Robust Energy Cost Optimization of Water Distribution System with Uncertain Demand

Alexander Goryashko* Arkadi Nemirovski†

Abstract

A methodology, based on the concept of Affinely Adjustable Robust Optimization, for optimizing daily operation of pumping stations is proposed, which takes into account the fact that a water distribution system in reality is unavoidably affected by uncertainties. For operation control, the main source of uncertainty is the uncertainty in the demand. Traditional methods for optimizing dynamical systems under uncertainty (Multistage Stochastic Programming) results in computationally intractable models already for small water distribution networks. The most popular optimization method for these problems is Dynamic Programming; however, in practice applications of this approach are restricted to networks with 1-2 pumping stations and/or 1-2 storages, because of severe computational difficulties arising in when state dimension of the controlled dynamical system exceeds 1-2. The approach presented in this paper provides a computationally tractable alternative to the outlined traditional methods in the cases when the problem under consideration, in the absence of uncertainty, can be formulated as a Linear Programming problem.

Key words: optimization of water supply networks, uncertain linear programming, robust optimization.

1 Introduction

Energy cost optimization of a water-supply network is a very important practical problem. For instance, the saving that would accrue from only a 5% reduction in the total power consumed in the US has been estimated at $48,000,000 per year (Tarquin and Dawdy 1989). So, there is no great surprise in the constant growth of the interest of practitioners and researches in optimizing the pump operations. Previous efforts have resulted in a range of mathematical tools for developing optimal pump schedules. One of the most popular tools is the Dynamic Programming approach originating from deMoyer and Horowitz (1975) and further developed in Jolland and Cohen (1980), Ferreira and Vidal 1984, Zessler and Shamir (1989), Lansey and Awumah (1994). Besides Dynamic Programming approach, basically all other known optimization techniques were used: Nonlinear Programming (Ormsbee et al. 1993, Pezeshk and Helweg 1993), Linear Programming (Jowit and

*corresponding author: Department of Applied Mathematics, Moscow State University of Printing Arts, Prianishnikova 2a, 127750 Moscow, Russia, ale_gory@rambler.ru

†School of Industrial and systems engineering, Georgia Institute of Technology, 765 Ferst Dr. NW, Atlanta GA 30332, USA, nemirovs@isyel.gatech.edu

Germanopulos (1992), Likeman (1993)), Mixed Integer Programming (Creasey 1998). The problem was attacked also with genetic algorithms (Simpson et al. 1999) and Simulated Annealing (Sakarya et al. 1999). The practical performance of the outlined techniques is rather restricted because of two well-known obstacles: nonlinear hydraulic characteristics of water distribution networks, on one hand, and large sizes of real-world networks (large number of nodes, numerous pumps of different types, etc.). As a result, the impact of optimization on operating real world water distribution systems remains rather restricted. This unpleasant fact is only partly caused by the outlined difficulties (nonlinearity and large sizes of the resulting optimization problems); another principal difficulty comes from *demand uncertainty*.

It is well known that the standard (deterministic) Mathematical Programming models can yield solutions which may become heavily infeasible as a result of data perturbations, even small. This, in particular, is true for water resources management models (Watkins and McKinney 1997). This is in sharp contrast with the fact that the major source of uncertainty in water distribution models, namely, the demand uncertainty, is completely ignored in the vast majority of literature on the optimal pump scheduling. As a rule, the authors work under the assumption that the demand pattern is known, perhaps from certain forecasting algorithm. Thus, it is implicitly suggested that feasibility properties of the optimal decision are not heavily affected by perturbations in the demand, either due to low perturbations of the demand in a real water distribution network, or due to high accuracy of the forecasting algorithms. In fact, the demand uncertainty in real water distribution networks is very high (see below for some real data), and it seems that realizations of the demand trajectories can hardly be predicted with high accuracy.

There exist approaches which handle uncertainty from the very beginning, notably Stochastic Programming. There is a large number of studies using the Stochastic Programming approach in water resource management (see detailed review in Watkins and McKinney 1997). In nearly all water resources applications of Stochastic Optimization, the models are aimed at optimizing the expected value of the objective function and do not allow to evaluate the trade-offs between the risks of infeasibility and the losses in optimality. Few exceptions are the multi-objective models of waste water treatment design (Uber et al. 1992), the hydropower operation model with penalizing violations of "soft" constraints (Jacobs et al. 1995), and optimization of operation in reservoirs with penalty for violation of constraints (Eiger and Shamir 1991).

In Watkins and McKinney (1997) it is shown that the new Robust Optimization methodology proposed in Mulvey et al. (1995) as applied to urban water transfer planning and ground-water quality management allows for finding solutions hedging against inherent parameter uncertainty.

Thus, at least in principle there are methods to solve pumping cost optimization problem with uncertain demand. Unfortunately, all proposed approaches to handling uncertainty, including (multistage) Stochastic Programming and Robust Optimization methodology (Mulvey et al. 1995), share a number of common disadvantages. Specifically, as far as computational schemes are concerned, all these approaches operate with a (finite) set of scenarios, which implies a highly nontrivial problem of how to build a "representative" sample of scenarios. When speaking about dynamical processes with not too small uncertainty, a small sample hardly will be representative, while a large sample can be very difficult to process numerically. All known attempts to resolve this difficulty require a lot of ad hoc decisions (the number of scenarios, the coefficients specifying the tradeoff between optimality and reliability, etc.) and are unable to *guarantee* the feasibility of the resulting solution at a given level of uncertainty even in the favourable cases when such a "reliable" solution does exist and is not very expensive. As a matter of fact, the traditional methodologies for handling uncertainty, as applied to water management problems, are unable to provide "reliable" solutions to problems of realistic sizes.

Over a decade ago, a new paradigm for optimization under uncertainty – the *Robust Counterpart* approach – was proposed and used in a number of applications (Ben-Tal and Nemirovski 1997, 1998 1999,2000,2001; El Ghaoui et al. 1997,1998; Ben-Tal, El Ghaoui et al. 2000, Chandrasekaran et al. 1998; for recent state of the art, see [1]; for the very first step on this direction, see Soyster 1973). This methodology seems to be attractive for water management problems as well.

The concept of Robust Counterpart is quite transparent. Let us focus on optimization programs in the form

$$\min_{x \in \mathbf{R}^n} \{f(x,\zeta) : F(x,\zeta) \leq 0\}, \tag{1}$$

where $x \in \mathbf{R}^n$ is the decision vector, ζ is the data and f, F are given functions (f is real-valued, F is vector-valued) of x depending on the data as on a parameter. An *uncertain* problem $\mathcal{P}$ is defined as a family of all problems (1) with the data belonging to a given *uncertainty set* $\mathcal{U}$. We assume that the "decision environment" is such that the constraints in (1) are *hard*, so that a meaningful candidate solution should satisfy the constraints *whatever is a realization of the constraint from the uncertainty set*. As about the "price" of a candidate solution x, we treat it in the same "worst-case-oriented" fashion, specifically, as the maximum $\max_{\zeta \in \mathcal{U}} f(x,\zeta)$ of the "actual price" $f(x,\zeta)$ over all possible realizations $\zeta \in \mathcal{U}$ of the data. With this "worst-case-oriented" approach it is natural to associate with the uncertain problem $\mathcal{P}$ its *Robust Counterpart*, which is the usual optimization problem

$$\min_{x \in \mathbf{R}^n} \left\{ \sup_{\zeta \in \mathcal{U}} f(x,\zeta) : F(x,\zeta) \leq 0 \quad \forall \zeta \in \mathcal{U} \right\}. \tag{2}$$

The optimal solution/value of the latter problem is called the *robust optimal* solution/value of the uncertain problem $\mathcal{P}$. A robust optimal solution, if exists, is "absolutely reliable" as far as the constraints are concerned (provided, of course, that the realizations of the data indeed reside in the uncertainty set). One may think that the "price" of this "absolute reliability" (the increase in the optimal value as compared to the "no uncertainty" case) normally should be pretty high. However, numerous examples considered in (Ben-Tal and Nemirovski, 1997,1998,1999,2000,2001; El Ghaoui et al. 1997,1998), same as the results we are about to present, demonstrate that the price of robustness can be quite small. A qualitative explanation to this phenomenon could be as follows. A practical optimization problem often possesses a "massive" set of nearly optimal solutions. Ignoring uncertainty and obtaining the decision by solving the problem with the "nominal" data, we usually end up with a point on the boundary of the above set, and such a point can become heavily infeasible for a problem with a slightly perturbed data. In contrast to this, the robust counterpart tries to choose the "most inner" solution from the above massive set, and such a solution is usually much more stable w.r.t. data perturbations than a boundary one.

In this paper, we intend to apply the Robust Counterpart methodology to the problem of optimizing operation of a water distribution system. The objective is to specify the pump-operation pattern, over a given T-period planning horizon, which results in the lowest possible energy cost. This objective can be stated as

$$\min \left\{ \sum_{i=1}^{I} \sum_{t=1}^{T} r_i(t) U_i(t) \right\}, \tag{3}$$

where

- I is the number of pumping stations, and T is the number of time periods, of equal duration, in the planning horizon,
- $r_i(t)$ is the tariff (per unit charge) at pumping station i in period t, $t = 1, ..., T$,

- $U_i(t)$ is the power consumption at pumping station n in period t.

In general, $U_i(t)$ is a nonlinear function of pressure head and pump discharge.

The above objective should be minimized under a system of constraints, specifically,

– linear equations for mass conservation at each node of the water network;

– nonlinear equations expressing energy conservation for each loop in the system;

– linear equations for tank levels, for each tank in the system;

– bounds on the maximal and the minimal tank levels, for each tank in the system during the entire planning horizon;

– linear constraints on the initial and (optionally) the final tank levels, for each tank in the system.

Thus, in general we are dealing with a complicated nonlinear optimization problem, which is highly demanding computationally already for a water distribution network with few tens of nodes, even provided that the demands at every node are known exactly.

There exist many ways to simplify the outlined original formulation, including hierarchical decomposition methods (for example, Coulbeck and Sterling 1978), restricting with simple operating rules (Moss 1979), simplifications allowing to decouple pumping station operation and the hydraulic considerations (Jowit and Germanopulos 1992), etc.

Our approach is as follows. We start with an extremely simple Linear Programming (LP) model of a water distribution network, where we circumvent completely the difficulties coming from nonlinear equations for energy conservation. In our LP model, we treat the demand as uncertain data element and apply to the resulting *uncertain* LP problem the recent "dynamical" version of the Robust Counterpart methodology (Ben-Tal et al. 2002). As a result, we get certain pump operation policy. Finally, we test this policy for hydraulic correctness by an extensive extended-period simulation.

The rest of the paper is organized as follows. In Section 1, we describe a simple pipeline water distribution network which can be naturally modelled by an LP model. In Section 2, we show how a more sophisticated hydraulic model (the AnyTown one) can be approximated by the LP model from Section 1. In Section 3, we apply to our LP model the Robust Counterpart methodology to get a kind of robust optimal pump control policy. We present numerical results for the AnyTown model and discuss their robustness properties vs. the demand uncertainty level. In Section 4 we demonstrate that the uncertainties in real life demands cannot be ignored, and investigate the hydraulic aspects of the robust optimal pump control policy we have built for the AnyTown model. Section 5 contains a discussion of several possibilities to use our approach in real-life water distribution networks and some concluding remarks.

2 A simple pipeline system and the LP model

Consider a simple water distribution system (Fig. 1) which includes several (in our example, two) upstream reservoirs, each with a pumping station, pipelines, a downstream reservoir (tank) and a distribution network. An upstream reservoir may be either a source of water or an intermediate reservoir. A pumping station allows for one or more combinations of parallel pumps; such a combination can be a source or a booster one, depending on the type of the corresponding upstream reservoir. As a rule, the main goal of simple models like the one we are discussing is to determine, over the 24-hour period, the minimum energy cost pumping stations schedule, assuming the 24-hour demand pattern known in advance (see, e.g., J.J.de Oliveira Sousa et al. 2001).

The mathematical model of the water distribution system with a single tank and I sources (cf. Fig. 1) is as follows. The planning horizon is T periods. On a period t:

- d_t is the total water demand during the period. All the demand must be satisfied;
- $v(t)$ is the volume of water in the tank at the beginning of the period ($v(1)$ is given);
- $p_i(t)$ is i-th supply of period t – the amount of water pumped into the system during the period from source # i in order to satisfy the demand of the period (and, perhaps, to fill the tank);
- $P_i(t)$ is the maximum allowed supply per period from source # i;
- $c_i(t)$ is the cost of pumping in a unit of water from source # i in period t (in reality, this cost includes two components: the price of the water depending on the specific source, and the cost of electricity required to pump the water from the source into the system).

Remark 2.1 *In reality, the "energetic" component of the cost C of pumping water within a given hour t is not exactly proportional to the flow rate p; the standard approximation is given by, for example, (Walski 1984)*

$$C = \sigma(t)\kappa e^{-1} H(p)p, \tag{4}$$

where $H(p)$ is the total dynamic head of the pump corresponding to the flow rate p, e is the efficiency of the pump, κ is an appropriate constant factor, and $\sigma(t)$ is the price of energy (H is measured in m, p - in m^3/h, e is dimensionless, kappa is measured in kwh/m^4, and sigma – in \$/kwh). In order to get a Linear Programming problem, we approximate, for every pump i, relation (4) *by a proportional dependence C on p: $C = c_i(t)p$.*

Other parameters of the problem are:

- V_{min} - the minimal allowed volume of the tank;
- V_{max} - the maximal storing capacity of the tank;
- Q_i - the maximal total amount of water which can be pumped into the system, over the entire planning horizon, from source # i.

Our goal is to minimize the total, over all sources and the entire planning period, electricity cost. The Linear Programming problem achieving this is as follows:

$$\begin{array}{lll}
\min\limits_{p_i(t),v(t),F} & F & \\
\text{s.t.} & \sum\limits_{t=1}^{T}\sum\limits_{i=1}^{I} c_i(t)p_i(t) \le F & \text{[cost account]} \\
& v(t+1) = v(t) + \sum\limits_{i=1}^{I} p_i(t) - d_t,\ t = 1,\dots,T & \text{[water balance equations]} \\
& V_{min} \le v(t) \le V_{max},\ t = 2,\dots,T+1. & \text{[bounds on tank's level]} \\
& 0 \le p_i(t) \le P_i(t),\ i = 1,\dots,I,\ t = 1,\dots,T & \text{[bound on pump's performance]} \\
& \sum\limits_{t=1}^{T} p_i(t) \le Q(i),\ i = 1,\dots,I & \text{[bound on source's capacity]}
\end{array} \tag{5}$$

It is convenient to eliminate the v-variables, thus coming to the inequality constrained problem:

$$\begin{array}{ll}\min\limits_{p_i(t),F} & F \\ \text{s.t.} & \sum\limits_{t=1}^{T}\sum\limits_{i=1}^{I} c_i(t)p_i(t) \le F \\ & 0 \le p_i(t) \le P_i(t),\ i=1,\ldots,I,\ t=1,\ldots,T \\ & \sum\limits_{t=1}^{T} p_i(t) \le Q(i),\ i=1,\ldots,I \\ & V_{min} \le v(1) + \sum\limits_{s=1}^{t}\sum\limits_{i=1}^{I} p_i(s) - \sum\limits_{s=1}^{t} d_s \le V_{max},\ t=1,\ldots,T. \end{array} \tag{6}$$

2.1 Modeling uncertainty

If the demand pattern $\{d_t\}_{t=1}^T$ were known in advance, the optimal pumping policy would be given by the optimal solution to the LP problem (6). Now assume that when solving the problem ("at time 0"), we do not know the future demands exactly; all we know is that the will-be demand pattern belongs to a given "stripe"

$$\mathcal{D} = \left\{ \{d_t\}_{t=1}^T : d_t \in [d_t^* - \theta d_t^*, d_t^* + \theta d_t^*],\ t=1,\ldots,T \right\}, \tag{7}$$

where $\{d_t^*\}_{t=1}^T$ is the "nominal demand", and θ, $0 \le \theta \le 1$, is the "demand uncertainty level" [1)] With uncertain demand given by (7), the LP problem (6) becomes uncertain, and we can (and will) apply to the problem the Robust Counterpart approach. However, the latter approach in its original form stems from the assumption that all components of the decision vector represent the "here and now" decisions which should be made when the problem is solved and thus cannot tune themselves to the actual values of the data. This assumption is definitely not true in the situation we are interested in, where the actual "here and now" decisions are only those $p_i(t)$ which correspond to the first period $t = 1$, while p's corresponding to the subsequent periods are typical "wait and see" decisions and as such can tune themselves, to some extent, to the actual data. Specifically, it is natural to assume that decisions on supplies $p_i(t)$ are made at the beginning of period t, and that we are allowed to make these decisions on the basis of demands d_r observed at periods $r \in I_t$, where I_t is a given subset of the segment $\{1, ..., t-1\}$. With this assumption, when solving the problem "at time 0" we are looking for $p_i(t)$ as for "decision rules" (*functions* of $\{d_r : r \in I_t\}$) rather than for "decisions" (*reals* independent of d's). In principle, the only restriction on the decision rules is that the corresponding realizations of decisions $p_i(t)$ should satisfy the constraints of (6) for every realization of the demands given by (7). It turns out, however, that allowing for decision rules to be of arbitrary structure, we end up with computationally intractable problem (the same phenomenon as in the multistage Stochastic Programming). In order to overcome this difficulty, in (Ben-Tal et al. (2002)) it is proposed to restrict the decisions rules to be *affine* functions of the corresponding portions of the data:

$$p_i(t) = \pi_{i,t}^0 + \sum_{r \in I_t} \pi_{i,t}^r d_r, \tag{8}$$

where the coefficients $\pi_{i,t}^r$ are our new decision variables. With this approach, (6) becomes an uncertain Linear Programming problem in variables $\pi_{i,t}^s$, F, and we can apply to this problem the

[1)] We restrict ourselves to model (7) in order to be as simple as possible; note that the approach in question is capable to handle much more complicated models of uncertainty, cf. Ben-Tal et al. (2002).

usual Robust Counterpart methodology, i.e., to require the constraints to be valid for all realizations of the demand given by (7). The resulting problem is

$$
\begin{array}{ll}
\min\limits_{\pi,F} & F \\
\text{s.t.} & \sum_{t=1}^{T}\sum_{i=1}^{I} c_i(t)\pi^0_{i,t} + \sum_{r=1}^{T}\left(\sum_{i=1}^{I}\sum_{t:r\in I_t} c_i(t)\pi^r_{i,t}\right) d_r - F \le 0 \\
& \pi^0_{i,t} + \sum_{r\in I_t}^{t} \pi^r_{i,t} d_r \le P_i(t),\ i=1,\dots,I,\ t=1,\dots,T \\
& \pi^0_{i,t} + \sum_{r\in I_t} \pi^r_{i,t} d_r \ge 0,\ i=1,\dots,I,\ t=1,\dots,T \\
& \sum_{t=1}^{T}\pi^0_{i,t} + \sum_{r=1}^{T}\left(\sum_{t:r\in I_t}\pi^r_{i,t}\right) d_r \le Q_i,\ i=1,\dots,I \\
& \sum_{s=1}^{t}\sum_{i=1}^{I}\pi^0_{i,s} + \sum_{r=1}^{t}\left(\sum_{i=1}^{I}\sum_{s\le t, r\in I_s}\pi^r_{i,s} - 1\right) d_r \le V_{max} - v(1) \\
& \quad t=1,\dots,T \\
& -\sum_{s=1}^{t}\sum_{i=1}^{I}\pi^0_{i,s} - \sum_{r=1}^{t}\left(\sum_{i=1}^{I}\sum_{s\le t, r\in I_s}\pi^s_{i,r} - 1\right) d_s \le v(1) - V_{min} \\
& \quad t=1,\dots,T, \\
& \forall\{d_t\in[d^*_t-\theta d^*_t, d^*_t+\theta d^*_t],\ t=1,\dots,T\}.
\end{array}
\tag{9}
$$

The reason for restricting ourselves with affine decision rules (8) is that the semi-infinite (i.e., with infinitely many linear constraints) problem (9) is equivalent to a usual LP program, specifically, the program

$$
\begin{array}{c}
\min\limits_{\pi,F,\alpha,\beta,\gamma,\delta,\zeta,\xi,\eta} \quad F \\
\sum_{i=1}^{I}\sum_{t:r\in I_t} c_i(t)\pi^r_{i,t} = \alpha_r,\ -\beta_r\le\alpha_r\le\beta_r,\ 1\le r\le T,\ \sum_{t=1}^{T}\sum_{i=1}^{I} c_i(t)\pi^0_{i,t} + \sum_{r=1}^{T}\alpha_r d^*_r + \theta\sum_{r=1}^{T}\beta_r d^*_r \le F; \\
-\gamma^r_{i,t}\le\pi^r_{i,t}\le\gamma^r_{i,t},\ r\in I_t,\ \pi^0_{i,t} + \sum_{r\in I_t}\pi^r_{i,t}d^*_r + \theta\sum_{r\in I_t}\gamma^r_{i,t}d^*_r \le P_i(t),\ 1\le i\le I, 1\le t\le T; \\
\pi^0_{i,t} + \sum_{r\in I_t}\pi^r_{i,t}d^*_r - \theta\sum_{r\in I_t}\gamma^r_{i,t}d^*_r \ge 0,\ \sum_{t:r\in I_t}\pi^r_{i,t} = \delta^r_i,\ -\zeta^r_i\le\delta^r_i\le\zeta^r_i,\ 1\le i\le I, 1\le r\le T, \\
\sum_{t=1}^{T}\pi^0_{i,t} + \sum_{r=1}^{T}\delta^r_i d^*_r + \theta\sum_{r=1}^{T}\zeta^r_i d^*_r \le Q_i,\ 1\le i\le I; \\
\sum_{i=1}^{I}\sum_{s\le t, r\in I_s}\pi^r_{i,s} - \xi^r_t = 1,\ -\eta^r_t\le\xi^r_t\le\eta^r_t,\ 1\le r\le t\le T, \\
\sum_{s=1}^{t}\sum_{i=1}^{I}\pi^0_{i,s} + \sum_{r=1}^{t}\xi^r_t d^*_r + \theta\sum_{r=1}^{t}\eta^r_t d^*_r \le V_{max} - v(1),\ 1\le t\le T, \\
\sum_{s=1}^{t}\sum_{i=1}^{I}\pi^0_{i,s} + \sum_{r=1}^{t}\xi^r_t d^*_r - \theta\sum_{r=1}^{t}\eta^r_t d^*_r \ge v(1) - V_{min},\ 1\le t\le T.
\end{array}
\tag{10}
$$

The approach we are advocating in this paper, in its simplest form, is to model a water distribution network by simple LP program (5), to treat the demand as uncertain and to pass from the resulting uncertain LP to its *Linearly Adjustable Robust Counterpart* (10). With this approach, the decision rules to be used to control the network are “robust optimal linear decision rules” (8) given by the optimal solution to (10). In the sequel, we refer to these rules as to the LRO (linear robust optimal) management policy.

3 A Case Study

The methodology. The water distribution network we deal with is too simple to be of practical interest. At the same time, we cannot directly apply the outlined methodology to a more realistic, highly nonlinear, hydraulic model of water distribution. What we intend to do is to use a two-stage Synthesis/Analysis approach as follows. At the first – the Synthesis – stage, we treat a given water distribution network as a simple network like the one depicted on Fig. 1. To this end, we replace all actual tanks with a single tank with the storage equal to the total storage of the actual tanks. Similarly, we model demands at numerous junctions of the actual system by a single demand junction with demand equal to the sum of actual demands at the junctions. Solving problem (10) associated with the simple resulting model of the network and a realistic uncertainty level, we end up with the LRO management policies for pumping stations. More exactly, we get the optimal values for the pump flow rates rather than the optimal (with respect to the energy) sequence of ”on” and ”off” states for specific pumps.

At the second – the Analysis – stage, we test the LRO policies yielded by the Synthesis stage by running a full extended-period simulation model. These simulations allow, in particular, to evaluate violations (if any) of the constraints on pressure values at every junction of the water network. Note that at the second stage we should, among other things, decide how to distribute the total demand of a period among different junctions; testing different distributions, we get a possibility to investigate the influence of fluctuations of demands in different junctions on the robustness and hydraulic properties of the proposed management policies.

The setup. We have carried out the outlined Synthesis/Analysis procedure for the AnyTown model (Thomas M.Walski et al., 1987) with slightly modified parameters (in particular, we are using the CMH units instead of the GPM units in the original model). The AnyTown water distribution network is shown on Fig. 2. The network includes 19 junctions, three tanks and one pumping station with three identical pumps. The planning period is $T = 24$ hours. The demand pattern corresponding to the AnyTown model is shown on Fig. 3 (left); at the Synthesis stage, we use it as the nominal demand $\{d_t^*\}_{t=1}^T$. The cost coefficients $c_i(t)$ are computed by approximating the function $H(p)p$ (corresponding to the AnyTown model) by linear function const $\cdot\, p$ (see Remark 2.1). The electricity tariff factors $\sigma(t)$, see (4), depicted at Fig. 3, right are by courtesy of the Israel Electric Company. At the Synthesis stage, we replace the AnyTown network with the single-source and single-tank network similar to the one depicted on Fig. 1. In our network, the maximal discharge $P_1(t)$ of the source is 5000 m^3/h for all $t = 1, ..., 24$, and the maximal total amount of the water pumped in over the 24-hour planning horizon is $Q_1 = 50000m^3$. The volume of water in the tank should be at least 1800 m^3 and at most 6560 m^3 (which corresponds to the totals of the minimal, respectively, the maximal volumes of the three tanks in the AnyTown model).

Calculating the LRO policy. We solved problem (10) for the outlined initial data using commercial code MOSEK (`http://www.mosek.com`). To get an impression of the result, we present at Fig. 3 a realization of random demand along with the corresponding trajectories of the tank levels, and the flow rates as given by the LRO policy.

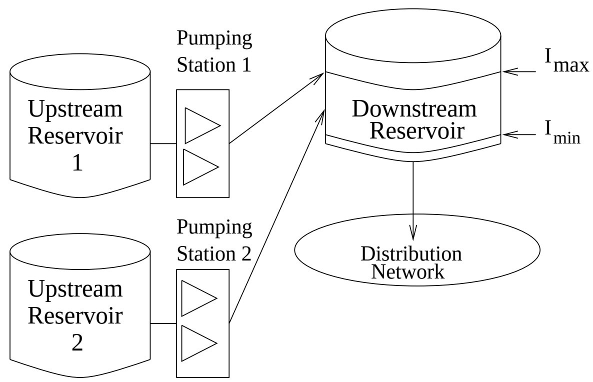


Figure 1: A simple pipeline network

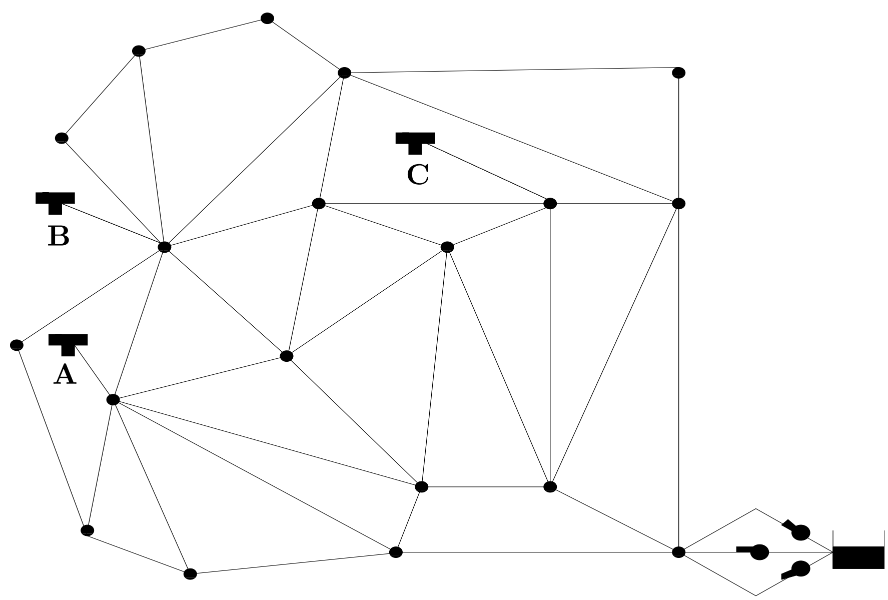
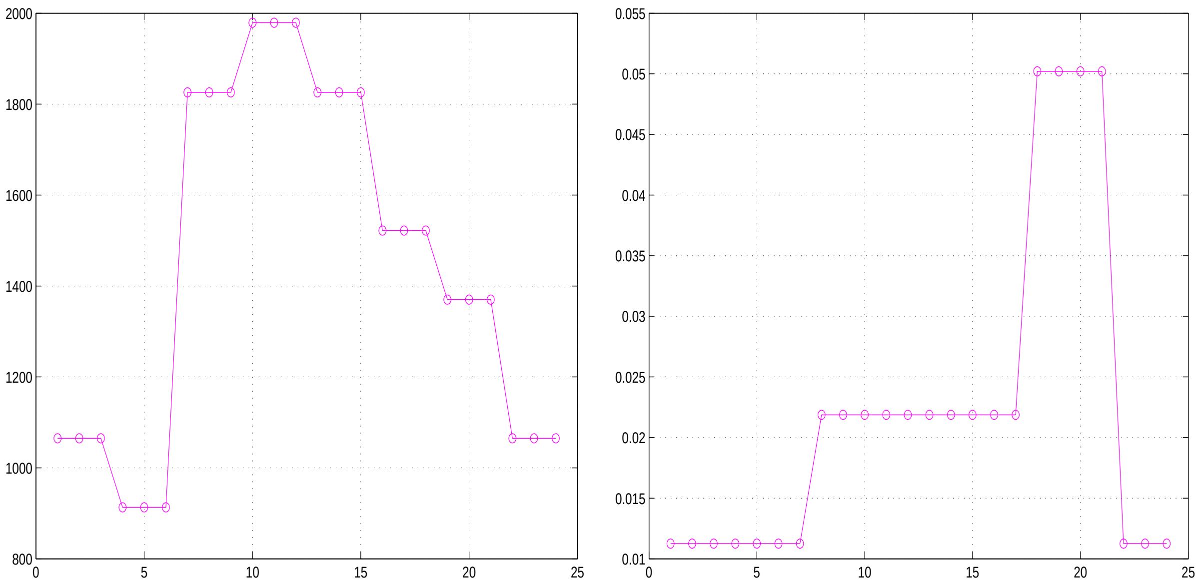
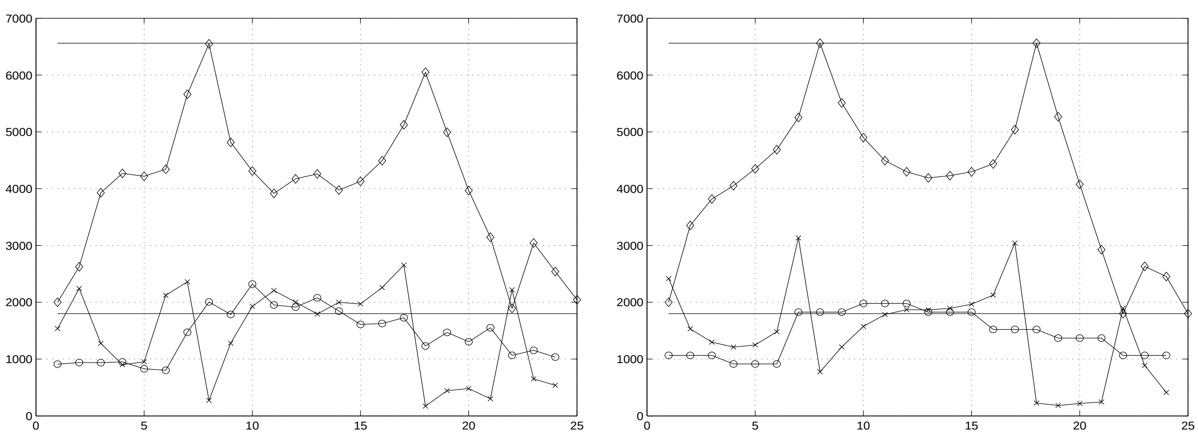


Figure 2: The AnyTown network



(a) (b)

Figure 3: Nominal demand pattern, m^3/h (a), and tariffs pattern, $\$\cdot h/m^3$ (b) vs. time



Policy for 20% uncertainty "Nominal policy" (no uncertainty)

o – demand, m^3/h; x – pump inflow, m^3/h; $\diamondsuit$ – water volume in the tank, m^3

horyzontal lines: bounds on water volume in the tank

Figure 4: Realizations of the LRO policy. It is profitable to accumulate water in the tank when the electricity tariff is low and to use these "savings" to cover as much demand at the "high tariff" hours as possible. This is exactly what the LRO policy does (compare the "low pump inflow" hours 18:00 – 21:00 on the plots with the "tariff peak" hours on Fig. 3 (b)).

The experiments at the Synthesis stage. In every one of the experiments during this stage, the corresponding LRO policy was tested against a given number (100) of simulations; in every one of the simulations, the actual demand d_t of period t was drawn at random, according to the uniform distribution, from the segment $[(1-\theta)d_t^*, (1+\theta)d_t^*]$, where θ was the "uncertainty level" characteristic for the experiment. The demands of distinct periods were independent of each other.

We have conducted two series of experiments:

1. The aim of the first series was to check the influence of the demand uncertainty θ on the total energy costs corresponding to the LRO management policy (recall that this is the policy (8) yielded by the optimal solution to problem (10)). We compared this cost to the "ideal" one, i.e., the cost we would have paid in the case when all the demands were known to us in advance and we were using the corresponding optimal management policy as given by the optimal solution of (5).

2. The aim of the second series of experiments was to check the influence of the "information basis" allowed for the management policy on the resulting management cost. Specifically, in our model as described in the previous section, when making decisions $p_i(t)$ at the beginning of period t, we can make these decisions depending on the demands of periods $r \in I_t$, where I_t is a given subset of the segment $\{1, 2, ..., t\}$. The larger are these subsets, the more flexible can be our decisions, and hopefully the less are the resulting costs. In order to quantify this phenomenon, we considered 7 "information bases" of the decisions:

   (a) $I_t^1 = \{1, ..., t-1\}$ (this *standard information base* seems to be the most natural: past is known, present and future are unknown);

   (b) $I_t^k = \{1, ..., t-k\}$, $k = 2, 3, 4, 6, 8$;

   (c) $I_t = \emptyset$ (i.e., no adjusting of future decisions to actual demands at all).

   Note that the poorest information base corresponds exactly to the policy yielded by the usual Robust Counterpart of our uncertain LP.

The results of our experiments are as follows:

1. The influence of the uncertainty level on the energy cost. Here we tested the LRO policy *with the standard information base* against different levels of uncertainty (from 5% to 30% with the step of 5%). For every uncertainty level, we have computed the average, over 100 simulations, energy cost given by the corresponding LRO policy. We saved the simulated demand trajectories and then used these trajectories to compute the ideal energy costs. The results are summarized in Table 3. As expected, the less is the uncertainty, the closer are our management costs to the ideal ones. What is surprising, is the pretty low "price of reliability": at the 20% uncertainty level, the average, over 100 simulations, energy cost for the LRO policy was just by 3.2% worse than the corresponding ideal cost; the similar quantity for 5%-uncertainty in the demand was was just 0.6%.

2. The influence of the information base on the energy cost. The influence of the information base on the performance of the LRO policy is displayed in Table 3. These experiments were carried out at the uncertainty level 20%. We see that the poorer is the information base of our management policy, the worse are the results yielded by this policy. In particular, it is a must to make the decisions depending on the demands in a not too distant past: at the 20% uncertainty

Uncertainty	LRO policy		Ideal cost		price of reliability
	Mean	Std	Mean	Std	
5%	873.7	6.1	864.6	5.5	1.0%
10%	884.8	11.3	864.1	10.9	2.3%
15%	895.2	16.8	867.1	15.9	3.5%
20%	898.7	22.6	861.3	22.3	4.0%
25%	923.2	25.3	873.8	26.3	6.8%
30%	893.1	32.8	862.8	32.7	7.0%

Table 1: Energy costs vs. uncertainty level, standard information base

information base for decision $p_i(t)$ is demand in periods	energy cost	
	Mean	Std
$1,\dots,t-1$	902	22.2
$1,\dots,t-2$	931	23.3
$1,\dots,t-3$	976	21.3
$1,\dots,t-4$	1013	26.2
$1,\dots,t-6$	1091	25.2
$1,\dots,t-8$	Infeasible	
$\emptyset$	Infeasible	

Table 2: The influence of the information base on the energy costs, uncertainty 20%

level, ignoring the demands of the past 7 periods, robust management becomes impossible[2)]. In particular, with this level of uncertainty, there does not exist a robust *non-adaptive* management policy: the usual *Robust Counterpart* of our uncertain LP is infeasible. In other words, in our illustrating example passing from a priori decisions yielded by RC to "adaptive" decisions yielded by AERC is indeed crucial. The influence of the information level at the management costs is depicted on Fig.3 3.

It is interesting to determine the maximal uncertainty level for which a *non-adaptive* robust policy (the one corresponding to $I_t = \emptyset$) is still possible. To get an answer, we computed the

[2)]This result is easy to predict in advance. Indeed, assume that the information base for the decision on $p(t)$ is $1, ..., t-k$, the uncertainty level is q, and let s, $s+1$,...,$s+k-1$ be a segment of k consecutive hours with total nominal demand d. Since the amount of water we are pumping in during these k hours is independent of the actual demand of the hours, and the latter can take all values between $d(1-q)$ and $d(1+q)$, the changes in the amount of water in the tank during our segment of k hours can be as large as $2qd$. If (9) is feasible, then these changes cannot be more than $V_{\max} - V_{\min}$, so that the necessary condition for (9) to be feasible, the information base at hour t being $1, ..., t-k$, is:

$$2q \max_{0 \le s \le t-k} \sum_{i=1}^{k} d^*_{s+i} \le V_{\max} - V_{\min}.$$

With our data, this condition is satisfied for $k \le 6$ and fails to be true for $k \ge 7$.

Uncertainty	LRO, base $1, ..., t-6$		LRO, base $1, ..., t-8$		LRO, base $\emptyset$	
	Mean	Std	Mean	Std	Mean	Std
25%	infeasible		infeasible		infeasible	
20%	1091	25.0	infeasible		infeasible	
15%	1041	18.5	1103	17.4	infeasible	
10%	977	11.2	1022	12.8	infeasible	
5%	923	6.0	944	6.7	1041	0.0

Table 3: Performance of policies with "poor" information bases

non-adaptive robust policy for uncertainty levels from 5% till 30% with the step 5%. The results are depicted in Table 3. We see that a non-adaptive management becomes possible at the uncertainty level 5% only, and the corresponding energy cost is by nearly 20% worse than the ideal one (compare with just 7%-nonoptimality, at the 30% uncertainty level, for the LRO with the standard information base, see Table 1).

4 EPANET simulation of the Linear Robust Optimal policy

We have seen that the LRO policy is completely acceptable, as far as the simplified linear model of the AnyTown network is concerned. It remains to investigate the most important issues – the hydraulic behaviour and the electric power consumption of the AnyTown network under the LRO-based operation.

Organization of the experiments. We have conducted an extended period simulation of the AnyTown hydraulic model by the EPANET toolkit in order to check whether the LRO operation policy ensures appropriate hydraulic behaviour of the network. The experiments were organized as follows.

1. The pumping station in the AnyTown model is modelled as a fictitious node with *negative* demand, and the net inflow $p(t)$ at this node is as prescribed by the LRO policy;

2. The energy cost is calculated as

$$\sum_{t=1}^{24} \sigma(t)\kappa e^{-1} H(t)p(t)$$

   (cf. (4)), with the head H as reported by the EPANET for the fictitious node and the electrical tariff $\sigma(t)$ as shown on Fig. 3 left. The efficiency coefficient e was set to 0.5.

3. For each day of the 200-day simulation period we have designed a *random* demand pattern. The hour demand multipliers $m(t)$ from the AnyTown deterministic demand pattern are perturbed at random according to the uniform distribution on the segment $[(1-\delta)*m(t),(1+\delta)*m(t)]$, where the uncertainty level δ was varied from experiment (i.e., 200-day simulation) to experiment in the range [0,0.5]. Besides randomly perturbing *common for all junctions* hour demand pattern, we perturbed at random, on a day-by-day basis, the *basic* demands

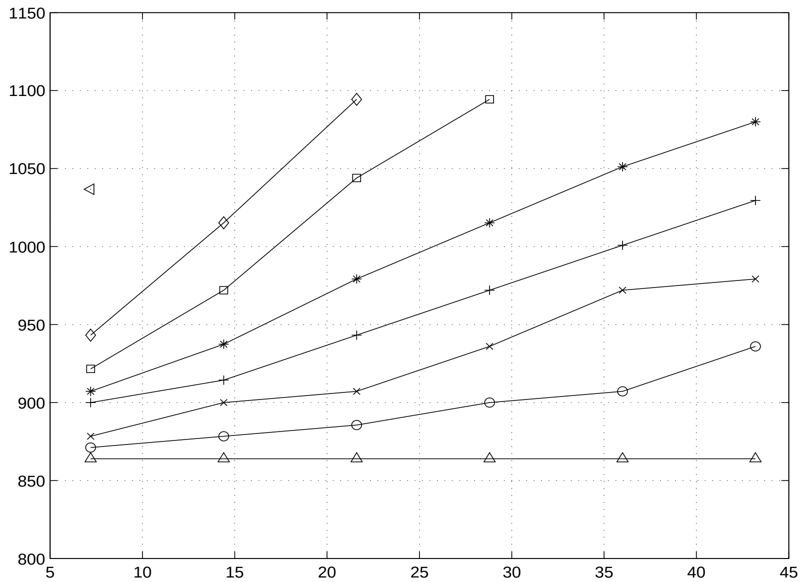


Figure 5: Management costs ($) vs. uncertainty levels (%)
Δ – ideal cost o – base I_t^1 x – base I_t^2 $+$ – base I_t^3
$*$ – base I_t^4 $\square$ – base I_t^6 $\diamond$ – base I_t^8 $\triangleleft$ – base $\emptyset$

at the junctions. The perturbations of the basic demands were independent of each other and uniformly distributed in the 40%-neighbourhoods of the nominal demands' values. Thus, there were two sources of uncertainty: the temporal and the spatial ones.

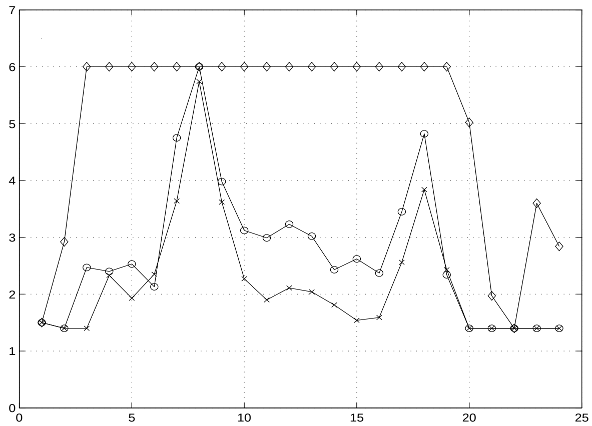
4. The LRO policy used in the EPANET simulation of the AnyTown model was given by (10), where the "uncertainty level" θ was set to 20%. In the first series of experiments, the temporal and the spatial uncertainties were adjusted to yield no more than 20% uncertainty in the resulting total demands. Then the levels of temporal and spatial uncertainties (and, consequently, the uncertainty in total demands) were gradually increased in order to find out when the LRO operation policy will be "crushed".

5. During simulations, we record the warning messages issued by EPANET (like "negative pressure", "hydraulic instability", etc.);

Simulation results for the AnyTown hydraulic model. To give an impression of the LRO policy, we present on Fig.6 4 a sample 24-hour pattern of tanks' levels. The quantitative simulation results are as follows:

1. For the AnyTown water network, the LRO operation policy at the 20% uncertainty level is absolutely reliable from the hydraulic viewpoint (in course of our simulations, we never observed a *single* EPANET warning).

2. When increasing the level of uncertainty beyond 20%, the first EPANET warnings appear when the perturbations in the nominal demand become as large as 50%. Specifically, the following model of perturbations was used:



o – tank A; x – tank B; ◊ – tank C

Figure 6: A sample pattern of tank levels for the LRO policy in the AnyTown model (demand uncertainty 20%).

ϵ \ k	0.45	0.5	0.6	0.8	0.9
0.01	0	2	2	2	2
0.1	0	11	11	11	12
0.2	0	8	8	8	13
0.4	0	16	16	16	32
0.6	0	21	21	19	60

Table 4: Number of days with negative pressure vs. probability of large fluctuation (ϵ) and magnitude of large fluctuation (k).

- with probability $1 - \epsilon$, the perturbation in the common for all junctions demand pattern for a particular hour t was drawn from the uniform distribution on $[0.8d_t^*, 1.2d_t^*]$ (recall that d_t^* is the nominal demand pattern);
- with complementary probability ϵ, the hour demand pattern at hour t was chosen as $(1 \pm \kappa)d_t^*$, with "big" κ and equal probabilities for positive and negative perturbations.

In our experiments, we played with both ϵ ("probability of large fluctuation") and κ ("magnitude of large fluctuation"), trying to understand when the LRO operation policy (computed under the assumption that the uncertainty never exceeds 20% of the nominal demand) fails to be hydraulically valid. Surprisingly, the "crush" of the LRO policy turned out to depend solely on κ: with $\kappa = 0.45$, no warnings were observed even with the probability of large fluctuation $\epsilon = 1$, while with $\kappa = 0.5$, negative pressures were observed already for $\epsilon = 0.01$, and their frequency, as it should be, increased nearly proportionally to ϵ (see Table 4).

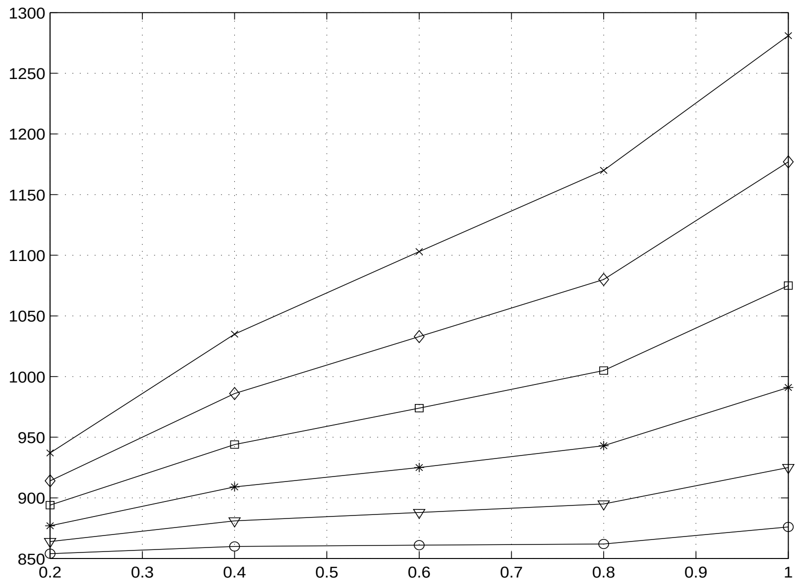


Figure 7: Energy cost for the AnyTown model vs. probability of large fluctuation ϵ
o: $k = 0.2$ ∇: $k = 0.3$ $*$: $k = 0.4$ $\Box$: $k = 0.5$ $\diamond$: $k = 0.6$ $\times$: $k = 0.7$

3. The average energy cost grows significantly with ϵ and κ even in the range $\kappa < 0.5$, as can be seen from Figure 3.

5 Summary and Conclusion

Optimization of design and operation in water distribution systems is a very suitable field for the robust optimization methodology. Indeed, almost all relevant data - actual diameters and lengths of the pipes, electricity and hydraulic characteristics of the pumps, not speaking about the demands - are heavily affected by uncertainty. For example, the variability of the actual demand in Hadera (Israel), i.e., the ratio of the standard deviation to the mean, both conditioned to a particular hour within a day and a particular "type of the day" (day off/working day), varies from 5.5% (3^{00}-4^{00} am, working day) to 37% (5^{00}-6^{00} pm, working day). (We acknowledge the courtesy of E. Salomons who provided us with the corresponding measurements, taken with 15 min. intervals in the period from 04/01/2001 till 12/30/2001). These data make our assumption on the 20% demand uncertainty quite reasonable.

In which cases the methodology described above would be most promising?

1. It is well known that some real-world water distribution network have simple layout (single tank and several sources with a pumping station for each source). For these networks the LRO policy could be used quite straightforwardly.

2. The LRO policy could be used for performance evaluation both for optimal control of an existing water distribution network and when designing a new network;

3. If and when the simplifying assumptions which we made in order to reduce a real hydraulic

model to a linear program are justified, the Robust Optimization Methodology could be used straightforwardly.

Consider the last statement in more detail. Linear programming models of real-world water distribution network are not unusual. The assumptions on the hydraulic characteristic of a network allowing to reduce the optimal pump scheduling problem to a linear program are formulated, e.g., in (Jowit and Germanopulos 1992). The assumptions are as follows:

- Water distribution network is a well-designed network, where internal network pressure remains within acceptable bounds for allowable service reservoir storage fluctuations;
- There is a large head lift for each pump station compared to the network nodal head changes induced by pumps/valves switchings elsewhere in the system;
- A flow a given pumping station will deliver to different zones depends solely on zonal consumer demands, and not on the changes in the network head/flow pattern caused by pumps/valves switchings elsewhere in the system.

From the discussion in (Jowit and Germanopulos 1992) it is clear that these assumptions are valid in many practical situations. It is worth of noting that the first assumption above is a proper explanation of the hydraulic stability in the AnyTown simulation. If and when the controlling setting can keep storage fluctuations within reasonable bounds, a "well-designed" water network has no problems with nodal pressures. Vice versa, it is natural to define a "well-designed water network" as a network where internal pressures can be kept within acceptable bounds at a reasonable energy cost, *for all demand fluctuations of allowed magnitude.* With this is mind, the outlined methodology can be used at the design stage in order to evaluate the quality of a would-be network.

It is known that besides the costs associated with energy consumption, the energy bill can include the cost associated with the maximum power drawn. As usual, the maximum power drawn is defined as the maximum power consumption registered during any 15-minute period, during the billing period. If this term is included in model's objective function as such, the objective function becomes nonlinear. However, in our LP model we could impose upper bounds on the maximum power drawn via the restrictions on the pump performance, see (5).

Acknowledgments. The authors are greatly indebted to Prof. Uri Shamir, Head of the Technion Water Research Institute, and Dr. Elad Salomons, for stimulating discussions.

References

[1] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, *Robust Optimization*, Princeton University Press, 2009.

[2] A. Ben-Tal, A. Nemirovski, "Stable Truss Topology Design via Semidefinite Programming" – *SIAM Journal on Optimization* **7** (1997), 991-1016.

[3] A. Ben-Tal, A. Nemirovski, "Robust solutions to uncertain linear programs" – *OR Letters* **25** (1999), 1-13.

[4] A. Ben-Tal, A. Nemirovski, "Robust Optimization — Methodology and Applications", *Mathematical Programming* Series B, **92** (2002), 453–480.

[5] A. Ben-Tal, A. Nemirovski, "Robust Convex Optimization" – *Mathematics of Operations Research* **23** (1998).

[6] A. Ben-Tal, A. Nemirovski, "Robust solutions of Linear Programming problems contaminated with uncertain data" – *Mathematical Programming* **88** (2000), 411-424.

[7] A. Ben-Tal, A. Goryashko, E. Gustlizer, A. Nemirovski (2002), "Adjustable Robust Counterparts of uncertain Linear Programming problems" – Research Report #2/2002, June 2002, Minerva Optimization Center, Faculty of Industrial Angineering and Management, Technion - Israel Institute of Technology, Technion, Technion City, Haifa 32000, Israel.

[8] R. DeMoyer, L. Horowitz, *A system approach to water distribution modeling and control*, Lexington Books, Lexington, Mass., 1975

[9] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, "Robust Semidefinite Programming" – in: R. Saigal, H. Wolkowitcz, L. Vandenberghe, Eds. *Handbook on Semidefinite Programming*, Kluwer Academis Publishers, 2000.

[10] J. D. Creasy "Pump scheduling in water supply: more than a mathematical problem", Computer application in water supply, vol. 2, (1998), B. Coulbeck and C. H. Orr, Eds., Research Studies Press Ltd. 279-289

[11] J. A. S. Ferreira, R. V. V. Vidal "Optimization of a pump=pipe system by dynamic programming.", *Engineering Optimization*, vol. 7, (1984), 241-251

[12] S. Chandrasekaran, G. H. Golub, M. Gu, A. H. Sayed, "Parameter estimation in the presence of bounded data uncertainty" – *Journal on Matrix Analysis and Appl.* **19** (1998), 235-252.

[13] B. Coulbeck, M. J. H. Sterling "Optimized control of water distribution systems", *Proc., IEE*, London, England, 125 (9), (1978), 1039-1044

[14] G. Eiger, U. Shamir, "Optimal operation of reservoirs by stohastic programming", *Engrg. Optimization*, 17, (1991), 293-312

[15] L. El-Ghaoui, H. Lebret, "Robust solutions to least-square problems with uncertain data matrices" – *SIAM J. of Matrix Anal. and Appl.* **18** (1997), 1035-1064.

[16] L. El-Ghaoui, F. Oustry, H. Lebret, "Robust solutions to uncertain semidefinite programs" – *SIAM J. on Optimization* **9** (1998), 33-52.

[17] Jacobs, J. et al. "SOCRATES: a system for scheduling hydroelectric generation under uncertainty", *Ann. of Operations Res.* , 59, (1995), 99-134

[18] G. Jolland., G. Cohen, "Optimal control of water distribution network by two multilevel methods", *Automatica*, vol. 16 (1980), 83-88

[19] P. W. Jowitt, G. Germanopoulos, "Optimal Pump Scheduliing in Water-Supply Networks", *Journal of Water Resources Planning and Management*, vol. 118, 4, (1992), 406-422.

[20] K. E. Lansey, K. Awumah, "Optimal pump operations considering pump switches", *Journal of Water Resources Planning and Management*, vol. 120, 1, (1994), 241-251.

[21] M. Likeman, "Constraint satisfaction methods in water supply scheduling", *Integrated computer applications in water supply*, vol. 1, 1993, B. Coulbeck, Ed., Research Studies Press Ltd., 213-225.

[22] S. M. Moss, "On-line optimal control of a water supply network", *PhD thesis, University of Cambridge, Cambridge, England* (1979)

[23] J. M. Mulvey, R. J. Vanderbei, S. A. Zenios, "Robust optimization of large-scale systems", *Operations Research* **43** (1995), 264-281.

[24] L. E. Ormsbee, L. S. Reddy, D. V. Chase, "Comparison of three nonlinear control algorithms for the optimal operation of water supply pumping systems." – in: B. Coulbeck, Ed., Integrated computer applications in water supply, Vol. 1, Research Studies Press Ltd., 1993, 259-271.

[25] S. Pezeshk, O. J. Helweg, "Optimization of pumping costs for well-field water distribution systems." – in: B. Coulbeck, Ed., *Integrated computer applications in water supply*, Vol. 1, Research Studies Press Ltd, 1993, 227-238.

[26] A. B. Sakarya, F. E. Goldman, L. W. Mays, "Models for the optimal scheduling of pumps to meet water quality." – in: D. A. Savic and G. A. Walters, Eds., Water industry systems: modeling and optimization applications, Vol. 2, Reserch Studies Press Ltd. 1999, 379-391

[27] A. R. Simpson, D. C. Sutton, D. S. Keane, S. J. Sheriff, "Optimal control of pumping at a water filtration plant using genetic algorithms", Water industry systems: modeling and optimization applications, 1999, D. A. Savic and G. A. Walters, Eds., Vol. 2, Reserch Studies Press Ltd. 407-415

[28] J. J. de Oliveira Sousa, M da Conceicao Cunha and J. A. de Almedia Sa Marques, "Optimal Pumping Scheduling Model for Energy Cost Minimization", Water Software Systems: theory and applications Volume two, 2001, Edited by B. Ulanicki, B. Coulbeck, J. Rance, Reserch Studies Press Ltd. 99-109

[29] A. L. Soyster, "Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming". - *Operations Research* (1973), 1154-1157.

[30] A. J. Tarquin, J. Dawdy, "Optimal pump operation in water distribution", *Journal Hydr. Engrg.,* ASCE, **115** (1989), 158-168.

[31] J. G. Uber, E. D. Brill Jr., J. T. Pfeffer, "Use of mathematical programming methods for complex system.". -*Journal of Water Resources Planning and Management* **118** (1992), 281-294.

[32] T. M. Walsky et al. "Battle of the Network Models: Epilogue". -*Journal of Water Resources Planning and Management*, vol. 113, 2, (1987), 191-203.

[33] T. M. Walsky. Analysis of Water Distribution System – Van Nostrand Reinhold Company, NY, 1984.

[34] D. W. Watkins Jr., D. C. McKinney, "Finding Robust Solution to Water Resources Problems", *Journal of Water Resources Planning and Management*, vol. 123, 1, (1997), 49-58.

[35] U. Zessler, U. Shamir "Optimal operation of water distribution systems", *Journal of Water Resources Planning and Management*, vol. 115, 6, (1989), 735-752.