IMSOR
THE INSTITUTE OF MATHEMATICAL STATISTICS
AND OPERATIONS RESEARCH

The Technical University of Denmark
DK-2800 Lyngby – Denmark

J. No. x544
1993-08-23
OPP

# Scenario Analysis and the Progressive Hedging Algorithm - Simple Numerical Examples

**Olafur P. Palsson**
**Hans F. Ravn**

**TECHNICAL REPORT**
**No. 11/1993**

ISSN 0906-9992

Trykt af IMSOR - DTH

# Scenario Analysis and the Progressive Hedging Algorithm - Simple Numerical Examples

Olafur P. Palsson*
Hans F. Ravn*

**Abstract**

The Progressive Hedging Algorithm (PHA) is a new method which may be applied to multiperiod optimization problems under uncertainty. It was first presented in Rockafellar and Wets (1991). In their paper no examples are given and the main emphasize is on the proofs and convergence analysis. Therefore, in order to understand the idea behind scenario analysis and PHA very simple numerical examples are given in this paper and the results are compared to results derived from classical stochastic optimization methods.

**Key Words:** Progressive Hedging; Scenario Analysis; Stochastic Optimization.

*The Institute of Mathematical Statistics and Operations Research, The Technical University of Denmark, DK-2800 Lyngby, Denmark.

# 1 Introduction

Many optimization problems with practical origin incorporate stochastic elements. In general it is very difficult to solve stochastic problems. This is due to either the mathematical properties (in particular if the space for stochasticity is infinite) or to the computational burden due to the combinatorial explosion (in the case with discrete stochastic events). Only few specific models can be solved easily, in particular the unconstrained Quadratic-Linear model with normally distributed stochastic elements. For an introduction to the classical material the readers are referred to Jensson (1975) (for a Mathematical Programming approach) and to Bertsekas (1987) (for the Dynamic Programming approach).

Recently, stochastic optimization problems have received renewed interest. One of the promising approaches is based on scenario analysis, which is a stochastic programming technique employing discrete scenarios with known probabilities, usually covering several time periods. Figure 1 shows a scenario tree with eight scenarios. Each of these eight scenarios are shown in Fig. 2.

The Progressive Hedging Algorithm (PHA) is a method to solve scenario problems. The PHA was first presented in Rockafellar and Wets (1991). Since then this method has been studied in several papers, e.g., Robinson (1991), Dembo (1991), Helgason and Wallace (1991), Wallace and Helgason (1991), and Palsson and Ravn (1993). In Helgason and Wallace (1991) it is shown how PHA can be combined with approximate solution of the individual scenario problems, resulting in a computationally efficient algorithm where two individual Lagrangian-based procedures are merged into one. This method is used to study examples from fisheries management. In Wallace and Helgason (1991) the structural properties of the PHA is discussed. In Palsson and Ravn (1993) the PHA is applied to the management of a heat storage.

The purpose of this paper is to give an easy introduction to the PHA. The paper starts with the description of the scenario analysis in connection with the decision problem. Then the PHA and some definitions needed are given, simple numerical examples are studied, and finally some concluding remarks are given.

# 2 Scenario Analysis

The present paper considers decision problems that are sequential in nature. This means that the decision vector $X \in \mathbb{R}^n$ is departed: $X = (x_1', x_2', \ldots, x_T')'$. Thus,

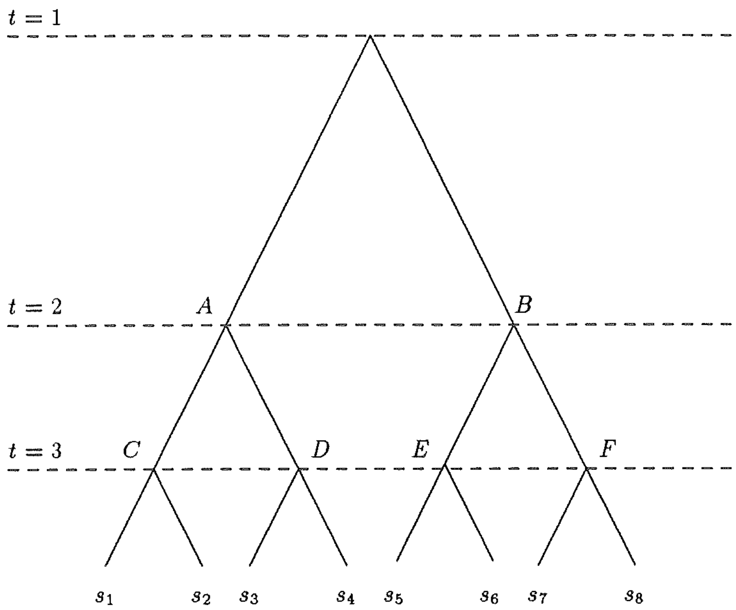


**Fig. 1 A scenario tree with eight scenarios. There are three time periods, i.e., $T=3$, and the stochastic variable can take two values each time.**

there are $T$ *stages* and at each stage $t$ one decision $x_t$ must be made, first $x_1$, then $x_2$, etc. A sequence of decisions is called a *policy.*

The information structure is also sequential. Thus, each time a decision is made some information is gained.

The information structure is described by scenarios. The uncertainty about parameters or components of the system is modeled by a small number of subproblems. Each version is called a *scenario.* The set of scenarios is denoted by $S$.

Each scenario $s$ is supposed to take place with probability $p_s$:

$$p_s > 0, \qquad \sum_{s \in S} p_s = 1 \tag{1}$$

Eight scenarios are indicated in Fig. 1. The sequential nature of the problem is indicated by the sequence of stages $t=1$, $t=2$ and $t=3$. Thus, $x_t$ must be chosen three times. Information is also gained three times, after a decision is made. Thus, at the time $x_1$ must be chosen, it is not known which of the eight scenarios, $s_1, \ldots, s_8$, will take place. The choice of $x_1$ must be made, and then the state will be in one of the two points $A$ or $B$. If the actual scenario was $s_1$, $s_2$, $s_3$ or $s_4$ the state will be in $A$, if the actual scenario was $s_5$, $s_6$, $s_7$ or $s_8$ the

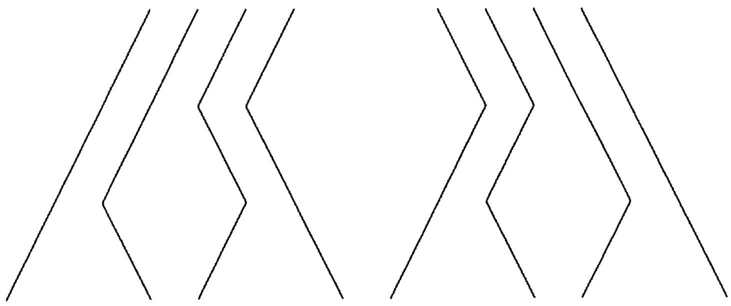

**Fig. 2 The eight scenarios from the scenario tree shown in Fig. 1.**

state will be in $B$.

It is assumed that after $x_1$ is chosen it can be observed, whether the state is in $A$ or $B$, and therefore information is gained.

Now $x_2$ must be chosen, and then the state will end up in $C$ or $D$ (if the state is in $A$) or $E$ or $F$ (if the state is in $B$), and so on.

This decision process has two kinds of requirement to the choice of $x_t$ at each stage. First, there is the requirement that $x_t$ must be *admissible.* This means that $x_t$ must be chosen from some set $C_s$.

Second there is the requirement that $x_t$ be *implementable.* This refers to the information structure. Thus, when $x_t$ is chosen at stage $t$ only the knowledge which has been gained until then is available.

This can be formulated more specifically as follows. As seen from Fig. 1 - 2 there are relations between the scenarios. Thus, at stage $t = 2$ the scenarios fall in two groups: one contains scenarios $s_1, \ldots, s_4$ the other scenarios $s_5, \ldots, s_8$. Similarly at stage $t = 3$ there are four groups of scenarios, each group with two scenarios.

Such a group is denoted as a *scenario bundle.* As seen, the scenario bundles at time $t$ are such that the scenarios in each bundle are observationally indistinguishable at time $t$.

Therefore, the requirement of implementability can now be stated as follows:

> *The policy is implementable if for all $t$, $x_t$ is dependent on the scenario bundle but not on the particular scenario in the bundle.*

At any time $t$, it is known what scenario bundle is relevant for the decision. Thus, at stage $t = 2$ on Fig. 1 it is known if the relevant scenario bundle contains $s_1, \ldots, s_4$ (this is the case if the state is at $A$) or $s_5, \ldots, s_8$ (if the state is at $B$).

The decision problem can now be formulated as follows. There is a sequential decision and information gaining process, defined by scenarios $s \in S$ with probabilities $p_s$, see Eq. (1). To each decision $X \in \mathbb{R}^n$ and each scenario $s \in S$ there is a constant set $C_s$ and an objective function value $f_s(X(s))$ can be calculated.

The purpose is to minimize the expectation, i.e., to solve

$$\min \sum_{s \in S} p_s f_s(X(s)) \tag{2}$$

subject to

$$X(s) \in C_s \qquad (X(s) \text{ is admissible}) \tag{3}$$

and

$$X(s) \text{ is implementable} \tag{4}$$

This problem is difficult to solve in general. The algorithm described below is based on the assumption that if it is known which scenario $s \in S$ is relevant then the optimal solution to this *deterministic* problem:

$$\min f_s(X(s)) \tag{5}$$

$$X(s) \in C_s \tag{6}$$

is easily found for any $s \in S$.

If this is true the following problem can also easily be solved

$$\min f_s(X(s)) + X'(s)W + \frac{1}{2}r||X(s) - \hat{X}||^2 \tag{7}$$

$$X(s) \in C_s \tag{8}$$

for any $s \in S$. In Eqs. (7) - (8) $\hat{X}$ is a given value in $\mathbb{R}^n$, $W$ is a given value in $\mathbb{R}^n$, $X'W$ is the scalar product of $X$ and $W$, $r \in \mathbb{R}$, $r > 0$, and $|| \cdot ||$ is the Euclidean norm.

The relation between the problems in Eqs. (5) - (6) and in Eqs. (7) - (8) may be interpreted as follows.

To each of the potentially relevant bundles at stage $t$ the policy may be calculated

$$\hat{X}_t = \sum p_s X_t(s) / \sum p_s \tag{9}$$

In this, the summation are over the scenarios in the relevant bundle, and $X_t(s)$ is the solution at time $t$ to a problem like Eqs. (7) - (8). It is seen that $\hat{X}_t$ in Eq. (9) can be interpreted as a conditional expectation, relative to the relevant scenario bundle, i.e., relative to the information available. It is also observed that by this definition $\hat{X}_t$ is implementable. Moreover, if $C_s$ is convex and independent of the particular scenario in the bundle, then $\hat{X}_t$ is admissible.

The vector $W \in \mathbb{R}^n$ in Eq. (7) may now be interpreted as a Lagrange multiplier vector relative to the constraint in Eq. (9) (or similarly Eq. (4)) and the last term in Eq. (7) may be interpreted as a penalty term which is introduced in order to attain convergence stability in an algorithmic sense.

In summary the project can be stated as follows: the decision problem in Eqs. (2) - (4) is solved by solving the following problem

$$\min \sum p_s f_s(X(s)) \tag{10}$$

$$X(s) \in C_s \tag{11}$$

$$X(s) = \hat{X} \tag{12}$$

where $\hat{X}$ is defined in Eq. (9), and Eq. (12) ensures that the solution is implementable.

In Rockafellar and Wets (1991) an algorithm is given for solution of the problem in Eqs. (10) - (12). This is described in the next section.

# 3 Progressive Hedging Algorithm

In this section the PHA is given as in Rockafellar and Wets (1991) and just the definitions needed are given. The readers are referred to Rockafellar and Wets (1991) for the details.

PROGRESSIVE HEDGING ALGORITHM. In iteration $\nu$ (where $\nu = 0, 1, \ldots$) one has an *admissible* but not necessarily *implementable* policy $X^\nu \in \boldsymbol{C}$ and a *price system* $W^\nu \in \boldsymbol{M}$. (Initially one can take $X^0$ to be the policy obtained by letting $X^0(s)$ be for each scenario $s \in S$ an optimal solution to the given scenario subproblem $(\boldsymbol{P}_s)$. One can take $W^0 = 0$.)

**Step 1.** Calculate the policy $\hat{X}^\nu = JX^\nu$, which is implementable but not necessarily admissible. (If ever one wishes to stop, this policy $\hat{X}^\nu$ is to be offered as the best substitute yet available for a solution to $\boldsymbol{P}$.)

**Step 2.** Calculate as $X^{\nu+1}$ an (approximately) optimal solution to the subproblem

$$(\boldsymbol{P}^\nu) \quad \text{minimize}[F(X) + < X, W^\nu > + \tfrac{1}{2}r||X - \hat{X}^\nu||^2] \quad \text{over all } X \in \boldsymbol{C}$$

This decomposes into solving (approximately) for each scenario $s \in S$ the subproblem

$$(\boldsymbol{P}^\nu_s) \quad \text{minimize}[f_s(x) + x \cdot W^\nu(s) + \tfrac{1}{2}r|x - \hat{X}^\nu(s)|^2] \quad \text{over all } x \in C_s$$

in order to get $X^{\nu+1}(s)$. The policy $X^{\nu+1}$ will again be admissible but not necessarily implementable.

**Step 3.** Update from $W^\nu$ to $W^{\nu+1}$ by the rule $W^{\nu+1} = W^\nu + rKX^{\nu+1}$. The price system $W^{\nu+1}$ will again be in $\boldsymbol{M}$. Return to Step 1 with $\nu$ replaced by $\nu + 1$. $\square$

The definitions and some comments are given in order of appearance:

- $\nu$ is an iteration number,
- $X^\nu$ is the decision to be made at iteration $\nu$, $X^\nu \in \mathbb{R}^n$,
- $\boldsymbol{C}$ is a set of admissible polices, $\boldsymbol{C} \subseteq \mathbb{R}^n$,
- $W^\nu$ is a price system (Lagrange multiplier), $W^\nu \in \boldsymbol{M} \subset \mathbb{R}^n$,
- $\boldsymbol{M}$ is a set, $\boldsymbol{M} \subset \mathbb{R}^n$. $\boldsymbol{M}$ is such that for each scenario bundle $A$ at stage $t$, $\sum_{s \in A} p_s W_t(s) = 0$. The algorithm secures that this is fulfilled, if $W^0$ fulfills it; the initial choice $W^0 = 0$ fulfills it,
- $X(s)$ is the decision to be made given the scenario $s$,
- $s$ is the scenario,
- $S$ is the scenario set,
- $\boldsymbol{P}_s$ is the scenario subproblem,
- $\hat{X}$ is a conditional expectation, defined as in Eq. (9) using $X = X^\nu$. Observe that if $C_s$ is convex for all $s \in S$ then $\hat{X}$ is admissible,
- $J$ is an aggregation operator (or conditional expectation operator relative to the given information structure and values $p_s$, i.e., $JX^\nu$ defines $\hat{X}$ as in Eq. (9) using $X = X^\nu$, where $p_s$ is the probability attached to $s$),
- $\boldsymbol{P}$ is the scenario problem,

- $F(X)$ is equal to $E\{f_s(X(s))\}$, where $E$ denotes the expectation value, see Eq. (10),
- $<,>$ is the inner product on Euclidean vector space,
- $r$ is a penalty parameter ($> 0$),
- $||\cdot||$ is a norm,
- $f_s(x)$ is the objective function value for the scenario subproblem $s$ and decision $X$,
- $x$ is the decision to be made for the scenario subproblem,
- $|\cdot|$ is the ordinary Euclidean norm ($|x| = \sqrt{\sum x_i^2}$),
- $C_s$ is a set of admissible polices for the scenario $s$, $C_s \subseteq \mathbb{R}^n$,
- $K$ is an operator ($KX = X - \hat{X}$), such that $KX^{\nu+1}$ in step 3 is equal to $X^{\nu+1} - \hat{X}^{\nu+1}$.

**Remark:** As noted, $\hat{X}^{\nu+1}$ is needed in the update of $W^\nu$ to $W^{\nu+1}$, but in the way the algorithm is written in Rockafellar and Wets (1991), $\hat{X}^{\nu+1}$ is not available at that moment (first at Step 1 in the next iteration). Therefore, it is recommended that the algorithm is rewritten as:

**Step 0.** Initialize.

**Step 1.** Solve (approximately) for each scenario $s \in S$ the subproblem (in order to get $X^{\nu+1}(s)$).

**Step 2.** Calculate the policy $\hat{X}^{\nu+1} = JX^\nu$.

**Step 3.** Update from $W^\nu$ to $W^{\nu+1}$. Return to Step 1 with $\nu$ replaced by $\nu + 1$. □

In Rockafellar and Wets (1991) it is shown, that if for any $s \in S$, $C_s$ is convex and $f_s$ is convex, and some regularity conditions apply then the algorithm will converge to an optimal solution to Eqs. (10) - (12). The regularity conditions are what can be expected from basic mathematical programming theory. The rate of convergence is linear. It is also shown that the algorithm will converge if Eqs. (7) - (8) are solved only approximately, and the error bound is specified.

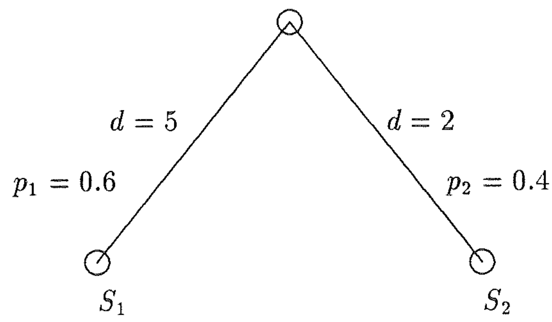


**Fig. 3 A small scenario tree.**

# 4 Numerical Example

In this section the PHA is used on simple stochastic optimization problems, the original algorithm is used.

## 4.1 Example 1

The problem is to find

$$\min_{x} E[(x-d)^2] \tag{13}$$

subject to the constraint

$$3 \leq x \leq 6, \; x \in \mathbb{R} \tag{14}$$

$d$ can takes two values $[5, 2]$ with probabilities $[0.6, 0.4]$, respectively. Thus, there are two scenarios, one with $d = 5$ and one with $d = 2$. This is illustrated in Fig. 3. It is observed that this problem is not sequential (except in a trivial sense) since only one choice is made.

It is clearly seen from Eqs. (13) - (14), that the optimal solutions for the deterministic problem corresponding to each of the two scenarios are $X^*(s_1) = 5$ and $X^*(s_2) = 3$, respectively. Now the Progressive Hedging Algorithm is applied to this problem.

### 4.1.1 PHA

$\nu = 0$:

$X^0 = (5, 3)$ (i.e. $X^0(s_1) = x_1^0 = 5$ and $X^0(s_2) = x_2^0 = 3$)
$W^0 = (0, 0)$ and $r = 1$ (chosen).

STEP 1:
$\hat{X}^0 = p_{s,1}X^0(s_1) + p_{s,2}X^0(s_2) = 0.6 \times 5 + 0.4 \times 3 = 4.2$
(this is admissible, i.e., $3 \leq \hat{X}^0 \leq 6$)

STEP 2:
$s_1 : \min[(x_1 - 5)^2 + x_1 0 + \frac{1}{2}1(x_1 - 4.2)^2] \Rightarrow X^1(s_1) = 4.73$
$s_2 : \min[(x_2 - 2)^2 + x_2 0 + \frac{1}{2}1(x_2 - 4.2)^2] \Rightarrow X^1(s_2) =$ ~~2.73~~
the solution must be admissible, i.e., $X^1(s_2) = 3$

STEP 3:
$W^1 = W^0 + r(X^1 - \hat{X}^1) = (0, 0) + 1(4.73 - 4.04, 3.0 - 4.04)$
$= (0.69, -1.04)$

$\nu = 1$:

STEP 1:
$\hat{X}^1 = p_{s,1}X^1(s_1) + p_{s,2}X^1(s_2) = 0.6 \times 4.73 + 0.4 \times 3.0 = 4.04$

STEP 2:
$s_1 : \min[(x_1 - 5)^2 + x_1 0.69 + \frac{1}{2}1(x_1 - 4.04)^2] \Rightarrow X^2(s_1) = 4.45$
$s_2 : \min[(x_2 - 2)^2 + x_2(-1.04) + \frac{1}{2}1(x_2 - 4.04)^2] \Rightarrow X^2(s_2) = 3.03$

STEP 3:
$W^2 = (0.69, -1.04) + 1(4.45 - 3.88, 3.03 - 3.88) = (1.26, -1.89)$

$\nu = 2$:

STEP 1:
$\hat{X}^2 = p_{s,1}X^2(s_1) + p_{s,2}X^2(s_2) = 0.6 \times 4.45 + 0.4 \times 3.03 = 3.88$

STEP 2:
$s_1 : \min[(x_1 - 5)^2 + x_1 1.26 + \frac{1}{2}1(x_1 - 3.88)^2] \Rightarrow X^3(s_1) = 4.21$
$s_2 : \min[(x_2 - 2)^2 + x_2(-1.89) + \frac{1}{2}1(x_2 - 3.88)^2] \Rightarrow X^3(s_2) = 3.26$

STEP 3:

**Table 1 Results from the iterations.**

| $\nu$ | $X(s_1)$ | $X(s_2)$ | $\hat{X}$ | $w_1$ | $w_2$ | $\hat{z}$ | $\hat{z}'$ |
|---|---|---|---|---|---|---|---|
| 0 | 5.00 | 3.00 | 4.20 | 0.00 | 0.00 | 0.40 | 2.32 |
| 1 | 4.73 | 3.00 | 4.04 | 0.69 | -1.04 | 0.44 | 2.22 |
| 2 | 4.45 | 3.03 | 3.88 | 1.26 | -1.89 | 0.60 | 2.17 |
| 3 | 4.21 | 3.26 | 3.83 | 1.64 | -2.46 | 1.01 | 2.16 |
| 4 | 4.06 | 3.43 | 3.81 | 1.89 | -2.84 | 1.35 | 2.16 |
| 5 | 3.97 | 3.55 | 3.80 | 2.06 | -3.09 | 1.60 | 2.16 |
| 6 | 3.91 | 3.63 | 3.80 | 2.18 | -3.26 | 1.77 | 2.16 |
| 7 | 3.88 | 3.69 | 3.80 | 2.25 | -3.38 | 1.90 | 2 16 |
| 8 | 3.85 | 3.73 | 3.80 | 2.30 | -3.45 | 1.98 | 2.16 |
| 9 | 3.83 | 3.75 | 3.80 | 2.33 | -3.50 | 2.04 | 2.16 |

$$W^3 = (1.26, -1.89) + 1(4.21 - 3.83, 3.26 - 3.83) = (1.64, -2.46)$$

$\nu = 3$:

STEP 1:

$$\hat{X}^3 = p_{s,1}X^3(s_1) + p_{s,2}X^3(s_2) = 0.6 \times 4.21 + 0.4 \times 3.26 = 3.83$$

STEP 2:

$$s_1 : \min[(x_1 - 5)^2 + x_1 1.64 + \tfrac{1}{2}1(x_1 - 3.83)^2] \Rightarrow X^4(s_1) = 4.06$$
$$s_2 : \min[(x_2 - 2)^2 + x_2(-2.46) + \tfrac{1}{2}1(x_2 - 3.83)^2] \Rightarrow X^4(s_2) = 3.43$$

STEP 3:

$$W^4 = (1.64, -2.46) + 1(4.06 - 3.83, 3.43 - 3.83) = (1.89, -2.84)$$

The results from 9 iterations are found in Table 1. The table also shows the results for the objective function, i.e. $\hat{z} = p_{s,1}(X(s_1) - d_1)^2 + p_{s,2}(X(s_2) - d_2)^2$ and $\hat{z}' = p_{s,1}(\hat{X} - d_1)^2 + p_{s,2}(\hat{X} - d_2)^2$. Observe that $\sum p_s w_s \simeq 0$, i.e., $0.6 \cdot 2.33 + 0.4 \cdot (-3.50) \simeq 0$.

### 4.1.2 Mathematical Programming Approach

The stochastic problem could also be solved by solving

$$\min_x [p_{s,1}(x - d_1)^2 + p_{s,2}(x - d_2)^2] = \min_x [0.6(x - 5)^2 + 0.4(x - 2)^2] \qquad (15)$$

with the constraint in Eq. (14). This gives $x^* = 3.8$, i.e., the same solution as before.

### 4.1.3 Choice of the Penalty Parameter

In Helgason and Wallace (1991) it is pointed out that the choice of the penalty parameter $r$ is of vital importance and the conclusion drawn there is that the penalty should be as small as possible, provided it is large enough to guarantee convergence.

The error can be measured as the expected value (with respect to the distribution of the scenarios), see Helgason and Wallace (1991)

$$m^\nu = E\left[\sum_{t=0}^{T}\left\{||\hat{X}_t^\nu(s) - \hat{X}_t^{\nu-1}(s)||^2 + \frac{1}{r^2}||W_t(s)^\nu - W_t(s)^{\nu-1}||^2\right\}\right] \qquad (16)$$

Hence the termination criterion is

$$m^\nu \leq \varepsilon \qquad (17)$$

For the simple problem in Example 1 results for various values of the penalty parameter, $r$, were investigated. For this case the error measure is simply

$$m^\nu = (\hat{X}^\nu - \hat{X}^{\nu-1})^2 + \frac{1}{r^2}(p_{s,1}(w_1^\nu - w_1^{\nu-1})^2 + p_{s,2}(w_2^\nu - w_2^{\nu-1})^2) \qquad (18)$$

The results are demonstrated in Table 2. It is seen that $m^\nu$ converges faster for large values of $r$, but it converges to a wrong value of $\hat{X}$. This is explained by the fact that for large values of $r$ the solutions of the subproblems are pressed together very quickly and then $\hat{X}$ converges slowly to the right value. On the other hand, for lower values of $r$, the solutions of the subproblems converge to the right solutions from both sides. This is illustrated in Fig. 4 – 6.

**Table 2 Number of iterations to fulfill the criterion in Eq. (17) for different $r$ values and the corresponding implementable solution, $\hat{X}$.**

| | $\varepsilon = 5 \cdot 10^{-3}$ | | $\varepsilon = 5 \cdot 10^{-5}$ | |
|---|---|---|---|---|
| $r$ | $\nu$ | $\hat{X}^{\nu}$ | $\nu$ | $\hat{X}^{\nu}$ |
| 0.1 | 75 | 3.80 | 123 | 3.80 |
| 0.5 | 17 | 3.80 | 28 | 3.80 |
| 1.0 | 10 | 3.80 | 16 | 3.80 |
| 2.0 | 6 | 3.81 | 10 | 3.81 |
| 3.0 | 5 | 3.85 | 9 | 3.81 |
| 4.0 | 5 | 3.86 | 10 | 3.81 |
| 5.0 | 5 | 3.90 | 11 | 3.81 |
| 10.0 | 4 | 4.03 | 15 | 3.83 |
| 100.0 | 2 | 4.19 | 8 | 4.15 |

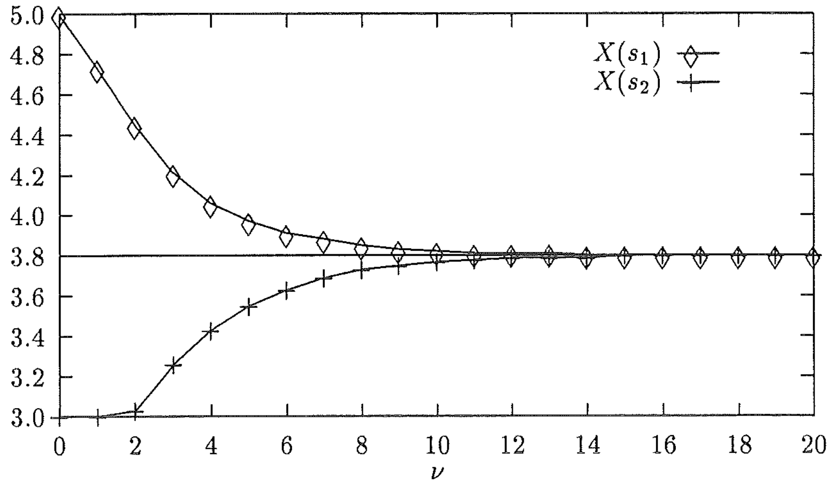


**Fig. 4 Convergence of the subproblems.** $r = 1.0$.

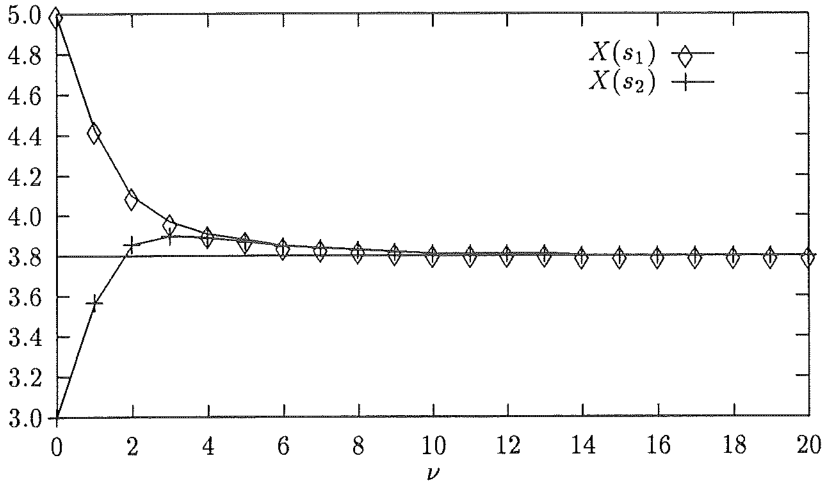


**Fig. 5 Convergence of the subproblems.** $r = 5.0$.

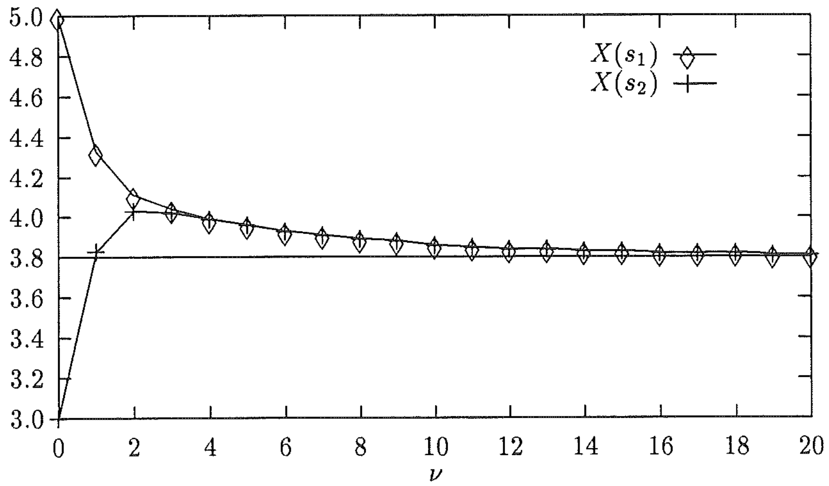


**Fig. 6 Convergence of the subproblems.** $r = 10.0$.

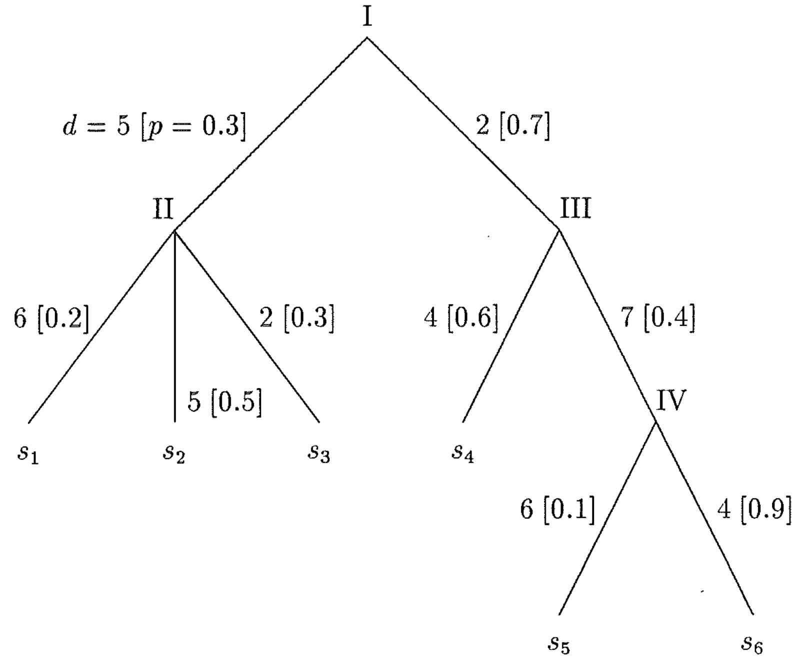


**Fig. 7 The scenario tree used in Example 2, the probabilities are in parentheses $[p]$ and $d$ is the stochastic variable.**

## 4.2 Example 2

The problem is now to find

$$\min_x E\left[\sum_{t=1}^{3}\left\{\frac{1}{2}(x_t - x_{t-1})^2 + (x_t - d_t)^2\right\}\right] \tag{19}$$

subject to the constraints

$$3 \le x_t \le 6, \qquad 1 \le t \le T \tag{20}$$

$d$ and $p$ take the values shown in Fig. 7. $x_0 = 4$ and $T$ is equal to two or three, i.e., $T = 2$ for scenarios $s_1$ to $s_4$ and $T = 3$ for scenarios $s_5$ and $s_6$. This problem is sequential. It is observed that at time $t = 1$ there is one scenario bundle containing all 6 scenarios. At time $t = 2$ there are two bundles, the bundle containing $s_1$, $s_2$ and $s_3$, and the bundle containing $s_4$, $s_5$ and $s_6$. At time $t = 3$ there are 5 bundles, 4 of them are "empty", and the last one contains the two scenarios $s_5$ and $s_6$.

Clearly this example is more complex than Example 1, therefore, not all intermediate results will be shown.

**Table 3 Initial solutions. $*$ means that the solution lies on the lower boundary.**

| $S$ | $X^0(s)$ | | | $W^0(s)$ | | |
|---|---|---|---|---|---|---|
| | $t=1$ | $t=2$ | $t=3$ | $t=1$ | $t=2$ | $t=3$ |
| $s_1$ | 4.91 | 5.64 | | 0.0 | 0.0 | |
| $s_2$ | 4.73 | 4.91 | | 0.0 | 0.0 | |
| $s_3$ | 4.25 | 3.00* | | 0.0 | 0.0 | |
| $s_4$ | 3.00* | 3.67 | | 0.0 | 0.0 | |
| $s_5$ | 3.46 | 5.85 | 5.95 | 0.0 | 0.0 | 0.0 |
| $s_6$ | 3.37 | 5.46 | 4.49 | 0.0 | 0.0 | 0.0 |

### 4.2.1 PHA

First the optimal solutions to the given scenario subproblems are found and the price system is set to zero. This is shown in Table 3. In this example $r$ is chosen equal to 1.

$\underline{\nu = 0:}$

**STEP 1**

The conditional expectations for the four decision states, labeled I, II, III, and IV in Fig. 7 are obtained from data in Tables 3 and 4

$$\begin{aligned}
\hat{X}_I^0 &= 0.06 * 4.91 + 0.15 * 4.73 + 0.09 * 4.25 \\
&\quad +0.42 * 3.00 + 0.028 * 3.46 + 0.252 * 3.37 = 3.59 \\
\hat{X}_{II}^0 &= 0.20 * 5.64 + 0.50 * 4.91 + 0.30 * 3.00 = 4.48 \\
\hat{X}_{III}^0 &= 0.60 * 3.67 + 0.04 * 5.85 + 0.36 * 5.46 = 4.40 \\
\hat{X}_{IV}^0 &= 0.10 * 5.95 + 0.90 * 4.49 = 4.63
\end{aligned}$$

**Table 4 The probabilities, see also Fig. 7.**

| $S$ | $p(s)$ | | | $\hat{X}$ | | |
|---|---|---|---|---|---|---|
| | $t=1$ | $t=2$ | $t=3$ | $t=1$ | $t=2$ | $t=3$ |
| $s_1$ | 0.06 | 0.20 | | I | II | |
| $s_2$ | 0.15 | 0.50 | | I | II | |
| $s_3$ | 0.09 | 0.30 | | I | II | |
| $s_4$ | 0.42 | 0.60 | | I | III | |
| $s_5$ | 0.028 | 0.04 | 0.1 | I | III | IV |
| $s_6$ | 0.252 | 0.36 | 0.9 | I | III | IV |

## STEP 2 & 3

As an example, solution of subproblem $s_1$ is found as follows ($x_1 = X(s_1)_{t=1}$, $x_2 = X(s_1)_{t=2}$, $w_1 = W(s_1)_{t=1}$, and $w_2 = W(s_1)_{t=2}$)

$$\begin{aligned} f &= \frac{1}{2}(x_1 - x_0)^2 + (x_1 - d_1)^2 + \frac{1}{2}(x_2 - x_1)^2 + (x_2 - d_2)^2 \\ &\quad + x_1 w_1 + x_2 w_2 + \frac{1}{2}r(x_1 - \hat{X}_I)^2 + \frac{1}{2}r(x_2 - \hat{X}_{II})^2 \end{aligned} \tag{21}$$

$\partial f / \partial x_1 = 0$ and $\partial f / \partial x_2 = 0$ results in

$$\begin{aligned} (4 + r)x_1 - x_2 &= x_0 + 2d_1 - w_1 + r\hat{X}_I \\ -x_1 + (3 + r)x_2 &= 2d_2 - w_2 + r\hat{X}_{II} \end{aligned} \tag{22}$$

respectively. This can easily be solved, the solutions are $x_1 = 4.57$ and 5.26, which also satisfy Eq. (20).

The price system is updated as

$$\begin{aligned} w_1^1 &= w_1^0 + r(x_1^1 - \hat{X}_I^1) = 0.0 + 1(4.57 - 3.54) = 1.03 \\ w_2^1 &= w_2^0 + r(x_2^1 - \hat{X}_{II}^1) = 0.0 + 1(5.26 - 4.37) = 0.89 \end{aligned} \tag{23}$$

The results for the other subproblems are found in Table 5.

<u>$\nu = 5$:</u>

**Table 5** $X^1(s)$ and $W^1(s)$.

| $S$ | $X^1(s)$ | | | $W^1(s)$ | | |
|---|---|---|---|---|---|---|
| | $t=1$ | $t=2$ | $t=3$ | $t=1$ | $t=2$ | $t=3$ |
| $s_1$ | 4.57 | 5.26 | | 1.03 | 0.89 | |
| $s_2$ | 4.47 | 4.74 | | 0.93 | 0.37 | |
| $s_3$ | 4.15 | 3.16 | | 0.61 | -1.21 | |
| $s_4$ | 3.09 | 3.85 | | -0.45 | -0.51 | |
| $s_5$ | 3.34 | 5.35 | 5.41 | -0.20 | 0.96 | 0.93 |
| $s_6$ | 3.30 | 5.14 | 4.37 | -0.25 | 0.75 | -0.10 |

**Table 6** $X^6(s)$ and $W^6(s)$.

| $S$ | $X^6(s)$ | | | $W^6(s)$ | | |
|---|---|---|---|---|---|---|
| | $t=1$ | $t=2$ | $t=3$ | $t=1$ | $t=2$ | $t=3$ |
| $s_1$ | 3.83 | 4.42 | | 3.08 | 2.62 | |
| $s_2$ | 3.84 | 4.27 | | 2.90 | 1.08 | |
| $s_3$ | 3.87 | 3.81 | | 2.35 | -3.55 | |
| $s_4$ | 3.41 | 4.32 | | -1.30 | -1.57 | |
| $s_5$ | 3.36 | 4.72 | 4.57 | -0.98 | 2.68 | 2.74 |
| $s_6$ | 3.37 | 4.74 | 4.26 | -1.01 | 2.32 | -0.31 |

**STEP 1**

$\hat{X}_I^5 = 3.53$, $\hat{X}_{II}^5 = 4.16$, $\hat{X}_{III}^5 = 4.48$ and $\hat{X}_{IV}^5 = 4.29$

**STEP 2 & 3**

$X^6(s)$ and $W^6(s)$ are shown in Table 6.

**Table 7 The conditional expectation, $\hat{X}$, for the four decision stages.**

| $\nu$ | $\hat{X}_I$ | $\hat{X}_{II}$ | $\hat{X}_{III}$ | $\hat{X}_{IV}$ |
|---|---|---|---|---|
| 0 | 3.59 | 4.48 | 4.40 | 4.63 |
| 1 | 3.54 | 4.37 | 4.39 | 4.48 |
| 2 | 3.53 | 4.29 | 4.41 | 4.40 |
| 3 | 3.53 | 4.24 | 4.44 | 4.35 |
| 4 | 3.53 | 4.19 | 4.46 | 4.32 |
| 5 | 3.53 | 4.16 | 4.48 | 4.29 |
| 10 | 3.53 | 4.08 | 4.54 | 4.25 |
| 15 | 3.53 | 4.05 | 4.55 | 4.23 |
| 20 | 3.53 | 4.05 | 4.55 | 4.23 |

**Summary:** Table 7 shows the conditional expectations for the four decision stages. It is noted that $\hat{X}_I$, the first decision, converges already after 2 iterations while the others converges much slower or after approximately 15 iterations.

Table 7 also shows that the first decision is 3.53, then if $d_1$ happens to be equal $x_1 = 5$ then the second decision is 4.05, else it is 4.55 and finally if $d_1 = 2$ and $d_2 = 7$ then the optimal decision is 4.23.

### 4.2.2 Classical Stochastic Optimization

**Dynamic Programming:** The stochastic problem may be solved by Dynamic Programming as follows.

The expected value at stage $t = 3$, state $IV$ is

$$V(3, IV, x_3, x_2) = \frac{1}{2}(x_3 - x_2)^2 + 0.1(x_3 - 6)^2 + 0.9(x_3 - 4)^2 \tag{24}$$

Differentiating and equating to zero gives the optimal $x_3$:

$$\frac{\partial V(3, IV, x_3, x_2)}{\partial x_3} = (x_3 - x_2) + 0.2(x_3 - 6) + 1.8(x_3 - 4) = 0 \tag{25}$$

or

$$x_3 = \frac{x_2}{3} + 2.8 \tag{26}$$

Here and in the sequel Eq. (20) is disregarded, which will be fulfilled. Thus the optimal expected value $V$ is:

$$V(3,IV,x_2) = \frac{1}{9}\left[\frac{1}{2}(8.4-2x_2)^2 + 0.1(x_2-9.6)^2 + 0.9(x_2-3.6)^2\right] \qquad (27)$$

The expected value at stage $t=2$, state $II$ is

$$V(2,II,x_2,x_1) = \frac{1}{2}(x_2-x_1)^2 + 0.2(x_2-6)^2 + 0.5(x_2-5)^2 + 0.3(x_2-2)^2 \qquad (28)$$

In the same way as before the optimal $x_2$ is:

$$x_2 = \frac{1}{3}[x_1+8.6] \qquad (29)$$

and the optimal expected value V:

$$V(2,II,x_1) = \frac{1}{9}\left[\frac{1}{2}(8.6-2x_1)^2 + 0.2(x_1-9.4)^2 + 0.5(x_1-6.4)^2 + 0.3(x_1+2.6)^2\right] \qquad (30)$$

The expected value of stage $t=2$, state $III$ is, using optimal decision at stage $t=3$ (see Eq. (27))

$$\begin{aligned} V(2,III,x_2,x_1) &= \frac{1}{2}(x_2-x_1)^2 + 0.6(x_2-4)^2 + 0.4(x_2-7)^2 \\ &\quad + 0.4V(3,IV,x_2) \qquad (31) \end{aligned}$$

This gives the optimal $x_2$:

$$x_2 = \frac{1}{9.8}[3x_1+34.56] \qquad (32)$$

and the optimal expected value $V$:

$$\begin{aligned} V(2,III,x_1) &= \qquad (33) \\ &\frac{1}{9.8^2}\left\{\frac{1}{2}(34.56-6.8x_1)^2 + 0.6(3x_1-4.64)^2 + 0.4(x_1-34.04)^2\right. \\ &+ \left. 0.4\frac{1}{9}\left[\frac{1}{2}(13.2-6x_1)^2 + 0.1(3x_1-59.52)^2 + 0.9(3x_1-0.72)^2\right]\right\} \end{aligned}$$

Finally, the expected value of stage $t=1$, state $I$ is, using the optimal decision at stage $t=2$ (see Eqs. (30) and (33))

$$\begin{aligned} V(1,I,x_1) &= \frac{1}{2}(x_1-4)^2 + 0.3(x_1-5)^2 + 0.7(x_1-2)^2 \\ &\quad + 0.3V(2,I,x_1) + 0.7V(2,II,x_1) \qquad (34) \end{aligned}$$

This gives the optimal $x_1 = 3.5620$, Eqs. (29) and (32) gives the optimal $x_2^{II} = 4.0540$ and $x_2^{III} = 4.6169$, respectively. Finally $x_2^{III}$ and Eq. (26) gives the optimal $x_3 = 4.3390$.

**Mathematical Programming Method:** An alternative solution to the stochastic problem is as follows.

The expected value is

$$
\begin{aligned}
&V(x_1, x_2^{II}, x_2^{III}, x_3) = \\
&\quad \frac{1}{2}(x_1-4)^2 + 0.3(x_1-5)^2 + 0.7(x_1-2)^2 \\
&+\ 0.3\left[\frac{1}{2}(x_2^{II}-x_1)^2 + 0.2(x_2^{II}-6)^2 + 0.5(x_2^{II}-5)^2 + 0.3(x_2^{II}-2)^2\right] \\
&+\ 0.7\left[\frac{1}{2}(x_2^{III}-x_1)^2 + 0.6(x_2^{III}-4)^2 + 0.4(x_2^{III}-7)^2\right] \\
&+\ 0.7\cdot 0.4\left[\frac{1}{2}(x_3-x_2^{III})^2 + 0.1(x_3-6)^2 + 0.9(x_3-4)^2\right]
\end{aligned} \tag{35}
$$

$\partial V/\partial x_1 = \partial V/\partial x_2^{II} = \partial V/\partial x_2^{III} = \partial V/\partial x_3 = 0$ gives

$$
\begin{array}{rrrrcl}
4.0x_1 & -0.3x_2^{II} & -0.70x_2^{III} & & = & 9.8 \\
-0.3x_1 & +0.9x_2^{II} & & & = & 2.58 \\
-0.7x_1 & & +2.38x_2^{III} & -0.28x_3 & = & 7.28 \\
 & & -0.28x_2^{III} & +0.84x_3 & = & 2.352
\end{array} \tag{36}
$$

respectively. The solutions are easily found as: $x_1 = 3.5620$, $x_2^{II} = 4.0540$, $x_2^{III} = 4.6169$ and $x_3 = 4.3390$. It is verified that these solutions satisfy Eq. (20). These are exactly the same as before.

### 4.2.3 Note (Choice of the Penalty Parameter)

It is noted that there are some discrepancies between these optimal solutions and the solutions obtained by the PHA, found in Table 7.

Table 8 compares the PHA solutions for various penalty $r$. In all cases $|\hat{X}^{\nu+1} - \hat{X}^{\nu}| < 5\cdot 10^{-8}$ at least. It is seen that for decreasing $r$ the solutions get closer to the exact ones. For $r = 0.001$ the solutions are almost exactly the same as those obtained by the classical methods.

# 5 Discussion

It has been demonstrated that the scenario technique is feasible and allows the modeling of stochastic elements. The PHA is explained and illustrated be nu-

**Table 8 The conditional expectation, $\hat{X}$, for the four decision stages.**

| $r$ | $\hat{X}_I$ | $\hat{X}_{II}$ | $\hat{X}_{III}$ | $\hat{X}_{IV}$ |
|---|---|---|---|---|
| 1.0 | 3.5304 | 4.0435 | 4.5557 | 4.2281 |
| 0.1 | 3.5572 | 4.0524 | 4.6079 | 4.3229 |
| 0.01 | 3.5615 | 4.0538 | 4.6156 | 4.3373 |
| 0.005 | 3.5618 | 4.0539 | 4.6165 | 4.3381 |
| 0.001 | 3.5620 | 4.0541 | 4.6168 | 4.3389 |

merical examples thus it is very easy to understand and implement. The only requirement is that the scenario subproblems, which are deterministic optimization problems, can be solved efficiently.

It will also be clear that the weak part of the PHA is the rate of convergence. The penalty parameter $r$ must be chosen with great care. This further indicates that the number of scenarios should be kept at a reasonable small number.

However, there are no other easy ways. Dynamic Programming and Mathematical Programming have been illustrated (Section 4.2.2). These techniques were easy to apply and the optimal solution were found directly. But the reason for this success was that the inequality constraints were not active. In general these techniques are difficult (or even impossible) to apply in case of constrained problems.

Conclusion: scenario analysis and the PHA may be an attractive alternative to the classical methods.

# Acknowledgement

This work has been supported by the District Heating Program of Energy Research Cooperation at the Nordic Council of Ministers and the Danish Energy Research Program - 1323/91-0010 which is hereby gratefully acknowledged.

# References


Bertsekas, D. P., (1987): *Dynamic Programming, Deterministic and Stochastic Models*, Prentice-Hall, Englewood Cliffs, NJ.

Dembo, R. S., (1991): "Scenario Optimization," *Annals of Operations Research*, 30:63–80.

Helgason, Th., and S. Wallace (1991): "Approximate Scenario Solutions in the Progressive Hedging Algorithm," *Annals of Operations Research*, 31:425–444.

Jensson, P., (1975): "Stochastic Programming, (in Danish: Stokastisk Programmering)," PhD thesis, IMSOR, The Technical University of Denmark.

Palsson, O. P., and H. F. Ravn (1993): "Stochastic Heat Storage Problem – Solved by the Progressive Hedging Algorithm," Technical Report, IMSOR, The Technical University of Denmark.

Robinson, S. M., (1991): "Extended Scenario Analysis," *Annals of Operations Research*, 31:385–398.

Rockafellar, R. T., and R. J.-B., Wets (1991): "Scenarios and Policy Aggregation in Optimization under Uncertainty," *Mathematics of Operations Research*, 16(1):119–147.

Wallace, S., and Th. Helgason (1991): "Structural Properties of the Progressive Hedging Algorithm," *Annals of Operations Research*, 31:445–456.

## OTHER IMSOR TECHNICAL REPORTS 1993

Melgaard, H., og H. Madsen:
CTLSM Continuous time linear stochastic modeling version 2.6 january 1993.
Nr. 1/1993. 53 pp.

Madsen, O.B.G., H. Ravn og J.M. Rygaard:
REBUS Program til dynamisk rutelægning for Københavns Brandvæsen.
Nr. 2/1993. 60 pp.

Palsson, O.P., H. Madsen og H.T. Søgaard:
Generalized predictive control for non-stationary systems.
Nr. 3/1993. 19 pp.

Palsson, O.P., H. Madsen og H.T. Søgaard:
Application of predictive control in district heating systems.
Nr. 4/1993. 20 pp.

Vidal, R.V.V., og L. Sørensen:
A critical assessment of the Danish water action plan.
Nr. 5/1993. 17 pp.

Vidal, R.V.V., og L. Sørensen:
Evaluation of integrated environmental models - some methodological issues.
Nr. 6/1993. 19 pp.

Kohl, N.:
Description of a bundle algorithm for non-smooth optimization.
Nr. 7/1993. 17pp.

Thyregod, P., og B.F. Nielsen:
Trends in marine losses and major casualties 1984 -1992.
Nr. 8/1993. 18 pp.

Iskander,M.G., og S.K.Jacobsen:
A computer-based modeling system for locational decision support.
Nr. 9/1993. 26 pp.

Madsen, H., K. Sejling, H.T. Søgaard, og O.P. Palsson:
On flow and supply temperature control in district heating systems.
Nr. 10/1993. 19 pp.

Palsson, O.P., og H.F. Ravn:
Scenario analysis and the progressive hedging algorithm - simple numerical examples.
Nr. 11/1993. 26 pp.