourier metrics – scenario reduction – transportation problem – electrical load scenario tree

1. Introduction

Various important real-life decision problems can be formulated as convex stochastic programs which can be mostly written in the form

$$\min_{x \in X} \mathbb{E}_P f(\omega, x) = \int_\Omega f(\omega, x) P(d\omega). \tag{1}$$

Here, $X \subset \mathbb{R}^n$ is a given nonempty convex closed set, Ω a closed subset of $\mathbb{R}^s$ and $\mathcal{B}$ the Borel σ-field relative to Ω, the function f from $\Omega \times \mathbb{R}^n$ to the extended reals $\overline{\mathbb{R}}$ is measurable with respect to ω and lower semicontinuous and convex with respect to x, and P a fixed probability measure on $(\Omega, \mathcal{B})$, i.e., $P \in \mathcal{P}(\Omega)$, with $\mathbb{E}_P$ denoting expectation with respect to P. This formulation covers two- and multi-stage stochastic programs with recourse. In these cases, X is the set of feasible first-stage decisions and the function values $f(\omega, x)$ evaluate the best possible outcomes of decisions x in case that ω is observed.

J. Dupačová: Charles University Prague, Department of Probability and Mathematical Statistics, 18675 Prague 8, Czech Republic, e-mail: dupacova@karlin.mff.cuni.cz

N. Gröwe-Kuska: Humboldt-University Berlin, Institute of Mathematics, 10099 Berlin, Germany, e-mail: nicole@mathematik.hu-berlin.de

W. Römisch: Humboldt-University Berlin, Institute of Mathematics, 10099 Berlin, Germany, e-mail: romisch@mathematik.hu-berlin.de

Mathematics Subject Classification (1991): 90C15, 90C31

Optimal scenario reduction

The optimal scenario reduction problem

$$\min_{Q\in\mathcal{P}_n(\Xi)} \zeta_r(P,Q)$$

with $P \in \mathcal{P}_N(\Xi)$, $N > n$, can be decomposed into finding the optimal scenario set J to be deleted and into determining the optimal new probabilities given J.

Let P have scenarios ξ_i with probabilities p_i, $i = 1, \ldots, N$, and Q being supported by a given subset of scenarios ξ_j, $j \notin J \subset \{1, \ldots, N\}$, $|J| = N - n$.

The best approximation of P with respect to ζ_r by such a distribution Q exists and is denoted by Q^*. It has the distance

$$D_J := \zeta_r(P,Q^*) = \min_{Q\in\mathcal{P}_n(\Xi)} \zeta_r(P,Q) = \sum_{i\in J} p_i \min_{j\notin J} \hat{c}_r(\xi_i,\xi_j)$$

and the probabilities $q_j^* = p_j + \sum_{i\in J_j} p_i$, $\forall j \notin J$, where

$J_j := \{i \in J : j = j(i)\}$ and $j(i) \in \arg\min_{j\notin J} \hat{c}_r(\xi_i,\xi_j)$, $\forall i \in J$

(optimal redistribution).

Determining the optimal index set J with prescribed cardinality $N-n$ is, however, a combinatorial optimization problem: (n-median problem)

$$\min\{D_J : J \subset \{1, ..., N\}, |J| = N - n\}$$

Hence, the problem of finding the optimal set J for deleting scenarios is $\mathcal{NP}$-hard and polynomial time algorithms are not available.

First idea: Reformulation as linear mixed-integer program

$$\min \quad \frac{1}{n} \sum_{i,j=1}^{N} p_j x_{ij} \hat{c}_r(\xi_i, \xi_j) \qquad \text{s.t.}$$

$$\sum_{j=1, j\neq i}^{N} x_{ij} + y_i = 1 \quad (i = 1, \ldots, N), \quad \sum_{i=1}^{N} y_i = n\,,$$

$$x_{ij} \le y_i \quad 0 \le x_{ij} \le 1 \quad (i, j = 1, \ldots, N)\,,$$

$$y_i \in \{0, 1\} \quad (1, \ldots, N).$$

and application of standard software or of specialized algorithms.

Solution: $x_{ij} = \begin{cases} \frac{\min_{i\in J} \hat{c}_r(\xi_i,\xi_j)}{n\hat{c}_r(\xi_i,\xi_j)} & , i \notin J, j \in J \\ 0 & , \text{else.} \end{cases} \qquad y_i = \begin{cases} 1 & , i \notin J \\ 0 & , i \in J. \end{cases}$

Fast reduction heuristics

Second idea: Application of (randomized) greedy heuristics.

Starting point $(n = N-1)$: $\min_{l\in\{1,\dots,N\}} p_l \min_{j\neq l} \hat{c}_r(\xi_l, \xi_j)$

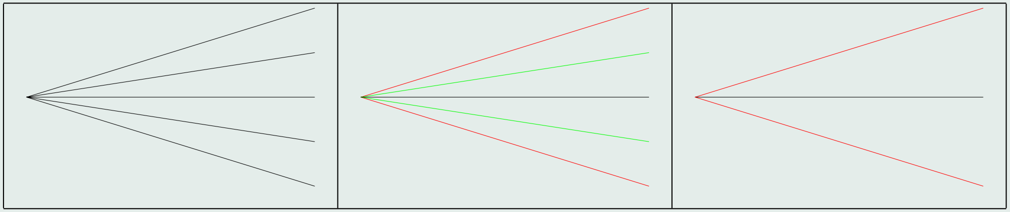
Algorithm 1: (Backward reduction)

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

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

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

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



Starting point ($n = 1$): $\min\limits_{u \in \{1,\ldots,N\}} \sum\limits_{k=1}^{N} p_k \hat{c}_r(\xi_k, \xi_u)$

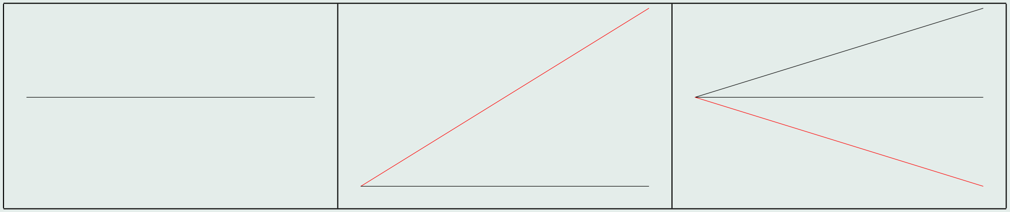
Algorithm 2: (Forward selection)

Step [0]: $J^{[0]} := \{1, \ldots, N\}$.

Step [i]: $u_i \in \arg \min\limits_{u \in J^{[i-1]}} \sum\limits_{k \in J^{[i-1]} \setminus \{u\}} p_k \min\limits_{j \notin J^{[i-1]} \setminus \{u\}} \hat{c}_r(\xi_k, \xi_j),$

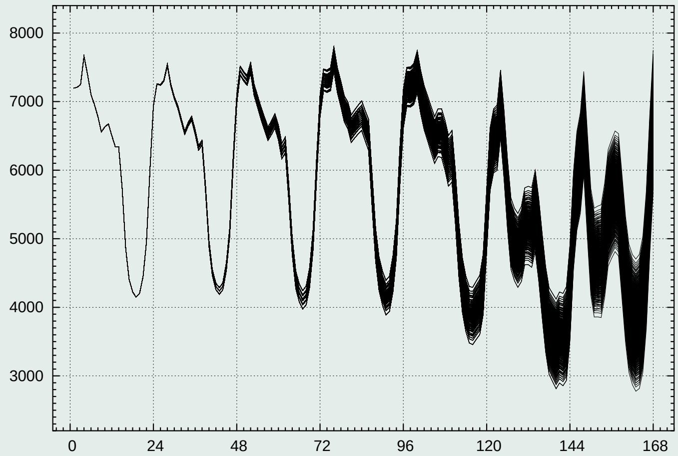
$J^{[i]} := J^{[i-1]} \setminus \{u_i\}$.

Step [n+1]: Optimal redistribution.

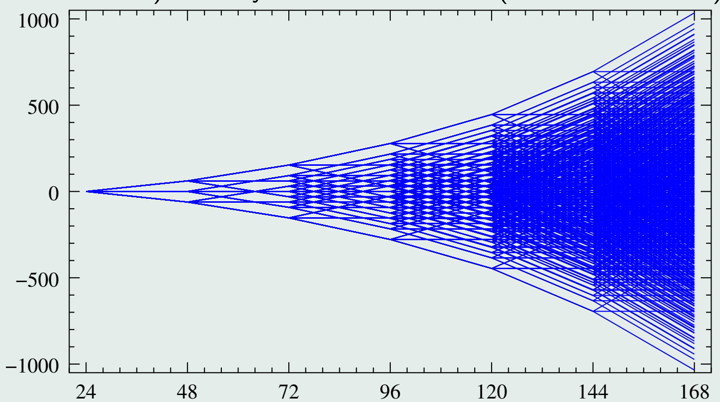


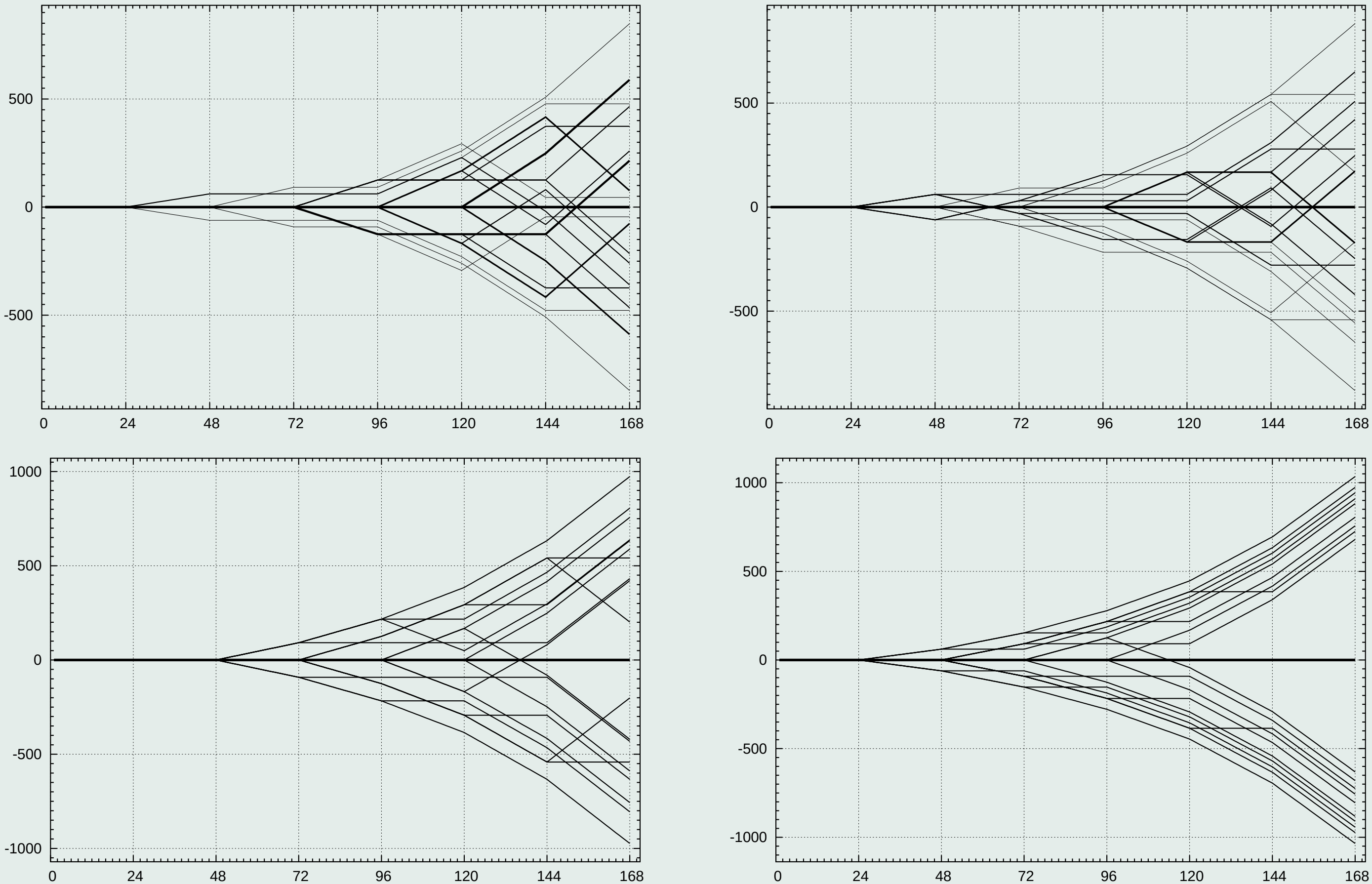
Example: (Electrical load scenario tree)

Ternary load scenario tree (N=729 scenarios)



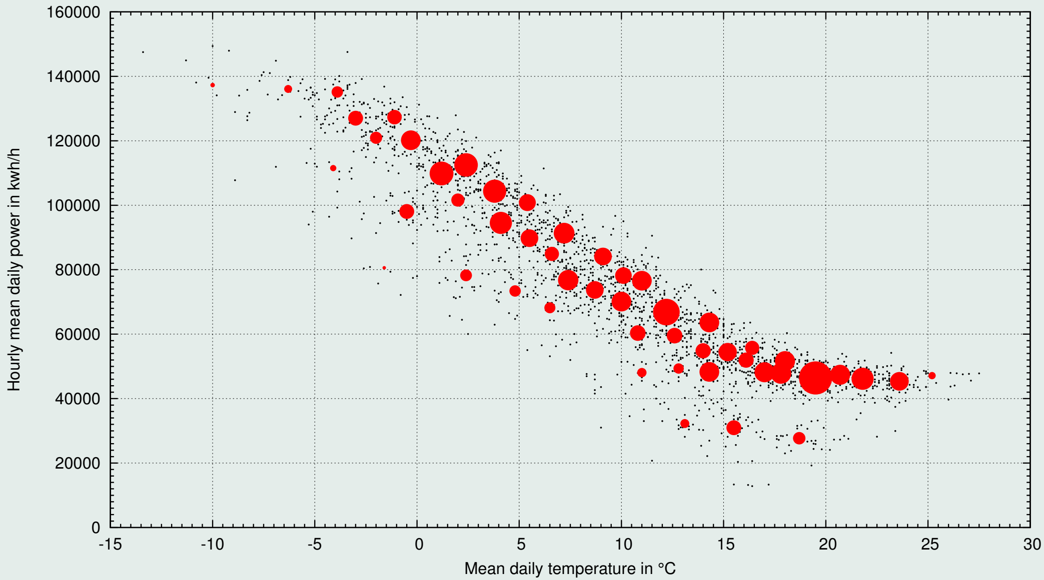
(Mean shifted) Ternary load scenario tree (N=729 scenarios)





Reduced load scenario trees with respect to the Fortet-Mourier distances ζ_r, $r = 1, 2, 4, 7$ and $n = 20$ (starting above left) (Heitsch-Römisch 07)

Application: Optimization of gas transport in a huge transportation network including hundreds of gas delivery nodes. A stationary situation is considered; more than 8 years of hourly data available at all delivery nodes; multivariate probability distribution for the gas output in certain temperature classes is estimated; 2340 samples based on randomized Quasi-Monte Carlo methods are generated and later reduced by scenario reduction to 50 scenarios.



(in: Koch, T., Hiller, B., Pfetsch, M. E., Schewe, L. (Eds.): Evaluating Gas Network Capacities, SIAM-MOS Series on Optimization, Philadelphia, 2015, Chapter 14, 295–315.)

Conclusions and outlook

- There exist reasonably fast heuristics for scenario reduction in linear two-stage stochastic programs. (Heitsch-Römisch 03).
- It may be worth to study and compare exact solution methods with heuristics.
- It is desirable to study scenario reduction based on the minimal function class $\mathcal{F} = \{\Phi(q(\cdot), h(\cdot) - T(\cdot)x) : x \in X\}$.
- Recursive application of the heuristics apply to generate scenario trees for multistage stochastic programs (Heitsch-Römisch 09).
- Heuristics for scenario reduction and scenario tree generation were implemented by H. Heitsch in GAMS/SCENRED 2.0.
- For scenario tree reduction the heuristics have to be modified.
- For mixed-integer two-stage stochastic programs and programs with chance constraints heuristics exist, but are based on different distances (discrepancies). They are more expensive and so far restricted to moderate dimensions.
- **Jitka's initial input was important** for all the further developments.

References

Avella, P., Sassano, A, Vasil'ev, I.: Computational study of large-scale p-median problems, *Mathematical Programming* 109 (2007), 89–114.

Dupačová, J., Gröwe-Kuska, N., Römisch, W.: Scenario reduction in stochastic programming: An approach using probability metrics, *Mathematical Programming* 95 (2003), 493–511.

Dupačová, J., Römisch, W.: Quantitative stability of scenario-based stochastic programs, in: Prague Stochastics '98 (M. Huškova et al. eds.), JCMF, Prague 1998, 119-124.

Heitsch, H., Römisch, W.: Scenario reduction algorithms in stochastic programming, *Computational Optimization and Applications* 24 (2003), 187–206.

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* 118 (2009), 371–406.

Henrion, R., Küchler, C., Römisch, W.: Scenario reduction in stochastic programming with respect to discrepancy distances, *Computational Optimization and Applications* 43 (2009), 67–93.

Li, Z., Floudas, C. A.: Optimal scenario reduction framework based on distance of uncertainty distribution and output performance:I. Single reduction via mixed integer linear optimization, *Computers and Chemical Engineering* 70 (2014), 50–66.

Römisch, W.: Stability of Stochastic Programming Problems, in: *Stochastic Programming* (A. Ruszczyński, A., Shapiro, A. eds.), Handbooks in Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 483–554.