International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

Cost-effective Sulphur Emission Reduction under Uncertainty

Altman, A. and Ruszczynski, A.

IIASA Working Paper

November 1993

Working Paper

Cost-effective Sulphur Emission Reduction under Uncertainty

Anna Altman
Andrzej Ruszczyński

WP-93-62
November 1993

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Cost-effective Sulphur Emission Reduction under Uncertainty

Anna Altman
Andrzej Ruszczyński

WP-93-62
November 1993

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Cost-effective Sulphur Emission Reduction under Uncertainty

Anna Altman* Andrzej Ruszczyński†

November 1993

Abstract

The problem of reducing SO_2 emissions in Europe is considered. The costs of reduction are assumed to be uncertain and are modeled by a set of possible scenarios. A mean-variance model of the problem is formulated and a specialized computational procedure developed. The approach is applied to the transboundary air pollution model with real-world data.

1 Introduction

Reducing the pollution of the environment becomes one of the challenges of the present time in the industrial countries, and especially in Europe. It is commonly agreed that action should be undertaken to stop the growth of the emission and eventually achieve a substantial reduction of depositions.

One of the issues that attracts attention of researchers and decision-makers is emission of sulphur dioxide to the atmosphere, which has a damaging effect to the environment through acid rains. Clearly, this is an international problem, because the air pollution can move across the borders and damage the environment in other countries. Therefore it is necessary to look for a common European solution of this problem.

There are many ways to approach such problems. One would be to reduce the emissions uniformly over the continent (e.g. by 30%). This, however, may prove prohibitively expensive. On the other hand, it might be possible to achieve the same reduction of depositions by a non-uniform reduction of emissions in a more cost-effective way.

To investigate this possibility, a model has been developed at IIASA (see [AKS93] and references therein) to describe the relation between emissions and depositions at various regions. There is also a model of costs of reducing emissions at different locations. This allows to formulate an optimization problem of finding the cheapest way to meet environmental standards.

*Systems Research Institute, Warsaw, Poland.

†International Institute for Applied Systems Analysis, Laxenburg, Austria.

However, there are many uncertainties in the problem due to inaccurate or missing data, unknown future energy policies for the countries, etc. The purpose of this paper is to formalize the problem of reducing emissions in the presence of uncertainty, to develop a specialized solution procedure and to apply it to the real-world data available so far.

In Section 2 we recall the deterministic formulation of the problem following [AKS93]. In Section 3 we develop a mean-variance model for the problem under uncertainty. The uncertainty is modeled by a number of scenarios of future costs of emission reduction.

Section 4 is devoted to the development of a specialized computational procedure for solving the problem under consideration. The algorithm is a version of a primal-dual logarithmic barrier method.

In section 5 we report computational results obtained for six different scenarios proposed by the modelers. In the last section we present our conclusions and give propositions for the future work.

2 The deterministic problem

There are K countries in our model. For each country k we denote by x^k the level of emissions, which will be our decision variable. We assume that there are functions $f_k(x^k)$, that express the cost of reducing emissions to levels $x^k, k = 1, \ldots, K$. The functions are assumed to be convex and piecewise linear (CPL). Our objective is to minimize the total cost $\sum_{k=1}^{K} f_k(x^k)$ subject to some environmental constraints and additional policy restrictions.

The environmental constraints are described by a vector $b \in \mathcal{R}^m$ of maximal grid depositions, where m is the number of reception areas, and by an array $T \in \mathcal{R}^{m*K}$ defining dependence between emissions $x = (x^1, \ldots, x^K)$ and depositions $y = (y_1, \ldots, y_m)$ by $y = Tx$.

The whole problem can be formulated as

$$\begin{aligned} & \min \sum_{k=1}^{K} f_k(x^k) \\ \text{subject to} \quad & Tx \leq b, \\ & l \leq x \leq u. \end{aligned} \tag{1}$$

Vectors $l, u \in \mathcal{R}^m$ are policy constraints given by policy-makers.

Every function $f_k(x^k)$ is defined on an interval $[x^k_{J(k)}, x^k_0]$ which can be divided into subintervals $[x^k_j, x^k_{j-1}], j = 1, \ldots, J(k)$, such that $f_k(x^k)$ is linear in each of them (note that the break points $x^k_j, j = 0, \ldots, J(k)$ are numbered from the largest to the smallest one). The unit cost of emission reduction in the j-th interval will be denoted by $c_{j,k}$. Formally

$$c_{j,k} = \frac{f_k(x^k_j) - f_k(x^k_{j-1})}{x^k_{j-1} - x^k_j}.$$

By the convexity of f_k,

$$c_{1,k} \leq c_{2,k} \leq \ldots \leq c_{J(k),k}. \tag{2}$$

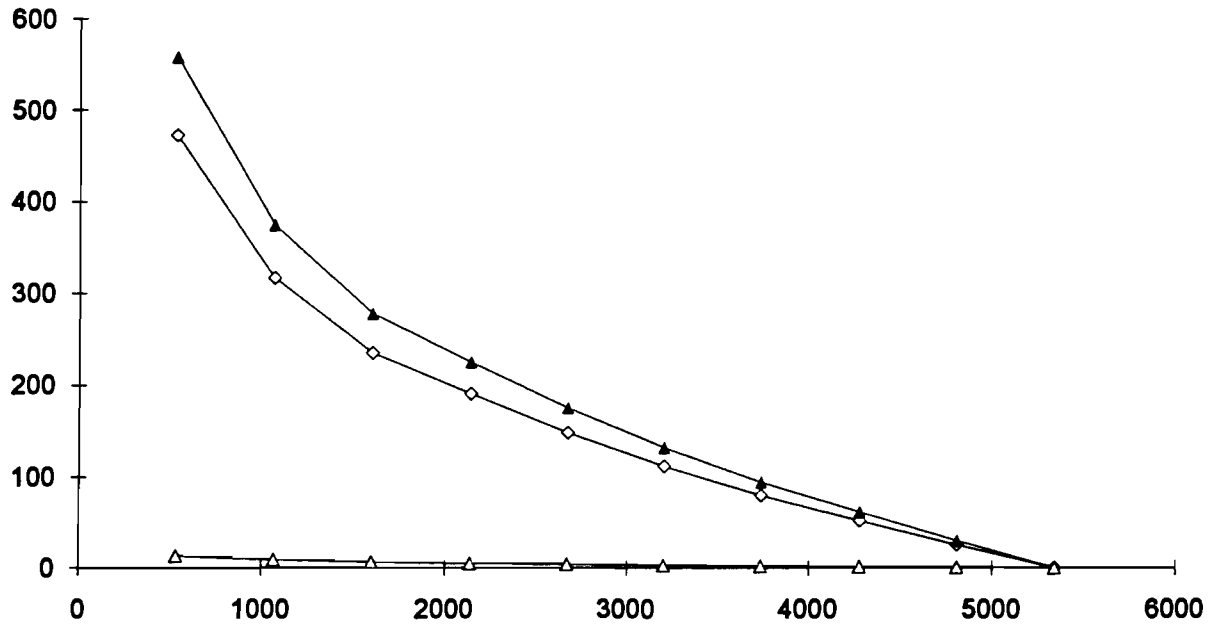

Figure 1: Examples of cost curves

This is illustrated in Fig. 1.

It is convenient to rewrite (1) as a linear problem by introducing (for each k) new variables $d_k(j), j = 1, \ldots, J(k)$ in such a way that

$$0 \leq d_k(j) \leq \bar{d}_k(j),$$

where $\bar{d}_k(j) = x^k_{j-1} - x^k_j$, and

$$d_k(j) = \begin{cases} \bar{d}_k(j) & \text{if} \quad x^k \leq x^k_j, \\ x^k_{j-1} - x^k & \text{if} \quad x^k \in (x^k_j, x^k_{j-1}), \\ 0 & \text{if} \quad x^k \geq x^k_{j-1}. \end{cases} \tag{3}$$

Then we can express emissions as

$$x^k = x^k_0 - \sum_{j=1}^{J(k)} d_k(j). \tag{4}$$

We can interpret variables $d_k(j)$ as successive reductions of the emission x^k starting from the maximum level x^k_0 and moving down through the break points of the function $f_k(x^k)$. Under (3), the total cost can be expressed as

$$\sum_{k=1}^{K} f_k(x^k) = \sum_{k=1}^{K} \sum_{j=1}^{J(k)} c_{j,k} d_k(j),$$

and it is a linear function.

Problem (1) can be reformulated as a linear programming problem

$$\begin{aligned} & \min \sum_{k=1}^{K} \sum_{j=1}^{J(k)} c_{j,k} d_k(j), \\ \text{subject to} \quad & Tx \leq b, \\ & x^k = x^k_0 - \sum_{j=1}^{J(k)} d_k(j), \\ & l \leq x \leq u, \\ & 0 \leq d_k(j) \leq \bar{d}_k(j). \end{aligned} \tag{5}$$

It is interesting to observe that we need not include condition (3) explicitly to problem statement (5). Condition (2) with strict inequalities immediately implies that the solution of (5) must satisfy (3).

We can solve (5) using any linear programming package.

3 The mean-variance model

Unfortunately, the costs of emission control are not deterministic quantities. There are many possible scenarios of energy production, consumption, fuel characteristics and

installed emission control measures. Therefore our problem is a decision problem with uncertainty.

There are many ways to formalize such decision problems. We can, e.g., use the worst-case approach and require the decision to be best for the worst possible conditions. This usually leads to very conservative and expensive solutions.

An approach that found many successful applications is to model uncertain quantities by random variables. Then we can use various concepts of the theory of probability to express our objectives and constraints. This leads to *stochastic programming models.*

In our case only the costs are uncertain, and that can be modeled by assuming that the unit costs $c_{j,k}$ are random (but still satisfy (2)). To be even more specific we shall restrict our considerations to the case of finitely many *scenarios* $s = 1, \ldots, S$. Each scenario s has some probability $p(s)$, such that $\sum_{s=1}^{S} p(s) = 1$, and is characterized by a collection of unit costs in subintervals

$$c_{1,k,s} \leq c_{2,k,s} \leq \ldots \leq c_{J(k),k,s} \; k = 1, \ldots, K.$$

Nevertheless, we still have many possibilities of expressing our objective. The simplest solution would be to minimize the expected cost

$$E = \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} c_{j,k,s} d_k(j). \tag{6}$$

This is equivalent to solving the problem with one average scenario with unit costs

$$\bar{c}_{j,k} = \sum_{s=1}^{S} p(s) c_{j,k,s}.$$

A significant drawback associated with the expected value approach is that it essentially ignores uncertainty of the cost.

Another possibility would be to define a nonlinear utility function and to optimize its expected value. Unfortunately we do not have a clear idea of such a function in our case.

Therefore we decide to use the mean-variance approach to our decision problem. With such an approach the quality is measured by two outcomes: the mean value (6) and the (weighted) variance

$$V = \sum_{s=1}^{S} \sum_{k=1}^{K} w_k^2 [\sum_{j=1}^{J(k)} c_{j,k,s} d_k(j) - e_k]^2, \tag{7}$$

where

$$e_k = \sum_{s=1}^{S} p(s) \sum_{j=1}^{j(k)} c_{j,k,s} d_k(j)$$

denotes the expected value of the cost for the k-th country and w_k's are some weighting coefficients.

The variance will be used to measure the risk associated with a decision. The weighting coefficients w_k can be used to bring the variance components associated

with different countries to some common measure. In particular we could make w_k inversely proportional to the GDP of the k-th country, which would measure the risk relative to the economic strength of the country rather than in absolute terms.

Both outcomes are used to form a composite objective

$$G = E + \alpha V, \tag{8}$$

where $\alpha > 0$ is a user-defined parameter.

The main idea of the mean-variance model is to replace the objective of (5) by the composite objective (8). The constraints remain unchanged.

By varying α one can generate a family of solutions with different trade-offs between the expected cost and the variance of the cost.

Let us formalize the mean-variance approach in our case. We introduce the variables

$$q_{k,s} = \sum_{j=1}^{J(k)} c_{j,k,s} d_k(j). \tag{9}$$

Then we can rewrite (6) and (7) as

$$E = \sum_{s=1}^{S} p(s) \sum_{s=1}^{K} q_{k,s},$$

$$V = \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} [w_k(q_{k,s} - e_k)]^2.$$

The mean-variance problem can be now formulated as follows

$$\min \sum_{s=1}^{S} p(s) \sum_{s=1}^{K} q_{k,s} + \alpha \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} [w_k(q_{k,s} - e_k)]^2 \tag{10}$$

subject to

$$Tx \le b,$$

$$l \le x \le u,$$

$$x^k = x_0^k - \sum_{j=1}^{J(k)} d_k(j), \text{ for every } k = 1, \ldots, K$$

$$q_{k,s} = \sum_{j=1}^{J(k)} c_{j,k,s} d_k(j) \text{ for every } k = 1, \ldots, K \text{ and } s = 1, \ldots, S,$$

$$0 \le d_k(j) \le \bar{d}_k(j) \text{ for every } k = 1, \ldots, K \text{ and } j = 1, \ldots, J(k).$$

To allow application of efficient computational techniques for solving quadratic problems we shall transform (10) to a problem with non-negative variables and with a separable quadratic part of the objective. First, we split $q_{k,s} - e_k$ into the positive and the negative parts, defining new variables $q_{k,s}^+$ and $q_{k,s}^-$ by

$$q_{k,s}^+ - q_{k,s}^- = q_{k,s} - e_k, \; q_{k,s}^+, q_{k,s}^- \ge 0. \tag{11}$$

The expected value and the variance have now the forms

$$E = \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} (q_{k,s}^{+} - q_{k,s}^{-} + e_k),$$

$$V = \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} [w_k (q_{k,s}^{+} - q_{k,s}^{-})]^2.$$

Unfortunately, the quadratic part of the objective is not separable now. But it can be proved (c.f. [Car92]) that the solution of our problem does not change if we replace (V) by a separable function:

$$V' = \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} w_k^2 [(q_{k,s}^{+})^2 + (q_{k,s}^{-})^2],$$

because at the solution only one of $q_{k,l}^{+}, q_{k,l}^{-}$ will be different from 0. Finally we arrive to the following problem:

$$\min \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} (q_{k,s}^{+} - q_{k,s}^{-} + e_k) + \alpha \sum_{s=1}^{S} p(s) \sum_{k=1}^{K} w_k^2 [(q_{k,s}^{+})^2 + (q_{k,s}^{-})^2] \qquad (12)$$

$$\begin{aligned} \text{subject to} \quad & Tx \leq b, \\ & l \leq x \leq u, \\ & q_{k,s}^{+} - q_{k,s}^{-} + e_k = \sum_{j=1}^{J(k)} c_{j,k,s} d_k(j), \\ & e_k, q_{k,s}^{+}, q_{k,s}^{-} \geq 0, \\ & 0 \leq d_k(j) \leq \bar{d}_k(j). \end{aligned}$$

The total number of constraints in this problem is equal to $m + K + S * K$. The total number of variables is equal to $K + \sum_{k=1}^{K} J(k) + 2 * S * K + K$.

4 Solution techniques

The problem (12), as we mentioned before, is a quadratic programming (QP) problem. We can rewrite it in the standard way. First, we define some constants and variables. The box constraints

$$l \leq x \leq u$$

for the problem (1) will be shifted to obtain lower bounds equal to zero. We define new variables $x' = x - l$, for which we have

$$0 \leq x' \leq u - l.$$

The entire vector of unknowns ξ is built from subvectors used before, i.e.

$$\xi = (x', t, d, q^{+}, q^{-}, \bar{e}).$$

There are only equality constraints in the standard form of QP problem. To satisfy this we introduce to the constraints $Tx \leq b$ a vector of slacks t and we obtain

$$Tx' + t = b',$$

where $t \in \mathcal{R}^m$ and $b' = b - Tl$.

For simplifying formulas we define a constant

$$SJ = \sum_{k=1}^{K} J(k).$$

A vector d is constructed from variables $d_k(j)$ defined in (3),

$$d = (d_1(1), \ldots, d_1(J(1)), \ldots, d_K(1), \ldots, d_K(J(K))) \in \mathcal{R}^{SJ}.$$

Next vectors q^+ and q^- are constructed from variables $q^+_{k,s}$ $q^-_{k,s}$ defined in (11),

$$q^+ = (q^+_{1,1}, \ldots, q^+_{K,1}, \ldots, q^+_{1,S}, \ldots, q^+_{K,S}) \in \mathcal{R}^{K*S},$$

$$q^- = (q^-_{1,1}, \ldots, q^-_{K,1}, \ldots, q^-_{1,S}, \ldots, q^-_{K,S}) \in \mathcal{R}^{K*S}.$$

A vector $\bar{e}$ is a vector of the expected values of costs, i.e.

$$\bar{e} = (e_1, \ldots, e_K) \in \mathcal{R}^K,$$

where e_k are defined in (7).

The dimension of the vector ξ will be denoted by n,

$$n = 2 * K + m + SJ + 2 * K * S.$$

The entire vector of the linear part of the objective is defined as follows:

$$g = (0_K, 0_m, 0_{SJ}, p_1 \mathbf{1}_K, \ldots, p_S \mathbf{1}_K, -p_1 \mathbf{1}_K, \ldots, -p_S \mathbf{1}_K, \mathbf{1}_K) \in \mathcal{R}^n.$$

We write 0 with a subscript for the vector of zeros and the subscript denotes dimension of the vector. In the same way we denote by $\mathbf{1}$ with a subscript the vector of ones and the subscript denotes the dimension.

The quadratic part of the objective is nonzero for terms including q^+ and q^- only. Formally we define the the quadratic matrix $Q \in \mathcal{R}^{n*n}$ in such way:

$$Q_{i,j} = \begin{cases} 2\alpha * p(s) w_k^2 & \text{if } i = j \text{ and the } i\text{-th component of } \xi \text{ is a } q^+- \text{ or } q^-\text{-} \\ & \text{variable and } s \text{ and } k \text{ denote the adequate number} \\ & \text{of scenario and country;} \\ 0 & \text{otherwise.} \end{cases}$$

Now we can write the full constraint matrix A,

$$A = \begin{pmatrix} T & I_m & 0 & 0 & 0 & 0 \\ I_K & 0 & E & 0 & 0 & 0 \\ 0 & 0 & C & -I_{K*S} & I_{K*S} & J \end{pmatrix}.$$

The submatrix T is the constraint matrix of the problem (1). The letter I with subscript denotes identity matrix, where subscript denotes its dimension. The matrix J is defined as

$$J = \begin{bmatrix} I_K \\ \vdots \\ I_K \end{bmatrix},$$

where submatrix I_K is repeated S times.

We can combine equations (9) and (11) and write dependencies between $q^+, q^-, \bar{e}$ and d in a matrix form, i.e.

$$q^+ - q^- = Cd - J\bar{e},$$

where

$$C = \begin{pmatrix} c_{1,1,1} \dots c_{J(1),1,1} & & \\ & \ddots & \\ & & c_{1,K,1} \dots c_{J(K),K,1} \\ \vdots & & \\ c_{1,1,S} \dots c_{J(1),1,S} & & \\ & \ddots & \\ & & c_{1,K,S} \dots c_{J(K),K,S} \end{pmatrix} \in \mathcal{R}^{(K*S)*SJ}.$$

The matrix $E \in \mathcal{R}^{K*SJ}$ is used to write in matrix form equations (4), i.e.

$$E = \begin{pmatrix} \overbrace{1 \dots 1}^{J(1)} & & 0 \\ & \ddots & \\ & & \overbrace{1 \dots 1}^{J(K)} \\ 0 & & \end{pmatrix}$$

and

$$x' + Ed = x'_0,$$

where $x'_0 = x_0 - l$.

The right hand side vector r is defined as follows

$$r = (b', x'_0, 0_{S*K}) \in \mathcal{R}^{m+K+S*K}.$$

Let us consider the vector ρ of upper bounds (all lower bounds are equal to zero). Only upper bounds for x' and d parts are less than infinity,

$$\rho_i = \begin{cases} u_i - l_i & \text{if the } i\text{-th component of } \xi \text{ is } x' \text{ variable,} \\ \bar{d}_k(j) & \text{if the } i\text{-th component of } \xi \text{ is } d \text{ variable,} \\ +\infty & \text{otherwise,} \end{cases}$$

where

$$i_1 = i - (K + m),$$

k is such that

$$\sum_{\kappa=0}^{k-1} J(\kappa) < i_1 \leq \sum_{\kappa=0}^{k} J(\kappa),$$

$$j = i_1 - \sum_{\kappa=0}^{k-1} J(\kappa).$$

We have not yet defined the number $J(0)$, it will be equal to zero.

Finally, we can formulate the complete QP problem in the standard form:

$$\begin{aligned} &\text{minimize } g^T\xi + \tfrac{1}{2}\xi^T Q\xi, \\ \text{subject to } &A\xi = r, \\ &\xi + \sigma = \rho, \\ &\xi, \sigma \geq 0. \end{aligned} \tag{13}$$

The dual of (13) is given by

$$\begin{aligned} &\text{maximize } r^T\lambda - \rho^T w - \frac{1}{2}\xi^T Q\xi, \\ \text{subject to } &A^T\lambda + z - w - Q\xi = g, \\ &\xi, z, w \geq 0, \end{aligned} \tag{14}$$

where $\lambda \in \mathcal{R}^M$ and $\xi, z, w \in \mathcal{R}^n$; by M we shall denote the number of constraints equal to $2 * K + m + SJ + 2 * S * K$.

For solving (13) and (14) we shall use the logarithmic barrier method. In such a method we augment the objective by adding to it a logarithmic barrier term, which yields

$$\begin{aligned} \text{minimize} \quad & g^T\xi + \frac{1}{2}\xi^T Q\xi - \mu\sum_{j=1}^{n}(\ln\xi_j + \ln\sigma_j), \\ \text{subject to} \quad & A\xi = r, \\ & \xi + \sigma = \rho, \\ & \xi, \sigma > 0 \end{aligned} \tag{15}$$

and an analogue of (14):

$$\begin{aligned} \text{maximize} \quad & r^T\lambda - \rho^T w - \frac{1}{2}\xi^T Q\xi + \mu\sum_{j=1}^{n}(\ln z_j + \ln w_j), \\ \text{subject to} \quad & A^T\lambda + z - w - Q\xi = g, \\ & \xi \geq 0, z, w > 0. \end{aligned} \tag{16}$$

The first order optimality conditions for (15) and (16) are

$$\begin{aligned} A\xi &= r, \\ \xi + \sigma &= \rho, \\ A^T\lambda + z - w - Q\xi &= g, \\ \Xi z &= \mu e, \\ \Sigma w &= \mu e, \\ \xi, \sigma, z, w &> 0, \end{aligned} \tag{17}$$

where Ξ, Σ, Z and W are diagonal matrices with the diagonal elements ξ_j, σ_j, z_j and w_j, respectively and $e \in \mathcal{R}^n$ is the vector of ones.

For solving (17) we use a quadratic analogue of Mehrotra's higher order method [Meh91]. This method computes a Taylor approximation of the optimal trajectory that starts at a given point and leads to the optimum of (13) and (14).

The Mehrotra method belongs to the class of continuation methods (c.f. [OrR70]). These methods are used for solving systems of nonlinear equations. In continuation methods the family of parametrized problems is considered. One element of the family is exactly our original problem (e.g. for parameter equal to 0). Other problems have a perturbated right hand side vector. For at least one value of the parameter we know the solution of the problem (e.g. for parameter equal to 1). Coming iteratively from the problem with known solution to the original one (i.e. changing the right hand side vector) we can find better approximation of the solution.

Let $\xi^0, \sigma^0, z^0, w^0 > 0$ and λ^0 be the current estimate of the solution of (13) and (14). Then

$$\begin{aligned} \pi_r &= A\xi^0 - r, \\ \pi_\rho &= \xi^0 + \sigma^0 - \rho, \\ \text{and } \pi_g &= A^T\lambda^0 + z^0 - w^0 - Q\xi^0 - g, \end{aligned}$$

are the resulting residuals in the primal and dual constraints. Next, we consider the parametric system of equations

$$\begin{aligned} A(\gamma) &= r + \phi_1(\gamma)\pi_r, \\ \xi(\gamma) + \sigma(\gamma) &= \rho + \phi_1(\gamma)\pi_\rho, \\ A^T\lambda(\gamma) + z(\gamma) - w(\gamma) - Q\xi(\gamma) &= g + \phi_1(\gamma)\pi_g, \\ \Xi(\gamma)z(\gamma) &= \phi_1(\gamma)\Xi^0 z^0 + \phi_2(\gamma)\mu e, \\ \Sigma(\gamma)w(\gamma) &= \phi_1(\gamma)\Sigma^0 w^0 + \phi_2(\gamma)\mu e, \\ \xi(\gamma), \sigma(\gamma), z(\gamma), w(\gamma) &> 0, \end{aligned} \tag{18}$$

where ϕ_1 and ϕ_2 are nonnegative functions determined on interval $[0,1]$ such that $\phi_2(0) = \phi_2(1) = 0, \phi_1(0) = 0, \phi_1(1) = 1$ and $\phi_1(\gamma) \in (0,1)$ for $\gamma \in (0,1)$. The system (18) for $\gamma = 1$ differs from the system (17) by the right hand side. For $\gamma = 0$ both system are identical.

Let $\Gamma(\gamma) = (\xi(\gamma), \sigma(\gamma), \lambda(\gamma), z(\gamma), w(\gamma))$ be the solution of (18) for given parameter γ. We let $\Gamma(1) = (\xi(1), \sigma(1), \lambda(1), z(1), w(1)) = (\xi^0, \sigma^0, \lambda^0, z^0, w^0)$, so (18) is satisfied for $\gamma = 1$. Thus $\Gamma(0)$ represents the solution of (18) for $\gamma = 0$, and solution of (17), and therefore approximated solutions of (13) and (14). The key point of Mehrotra's approach is to use local higher order information available at point $\Gamma(1)$ to construct a direction that well approximates the first point of trajectory Γ. In our implementation $f(\gamma) = \gamma(1-\gamma)^2$ and $g(\gamma) = \gamma$, which refers to Mehrotra's Algorithm II.

Since $\Gamma(\gamma)$ is a solution of (18) for a given γ, the appropriate higher order terms of Taylor polynomial approximation of correction $(\Delta\xi, \Delta\sigma, \Delta\lambda, \Delta z, \Delta w)$ to the current estimate $(\xi, \sigma, \lambda, z, w)$ result from the recursive differentiation of (18). The i-th order

term of the correction vector can be obtained from

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ -Q & A^T & 0 & I & -I \\ I & 0 & I & 0 & 0 \\ Z^0 & 0 & 0 & 0 & 0 \\ 0 & 0 & W^0 & 0 & \Sigma^0 \end{bmatrix} \begin{bmatrix} \Delta\xi^{(i)} \\ \Delta\lambda^{(i)} \\ \Delta\sigma^{(i)} \\ \Delta z^{(i)} \\ \Delta w^{(i)} \end{bmatrix} = \begin{bmatrix} \eta_1^{(i)} \\ \eta_2^{(i)} \\ \eta_3^{(i)} \\ \eta_4^{(i)} \\ \eta_5^{(i)} \end{bmatrix}, \tag{19}$$

where

$$\begin{aligned} \Delta\xi^{(i)} &= \frac{\xi^{(i)}}{i!}, \\ \Delta\sigma^{(i)} &= \frac{\sigma^{(i)}}{i!}, \\ \Delta\lambda^{(i)} &= \frac{\lambda^{(i)}}{i!}, \\ \Delta z^{(i)} &= \frac{z^{(i)}}{i!}, \\ \Delta w^{(i)} &= \frac{w^{(i)}}{i!} \end{aligned}$$

and

$$\begin{aligned} \eta_1^{(i)} &= \frac{1}{i!}\phi_1^{(i)}(1)\pi_r, \\ \eta_2^{(i)} &= \frac{1}{i!}\phi_1^{(i)}(1)\pi_g, \\ \eta_3^{(i)} &= \frac{1}{i!}\phi_1^{(i)}(1)\pi_\rho, \\ \eta_4^{(i)} &= \frac{1}{i!}[\phi_1^{(i)}(1)^0 z^0 + \phi_2^{(i)}(1)\mu e] - \sum_{i=1}^{i-1} \Delta\Xi^{(i)}\Delta z^{(i-l)}, \\ \eta_5^{(i)} &= \frac{1}{i!}[\phi_1^{(i)}(1)\Sigma^0 w^0 + \phi_2^{(i)}(1)\mu e] - \sum_{l=1}^{i-1} \Delta\Sigma^{(l)}\Delta w^{(i-l)}. \end{aligned} \tag{20}$$

For every i the matrix involved in all these linear systems is the same. The factorization of (20) is to be computed only once. For the linear case we can compute, c.f. [AlG93a], [AlG93b], the search directions in the primal and dual spaces as

$$\begin{aligned} d_\xi &= -\sum_{i=1}^{l_p}(-\gamma_p)^i \Delta\xi^{(i)}(1), \\ d_\sigma &= -\sum_{i=1}^{l_p}(-\gamma_p)^i \Delta\sigma^{(i)}(1), \\ d_\lambda &= -\sum_{i=1}^{l_d}(-\gamma_d)^i \Delta\lambda^{(i)}(1), \end{aligned} \tag{21}$$

$$d_z = -\sum_{i=1}^{l_d}(-\gamma_d)^i \Delta z^{(i)}(1),$$
$$d_w = -\sum_{i=1}^{l_d}(-\gamma_d)^i \Delta w^{(i)}(1),$$

where l_p and l_d are orders of Taylor polynomials in the primal and dual spaces, respectively. The parameters γ_p and γ_d in (21) are the largest numbers in $[0,1]$ for which

$$\xi - d_\xi \geq 0,$$
$$\sigma - d_\sigma \geq 0,$$
$$z - d_z \geq 0,$$
$$w - d_w \geq 0.$$

Unfortunately, for quadratic problems the use of (21) can cause the loss of feasibility in dual constraints. If we have a feasible solution $(\xi, \sigma, \lambda, z, w)$ then the new point $(\xi - d_\xi, \sigma - d_\sigma, \lambda - d_\lambda, z - d_z, w - d_w)$ need not be dual feasible as in the linear case. In a quadratic problem the primal variable ξ appears also in dual constraints, and then

$$A^T(\lambda - d_\lambda) + (z - d_z) - (w - d_w) - Q(\xi - d_\xi) = c + Qd_\xi.$$

To overcome this disadvantage we can follow [VaC93] and use $\gamma_p = \gamma_d$ and $l_p = l_d$ for primal and dual space. But it implies that $\gamma = \min(\gamma_p, \gamma_d)$ and $l = \min(l_p, l_d)$. This approach slows the method down.

In our computations we decided to use different γ and l for the primal and the dual space, similarly to the linear case. Our experiments show that with formulas (21) the algorithm works much faster.

After computing the search directions we define step factors f_p and f_d as in [Meh91] and we define new approximations of optimal point

$$\xi := \xi - f_p d_\xi,$$
$$\sigma := \sigma - f_p d_\sigma,$$
$$\lambda := \lambda - f_d d_\lambda,$$
$$z := z - f_d d_z,$$
$$w := w - f_d d_w.$$

Elimination of $\Delta\sigma^{(j)}, \Delta z^{(j)}$ and $\Delta w^{(j)}$ reduces (19) to

$$H \cdot \begin{bmatrix} \Delta\xi^{(i)} \\ \Delta\lambda^{(i)} \end{bmatrix} = \begin{bmatrix} -\Theta^{-1} & A^T \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta\xi^{(i)} \\ \Delta\lambda^{(i)} \end{bmatrix} = \begin{bmatrix} h^{(i)} \\ \eta_1^{(i)} \end{bmatrix}, \tag{22}$$

where

$$h^{(i)} = \eta_2^{(i)} - (\Xi^0)^{-1}\eta_4^{(i)} + (\Sigma^0)^{-1}(\eta_5^{(i)} - W^0\eta_3^{(i)})$$

and

$$\Theta = (Q + (\Xi^0)^{-1}Z^0 + (\Sigma^0)^{-1}W^0)^{-1}, \tag{23}$$

where $\eta_1^{(i)}, \eta_2^{(i)}, \eta_3^{(i)}, \eta_4^{(i)}, \eta_5^{(i)}$ are defined by (20).

Further, we reduce (22) to the normal equation form

$$(A\Theta A^T)\Delta\lambda^{(i)} = A\Theta h^i, \tag{24}$$

and we compute sparse Cholesky factorization of the positive definite matrix $A\Theta A^T$.

Now it is clear why it is convenient to have a diagonal matrix Q. With a diagonal Q all terms in (23) are diagonal and computing Θ is very easy. Furthermore the sparsity pattern for $A\Theta A^T$is the same for every Θ and we can use the same techniques of symbolic factorization as in the the linear case [DER86]. Hence the method used for solving (24) is the same as in linear case. We can also use the same techniques for finding starting points.

5 Numerical results

The methodology described in the paper has been applied to the transboundary air pollution (TAP) model developed by [AKS93].

There were 36 countries in the model ($K = 36$) and for each country we had 6 scenarios of cost curves ($S = 6$). Environmental constraints were imposed in 169 reception areas ($m = 169$). Together the problem (12) had 36 $x'-$variables, 169 $t-$variables, 236 $d-$variables, 216 both q^+- and q^--variables, and 36 $\bar{e}$ variables, the total number of variables was 909. There are 169 constraints connected with the vector b', 36 with vector x'_0 and 216 others, together we had 421 equality constraints.

The solution of the problem for various values of the parameter α in (8) is presented in Tables 1 and 2. We used rather large values of α to account for the difference of orders of magnitude between the unscaled expected cost and the normalized variance.

We see from the results that the solution does not significantly change, when more stress is put on the variance. Presumably, the scenarios (cost curves) have substantial similarities in their qualititative behaviour and differ mainly in the scale of costs rather than in the shape of the function.

It is also interesting to note that the numerical method suggested in the paper proved efficient for this class of problems. In Table 3 we summarize its performance for different values of the parameter α. The method never failed, although the required precision was very high (10^{-6}). Both the number of iterations and the execution time are rather low. All computations were done on a SUN Sparc 2 workstation.

6 Conclusions

The methodology used in this paper can be used not only for this specific transboundary air pollution problem, but also for a wide class of problems with uncertainty in the costs. Using a mean-variance model we can properly treat a non-deterministic problem with small data collections, when more sophisticated stochastic methods are not useful. The mean-variance model has a nice interpretation in terms of risk. Furthermore this method leads to quadratic programming problems, well described in the literature.

	$\alpha = 0$	$\alpha = 10^4$	$\alpha = 10^5$	$\alpha = 10^6$
Albania	0.63704E+02	0.63608E+02	0.63587E+02	0.63585E+02
Austria	0.38888E+02	0.38888E+02	0.38888E+02	0.38888E+02
Belgium	0.57276E+02	0.43902E+02	0.41583E+02	0.41085E+02
Bulgaria	0.25981E+03	0.25981E+03	0.25981E+03	0.25981E+03
Czechoslov.	0.38133E+03	0.41695E+03	0.41957E+03	0.41995E+03
Denmark	0.24088E+02	0.24088E+02	0.24088E+02	0.24088E+02
Finland	0.58153E+02	0.58153E+02	0.58153E+02	0.58153E+02
France	0.42800E+03	0.36986E+03	0.30413E+03	0.30177E+03
Germany-W.	0.25963E+03	0.25963E+03	0.25963E+03	0.25963E+03
Germany-E.	0.17000E+03	0.17000E+03	0.17000E+03	0.17000E+03
Greece	0.23545E+03	0.16482E+03	0.15123E+03	0.14895E+03
Hungary	0.48045E+03	0.46555E+03	0.46445E+03	0.46433E+03
Ireland	0.27268E+02	0.47992E+02	0.47757E+02	0.47958E+02
Italy	0.52973E+03	0.52339E+03	0.52570E+03	0.52576E+03
Luxembourg	0.34023E+01	0.48174E+01	0.49430E+01	0.49483E+01
Netherlands	0.53130E+02	0.53130E+02	0.53130E+02	0.53130E+02
Norway	0.15823E+02	0.15823E+02	0.15823E+02	0.15823E+02
Poland	0.41634E+03	0.57392E+03	0.60654E+03	0.60951E+03
Portugal	0.92994E+02	0.80659E+02	0.67832E+02	0.66636E+02
Romania	0.54639E+03	0.52621E+03	0.52510E+03	0.52501E+03
Spain	0.74977E+03	0.75060E+03	0.75156E+03	0.75163E+03
Sweden	0.39099E+02	0.39099E+02	0.39099E+02	0.39099E+02
Switzerland	0.23892E+02	0.21237E+02	0.21237E+02	0.21237E+02
Turkey	0.83348E+03	0.83348E+03	0.83348E+03	0.83032E+03
United-King.	0.73334E+03	0.73310E+03	0.74634E+03	0.74678E+03
Yugoslavia	0.27231E+03	0.42090E+03	0.42714E+03	0.42763E+03
Baltic-Sea	0.14652E+02	0.14652E+02	0.14652E+02	0.14652E+02
North-Sea	0.34652E+02	0.34652E+02	0.34652E+02	0.34652E+02
Atlan.-Ocean	0.63320E+02	0.63320E+02	0.63320E+02	0.63320E+02
Kola-Karelia	0.12664E+03	0.12664E+03	0.12664E+03	0.12664E+03
S.Petersburg	0.15334E+03	0.97485E+02	0.10236E+03	0.10269E+03
Baltic-reg.	0.14082E+03	0.14082E+03	0.14082E+03	0.14082E+03
Byelorussia	0.13500E+03	0.10800E+03	0.11028E+03	0.11132E+03
Ukraine	0.84778E+03	0.75583E+03	0.75583E+03	0.75583E+03
Moldavia	0.11563E+03	0.11563E+03	0.11563E+03	0.11563E+03
Rem.Eur.CIS	0.18642E+04	0.12812E+04	0.91461E+03	0.87483E+03

Table 1: Emissions of SO_2 (in kT of S)

	$\alpha = 0$	$\alpha = 10^4$	$\alpha = 10^5$	$\alpha = 10^6$
Expected Cost	0.3870E+04	0.4334E+04	0.4596E+04	0.4628E+04
Weighted Variance	5.4678	5.2158	5.2044	5.2033

Table 2: The values of the objectives

	$\alpha = 0$	$\alpha = 10^2$	$\alpha = 10^4$	$\alpha = 10^5$	$\alpha = 10^6$	$\alpha = 10^9$
Iterations	20	18	21	24	30	32
Time (s)	35.09	32.66	36.58	41.66	48.37	51.53

Table 3: Performance of the interior point method

The numerical method used here, the quadratic version of an interior point method HOPDM [AlG93a], is safe and works quite well. In this method the most recent computational techniques are implemented, such as symbolic Cholesky factorization of sparse matrices, splitting dense columns, minimum degree ordering and many others. The application problem described in the paper motivated the development of the general quadratic programming solver for sparse and large scale problems.

Unfortunately, not all scenarios available so far are complete and their set does not seem to be variable enough. Having a richer collection of data we might obtain more interesting results with greater role played by risk. Still, it is also possible that very tight deposition constraints do not leave much room for stochastic optimization in this case.

In the current model only costs are uncertain. But depositions depend on many uncertain factors, especially on weather. In a more sophisticated approach one might incorporate uncertainty into the constraints, which requires further research and cooperation of experts in both the environmental problems and optimization.

References

[AlG93a] A. Altman and J. Gondzio, *An efficient implementation of a higher order primal-dual interior point method for large sparse linear programs*, Archiv of Control Sciences (1993). To appear.

[AlG93b] ———, *HOPDM - a higher order primal-dual method for large scale linear programming*, European J. Oper. Res. **66**, no. 1 (1993) 159–160.

[AKS93] M. Amann, G. Klaassen and W. Schöpp, *Closing the gap between the 1990 deposition and the critical sulfur deposition values*, Tech. report, International Institute for Applied Systems Analysis, 2361 Laxenburg, Austria, 1993.

[Car92] T. J. Carpenter, *Practical interior point methods for quadratic programming*, Ph.D. thesis, Princeton University, 1992.

[DER86] I. S. Duff, A. M. Erisman and J. K. Reid, *Direct Methods for Sparse Matrices*, Claredon Press, Oxford, 1986.

[Meh91] S. Mehrotra, *Higher order methods and their performance*, Tech. Report 90-16R1, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, 1991.

[OrR70] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, London, 1970.

[VaC93] R. J. Vanderbei and T. J. Carpenter, *Symmetric indefinite systems for interior point methods*, Math. Programming **58** (1993) 1–32.