McMaster University

Advanced Optimization Laboratory

Title:

Second-Order Cone Optimization Approach to
Groundwater Quantity Management

Authors:

Ndambuki J. M., Terlaky T., Stroet C. B. M.
and Veling E. J. M.

AdvOl-Report No. 2005/9

June 2005, Hamilton, Ontario, Canada

Second-Order Cone Optimization Approach to Groundwater Quantity Management

Ndambuki J. M.[*], Terlaky T.[†], Stroet C. B. M.[‡] and Veling E. J. M.[§]

14-05-2005

Abstract

As demand for potable water increases due to either increase in world population, contamination of surface water bodies or both, groundwater aquifers are constantly being viewed as the ultimate alternative source for additional quantities. Moreover, as these aquifers get exploited, it is becoming mandatory to incoorporate management schemes to ensure sustainable exploitation of such sources. However, because of the complexity of the earth material that forms such aquifers, it has become evident that deterministic management approaches are not suitable for designing such management schemes. This has resulted in researchers developing stochastic management methodologies which recognize the fact that measurement data used to characterize the groundwater aquifers is not only scarce, but also uncertain due to, for example measurement errors. The most popular approach of addressing this uncertainty has been through the Monte Carlo approach and its variants. However, Monte Carlo approach is CPU intensive and therefore only a few realizations can be considered. In this paper, we introduce a new optimization technique which explicitly takes into account the uncertainty in the hydraulic conductivity of a groundwater aquifer. In this approach, we transform the uncertain groundwater management problem into a Second - Order Cone Optimization problem which is then solved. Furthermore, we extent the methodology to address both a single objective and multi-objective groundwater quantity optimization in the presence of uncertainty. Results from an hypothetical example demonstrate that solutions obtained through this approach are robust, meaning that small perturbations in the hydraulic conductivity values will not affect the optimal solutions computed.

[*]Corresponding author e-mail: jmndambuki@yahoo.co.uk, Dpartment of Civil and Structural Engineering, Moi University, P.O. Box 3900, ELDORET, KENYA

[†]Advanced Optimization Lab, Department of Computing and Software, Mcmaster University, Hamilton, ON, Canada

[‡]Delft University of Technology, Faculty of Civil Engineering and Geosciences, Stevinweg 1, NL-2628 CN Delft, The Netherlands

[§]Delft University of Technology, Faculty of Civil Engineering and Geosciences, Stevinweg 1, NL-2628 CN Delft, The Netherlands

1 Introduction

The desire for sustainable development is increasingly becoming the norm rather than the exception in water resources exploitation in general, and groundwater resources in particular. However, because of the uncertainty involved in specifying the groundwater aquifer flow parameters, optimal management schemes designed within such an environment are inherently uncertain, hence sensitive to perturbations in the input values of the flow parameters. Several approaches have evolved over time in an attempt to address this uncertainty, the ultimate goal of such approaches being the desensitization of the optimal solutions.

Probably the oldest technique of addressing the issue of uncertainty is sensitivity analysis. Within this approach, the optimization problem is solved in a deterministic manner, after which the input parameters are changed each at a time and the behavior of the optimal value function is observed. The main purpose of this analysis is to determine what is the effect of data/cost coefficient changes to the optimal value function. For example, *Aguado et al.*, [1977] used sensitivity analysis to examine the effect of hydraulic conductivity on the optimal solution of a model for site dewatering; *Bredehoeft and Young* [1983] investigated how variation in surface water supply affected investment in well capacity for an area where both groundwater and surface water was used for irrigation (see also *Gorelick* [1982]). Although this method gives an idea of the relative costs of the parameter errors, it does not address the very pertinent issue of how to incorporate uncertainty in parameters into the decision-making process.

Another approach of dealing with uncertainty is to replace the uncertain quantities with their expected values after which the problem is solved deterministically. Further, one may wish to consider some critical e.g., "worse-case" values and then proceed to solve the management problem within the framework of deterministic optimization. For a given decision variable, one calculates the worst that could happen in terms of all the objectives and then chooses a solution that minimizes the value of the worst-case loss. Although one can argue that, replacing the uncertain quantity with the worse-case value will guard the system against failure (by being extremely conservative), the question of how extreme a case should be considered is not easy to resolve. Moreover, replacing the parameters by their expected values may lead to decisions that are far from being optimal. However, if uncertainty in the parameters is reasonably small and does not critically affect the operation of the system, expected values can still give optimal solutions that are reasonable (*Willis and Yeh* [1987]).

Chance constrained technique is another methodology which has been applied in an attempt to resolve the issue of uncertainty (see *Tung* [1986], *Mayer* [1992], *Willis and Yeh* [1987], *Provencher and Burt* [1994], *Hantush and Marino* [1994], *Hantush and Marino* [1995], *Gorelick* [1990]). This approach treats the stochastic constraints in a probabilistic framework. This means that the fulfillment of the stochastic inequalities is based on pre-determined probability levels, hence optimal solutions are sought from the feasible set inscribed by these inequalities. An attribute associated with this approach is that the resulting optimization problems are of the same size as their deterministic equivalent, hence require more or less the same computational effort. Its shortcoming is that it can only control the probability of violating a given constraint without explicitly mitigating the extent and consequently the effects arising from such violations (*Wagner et al.*, [1992]).

The Monte Carlo method (also known as simulation method) is probably the most powerful and widely used (see e.g., *de Marsily* [1986], *Peck et al.*, [1988], and *Gelhar* [1993]). The popularity of this method is attributed to its ability to simultaneously work with many classes of uncertain parameters. In this method, one generates a large set of realizations of the uncertain parameter (for example, hydraulic conductivity), and for each realization, the forward problem is solved. Statistical analysis of the ensemble of the calculated dependent variable (for example, hydraulic heads) is carried out to determine the expected value and the variance for each location. The variance is used as a measure of uncertainty about the norminal values. Thus, increasing the number of simulations and computing the variance, it is possible to determine the number of simulations required to approach convergence. The computed nominal values are then used in the solution of the optimization problem. The main setback of this method is the number of realizations required to give reliable statistical measures. Typical number of realizations range between hundreds and thousands meaning more computer time, hence high expenses. For example, *van Leeuwen et al.*, [1998] used 1000 realizations, *Wagner et al.*, [1994] used 100 realizations, while *Wagner and Gorelick* [1989] used a total of 30 realizations (see also *Sykes et al.,* [1989], *Wagner et al.*, [1992]). Notable other applications of this method include the works of *Andricevic* [1990], *Andricevic and Kitanidis* [1990], *Kaunas and Haimes* [1985], *Hipel* [1994].

A variant of Monte Carlo approach was used by *Wagner et al.*, [1994] to solve an optimization problem whose objective was to minimize cost. In their work, they introduced the concept of recourse within the problem formulation so as to address the issue of financial consequences of violating the problem constraints. The same notion of recourse was used by *Ndambuki et al.*, [2000] in which they solved a multi-objective optimization problem coupling the optimization problem with a decision support system.

Although solutions obtained through the Monte Carlo approach and its variants are robust to some degree (robustness is dependent on the number of realizations considered), these approaches are CPU intensive, hence limitation of the number of realizations considered is unavoidable. In this paper, we introduce a new approach of solving groundwater quantity management problems whose input parameters (hydraulic conductivity) are uncertain. We transform the optimization problem with uncertain input parameters to a second-order cone optimization (SOCO) problem which is efficiently solved by recently developed interior-point methods (see e.g., *Sturm* [1999], *Lobo et al.*, [1998], *Andersen and Andersen* [1999], *Boyd et al.*, [1994]).

2 Robust optimization

To facilitate further discussion on our approach, we now introduce the concept of robust optimization. We can define robust optimization as an optimization process in which uncertainty resulting from data perturbations is explicitly taken into consideration and the resulting optimal solution is insensitive to data perturbations.

Consider a linear optimization (LO) problem of the form:

$$\underset{\mathbf{x}}{\text{minimize}} \;\; c^T\mathbf{x}, \tag{1}$$

subject to:

$$a_i^T\mathbf{x} \leq b_i, \;\; i = 1, \cdots, m, \tag{2}$$

where $c \in \mathbf{R}^n$, $\mathbf{x} \in \mathbf{R}^n$, $a_i \in \mathbf{R}^n$ and $b_i \in \mathbf{R}$. In such a problem, the parameters a_i, b_i and c may not be known with high precision. What may be known is a domain Ξ (uncertainty set) containing the actual unknown values. A realistic solution to this problem requires that all actual constraints are satisfied by the solution x whatever the actual values realized by the parameters from the uncertainty set, Ξ. This means that one is interested in identifying a solution set (robust feasible candidate solutions) satisfying all the realizations of the uncertain constraints. In the rest of this section, we review the robust optimization methodology developed in *Ben-Tal and Nemirovski* [1998], and *Lobo et al.*, [1998].

For exposition, assume that b_i and c are deterministic while a_i are uncertain parameters lying in given ellipsoids, Ξ_i, (uncertainty sets) defined as (see e.g., *Ben-Tal and Nemirovski* [1998], *Lobo et al.*, [1998]):

$$a_i \in \Xi_i = \{\overline{a}_i + Q_i u \mid \|u\| \leq 1\}. \tag{3}$$

where Q_i is an $n \times \xi$ matrix of perturbations (dimension of the perturbation matrix depends on the actual problem being solved) which describes the shape and size of the uncertainty in the response a_i, $u \in R^{\xi} : \|u\| \leq 1$ which enforces the perturbations to vary in ellipsoids, and $\overline{a}_i \in \mathbf{R}^n$ are the nominal responses and forms the center of the ellipsoids defined by equation (3). Since in the worst case scenario a robust solution requires satisfaction of all constraints for all possible outcomes of the uncertain parameters, a robust formulation of (1)-(2) can be written as (see *Lobo et al.*, [1998], *Ben-Tal and Nemirovski* [1998]):

$$\underset{\mathbf{x}}{\text{minimize}} \;\; c^T\mathbf{x}, \tag{4}$$

subject to:

$$a_i^T\mathbf{x} \leq b_i, \;\; \forall a_i \in \Xi_i, \, i = 1, \cdots, m. \tag{5}$$

Formulation (4)-(5) is now a "deterministic" optimization problem having infinitely many constraints. Using (3), we can say that a solution to the optimization problem defined by (4)-(5) is robust feasible if and only if for every $i = 1, \cdots, m$ we have

$$\left[\overline{a}_i^T\mathbf{x} + (Q_i u_i)^T\mathbf{x} - b_i\right] \leq 0, \forall u_i : u_i^T u_i \leq 1 \tag{6}$$

which is equivalent to:

$$\max_{u_i : u_i^T u_i \leq 1} \left[\overline{a}_i^T\mathbf{x} + (Q_i u_i)^T\mathbf{x} - b_i\right] \leq 0, \tag{7}$$

i.e.,

$$\overline{a}_i^T\mathbf{x} - b_i + \max_{u_i : u_i^T u_i \leq 1} (Q_i u_i)^T\mathbf{x} \leq 0, \tag{8}$$

or

$$\|Q_i^T \mathbf{x}\| \leq b_i - \overline{a}_i^T \mathbf{x}. \tag{9}$$

Constraint (9) is called a second-order cone constraint (also known as "Lorentz cone" or "Ice Cream cone" constraints) because the affinely defined variables $f = Q_i^T \mathbf{x} \in \mathbf{R}^{k-1}$ and $t = b_i - \overline{a}_i^T \mathbf{x} \in \mathbf{R}$ are constrained to belong to the second-order cone of dimension k defined by (see e.g., *Ben-Tal and Nemirovski* [1998], *Lobo et al.*, [1998], *Vanderbei and Yurtan* [1998]):

$$\mathcal{C}_k = \left\{ [f, t] \;\; f \in R^{k-1}, \;\; t \in R : \|f\| \leq t \right\} \tag{10}$$

A solution to formulation (4) is therefore robust feasible if and only if it satisfies the system of inequalities given in (9). Formulation (4) can, therefore be written as a Second-Order Cone Optimization (SOCO) problem, also known as Second-Order Conic Programming (SOCP) problem and frequently referred to as Conic Quadratic Programming Problem as follows:

$$\underset{\mathbf{x}}{\text{minimize}} \;\; c^T \mathbf{x}, \tag{11}$$

subject to:

$$\overline{a}_i^T \mathbf{x} + \|Q_i^T \mathbf{x}\| \leq b_i, \; i = 1, \cdots, m. \tag{12}$$

From now onwards, we shall refer to problems of the form given by formulation (11)-(12) as SOCO problems. Thus a SOCO problem is an optimization problem with linear objective and finitely many Second-Order Cone constraints. In general, SOCO problems are convex problems that include linear and quadratic programs as special cases. For example, when $k = 1$, the unit second-order cone is defined as (see equation (10)):

$$\mathcal{C}_1 = \{ \; t \in R : 0 \leq t\} \tag{13}$$

hence the constraints become linear and the original SOCO problem (formulation (11)-(12)) reduces to a linear optimization problem. Further, letting $\overline{a}_i$ in inequality (12) go to zero reduces the second-order cone constraints to quadratic and the problem becomes quadratically constrained.

It is worth to note that in the process of changing the uncertain LO problem (equation (1)) with ellipsoidal uncertainty to its robust counterpart (equation (11)-(12)), we end up with a SOCO problem whose constraints are now conic quadratic. Further, the norm terms in formulation (12) act as 'regularization terms' or 'penalizing terms', discouraging large values of decision variables, $\mathbf{x}$, in directions of high magnitudes of uncertainty. Indeed this norm term can be thought of as a tolerance margin which has to be maintained if the stochastic constraints are to be satisfied. Solving optimization problem defined by (11)-(12) results in a robust solution which should remain feasible for all possible data perturbations from the given ellipsoidal uncertainty sets.

2.1 Robust single objective groundwater optimization

To demonstrate the power of this novel methodology, we are going to solve a groundwater management problem where our objective is to minimize the operational cost of a groundwater aquifer subjected to some constraints. The constraints define the feasible domain of the optimization problem. We start by formulating the problem, and then follow with an application. Our optimization problem is defined as follows:

$$\underset{\mathbf{x}}{\text{minimize}} \ \left[Z = c^T\mathbf{x}\right], \tag{14}$$

subject to:

$$a_i\mathbf{x} \leq b_i, \ i = 1, \cdots, N_c, \tag{15}$$

$$e^T\mathbf{x} \geq W_d, \tag{16}$$

$$0 \leq \ x_j \leq U_j, \ j = 1, \cdots, N_w, \tag{17}$$

where $c \in \mathbf{R}^{N_w}$, $a_i \in \mathbf{R}^{N_w}$, $e \in \mathbf{R}^{N_w}$ are the problem parameters, $\mathbf{x} \in \mathbf{R}^{N_w}$ are the design variables and

- N_w is the number of pumping wells ($n \rightarrow N_w$ see inequality (2));
- N_c is the number of control points ($m \rightarrow N_c$ see inequality (2));
- c_j is the aggregated daily cost of pumping and transportation in monetary units (MU) per unit volume at cell j;
- x_j is the pumping rate in cell j;
- $a_i = (a_{i,1}, \cdots, a_{i,j}, \cdots, a_{i,Nw})^T$ is the total response at control point i due to pumping from all pumping wells;
- $a_{i,j}$ is the response at control point i due to pumping in cell j;
- b_i is the constraining value at control point i;
- W_d is total water demand;
- e is a vector of ones; and
- U_j is the maximum pumping rate allowed in cell j.

If we knew all the above problem parameters (i.e., c_j, a_i, b_i, W_d and U_j) with high precision, then the above formulated optimization model would be deterministic. If feasible, it would then result to an optimal solution which is the best among all other feasible alternatives. However, because of our limitation in precising the parameters, some of them will only be approximately known. This implies that an optimal solution obtained from such a model is, therefore, likely to be very sensitive to variations of such parameters. In our situation, we shall assume that all parameters except the responses, i.e., a_i (inequality (15)) are exactly known. Hence, the constraints defined using these parameters can be cast as second-order cone constraints. The original optimization problem (formulation (14)-(17)) is transformed into the following SOCO problem:

$$\underset{\mathbf{x}}{\text{minimize}}\ \left[Z = c^T\mathbf{x}\right], \tag{18}$$

subject to:

$$\overline{a}_i^T\mathbf{x} + \|Q_i^T\mathbf{x}\| \le b_i,\ i = 1, \cdots, N_c, \tag{19}$$

$$e^T\mathbf{x} \ge W_d, \tag{20}$$

$$0 \le\ x_j \le U_j,\ j = 1, \cdots, N_w, \tag{21}$$

where $\|Q_i^T\mathbf{x}\| = \{\mathbf{x}^T Q_i Q_i^T \mathbf{x}\}^{\frac{1}{2}}$ and $Q_i Q_i^T = D_i$ are $N_w \times N_w$ covariance matrices of the responses a_i and $\overline{a}_i$ are the nominal responses. To achieve different levels of robustness, we can scale the perturbation matrices Q_i by replacing Q_i in inequality (19) by ηQ_i where $\eta \ge 0$ is the scaling factor. We shall discuss the implications of this scaling later when discussing the results.

3 Multi-objective Optimization

In multi-objective optimization, one is interested in optimizing (minimizing or maximizing) several objectives simultaneously. The consideration of several objectives simultaneously leads to an optimization approach known as multi-objective or vector optimization. This approach recognizes the fact that not all the objectives can achieve their optimal values simultaneously unless the objectives are not competing (conflicting). This means that there is no unique solution to such problems. However one may establish a specific numeric goal (also known as aspiration level) for each of the objectives and then seek a solution that minimizes the sum of deviations of the objective functions from their respective goals. This solution process is known as goal programming. If we denote the aspiration levels of the objective functions $Z_k(x)$ by $\overline{Z}_k$ for $k = 1, \cdots, K$, then the task is to find a solution which minimizes the deviations (under-achievements), $d_k = \overline{Z}_k - Z_k$, for $k = 1, \cdots, K$.

Consider now the n-dimensional vector space $\mathbf{R}^n$. For any two points $\mathbf{r} = (r_1, \cdots, r_n)$ and $\bar{\mathbf{r}} = (\bar{r}_1, \cdots, \bar{r}_n)$, we can express the distance $\mathbf{d}(\mathbf{r}, \bar{\mathbf{r}})$ as the norm given by $\|\ \mathbf{d}\ \| = \|\mathbf{r} - \bar{\mathbf{r}}\|$. Thus the distance between the two points can be measured by the $L_p - metric$ (Holder's norm) as:

$$\|\mathbf{d}\|_p = \left(\sum_{k=1}^{n} |\ r_k - \bar{r}_k\ |^p\right)^{\frac{1}{p}}. \tag{22}$$

Let $\overline{\mathbf{Z}} = \left\{\overline{Z}_1, \cdots, \overline{Z}_K\right\}^T$ and $\mathbf{Z}(x) = \{Z_1(x), \cdots, Z_K(x)\}^T$. The distance between vectors $\overline{\mathbf{Z}}$ and $\mathbf{Z}(x)$ can be measured by the Holder's norm as:

$$\left\|\overline{\mathbf{Z}} - \mathbf{Z}(\mathbf{x})\right\|_p = \left(\sum_{k=1}^{K} |\ \overline{Z}_k - Z_k(\mathbf{x})\ |^p\right)^{\frac{1}{p}}. \tag{23}$$

To avoid biased solution and express the decision maker's (DM's) preferences towards the considered objectives, weighting, w can be applied to equation (23) which then becomes:

$$\left\|\overline{\mathbf{Z}}-\mathbf{Z}(\mathbf{x})\right\|_{w,p}=\left(\sum_{k=1}^{K} w_k \mid \overline{Z}_k - Z_k\left(x\right) \mid^p\right)^{\frac{1}{p}}. \tag{24}$$

By using the L_2- metric, in order to minimize the distance between $\overline{\mathbf{Z}}$ and $\mathbf{Z}(x)$ within the feasible set, Ω, one would then solve an optimization problem of the form:

$$\underset{\mathbf{x}\in\Omega}{\text{minimize}} \left\|\overline{\mathbf{Z}}-\mathbf{Z}\left(\mathbf{x}\right)\right\|_{w,2}, \tag{25}$$

Where $\mathbf{x}$ is the optimization variable, Ω is the feasible set and the other parameters are as defined before. By introducing a scalar deviational variable δ, problem (25) translates to the following:

$$\underset{\mathbf{x}\in\Omega}{\text{minimize}}\ \delta, \tag{26}$$

subject to:

$$\left\|\overline{\mathbf{Z}}-\mathbf{Z}\left(\mathbf{x}\right)\right\|_{w,2} \leq \delta, \tag{27}$$

which is a SOCO problem as long as Ω is defined by linear and/or second-order cone constraints. Note that inequality (27) implies that $\delta \geq 0$.

3.1 Robust multi-objective groundwater optimization

Having presented a general introduction on multi-objective optimization in the preceding sub-section, we now present a formulation of the multi-objective groundwater management problem which we seek to solve. The problem is similar to the single objective one (problem (18)-(21)) we have already formulated except for the fact that an extra objective function is included. In addition to the objective of minimizing the operational cost considered in formulation (18)-(21), we also wish to maximize the amount of water extracted from the groundwater aquifer. Clearly, the two objectives are not only conflicting (increasing the amount of water extracted would necessarily result in an increase in operational cost) but also non-commensurate (different units in the objective attributes) and hence the complexity involved in the solution process. Our multi-objective SOCO problem is as follows:

$$\underset{\mathbf{x}}{\text{minimize}} \left[Z_1\left(\mathbf{x}\right)=c^T\mathbf{x}\right], \tag{28}$$

$$\underset{\mathbf{x}}{\text{maximize}} \left[Z_2\left(x\right)=e^T\mathbf{x}\right], \tag{29}$$

subject to:

$$\overline{a}_i^T\mathbf{x}+\|Q_i^T\mathbf{x}\| \leq b_i,\ i=1,\cdots,N_c, \tag{30}$$

$$e^T\mathbf{x} \geq W_d, \tag{31}$$

$$0 \leq\ x_j \leq U_j,\ j=1,\cdots,N_w, \tag{32}$$

where all the variables and parameters are as defined before.

By introducing a deviational variable, δ, and considering the $L_2-metric$ (Euclidean distance) as the measure of closeness between the aspiration levels $\overline{Z}_1$ and $\overline{Z}_2$

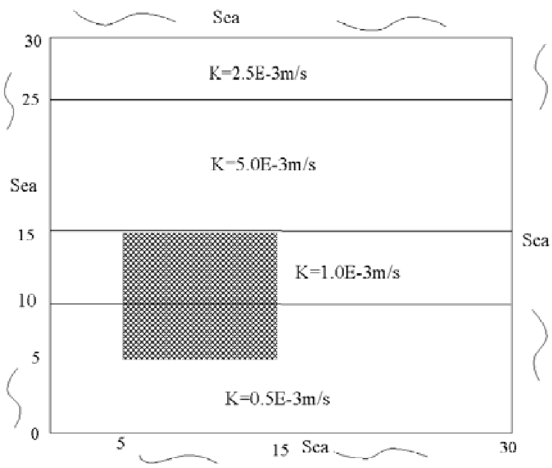

Figure 1: Ecological protection zone and conductivity zones

and the feasible objective region, the above multi-objective optimization problem (problem (28)-(32)) can be reformulated as:

$$\underset{\mathbf{x}}{\text{minimize}}\ \delta, \tag{33}$$

subject to:

$$\left\|\overline{\mathbf{Z}} - \mathbf{Z}(\mathbf{x})\right\|_{w,2} \leq \delta, \tag{34}$$

$$\overline{a}_i^T\mathbf{x} + \|Q_i^T\mathbf{x}\| \leq b_i,\ i = 1, \cdots, N_c, \tag{35}$$

$$e^T\mathbf{x} \geq W_d, \tag{36}$$

$$0 \leq\ x_j \leq U_j,\ j = 1, \cdots, N_w, \tag{37}$$

where δ is a scalar variable, $\overline{\mathbf{Z}} = (\overline{Z}_1, \overline{Z}_2)^T$ are the aspiration levels of objectives $\mathbf{Z}(\mathbf{x}) = (Z_1(\mathbf{x}), Z_2(\mathbf{x}))^T$ and objectives $\mathbf{Z}(\mathbf{x})$ are as defined by equations (28) and (29) respectively. The other parameters and variables are as defined before. To express DM's preferences, preference values can be included in inequality (34) when such preferences exists.

3.2 Application of the methodology to an example

We use a hypothetical example to demonstrate the applicability of the methodology outlined above. Our goal in this problem is to supply domestic water to a distribution center situated in the middle of a single confined square aquifer of thickness 35m. The aquifer is an island of dimensions 30km and with parameters as shown in Figures 1 and 2.

Objective The objective is to supply the required amount of water at the lowest possible cost while at the same time satisfying the hard (deterministic) constraints and ensuring robustness of the optimal solution.

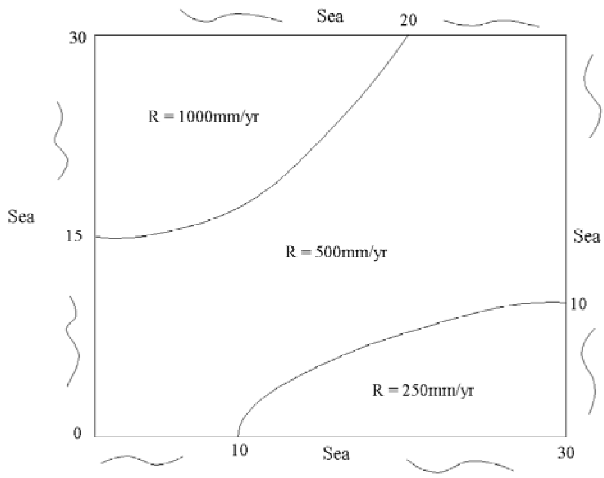

Figure 2: Recharge zones

Constraints The DM expresses constraints as follows:

- In the specified ecological protection zone, the minimum water level equals $5m$ above sea level.
- The hydraulic head in all the nodes except those in contact with the sea are bounded by the bottom of the aquifer.
- To avoid saltwater intrusion, head in cells next to the sea are not allowed to fall below $0.2m$ above sea level.
- There is minimum water demand of $3m^3/s$ at the distribution center which should be satisfied.
- Pumping rates from potential wells shown in Figure 3 are limited to a maximum yield of $1.5m^3/s$.
- Unit costs of exploitation (measured in Monetary Units, MU) defined at each potential well are calculated as a combination of the water pumping costs and water transport costs which depend on the distance from the well to the distribution center. The costs are assumed to be higher towards the boundaries. The aggregated unit cost coefficients at each cell take on values from $0.1MU$ at the center and increase at a rate of $0.1MU$ for every $200m$ distance.

The problem is to find the location of wells and the corresponding pumping rates, satisfying all or nearly all the constraints on the hydraulic heads and pumping rates, and the minimum demand at the distribution center. Moreover, the solution should be robust in an environment of uncertain spatial hydraulic conductivity values.

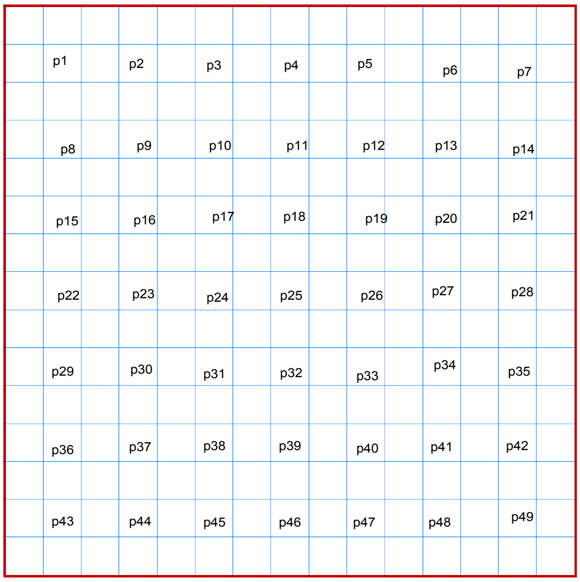

Figure 3: Location of potential pumping wells

3.3 Discussion of results

For the analysis of the example we have outlined, 20 realizations of the uncertain hydraulic conductivity, k, were generated using the zonal k values shown in Figure 1, a standard deviation (log) of 0.5 and a correlation length of 30000m in x-direction and 7500m in y -direction. These correlation lengths were chosen so as to replicate the hydraulic conductivity field of the example as used by *Gorelic* [1982] in a view to comparing the results. All the SOCO problems were solved using Sturm's SeDuMi package (*Sturm* [1999]).

3.3.1 Single objective optimization problem

In the case of the single objective optimization problem, two optimization scenarios which include a deterministic model where only the nominal values were considered (i.e., the perturbation matrices are all zero); and a robust model where both the nominal values and the perturbations were taken into account were analyzed. We seek to solve problem (18)-(21) already formulated. This problem has a total of 49 decision variables (our decision variables are concerned with the determination of location of boreholes and their corresponding pumping rates), 225 second-order cone constraints (coefficients of these constraints are uncertain, hence the constraints are stochastic), and 99 linear constraints (including nonnegativity constraints) which are assumed to be deterministic. Thus, our problem has a mixture of deterministic and stochastic constraints, where the majority of the constraints are stochastic.

Deterministic model This analysis was carried out so as to make comparison of the optimal strategies between a fully homogeneous scenario (where the hydraulic conductivity of the aquifer does not vary from place to place), a banded heteroge-

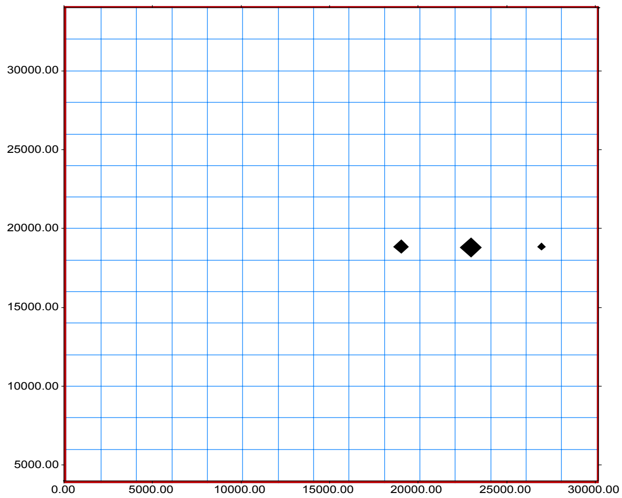

Figure 4: Homogeneous case

neous scenario (where different zones assume different hydraulic conductivity values and within each zone, the hydraulic conductivity value does not vary spatially) and a nominal scenario (where the mean responses, $\overline{a}_i$ computed from the 20 generated realizations of hydraulic conductivity values are used and the perturbation matrices, Q_i are assumed to be equal to zero). The results indicate that the homogeneous scenario under-estimates the cost of the project compared with the other two scenarios. The nominal scenario is the most expensive, while the cost of the optimal strategy of the heterogeneous scenario lies between that of the homogeneous and nominal ones (see Table 1). Figures 4, 5 and 6 clearly show the locations of the wells for the three optimal scenarios. The optimal scenario corresponding to the nominal case operates more wells (7) which are more spread out than the other two scenarios (compare Figures 4, 5 and 6). This tends to suggest that the higher the heterogeneity, the higher the number of active wells and consequently the higher the operational costs.

Table 1. Optimal scenarios

Scenarios	Cost (MU)	Active wells
Homogeneous	1.38	3
Heterogeneous	1.70	4
Nominal	1.95	7
Robust	2.07	16

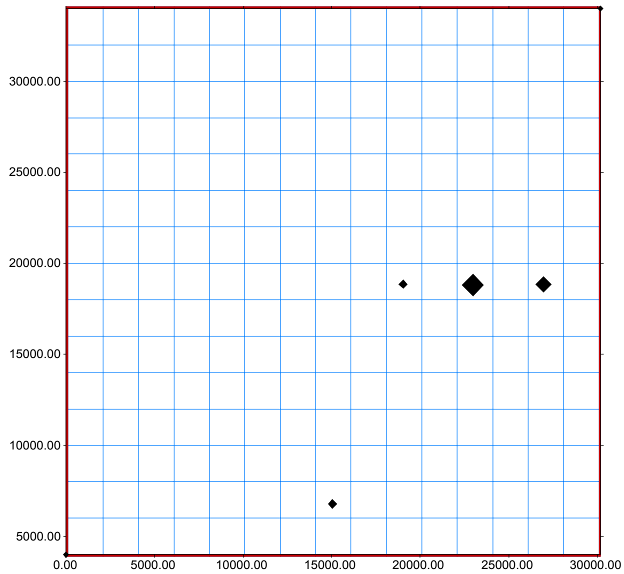

Figure 5: Heterogeneous case

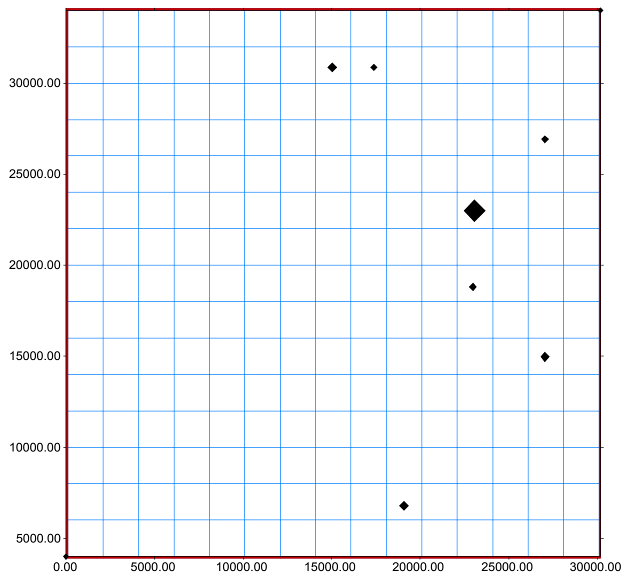

Figure 6: Nominal case

Table 2. Penalty costs for various models

Realization	Heterogeneous	Nominal	Robust
1.	51.732	5.088	3.954
2.	1.810	16.004	9.390
3.	8.790	3.341	1.283
4.	10.354	4.483	4.795
5.	44.333	9.547	8.095
6.	23.184	2.263	2.079
7.	4.288	0.046	0.133
8.	27.431	11.504	6.016
9.	41.255	6.494	4.470
10.	5.485	0.000	0.032
11.	55.603	3.070	2.048
12.	7.793	4.556	1.535
13.	4.717	10.566	4.473
14.	14.101	23.353	16.968
15.	3.545	1.060	1.119
16.	20.802	2.194	1.264
17.	2.883	3.962	3.117
18.	6.504	11.162	9.927
19.	2.559	1.875	1.828
20.	11.794	3.253	1.884
	$\sum = 348.961$	$\sum = 124.271$	$\sum 84.409$

Robust scenario The same problem was solved when considering perturbations. The optimal objective function value computed amounts to $2.07MU$. Figure 7 shows the optimal pumping strategy for this scheme. As expected, most of the wells are located in the north and east of the model domain away from the ecological protection zone. The robust optimal scheme operates more wells (16 wells) than any of the previously discussed cases (see Table 1). Moreover, this robust optimal solution is more expensive as it operates more wells than any of the other optimal solutions.

To establish how well the deterministic heterogeneous, nominal and robust optimal schemes (we excluded the deterministic homogeneous scheme because as was shown by *Ndambuki et al.*, [2000] it has the worst violation of constraints compared to the heterogeneou case) would perform in an environment of uncertainty, a post-optimality Monte Carlo analysis was performed using the 20 pre-generated hydraulic conductivity realizations (exactly the same realizations previously used to compute a_i and hence $\overline{a}_i$ and Q_i). Violations of the constraints were calculated for each optimal scheme as the hydraulic conductivity field (and hence the responses a_i) assumed different values. The violations were translated into penalty cost by using a linear penalty function: $\rho = qv$ where ρ is the penalty cost, q is the penalty parameter (in this case set to 0.2) and v is the amount by which the constraint is violated.

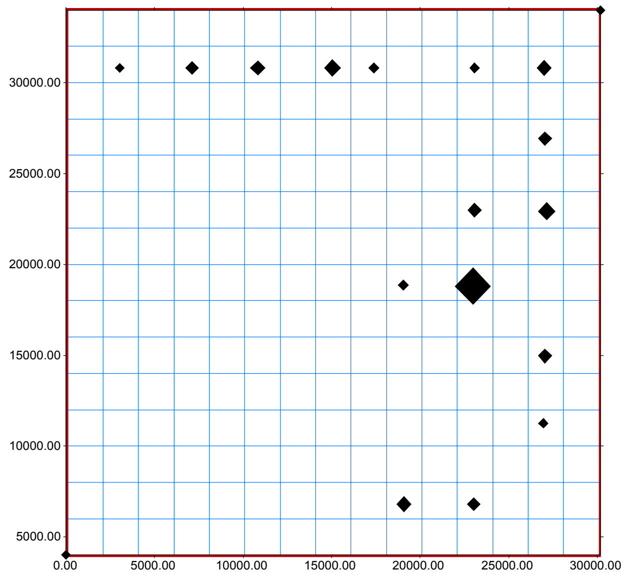

Figure 7: Robust case

The results show that the deterministic heterogeneous optimal solution is worse in terms of violation of constraints, hence more sensitive to parameter variations. To the contrary, the robust solution has the least violations (see Table 2). Note that in the case of the robust model, the perturbation matrices were scaled down and therefore the optimal robust solution will not respect all the scenarios coming from the Monte-Carlo simulations as would be expected. From Figure 8, it is apparent that for the robust solution, no dramatic violations are realized unlike in the other solutions. This is because the robust optimal solution is more cautious than the optimal solutions from the other two scenarios; meaning that the robust optimal solution will avoid stressing the aquifer system in directions in which the uncertainty is large. Further, it is interesting to note that for this robust solution, the objective function value of $2.07MU$ is not much different from that of the nominal case of $1.95MU$ (an increase of about 6%). However, the robustness of the robust solution has substantially increased by about 32% from that of the nominal one.

3.3.2 Multi-objective optimization problem

The multi-objective SOCO problem solved (formulation (33)-(37)) resulted in 50 decision variables (49 of them dealing with the location of pumping wells and their strengths and 1 variable measuring the deviation of the optimal solution from the aspiration levels), 101 linear deterministic constraints, 1 deterministic second-order cone constraint to minimize the deviation from the aspiration levels and 225 second-order cone constraints to capture the robustness (coefficients of these constraints are uncertain, hence stochastic). The targets (aspiration levels) for the two objective functions were computed as $\overline{Z}_1 = 2.07MU$ and $\overline{Z}_2 = 4.47m^3/s$. We then solved the multi-objective optimization problem defined by formulation (33)-(37)) for various levels of robustness. As in the case of the single objective SOCO problem already

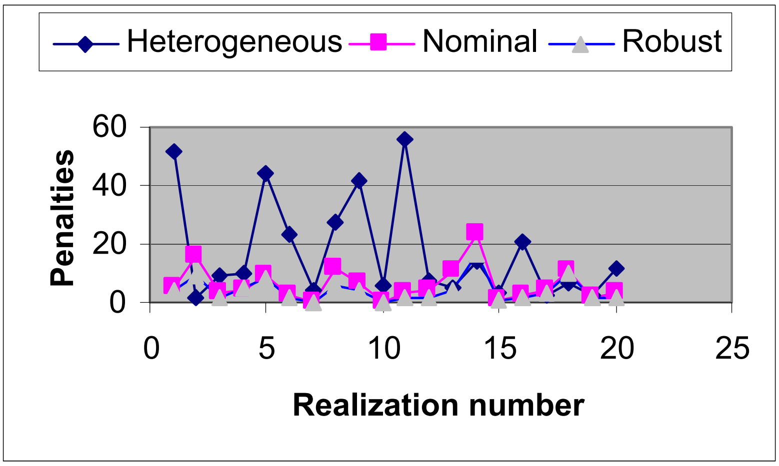

Figure 8: Comparison of solution robustness

discussed, a scaling factor η can be introduced in inequality (35) by replacing the perturbation matrices Q_i by ηQ_i where $\eta \geq 0$. If $\eta = 0$, it means that there is no uncertainty (the input parameters are exactly known), $\eta = 1$ means that the uncertainty is as given by the covariance matrices Q_i and for $\eta > 1$, it means that the uncertainty is higher than that given by the covariance matrices Q_i.

Table 3 shows the levels of robustness and the corresponding number of active wells and values of the operational cost (these results are for a guaranteed volume of water amounting to $3.0m^3/s$). The results show that as the level of robustness, η is increased from 0 to 0.3, the number of active wells increase and consequently the operational cost. This is because increasing the level of robustness implies increase in the volume of the ellipsoid and hence one has to search for a solution within a zone of increasingly high uncertainty. The consequence is that each active well will pump less and since the guaranteed volume of water must be realized, then more pumping wells will have to be mobilized resulting to higher operational costs.

Table 4 and Figure 9 shows how the deviational variable, δ, varies with the level of robustness (remember that the deviational variable is a measure of the discrepancy between the aspiration level, in this case the ideal vector, and the optimal or compromise solution realized). It is interesting to note that as the level of robustness is increased, the deviational variable increases too. This is because an increase in robustness means an increase in the volume of the ellipsoid which implies that the solution sought must guard against a wider range of uncertainty, hence more conservative. The consequence is that such a solution will lie somewhere in the interior of the feasible set. Thus, as the level of robustness is increased further, the solutions will come from increasingly deeper into the feasible set.

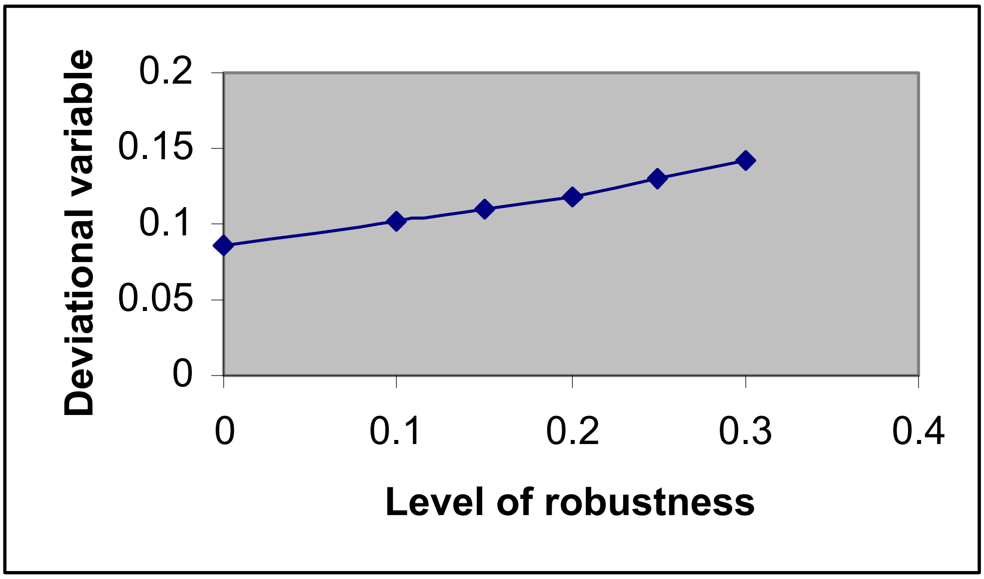

Figure 9: Robustness vs deviational variable

Table 3. Robustness vs Active wells and Cost

Level of robustness	Active wells	Cost (MU)
0.0	11	2.160
0.1	14	2.176
0.15	14	2.185
0.2	16	2.191
0.25	19	2.193
0.3	19	2.195

Table 4. Deviational variable vs robustness

Level of Robustness	Deviational variable, δ
0.0	0.086
0.1	0.103
0.15	0.111
0.2	0.119
0.25	0.130
0.3	0.143

Figures 10-12 show the optimal solutions corresponding to the 3 levels of robustness namely 0.0, 0.2, and 0.3. These levels of robustness were chosen to show how the optimal solutions evolve as the levels of robustness is increased. Pumping wells p_3, p_4, p_5, p_6, p_{14}, p_{20}, p_{27}, p_{35}, p_{47} and p_{48} are used 100% of the cases; p_2, p_7, and p_{21}, 83%; p_{42}, 67%; p_1 and p_{28}, 50%; p_{46} and p_{49}, 33%; while p_8, p_{26}, p_{34} and p_{41} are used 17% of the cases. The rest of the pumping wells are not used at all. Figure 13 shows how frequently each of the wells is used. This Figure (compare with Figure 3 which shows the location of pumping wells) shows that pumping wells lo-

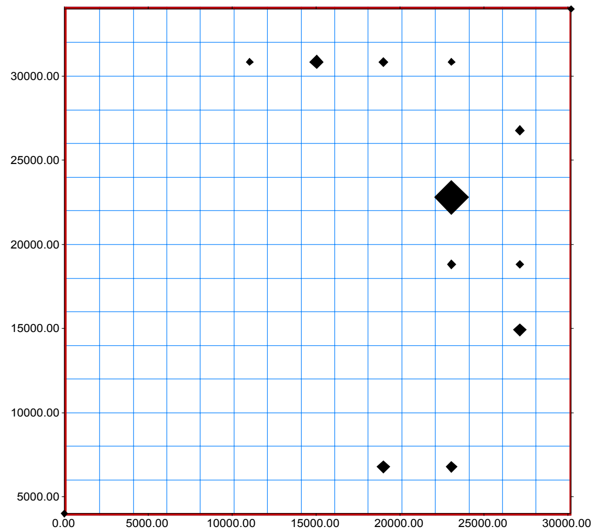

Figure 10: Optimal solution for 0.0 level of robustness

cated to the north and east of the model domain are used more often than pumping wells located elsewhere as expected (these pumping wells are located away from the ecological protection zone).

Compromise solutions corresponding to various quantities of water extracted are given in Table 5 (these solutions are for 0.2 level of robustness). It is apparent that as the quantity of water extracted is increased, the operational cost likewise increase. This reflects some trade-off between the two conflicting and non-commensurate objectives. The optimal schemes corresponding to the optimal solutions number 1, 3, and 5 are depicted in Figures 14-16. Similarly, these optimal solutions were chosen to show how the pumping strategies evolve under different objective trade-offs.

Table 5. Volume vs Cost

Optimal solution number	Volume (m^3/s)	Cost (MU)
1	3.0	2.191
2	3.15	2.234
3	3.25	2.348
4	3.3	2.419
5	3.45	2.670

The results indicate that as the water demand is increased from a minimum of $3.0m^3/s$ to a maximum of $3.45m^3/s$, the operational cost increases from a minimum of $2.191MU$ to a maximum of $2.670MU$. The number of active pumping wells range from a minimum of 14 wells to a maximum of 20 wells. Figure 17 gives an indication of how frequently each pumping well is used across the 5 optimal trade-off solutions given in Table 5.

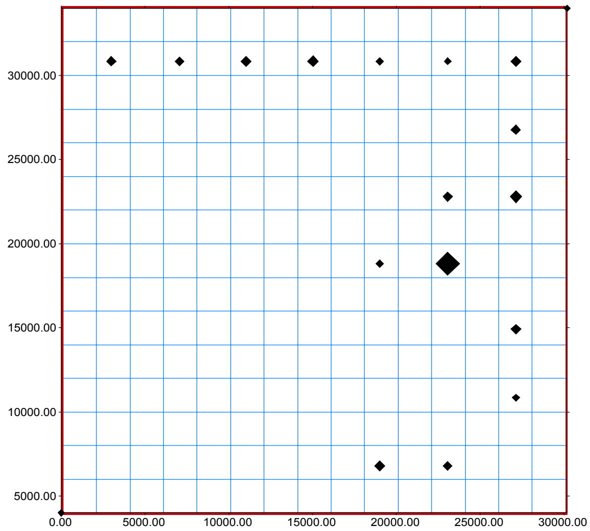

Figure 11: Optimal solution for 0.2 level of robustness

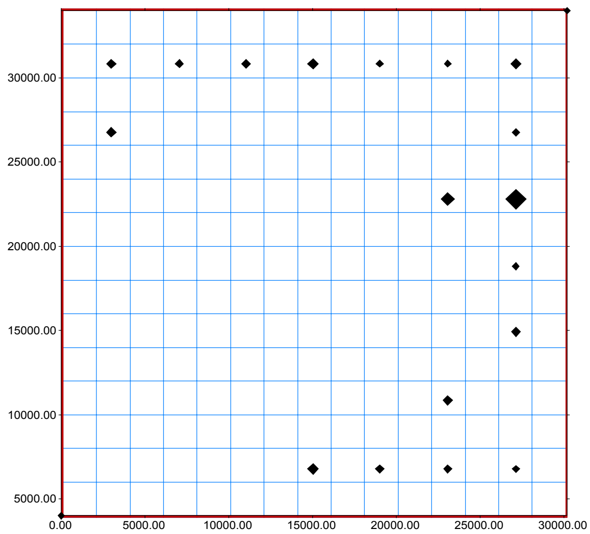

Figure 12: Optimal solution for 0.3 level of robustness

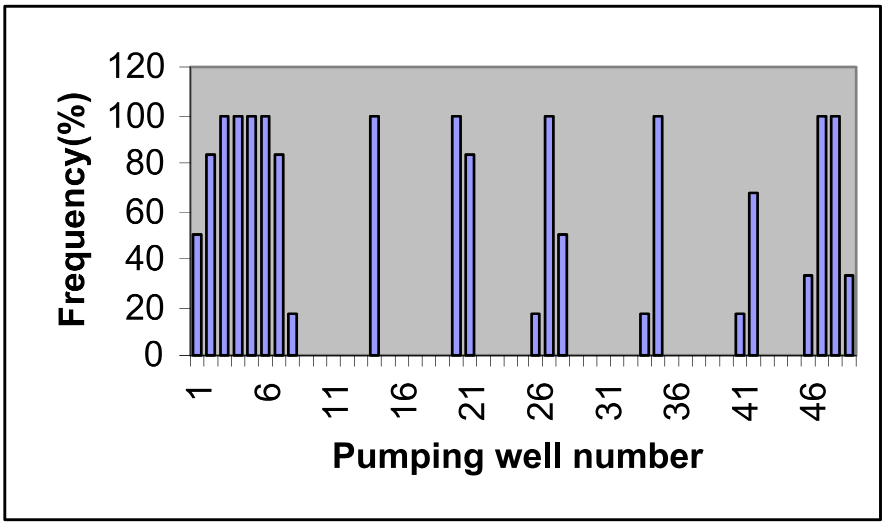

Figure 13: Well use frequency

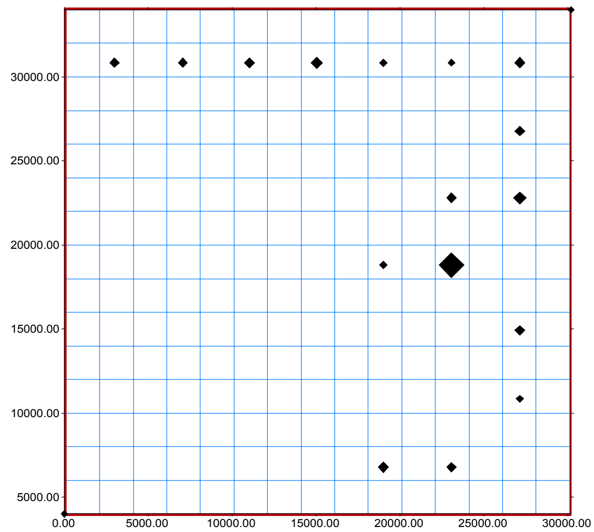

Figure 14: Optimal solution number1

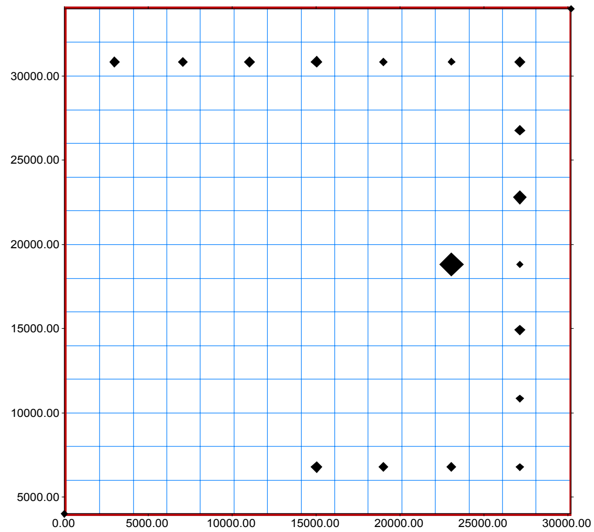

Figure 15: Optimal solution number 3

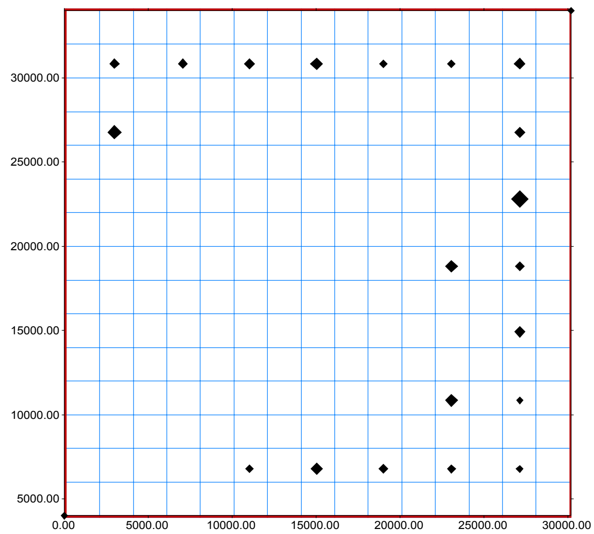

Figure 16: Optimal solution number 5

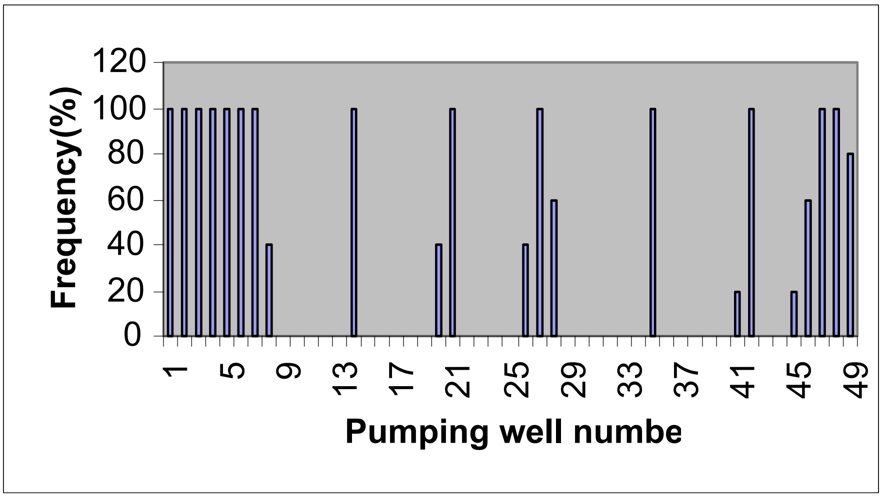

Figure 17: Well use frequency for trade-off solutions

3.4 Conclusions

We have shown that when confronted with uncertain linear optimization problems, such problems can be cast as Second-Order Cone Optimization (SOCO) problems . Further, we have presented a solution methodology for such problems and demonstrated its power through a hypothetical example. Results from the example indicate that the solutions are indeed more robust in comparison with solutions obtained when the parameters are assumed to be known without error. Moreover, the price to pay for a more robust solution (in terms of percentage increase in objective function value) is relatively smaller compared to the gains (in terms of percentage increase in robustness). This means that by approaching the uncertain problem through SOCO, one is assured of robust solutions which are unlikely to be adversely affected by minimal perturbations within the problem parameters. By recognizing that every system is fraught with uncertainty, it is the desire of every DM to be presented with such robust solutions.

We have also shown how one can formulate a multi-objective optimization problem whose parameters are uncertain as a SOCO problem. We have further demonstrated how one could increase or decrease the robustness of the solutions and therefore be able to choose an optimal scheme taking into account the values of the objectives considered and the level of robustness such a solution is able to guarantee. An advantage of this approach is that one does not have to consider a large number of realizations to derive reasonable statistics of the uncertain parameters (as is the case with Monte Carlo approach). Further, high levels of robustness can easily be achieved by increasing the volume of the uncertainty ellipsoids. Thus it is computationally advantageous to approach uncertain optimization problems through SOCO instead of Monte Carlo optimization approach.

Acknowledgement

This work was done when the first author, J. M. Ndambuki, was a guest for $1\frac{1}{2}$

years at the Department of Information Technology and Systems, Delft University of Technology, Netherlands. He wishes to express his appreciation to the Department and the SSOR group for the hospitality accorded to me.

The work of the second author was supported by grant No. 0048923 of the Canada Research Chair program and MITACS.

References

Aguado, E., N. Sitar, and I. Remson, Sensitivity analysis in aquifer studies, *Water Resour. Res.*, 13(4), 733-737, 1977.

Andersen, E. D., and K. D. Andersen, Exploiting parallel hardware to solve optimization problems, *SIAM News*, 32(4), 1999.

Andricevic, R., A real-time approach to management and monitoring of groundwater hydraulics, *Water Resour. Res.*, 26(11), 2747-2755, 1990.

Andricevic, R., and P. K. Kitanidis, Optimization of pumping schedule in aquifer remediation under uncertainty, *Water Resour. Res.*, 26(5) 875-885, 1990.

Ben-Tal, A., and A. Nemirovski, *Convex optimization in engineering*, Tecnion Israel Institute of Technology, 1998.

Boyd, S., L. Vandenberghe, and M. Grant, Efficient convex optimization for engineering design, in *Proceedings IFAC Symposium on robust control design*, Rio de-Janeiro, Brazil, 1994.

Bredehoeft, J. D., and R. A. Young, Conjuctive use of groundwater and surface water for irrigated agriculture: risk aversion, *Water Resour. Res.*, 19(5), 1111-1121, 1983.

de Marsily, G., Q*uantitative hydrogeology*, Academic Press Inc., 1986.

Gelhar, L. W., *Stochastic subsurface hydrology*, Prentice-Hall Inc., 1993.

Gorelick, S. M., A model for managing sources of groundwater pollution, *Water Resour. Res.*, 18(4), 773-781, 1982.

Gorelick, S. M., Large scale nonlinear deterministic and stochastic optimization, formulations involving simulation of subsurface contamination, *Mathematical Programming*, 48, 19-39, 1990.

Hantush, M. M., and M. A. Marino, Two-dimensional stochastic analysis and optimal estimation in aquifer: random recharge, *Water Resour. Res.*, 30(2), 559-569, 1994.

Hantush, M. M., and M. A. Marino, Continuous-time stochastic analysis of groundwater flow in heterogeneous aquifers, *Water Resour. Res.*, 31(3), 565-575, 1995.

Hipel, K. W., *Stochastic and statistical methods, in hydrology and environmental engineering* vol. 2, Kluwer Academic Publishers, 1994.

Kaunas, J. R., and Y. Y. Haimes, Risk management of groundwater contamination in multi-objective framework, *Water Resour. Res.*, 21(11), 1721-1730, 1985.

Lobo, M. S., L. Vandenberghe, S. Boyd, and H. Lebret, Applications of second-order cone programming, *Linear Algebra and its Applications*, 284, 193-228, 1998.

Mayer, J., Computational techniques for probabilistic constrained optimization, in K. Marti, editor, *Stochastic optimization, numerical methods and technical applications; Proceedings of a GAMM/IFIP-Workshop* held at the Federal Armed Forces University Munich, 1992.

Ndambuki, J. M., F. A. O. Otieno, C. B. M. Stroet, and E. J. M. Veling, Groundwater management under uncertainty: a multi-objective approach, *Water SA*, 26(1), 35-42, 2000.

Peck, A., S. Gorelick, G. de Marsily, F. Foster, and V. Kovalevsky, Consequences of spatial variability in aquifer properties and data limitations for groundwater modelling practice, Technical Report 175, IAHS, 1988.

Provencher, B., and O. Burt, Approximating the optimal groundwater pumping policy in a multi-aquifer stochastic use setting, *Water Resour. Res.*, 30(3), 833-843, 1994.

Sturm, J. F., *SeDuMi: a matlab toolbox for solving optimization problems over symmetric cones*, systems optimization, Maastricht University, 1.03 edition, http://sedumi.mcmaster.ca, 1999.

Sturm, J. F., Using SeDuMi 1.03, a Matlab toolbox for optimization over symmetric cones: in F. Potra, C. Roos, and T. Terlaky, editors, *Optimization Methods and Software*, vol. 11-12, 1999.

Sykes, J. F., B. E. Sleep, and N. R. Thimson, A moment method for calculating groundwater flow uncertainty: the analysis of a landfill, in G. Jousma, J. Bear, Y. Y. Haimes, and F. Walter, editors, *Groundwater contamination:use of models in decision- making, Proceedings of the International Conference on Groundwater Contamination*, Amsterdam, Netherlands, 219-227, 1989.

Tung, Y.-K., Groundwater management by chance constrained model, *Water Resour. Plan. and Man.*, 112(1), 1-19, 1986.

van Leeuwen, M., C. B. M. Stroet, A. P. Butler, and J. A. Tompkins, Stochastic determination of well capture zones, *Water Resour. Res.*, 34(9), 2215-2223, 1998.

Vanderbei, R. J., and H. Yurtan, Using LOQO to solve second-order cone programming problems, technical report SOR-98-9, statistics and operations research, Princeton University, 1998.

Wagner, B. J., and S. M. Gorelick, Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design, *Water Resour. Res.*, 2211-2225, 1989.

Wagner, J. M., U. Shamir, and D. H. Marks, Containing groundwater contamination: planning models using stochastic programming with recourse, *European Jour. Oper. Res.*, 77, 1-26, 1994.

Wagner, J. M., U. Shamir, and H. R. Nemati, Groundwater quality management under uncertainty: stochastic programming approaches and the value of information, *Water Resour. Res.*, 28(5), 1233-1246, 1992.

Willis, R., and W.-G. Yeh, *Groundwater systems planning and management,* Prentice-Hall Inc., 1987.