Water Resources Research

RESEARCH ARTICLE

10.1029/2023WR035508

Key Points:

- Optimal operation of water distribution systems under uncertainty is addressed with a dynamic, data-driven approach
- An adjustable robust optimization model is developed to provide a decision rule that adapts to the latest revealed information
- Comparative analysis of the method against traditional folding horizon methods demonstrates superior performance

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

G. Perelman,
gal-p@campus.technion.ac.il

Citation:

Perelman, G., & Ostfeld, A. (2023). Adjustable robust optimization for water distribution system operation under uncertainty. *Water Resources Research*, *59*, e2023WR035508. https://doi.org/10.1029/2023WR035508

Received 6 JUN 2023
Accepted 21 NOV 2023

Adjustable Robust Optimization for Water Distribution System Operation Under Uncertainty

Gal Perelman[1] and Avi Ostfeld[1]

[1]Faculty of Civil and Environmental Engineering, Technion—Israel Institute of Technology, Haifa, Israel

Abstract The optimal operation of water distribution systems (WDS) is a paramount objective for water utilities due to the substantial energy consumption associated with pumping. A major challenge in optimizing WDS operation is addressing uncertainties such as those related to consumer demands. Real-time operation under uncertainty necessitates a dynamic approach that can utilize the newly observed information and adjust the operational policy accordingly. This study presents an adjustable robust optimization (ARO) approach to tackle this challenge. Unlike static optimization methods, ARO generates a decision rule policy that is dynamically adjusted as new data becomes available and the operational horizon evolves, thereby ensuring adaptability to changing conditions. Furthermore, the study includes a quantitative analysis of typical demand uncertainty that supports the formulation of the ARO model. The proposed method is evaluated through two case studies and compared with traditional folding horizon approaches. The results indicate that the ARO method is competitive with traditional methods in terms of objective value and surpasses them in terms of robustness. An additional advantage of the method is its offline operation capability which enables it to produce decision rules independent of real-time programs. This feature facilitates various practical applications such as what-if analyses, maintenance work planning, and preparation for other special events.

Plain Language Summary The study is focused on developing strategies to optimize the operation of water distribution systems under uncertainty. Pumping and distributing water consumes a lot of energy, therefore water utilities strive to optimize the system's efficiency. This optimization problem involves uncertainties, such as consumer water demands. To address this challenge, a method is developed based on adjustable robust optimization (ARO) theory. Unlike conventional optimization methods that produce a fixed policy for the entire operational horizon, ARO generates an adjustable policy as a decision rule. It creates a dynamic decision-making policy that can adapt to new data and evolving conditions. The study also includes an analysis to quantify the typical uncertainty associated with water demand which supports the formulation of the ARO model and justifies parts of the theory behind it. The proposed method is tested on two case studies and compared with traditional approaches. Several sensitivity analyses were held to present the ARO advantages and compare it with other methods. The results showed that ARO performs competitively with other methods and outperforms some of them.

1. Introduction

Water distribution systems (WDS) are critical infrastructure assets as they play a vital role in facilitating various human activities. Their purpose is to ensure a continuous and reliable water supply with sufficient quantities, pressure, and quality. The operation of WDS consumes significant amounts of energy (Sharif et al., 2019) with substantial implications for both the economy and the environment (Dziedzic & Karney, 2015; Salomons & Housh, 2020). Consequently, it is essential to develop methodologies to optimize the operation of WDS. Indeed, the problem has been researched extensively in the past decades (Lansey, 2007; Mala-Jetmarova et al., 2017). While academic literature offers numerous approaches to minimize energy costs in WDS operation, these are rarely adopted by water utilities. Rao and Salomons (2007) identified several barriers to the practical integration of optimization models. Such models are often very complex and require a specialized technical skill set. Additionally, networks' complexity can limit the applicability, leading to oversimplified models on the one hand or to long run times and local minima on the other hand. Moreover, most techniques aim at solely minimizing cost, overlooking other objectives such as network performance and robustness. Robustness is usually associated with uncertain factors of the optimization problem and lately is recognized as a key factor in WDS management.

In recent years there has been a growing recognition that certain aspects of WDS management are subject to uncertainty, including aspects regarding the operation of WDS (Hutton et al., 2012; Maier et al., 2016). Sources of uncertainty in WDS operation have been reviewed by Dandy et al. (2022) and include consumer demands, system state (e.g., tank levels, failures), electricity prices, and hydraulic parameters (e.g., roughness coefficients).

A common strategy to address uncertainties involves forecasting uncertain factors and then solving a deterministic problem. However, this approach can result in a sub-optimal or even infeasible operation plan if the realized scenario differs from the estimated one. Estimating uncertain parameters can present further difficulties where not every uncertainty is predictable and in some cases, there is not available adequate data.

Another approach is to set a steady policy that is independent of the uncertain parameters, thereby guaranteeing the feasibility of the solution. In the realm of pump scheduling this can be done by setting tanks' levels as triggers for pumps to turn on and off (Housh & Salomons, 2019; Quintiliani & Creaco, 2019). Such an approach ties the schedule of each pump to a single tank and misses the system as a whole. Moreover, it struggles to adapt to the system's dynamic nature, reflected in the variability across different days, weeks, or months. Thus, frequently necessitates adjustments to the operational policy. One of the most prevalent strategies to tackle uncertainty is to reduce the uncertain space by delaying future decisions to the time these decisions need to be implemented (Castelletti et al., 2023). This framework, which is often referred to as folding horizon or model predictive control (MPC), involves repeatedly solving optimization problems with the operational horizon constantly shifting forward. After each optimization cycle, decisions for the current time are implemented, the horizon is shifted one time step forward, and a new solution is obtained before the next time step begins. The advantage of MPC lies in its ability to avoid implementing decisions at the edge of the horizon where forecasts are less reliable. Several studies suggested MPC for WDS operation. Fiorelli et al. (2013) optimized the operation of a small network in Luxembourg by incorporating a dynamic horizon that extends until the end of the current day along with a simple demand forecast that is based on a constant demand pattern. Another MPC study conducted by Grosso et al. (2017) employed ARIMA time-series analysis for demand forecasting. Acknowledging that forecasts themselves can contain errors that might lead to infeasibility, Grosso et al. (2017) extended their MPC model to incorporate forecast errors by introducing two different stochastic approaches. One approach involves chance constraint (CC) relaxation ensuring that the constraints' violation rate stays below a predefined threshold. The second approach employs a scenarios tree. Both methods are integrated into the MPC control loop such that a new uncertain problem is solved in each time step.

A different approach to tackling optimization problems under uncertainty is robust optimization (RO), introduced by Ben-Tal et al. (2009). RO adopts a worst-case min-max optimization framework such that a feasible solution is guaranteed for every scenario within a defined range commonly referred to as an uncertainty set. To mitigate the over-conservatism that stems from a worst-case approach, Ben-Tal et al. (2009) proposed using an Ellipsoid representation of the uncertainty set, effectively pruning the most extreme scenarios. RO offers several key advantages. First, the method does not require complete probabilistic knowledge of the random variables (e.g., probabilistic density functions), where uncertainty sets can be constructed based on mean and standard deviation (STD) alone. Additionally, RO can provide a global optimum solution within short run times even for large-scale problems. Several studies suggested RO to uncertain optimization problems in the water resources domain (Housh, 2011, 2017; Pankaj et al., 2022; G. Perelman et al., 2023; L. Perelman et al., 2013; Schwartz et al., 2016). However, RO has a static nature, such that all the decisions are made at the beginning of the operational horizon. As such, RO suffers from several drawbacks. One drawback is that as decisions are made far into the future, increasing levels of uncertainty must be considered. Second, the approach fails to incorporate newly revealed, relevant information contrasting with the MPC approach which constantly updates based on the latest information. To address these drawbacks, adjustable robust optimization (ARO) was introduced by Ben-Tal et al. (2004), and offers a dynamic RO approach that incorporates new realized data through the implementation of the solution. ARO is gradually being applied to a wide range of problems and demonstrating promising results.

The main feature of ARO is the ability to adjust certain decision variables after the revelation of part of the uncertainty. The method categorizes decision variables into two types: "here and now" and "wait and see." The latter type of variables depends on values of the realized uncertainty, or in other words, these variables are adjusted according to the new incoming data. ARO has been applied to various dynamic operation problems such as power grid extension (Moreira et al., 2015), renewable energy dispatch (Li et al., 2015), inventory management (See & Sim, 2009), hydro-electric reservoir operation (Pan et al., 2015), flood protection (Postek et al., 2019) and also for WDS operation (Goryashko & Nemirovski, 2014) and presented high performance in all of them.

10.1029/2023WR035508

Real-time operation of WDS is studied mainly through MPC frameworks, with and without uncertainty (Castelletti et al., 2023). ARO suggests a novel dynamic approach that is based on generating an optimized decision rule that can utilize real-time measured data. Moreover, the decision rule is designed to not only optimize operational costs but also to guarantee the feasibility of the solution against a range of scenarios. Decision rules can be a powerful tool for real-time control of large complex systems, they avoid heavy computations that must be done in short run times and are less expected to fail due to missing or faulty data. However, until now the use of decision rules in WDS operation is very limited and stayed around simple control schemes such as tank level triggers (Housh & Salomons, 2019; Quintiliani & Creaco, 2019). Here a more advanced decision rule is presented with the advantage of optimizing the systems as a whole, considering uncertainty, and without compromising the optimality of the solution.

The current study aims to develop an ARO based model for optimizing pump scheduling and minimizing energy costs. The proposed method is illustrated through two case studies and is compared with a traditional MPC approach to explore the respective advantages and disadvantages of the two methods regarding operation under demand uncertainty.

2. Methodology

Optimizing the operation of WDS entails solving a scheduling problem with the objective of minimizing operational costs (e.g., energy expenses). This is achieved by determining the optimal schedule for the network pumps and valves. The schedule must consider not only the immediate moment but also a future operational horizon. This is essential for adapting to dynamic conditions such as electricity prices, future water demands, the availability of water sources and pumps, etc. Typically, the operational horizon is discretized into equal intervals of a fixed duration (e.g., 1 hr). Model parameters are explicitly defined for each time step, and decisions can be made based on the same division. To account for the lack of complete knowledge regarding some of the parameters, random variables denoted ξ are introduced. The random variables can get different values to represent the range of different possible scenarios. It is noted that although ξ is referred to as a random variable, its probability density function (PDF) may not necessarily be known. However, PDFs are not required for the proposed method as will be explained below. A general statement of an uncertain optimization problem can be described as follows:

Let x be a vector of decision variables, and ξ a vector of uncertain parameters. Let f be the objective function to minimize, and g be a set of constraints. A general formulation of an uncertain optimization problem can then be represented as:

$$Z = \min_{x}\{f(x,\xi) : g(x,\xi) \le 0\} \tag{1}$$

As mentioned earlier, RO theory is grounded in considering the worst-case realization of ξ. To account for the worst-case in both the objective and the constraints, the functions f and g are maximized over the uncertainty space ξ. For that purpose, inner optimization problems are formulated, with ξ as the decision variables, where ξ is bounded within an uncertainty set U. Maximizing g over ξ ensures that the set of constraints will be satisfied for any decision x regardless of the actual realization of ξ. Similarly, maximizing f over ξ will yield a worst-case optimal objective that serves as an upper bound of the solution. This worst-case problem (Equation 2) is referred to as Robust Counterpart (RC) which is a deterministic optimization problem and thus can be solved with standard methods. For a more comprehensive overview of the RO theory, the reader is referred to Ben-Tal et al. (2009). For a more detailed description of the RO application for WDS optimization, see (G. Perelman et al., 2023).

$$Z = \min_{x}\left\{\max_{\xi\in U} f(x,\xi) : \max_{\xi\in U} g(x,\xi) \le 0\right\} \tag{2}$$

Problem (Equation 2) is a static RO where all the decision variables are determined at the beginning of the operational horizon. To integrate adjustable variables that can adapt to new information, the problem variables are replaced with affine functions. The purpose of this replacement is to distinguish between "here and now" decisions and "wait and see" decisions where the latter type depends on the observed values from previous time steps. ARO theory suggests limiting this dependency to affine functions in order to maintain the problem tractability (Yanıkoğlu et al., 2019). Here, linear decision rules (LDR) are used as affine functions to model the "wait and see" adjustable variables. Every decision variable from the original set is transformed into an LDR, which depends on the observed values of the uncertain parameter. For example, if x is a vector of decisions that need to be made for every period (t) of the problem, it will be reformulated as a series of constants (π) multiplied by the uncertainty of previous time steps until $t-1$:

$$x_t = \pi_t^0 + \sum_{r=1}^{t-1} \pi_t^r \xi_r \tag{3}$$

After the above transformation, Equation 3 introduces the LDR coefficient denoted as π as the new decision variables of the problem. Consequently, the optimization solution becomes a decision rule rather than a static decision, with π as the decision rule weights. The operational policy can be adjusted by the rule according to the obtained values of the uncertain parameter. By using a linear rule, the tractability of the original problem is preserved. In the case of linear problems, the final RC will be a linear or second order cone problem according to the uncertainty set (polyhedral and ellipsoid respectively) (Yanıkoğlu et al., 2019).

2.1. WDS Optimal Operation

The optimal operation of WDS can be formulated as a linear program (LP) based on the above theoretical framework. The following formulation is general for any system where examples for illustrative and real-life case studies are presented in Figures 3 and 6 respectively. The formulation takes into account both fixed speed pumps (FSP) and variable speed pumps (VSP). It is assumed that FSP pumping stations can be operated within a finite number of states (pump combinations), thus there exists a finite number of states the system can be in Jowitt and Germanopoulos (1992). Accordingly, a continuous decision variable is assigned to every pump station (PS) state at every time step, to represent the duration portion this state is operated. Only one state of every station can be operated at a given time. In the case of VSP, a continuous decision variable is assigned to every time step (t) to represent the flow of the VSP at that time step. Each FSP and VSP variable is associated with storage tanks as inflow and outflow according to the network topology. Let $x_{t,i}^{\mathrm{FSP}}$ be the portion of time step (t) where state (i) is operated; the flow and power of state (i) are noted as Q_i, P_i. Each PS consists of a subset of all its FSP states. Let $x_{t,j}^{\mathrm{VSP}}$ be the flow of VSP j at time step (t), the power of VSP pumps is a function of the VSP flow: $P_j\left(x_{t,j}^{\mathrm{VSP}}\right)$. Elec_t is the electricity tariff at time (t). $\overline{d}_{t,s}$ is the nominal demand at time (t) from tank (s), and $d_{t,s}$ is the actual demand such that the random deviation is defined as $\xi_{t,s} = \left(d_{t,s} - \overline{d}_{t,s}\right)$. The duration of each time step is noted Δt (1 hr), S is the set of all storage tanks, and T is the number of time steps. Pumps that deliver water into tank (s) are noted *Inflow* as follows: $\mathrm{x}_{\mathrm{t,i}}^{FSP} \in s_I$. Similarly, *Outflow* pumps are noted $x_{t,i}^{\mathrm{FSP}} \in s_O$.

2.1.1. Objective Function

The objective function of this problem is minimizing the cost of energy. In the below formulation, the objective function is transformed into a constraint by introducing a new variable τ. Equation 5 asserts that the total energy cost is smaller or equal to τ where according to the objective, τ is minimized. Accordingly, the actual term to minimize is the energy cost. The motivation for this transformation is to simplify the subsequent ARO formulation by employing a deterministic objective function that does not explicitly include the decision variables.

2.1.2. Constraints

Equations 6 and 7 describe the mass balance for each tank and each time step, constraining the tank volumes to be within the min and max allowable limits. Equation 8 ensures that the final volume in each tank is at least equal to the initial volume. Equation 9 restricts the duration of each decision variable to no longer than a single time step and stipulates that only one pump's combination can be active at any given moment. Considering the above notation, the optimal operation of WDS can be described as follows:

$$Z = \min_{x,\xi \in U} \tau \tag{4}$$

Subject to:

$$\sum_{t=1}^{T} \mathrm{Elec}_t \cdot \Delta t \left[\sum_{i=1} x_{t,i}^{\mathrm{FSP}} \cdot P_i + \sum_{j=1} P_j\left(x_{t,j}^{\mathrm{VSP}}\right) \right] \leq \tau \tag{5}$$

10.1029/2023WR035508

$$\sum_{t=1}^{\tilde{t}}\sum_{i\in s_O} x_{t,i}^{\mathrm{FSP}} Q_i - \sum_{t=1}^{\tilde{t}}\sum_{i\in s_I} x_{t,i}^{\mathrm{FSP}} Q_i + \sum_{t=1}^{\tilde{t}}\sum_{j\in s_O} x_{t,j}^{\mathrm{VSP}} - \sum_{t=1}^{\tilde{t}}\sum_{i\in s_I} x_{t,j}^{\mathrm{VSP}} \le v_{0,s} - v_{s,\min} - \sum_{t=1}^{\tilde{t}} d_{t,s}\Delta t \; \forall \tilde{t} = 1\ldots T,\ \forall s \in S,\ \forall \xi_t \in U \tag{6}$$

$$\sum_{t=1}^{\tilde{t}}\sum_{i\in s_I} x_{t,i}^{\mathrm{FSP}} Q_i - \sum_{t=1}^{\tilde{t}}\sum_{i\in s_O} x_{t,i}^{\mathrm{FSP}} Q_i + \sum_{t=1}^{\tilde{t}}\sum_{j\in s_I} x_{t,j}^{\mathrm{VSP}} - \sum_{t=1}^{\tilde{t}}\sum_{i\in s_O} x_{t,j}^{\mathrm{VSP}} \le v_{s,\max} - v_{0,s} + \sum_{t=1}^{\tilde{t}} d_{t,s}\Delta t \; \forall \tilde{t} = 1\ldots T,\ \forall s \in S,\ \forall \xi_t \in U \tag{7}$$

$$\sum_{t=1}^{T}\sum_{i\in s_O} x_{t,i}^{\mathrm{FSP}} Q_i - \sum_{t=1}^{T}\sum_{i\in s_I} x_{t,i}^{\mathrm{FSP}} Q_i + \sum_{t=1}^{T}\sum_{j\in s_O} x_{t,j}^{\mathrm{VSP}} - \sum_{t=1}^{T}\sum_{i\in s_I} x_{t,j}^{\mathrm{VSP}} \le v_{0,s} - v_{T,s} - \sum_{t=1}^{T} d_{t,s}\Delta t \; \forall s \in S,\ \forall \xi_t \in U \tag{8}$$

$$0 \le \sum_{i\in \mathrm{PS}} x_{t,i}^{\mathrm{FSP}} \le 1 \quad \forall t = 1\ldots T, \forall \mathrm{PS} \tag{9}$$

The inclusion of the auxiliary parameter $\tilde{t}$ in some of the constraints allows for the computation of cumulative flows and demands, which in turn results in the tank volumes at each time step (t). The purpose of this parameter is to sum variables from the optimization start ($t = 0$) to every time step until the optimization end ($t = T$). This notation supports the calculation of accumulated flows and demands that dictate the tanks' volumes at time step (t).

Next, the problem is converted to its adjustable form by replacing the decision variables with LDR. The LDR presented in Equation 3 is initially set up to depend on a single uncertain parameter. However, in real problems, usually there are multiple uncertain factors such as multiple consumers in WDS where each consumer has different uncertainty characteristics. It is desired that the decision variables will adopt information from all the consumers in the network, even those that are not topologically linked to a specific decision variable (pump). This is because the network is operated as an integrated system where different components are affecting each other. Therefore, a new formulation for LDR is required that can accommodate multiple uncertainties. We propose a 'multi-uncertainty LDR' that allows decision variables to be functions of all identified points of uncertainty across the network:

$$x_{t,i}(d) = \pi_{t,i}^{0} + \sum_{s=1}^{S}\sum_{r=1}^{t-1} \pi_{t,i}^{r}\left(d_{r,s} - \overline{d}_{r,s}\right) \tag{10}$$

$$d_s \in U_s, \quad \forall s \in S \tag{11}$$

With the improved LDR (Equation 10), the "wait and see" operation decisions for a given pump are adjusted not just by information from consumers in the discharge of the pump but according to data arriving from all the parts of the network. Thus, maximizing the utility of the revealed data. For ease of notation, and to mark that from now on, the decision variables x are dependent on the actual demands d, they will be noted $x_{t,i} = x_{t,i}(d)$ according to the LDR presented in Equation 10. To derive the deterministic RC problem, the decision variables from the problem in Equations 4–9 are replaced with the adjusted variables.

$$Z = \min_{x,d\in U} \tau \tag{12}$$

Subject to:

$$\max_{d\in U}\left[\sum_{t=1}^{T} \mathrm{Elec}_t \cdot \Delta t\left(\sum_{i=1} x_{t,i}^{\mathrm{FSP}}(d)\cdot P_i + \sum_{j=1} P_j\left(x_{t,j}^{\mathrm{VSP}}(d)\right)\right)\right] \le \tau \tag{13}$$

$$\max_{d\in U}\left[\sum_{t=1}^{\tilde{t}}\left(-\sum_{i\in s_I} x_{t,i}^{\mathrm{FSP}}(d)Q_i + \sum_{i\in s_O} x_{t,i}^{\mathrm{FSP}}(d)Q_i - \sum_{j\in s_I} x_{t,j}^{\mathrm{VSP}}(d) + \sum_{i\in s_O} x_{t,j}^{\mathrm{VSP}}(d) + d_{t,s}\Delta t\right)\right] \le v_{0,s} - v_{s,\min} \forall \tilde{t} = 1\ldots T, \forall s \in S, \forall d \in U \tag{14}$$

$$\max_{d\in U}\left[\sum_{t=1}^{\tilde{t}}\left(\sum_{i\in s_I}x_{t,i}^{\text{FSP}}(d)Q_i-\sum_{i\in s_O}x_{t,i}^{\text{FSP}}(d)Q_i+\sum_{j\in s_I}x_{t,j}^{\text{VSP}}(d)-\sum_{i\in s_O}x_{t,j}^{\text{VSP}}(d)-d_{t,s}\Delta t\right)\right]\leq -v_{0,s} + v_{s,\max}\ \forall\tilde{t}=1\ldots T,\ \forall s\in S,\ \forall d\in U \tag{15}$$

$$\max_{d\in U}\left[\sum_{t=1}^{T}\left(-\sum_{i\in s_I}x_{t,i}^{\text{FSP}}(d)Q_i+\sum_{i\in s_O}x_{t,i}^{\text{FSP}}(d)Q_i-\sum_{j\in s_I}x_{t,j}^{\text{VSP}}(d)+\sum_{i\in s_O}x_{t,j}^{\text{VSP}}(d)+d_{t,s}\Delta t\right)\right]\leq v_{0,s} + v_{T,s}\ \forall s\in S,\ \forall d\in U \tag{16}$$

$$0\leq\sum_{i\in\text{PS}}x_{t,i}^{\text{FSP}}(d)\leq 1\quad\forall t=1\ldots T,\forall\text{PS} \tag{17}$$

We note that all the constraints of the above problem (Equations 12–17) have a similar general form as presented in Equation 18. With respect to the inner max problems of the constraints, the LDR coefficients, $\pi_{t,i}^{r}$, are constants while the uncertain parameter d is the decision variable. Therefore, the LDR coefficients are noted a in Equation 18.

$$\max_{d\in U}\left[a^0+\sum_t a^t d_t\right]\leq 0 \tag{18}$$

The solution to the maximization problem (Equation 18) depends on the uncertainty set U which the uncertain demands (d) are drawn from. In this paper, two types of uncertainty sets were examined. A conservative box uncertainty set (Equation 19), and an Ellipsoid uncertainty set (Equation 20) assume that correlations exist between the different components of the uncertainty.

$$\max_{d\in U}\left[a^0+\sum_t a^t d_t\right]\leq 0, U=\left\{d\ :\ \bar{d}+\delta\xi,\ \|\xi\|_\infty\leq\Omega\right\} \tag{19}$$

$$\max_{d\in U}\left[a^0+\sum_t a^t d_t\right]\leq 0, U=\left\{d\ :\ \bar{d}+\delta\xi,\ \|\xi\|_2\leq\Omega\right\} \tag{20}$$

A mathematical description of the uncertainty set is detailed in Equations 19 and 20. Where δ is a mapping matrix derived from the covariance matrix of the random demands d. δ symbolizes the level of uncertainty and the proportions between different entries in the demands vector. Ω is a parameter of the model that can be tuned by the user and stands for robustness, or in other words how extreme the scenarios that the model is required to be immunized against. An equivalent way to describe the meaning of δ and Ω is such that δ is the STD and Ω is the number of STDs around the mean values included in the uncertainty set. Further explanations on constructing the uncertainty set are below. The solution for problems (Equations 19 and 20) according to (A. Ben-Tal & Nemirovski, 1999) presented in Equations 21 and 22 respectively:

$$a^0+\sum_t a^t\bar{d}_t+\Omega|a| \tag{21}$$

$$a^0+\sum_t a^t\bar{d}_t+\Omega\|\delta a^t\|_2 \tag{22}$$

By replacing all the LDR noted as $x_{t,i}(d)$ in Equations 12–17 with the deterministic LDR in Equations 21 and 22 the RC form of the problem is obtained. The RC is linear for the box uncertainty set and a second order convex problem for the Ellipsoid set. As such, it remains tractable for both uncertainty sets.

The described above formulation is based on earlier work of Goryashko and Nemirovski (2014), however, it incorporates several significant enhancements. First, the proposed formulation better represents the nonlinear hydraulics relationships between pumps' flow, head, and efficiency. An Ellipsoid uncertainty set is used for most experiments, which is less conservative than box uncertainty sets and describes better the uncertainty related to water consumption (see below). Moreover, the method includes multi-uncertain LDR allowing the "wait and see" adjustable variables to utilize network-wide information and improve the adjustability of the decision variables.

10.1029/2023WR035508

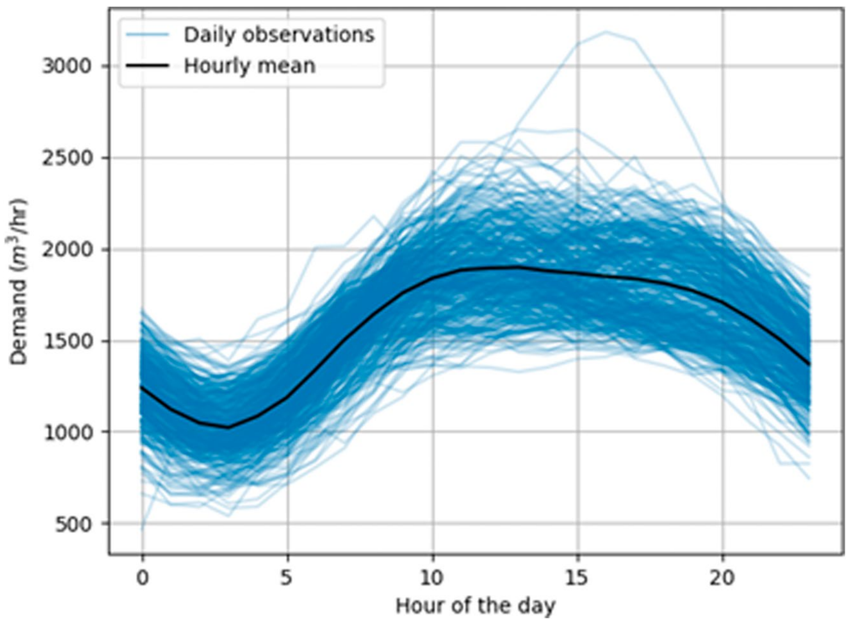

Figure 1. Daily demand pattern, each line represents the hourly values of a different day.

2.2. Uncertainty Quantification

Uncertainty surrounding consumer demands is probably the most common type of uncertainty in WDS management. To address this issue in the optimization problem, it is necessary to first quantify and characterize it. Previous work by Housh et al. (2011) already proved mathematically why Ellipsoid uncertainty sets are equivalent to multivariate normal distribution and provided justification for the use of Ellipsoid in RO problems. Here the same claim is justified in another way by analyzing a real data set of observed demands over 1 year. To quantify the uncertainty, the analysis examined not the demand values themselves but the deviations from known nominal values. These deviations define the range of plausible scenarios, where the probability for each scenario can be estimated based on the density of the deviations. The analyzed data set has a periodic daily pattern as shown in Figure 1. Accordingly, the nominal values are chosen to be the mean values across the same hour of different days.

To investigate the nature of uncertainty, which in this case reflects in the deviations from nominal values, the differences between the observed and the hourly means were calculated for each hour of the day. This resulted in 24 series of values, with each series representing the deviation of a specific hour of the day over 365 days. Due to the large dimensions of the uncertain parameter (24×1, or $24 \times n$ in the case of n consumers), visualizing the full uncertainty space becomes impractical. Therefore, the uncertainty space was decomposed and analyzed in pairs, with each hour examined individually against all other hours of the day. To estimate the correlation between a pair of 2 hr, confidence intervals were calculated using a 2D normal distribution of the two series. If the observed data fell within the confidence intervals, it indicated that the uncertainty follows a normal distribution. A normal distribution implies that the probability of getting an extreme value is decreasing exponentially toward the edges of the distribution. In a multivariate normal distribution, the joint probability of extreme value in multiple coordinates is extremely low. In terms of water consumption, this suggests that extreme values are obtained through all hours of the day, which is a very rare event. Figure 2 shows the relationships between pairs of different hours of the day. The diagonal subplots display the histograms of the deviations in the ith hour, while all other subplots show the deviation in the ith hour plotted against the deviations in the jth hour. It can be seen that the values are falling within ellipses where each ellipse represents one STD of the 2D normal distribution. Most values lie within the two first STDs. Additionally, there is a notable linear correlation between consecutive hours (as seen in the subplots to the right of the histograms). The correlation weakens as the time difference between hours increases, which is expected. However, even for later hours, the relationship still follows an elliptical shape. The complete information about the described correlation is captured in the mapping matrix introduced in Equation 20. The mapping matrix controls how the

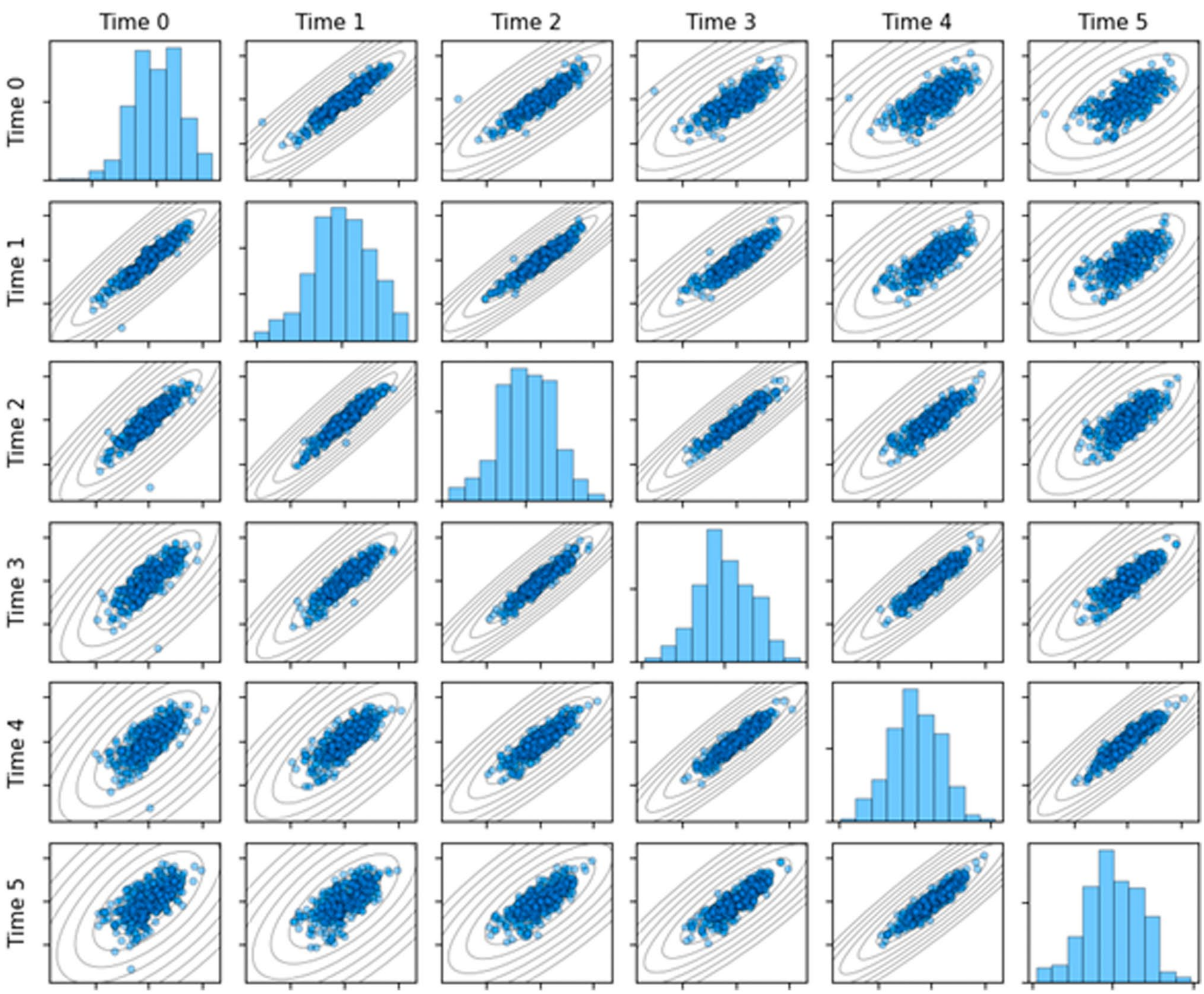

Figure 2. Visualization of deviation from nominal demand values.

uncertainty is distributed across the random vector to obtain the worst-case scenario. Solving the optimization problem subject to the worst case is what guarantees the feasibility of the solution for any scenario within the uncertainty set.

3. Application Examples

The proposed methodology is implemented in two case studies. The first case study is a small illustrative network to demonstrate the methodology in detail, then a real-life network is used to demonstrate the applicability in large networks and for sensitivity analysis. The method was coded in Python using RSOME (Robust Stochastic Optimization Made Easy), a Python package for RO (Chen et al., 2020). The optimization problems were solved with Gurobi 10.0.1 solver (Gurobi Optimization, 2023).

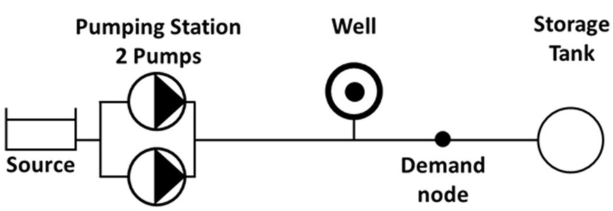

Figure 3. Illustrative network layout.

3.1. Case Study 1—Illustrative Network

An illustrative network is presented in Figure 1. The network consists of a single aggregative consumer supplied by two sources, a PS with two pumps and a well. The water sources and demand are regulated by a single tank

Table 1
Pump Station Hydraulic Operational States

Flow (m^3/hr)	Mean power (kWatt)	STD power (kWatt)	S. Energy (kWatt-hr/m^3)	Unit 1	Unit 2
250	100	10	0.4	1	0
250	95	10	0.38	0	1
400	172	10	0.43	1	1

with a volume of 3,000 m^3. The min and max allowed volumes are 500 and 2,800 m^3 respectively, and the initial volume is 1,500 m^3.

The network is optimized for 24 hr period to find the minimum operational energy costs. Electricity tariffs contain two rates: on-peak and off-peak with respective costs of 1.25 and 1 € per kWatt-hr occur during working hours (08:00–17:00) and the rest of the day (18:00–08:00) respectively. Operational points for the PS and well are detailed in Tables 1 and 2.

The uncertainty set is built based on different levels of uncertainty to analyze the tradeoff between robustness and cost which is referred to as the "price of robustness." The level of uncertainty is noted by θ, and it represents the possible deviation from the nominal values. In this study, different levels of uncertainty were examined between 0% and 25% with 5% steps. For a given uncertainty level θ, the covariance matrix is structured as follows:

$$\sigma = \begin{bmatrix} \theta & 0 & \dots & 0 \\ 0 & \theta & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \theta \end{bmatrix} \tag{23}$$

$$\rho = \begin{bmatrix} 1 & r_{1,2} & \dots & r_{1,T} \\ r_{2,1} & 1 & & \\ \vdots & & \ddots & \vdots \\ r_{T,1} & \dots & \dots & 1 \end{bmatrix} \tag{24}$$

$$\Sigma = \sigma\rho\sigma^T \tag{25}$$

$$\Sigma = \delta\delta^T \tag{26}$$

where $r_{i,j}$ are correlation coefficients between different consumers and different time steps. ρ is a correlation matrix, Σ is the covariance matrix and δ is the mapping matrix used to define the uncertainty set, see Equation 20. For more detailed information on how the uncertainty set is constructed the reader is referred to (G. Perelman et al., 2023).

The network was analyzed by first solving a deterministic problem such that the uncertainty level equals 0. Next, increasing levels of uncertainty were examined with several uncertainty sets. These included a box uncertainty set and two Ellipsoid sets with robustness values of $\Omega = 1$ and $\Omega = 2$. Two problems were solved for each uncertainty configuration. A static RO problem that yields the worst-case optimal policy and an ARO where the objective value is dependent on the observed demands. For the ARO problems, the nominal objective value was calculated based on the nominal demands where the nominal solution is equivalent to the expected value of the solution. The results for all the experiments are presented in Table 3, the values inside the parentheses mark the deviation from the deterministic problem. It is evident that the adjusted solution is clearly superior to the static one. The static worst-case solution deviates from the deterministic in the range of 2%–13.6% compared to less than 2.8% in all the nominal adjusted solutions. Moreover, the proportion between uncertainty level to increase in cost is more moderate in the adjusted solution than in the worst-case solutions. The most conservative uncertainty set (box with an uncertainty level of 25%) results in an objective value that is only 2.8% larger than the deterministic solution. Furthermore, a static RO policy is infeasible for box uncertainty sets with most uncertainty levels and for some of the ellipsoid uncertainty sets, while ARO was able to find a feasible optimal policy for all uncertainty configurations.

Table 2
Well Hydraulic Operational States

Flow (m^3/hr)	Mean power (kWatt)	STD power (kWatt)	S. Energy (kWatt-hr/m^3)	Well
300	126	5	0.42	1

In ARO solutions, at every hour new operational decisions can be made where the decisions depend on data realized in previous time steps. As such'

10.1029/2023WR035508

Table 3
Experimental Results for the Illustrative Network

Uncertainty (%)	Worst-case: RO			Nominal: ARO		
	Box	Ellipsoid Ω = 1	Ellipsoid Ω = 2	Box	Ellipsoid Ω = 1	Ellipsoid Ω = 2
0	1905.8 (0%)	1905.8 (0%)	1905.8 (0%)	1905.8 (0%)	1905.8 (0%)	1905.8 (0%)
5	2103.5 (10.4%)	1944.3 (2.0%)	1985.5 (4.2%)	1909.4 (0.2%)	1908.6 (0.1%)	1911.3 (0.3%)
10	Inf	1985.5 (4.2%)	2075.3 (8.9%)	1918.1 (0.6%)	1911.3 (0.3%)	1919.3 (0.7%)
15	Inf	2030.4 (6.5%)	2165.2 (13.6%)	1930.5 (1.3%)	1914.7 (0.5%)	1928.6 (1.2%)
20	Inf	2075.3 (8.9%)	Inf	1944.8 (2.0%)	1919.3 (0.7%)	1938.0 (1.7%)
25	Inf	2120.2 (11.2%)	Inf	1959.2 (2.8%)	1924.0 (1.0%)	1949.7 (2.3%)

the obtained solutions are not decision variables values but the coefficients of LDR. For example, the optimal solution for the first 12 hr of the well operation ($x^{\mathrm{FSP}}_{t,\mathrm{well}}$) is presented in Equation 27, where the full 24 hr solution is attached as Supporting Information S1. The solution is structured such that each decision is dependent only on previous periods, hence the upper triangular of the matrix contains only zeros. For example, operating the well at time $t = 2$ will be according to the demand value observed in time $t = 1$: $x^{\mathrm{FSP}}_{t=2,\mathrm{well}} = 0.11 + 10^{-3} \cdot 8 \cdot \xi_1$. Moreover, it can be seen that the weights of the latest data are larger than the weights of former time steps. Meaning that more importance is given to the latest observations. At times $t = 8$–12 the well is not operated at all which makes sense since the high-peak tariff starts at 08:00, and therefore all the coefficients equal 0.

$$x^{\mathrm{FSP}}_{t,\mathrm{well}} = \begin{bmatrix} 0.53 \\ 0.11 \\ 0.13 \\ 0.15 \\ 0.19 \\ 0.25 \\ 0.36 \\ 0.52 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 10^{-3} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 11 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 15 & 13 & 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 19 & 17 & 20 & 26 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 22 & 20 & 24 & 32 & 50 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 16 & 16 & 20 & 28 & 43 & 80 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 9 & 12 & 16 & 25 & 40 & 109 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \\ \xi_{12} \end{bmatrix} \tag{27}$$

Next, the ARO method was compared with other dynamic approaches based on a folding horizon strategy (MPC). As explained in the introduction section, MPC models solve optimization problems repeatedly while updating the input parameters with the latest observed data. In this study, two MPC models were examined: a deterministic model, and an uncertain model. The deterministic model is based on the standard LP formulation, while the uncertain model was formulated with a static RO. Both models are based on the formulation in Equations 4–9. In the deterministic model, demands are assumed to be certain whereas in the RO model demands are maximized to be at the worst-case scenario. In every time step, the demands for the next horizon were predicted using a simplistic method of assigning the nominal (mean) value for each hour. To evaluate the performance of each method, Monte Carlo simulations were conducted with 1,000 random samples drawn from multivariate normal distribution with the same characteristics of the uncertainty set used in the RO and ARO methods. The results of the MPC and the ARO methods are presented in Figure 4.

The results depicted in Figure 4 demonstrate that the ARO method performs competitively when compared to traditional folding horizon (MPC) methods. Although the robust MPC yields the lowest mean value, surpassing

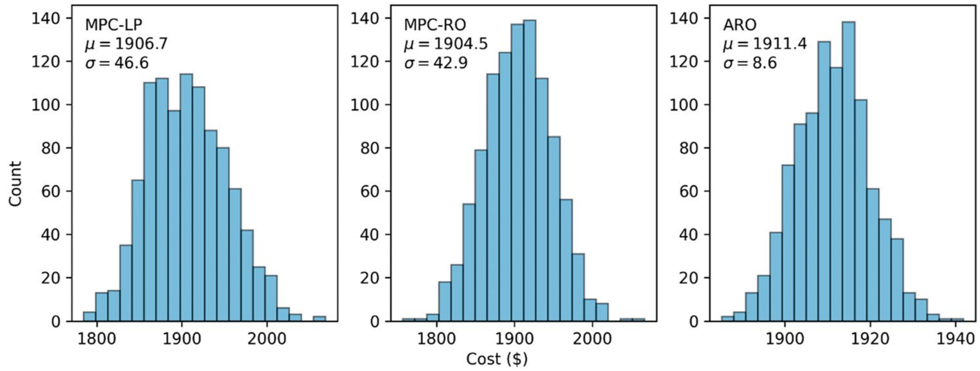

Figure 4. Comparison of adjustable robust optimization method with conventional model predictive control for the illustrative example.

even a deterministic (perfect knowledge) solution, the difference between the three approaches is 0.2% which is negligible from a practical perspective. Nevertheless, ARO outperforms MPC in terms of STD. While MPC solutions range from 1,750 to 2,070, the range of the ARO solutions is narrower, spanning from 1,880 to 1,940. Consequently, ARO suggests a more robust approach that is less susceptible to the realization of uncertain parameters. Another advantage of ARO is that it provides an offline solution as a decision rule prior to any decision implementation. This implies that ARO offers an operation policy that can be employed in advance to examine extreme scenarios, plan maintenance works, and other network management activities without compromising on performance that is typically achieved only by real-time methods.

An important feature of the proposed ARO approach is its dependency on real-time observed data. Therefore, any delays in data receiving might affect the optimality of the solution. Such delays can result from sensors' time sampling, synchronization frequency, data verifications, and other data management procedures. To investigate the importance of the data availability a sensitivity analysis was held where the LDR coefficients of the recent k periods were assigned with a value of 0. For that, Equation 10 is changed such that the inner summation runs from the optimization first step until the previous $k+1$ step. The coefficients that are multiplied by the demands of the latest k steps are zeros such that these demand values cannot affect the solution.

$$x_{t,i}(d) = \pi^0_{t,i} + \sum_{s=1}^{S} \sum_{r=1}^{t-(1+k)} \pi^r_{t,i}\left(d_{r,s} - \overline{d}_{r,s}\right) \tag{28}$$

The analysis was held with an ellipsoid uncertainty set, the same levels of uncertainty used in the standard solution, and data latency of 0–22 hr ($k = 0,1\ldots,22$). A latency of 0 hr is equivalent to the ARO solutions in Table 3 where it is assumed that all the information from previous steps is available. A latency of 23 hr means reducing the ARO to its worst-case version which is equivalent to RO. It is expected that the performance of the method will deteriorate with the decrease in available data. Figure 5 shows that the results match the intuitive expectations where an increase in the information gap results in higher energy costs. It is noted that any additional hour of latency in the range of 6 hr before the current time step has a significant impact on the solution. Where in the range of above 6 hr, the marginal latency has a smaller effect. This conclusion is consistent with the LDR coefficients presented in Equation 27 and Supporting Information S1 where the latest coefficients get the larger weights and have the greatest importance for the solution.

3.2. Case Study 2—Sopron Network

A second case study is a real-life network of the city of Sopron, Hungary (Selek et al., 2012). The water sources of the network are five wells equipped with VSP. The network consists of eight pressure zones (tanks), and eight pump stations with parallel FSP. In five out of the eight pressure zones there is demand while the other three pressure zones are used to regulate flow from the wells before pumped into the network. Each of the FSP pump stations gets power from a different power station such that the available power is constrained. Moreover, the problem includes additional constraints that were not part of the illustrative example. Each of the wells has a min and max required volume to pump within 24 hr of operation. The changes in VSP flows are limited such that

10.1029/2023WR035508

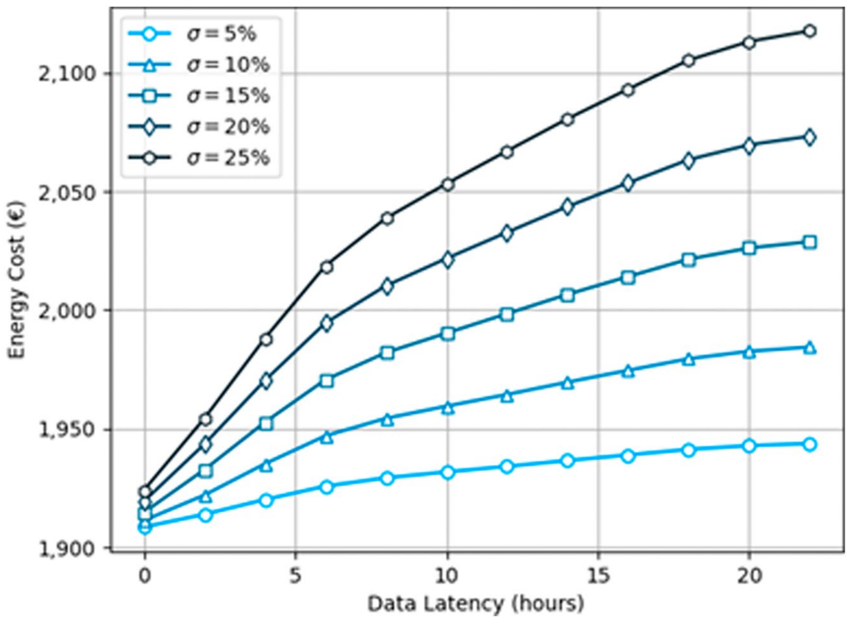

Figure 5. Data latency analysis for the illustrative example.

flows are constant within the electricity tariff periods. To address the additional constraints the mathematical formulation (Equations 12–17) was extended as follows:

$$x_{t,j}^{\mathrm{VSP}}(d) = x_{t+1,j}^{\mathrm{VSP}}(d),\ \forall t \in \mathrm{TP},\ \forall j \in \mathrm{VSP} \tag{29}$$

$$V_{j,\min} = \sum_{t=1}^{T} x_{t,j}^{\mathrm{VSP}}(d) \cdot \Delta t \leq V_{j,\max}, \forall j \in \mathrm{VSP} \tag{30}$$

$$\sum_{i \in \mathrm{ps}} x_{t,i}^{\mathrm{FSP}}(d) \cdot P_i \leq \mathrm{PS}_{t,\max},\ \forall t = 1 \ldots T,\ \forall \mathrm{ps} \in \mathrm{PS} \tag{31}$$

Where TP is tariff periods, meaning time steps where the tariff is not changed. Within these periods, the flow of VSPs is constant. $V_{j,\min}$ and $V_{j,\max}$ are the total daily required volume of wells to meet hydrological limitations. ps is a single power station and PS is the set of all power stations. The total power consumed from a power station at time (t) must be lower than the power station capacity at time (t). The full description of the second case study and all its detailed data can be found in Selek et al. (2012). The network topology is presented in Figure 6.

The uncertainty sets for Sopron network constructed with the demands detailed in Selek et al. (2012) as the nominal values and different levels of uncertainty similar to the illustrative example. The results for Sopron network are presented in Table 4.

The nature of the results is very similar to the illustrative network. The proportion between the level of uncertainty to the increase in energy expenses is consistent with the range obtained in the illustrative network. The worst-case policy yields solutions with a 2%–9.2% increase compared to the deterministic solution. In the adjustable approach, the impact of uncertainty on the operational costs ranges from 0.2% to 4.3% increase. Compared to the illustrative network, here, the uncertainty consequence is more dramatic where the increase in energy cost is almost double than in the illustrative network. ARO shows its great advantage compared to the static RO. While the static robust policy reaches 9.2% more than the deterministic, and while no robust policy was found for uncertainty levels larger than 20% (10%) for Ellipsoid with $\Omega = 1$ ($\Omega = 2$), the ARO method exceeds only 4.3% from the deterministic in the most extreme ellipsoid uncertainty set and was able to find an optimal policy for all Ellipsoid uncertainty sets.

Similar to the illustrative network, the method performance for Sopron network was compared with the MPC approach, the results are presented in Figure 7. Here the advantages of ARO are more prominent as it

10.1029/2023WR035508

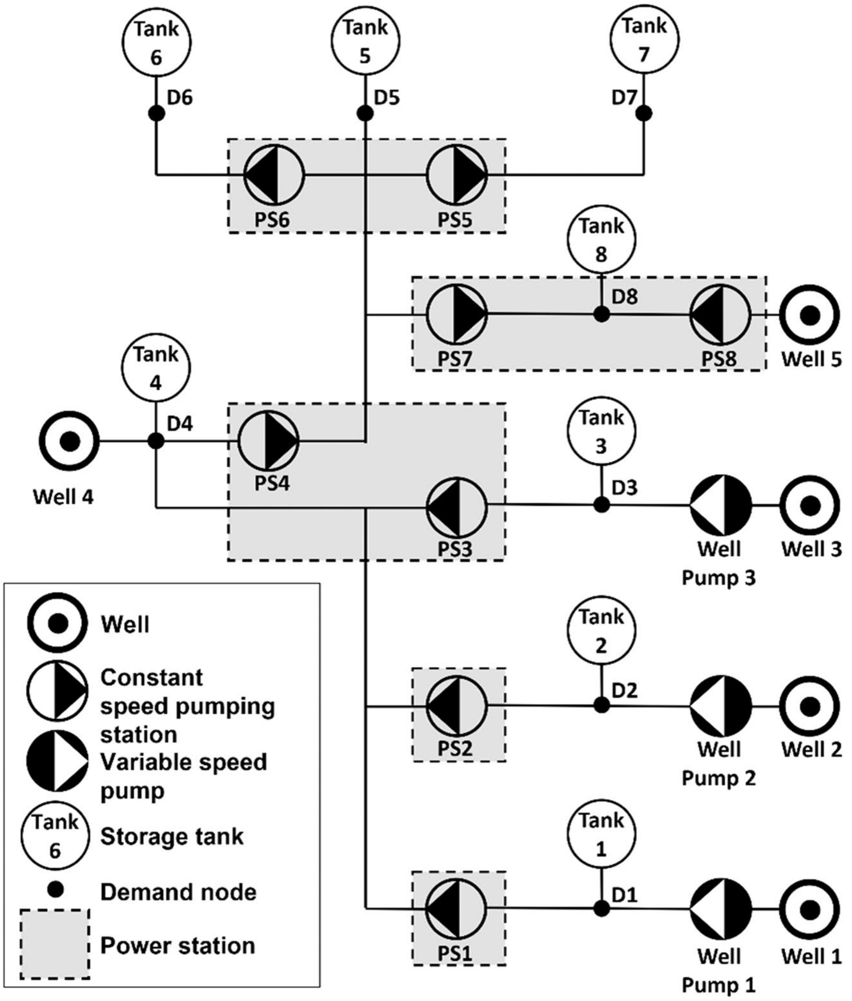

Figure 6. Sopron network layout.

outperforms the traditional folding horizon in both expected costs and STD. The improvement in expected value is not dramatic, 1.3% compared to the deterministic MPC and only 0.3% compared to the uncertain MPC. However, ARO solutions are much more stable as the STD of the results is only 15 which is 0.2% of the expected value. The low STD value indicates the robustness of the method. Despite the size of the system and the large number of control elements, the method achieves stable results that are not too sensitive to the scenarios' realizations.

Table 4
Experimental Results for the Sopron Network

	Worst-case: RO			Nominal: ARO		
Uncertainty (%)	Box	Ellipsoid $\Omega = 1$	Ellipsoid $\Omega = 2$	Box	Ellipsoid $\Omega = 1$	Ellipsoid $\Omega = 2$
0	6685.5 (0%)	6685.5 (0%)	6685.5 (0%)	6685.5 (0%)	6685.5 (0%)	6685.5 (0%)
5	Inf	6827.6 (2.1%)	6980.2 (4.4%)	6741.9 (0.8%)	6699.5 (0.2%)	6718.6 (0.5%)
10	Inf	6980.2 (4.4%)	7303.9 (9.2%)	6832.7 (2.2%)	6718.6 (0.5%)	6768.9 (1.2%)
15	Inf	7140.9 (6.8%)	Inf	Inf	6742.5 (0.9%)	6825.9 (2.1%)
20	Inf	7303.9 (9.2%)	Inf	Inf	6768.7 (1.2%)	6893.5 (3.1%)
25	Inf	Inf	Inf	Inf	6796.5 (1.7%)	6976.3 (4.3%)

10.1029/2023WR035508

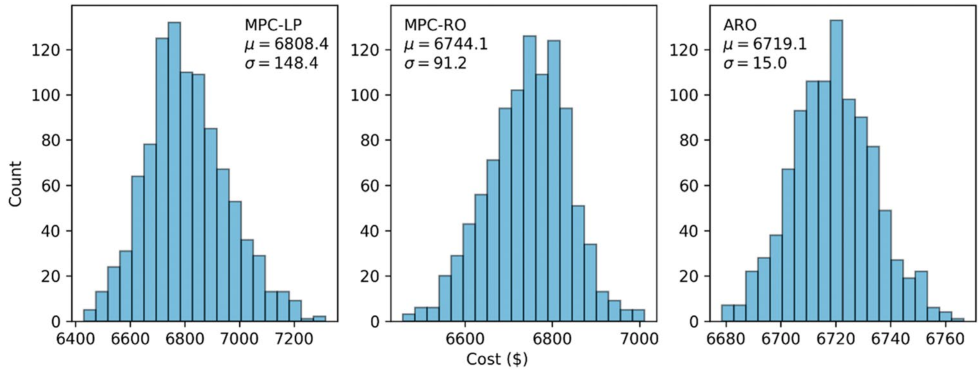

Figure 7. Comparison of adjustable robust optimization method with conventional model predictive control for the Sopron network.

4. Discussion

Demand uncertainty can be found in all WDS optimization problems, specifically in pump scheduling. In light of this uncertainty, robustness is another crucial objective in addition to minimizing operational costs. In this study, robustness is defined as the probability that all the problem constraints are satisfied. The use of RO based methods guarantees the feasibility of the operational policy against any scenario within a predefined range, determined by the desired probability level of constraint satisfaction. By conducting an analysis similar to the one that is presented in Figures 1 and 2, water utilities can design Ellipsoid uncertainty sets that guarantee operational robustness within a specific confidence interval. For example, an uncertainty set encompassing 95% of the scenarios (95% of the dots in Figure 2) will result in a policy that will be feasible for at least 95% of the days. The results of our study indicate that achieving such robustness incurs a relatively small increase in operational costs, approximately 2%–4%. Furthermore, the operational costs remain stable and close to expected values, even in extreme scenarios. Additionally, as mentioned above the proposed method offers another indirect robustness of the system management by the fact that the solution is obtained offline even though it carries real-time models' advantages. By optimizing the decision rule offline the optimal policy is obtained in advance before any information is discovered. This can improve the robustness of the system as it allows a more in-depth analysis of the results and planning toward special events and activities in the network.

Previous studies reported did not suggest a real-time solution which also considers uncertainty. Moreover, studies that addressed uncertainties suffer from inapplicability for different reasons. For example, chance constrained (CC) methods are highly complex to solve and in doubt if can be applied to real networks. Sampling methods and evolutionary algorithms require long run times and global optimum is not necessarily obtained. From a practical perspective, ARO suggests a solution that answers these difficulties without compromising optimality.

5. Conclusions

In this study, an ARO is proposed for optimizing the operation of WDS under demand uncertainties. Also, a typical demand data set is analyzed to characterize and quantify this kind of uncertainty and support the construction of uncertainty sets on which ARO is based. An important conclusion from this analysis is that consumer demands are temporally correlated which justifies the use of ellipsoid uncertainty sets in RO and ARO models. The proposed method is a dynamic RO where its main advantage lies in its ability to be adjusted in real-time according to observed data. The solution provided by the method is not a static operational plan but a decision rule that allows adjustments during its implementation as a response to the uncertainty realization. The method was compared with static RO and also with traditional dynamic approaches of MPC and exhibited superior performance. In terms of expected energy costs, ARO presented comparable performance to the MPC approaches. In terms of stability, ARO performance was less fluctuated with smaller deviations from its nominal expected value. Another advantage of the proposed method is the form of the solution as a decision rule rather than a fixed operation schedule. The rule allows operators to analyze the response to various future events and plan ahead. Since the method depends on real-time data, a sensitivity analysis was conducted to examine the impact of data latency. The results of this sensitivity analysis represent the price of the gap in knowledge. Thus, it is a useful

10.1029/2023WR035508

tool for stakeholders to evaluate the worth of investments in improving the data availability. To conclude, ARO suggests a practical method to optimize the operation of WDS while considering uncertainty. By utilizing real-time observed data, the method achieves optimal solutions that can compete with deterministic methods and yet still hold the advantage of robustness against a range of uncertain scenarios.

Data Availability Statement

The data and code used in this research are available on GitHub (https://github.com/GalPerelman/wds-aro) (G. Perelman, 2023). The optimization models were formulated with RSOME (Robust Stochastic Optimization Made Easy), a Python package for RO (Chen et al., 2020).

Acknowledgments
This research work was made possible with special assistance from the JNF, Israel. This research work was supported by The Bernard M. Gordon Center for Systems Engineering at the Technion.

References

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*. Princeton university press.

Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, *99*(2), 351–376. https://doi.org/10.1007/S10107-003-0454-Y/METRICS

Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, *25*(1), 1–13. https://doi.org/10.1016/S0167-6377(99)00016-4

Castelletti, A., Ficchì, A., Cominola, A., Segovia, P., Giuliani, M., Wu, W., et al. (2023). Model predictive control of water resources systems: A review and research agenda. *Annual Reviews in Control*, *55*, 442–465. https://doi.org/10.1016/J.ARCONTROL.2023.03.013

Chen, Z., Sim, M., & Xiong, P. (2020). Robust stochastic optimization made easy with RSOME. *Management Science*, *66*(8), 3329–3339. https://doi.org/10.1287/MNSC.2020.3603

Dandy, G., Wu, W., Simpson, A., & Leonard, M. (2022). A review of sources of uncertainty in optimization objectives of water distribution systems. *Water*, *15*(1), 136. https://doi.org/10.3390/W15010136

Dziedzic, R., & Karney, B. W. (2015). Energy metrics for water distribution system assessment: Case study of the Toronto network. *Journal of Water Resources Planning and Management*, *141*(11), 04015032. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000555

Fiorelli, D., Schutz, G., Metla, N., & Meyers, J. (2013). Application of an optimal predictive controller for a small water distribution network in Luxembourg. *Journal of Hydroinformatics*, *15*(3), 625–633. https://doi.org/10.2166/hydro.2012.117

Goryashko, A. P., & Nemirovski, A. S. (2014). Robust energy cost optimization of water distribution system with uncertain demand. *Automation and Remote Control*, *75*(10), 1754–1769. https://doi.org/10.1134/S000511791410004X

Grosso, J. M., Velarde, P., Ocampo-Martinez, C., Maestre, J. M., & Puig, V. (2017). Stochastic model predictive control approaches applied to drinking water networks. *Optimal Control Applications and Methods*, *38*(4), 541–558. https://doi.org/10.1002/OCA.2269

Gurobi Optimization, L. (2023). Gurobi optimizer reference manual.

Housh, M. (2011). Optimal multi-year management of regional water resources systems under uncertainty.

Housh, M. (2017). Non-probabilistic robust optimization approach for flood control system design. *Environmental Modelling & Software*, *95*, 48–60. https://doi.org/10.1016/J.ENVSOFT.2017.05.003

Housh, M., Ostfeld, A., & Shamir, U. (2011). Optimal multiyear management of a water supply system under uncertainty: Robust counterpart approach. *Water Resources Research*, *47*(10), 1–15. https://doi.org/10.1029/2011WR010596

Housh, M., & Salomons, E. (2019). Optimal dynamic pump triggers for cost saving and robust water distribution system operations. *Journal of Water Resources Planning and Management*, *145*(2), 04018095. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001028

Hutton, C. J., Kapelan, Z., Vamvakeridou-Lyroudia, L., Savic´, D. A., Savic´, S., & Asce, A. M. (2012). Dealing with uncertainty in water distribution system models: A framework for real-time modeling and data assimilation. *Journal of Water Resources Planning and Management*, *140*(2), 169–183. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000325

Jowitt, P. W., & Germanopoulos, G. (1992). Optimal pump scheduling in water-supply networks. *Journal of Water Resources Planning and Management*, *118*(4), 406–422. https://doi.org/10.1061/(asce)0733-9496(1992)118:4(406)

Lansey, K. E. (2007). The evolution of optimizing water distribution system applications. In *8th Annual Water Distribution Systems Analysis Symposium 2006* (p. 5). https://doi.org/10.1061/40941(247)5

Li, Z., Wu, W., Zhang, B., & Wang, B. (2015). Adjustable robust real-time power dispatch with large-scale wind power integration. *IEEE Transactions on Sustainable Energy*, *6*(2), 357–368. https://doi.org/10.1109/TSTE.2014.2377752

Maier, H. R., Guillaume, J. H. A., van Delden, H., Riddell, G. A., Haasnoot, M., & Kwakkel, J. H. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling & Software*, *81*, 154–164. https://doi.org/10.1016/J.ENVSOFT.2016.03.014

Mala-Jetmarova, H., Sultanova, N., & Savic, D. (2017). Lost in optimisation of water distribution systems? A literature review of system operation. *Environmental Modelling and Software*, *93*, 209–254. https://doi.org/10.1016/j.envsoft.2017.02.009

Moreira, A., Street, A., & Arroyo, J. M. (2015). An adjustable robust optimization approach for contingency-constrained transmission expansion planning. *IEEE Transactions on Power Systems*, *30*(4), 2013–2022. https://doi.org/10.1109/TPWRS.2014.2349031

Pan, L., Housh, M., Liu, P., Cai, X., & Chen, X. (2015). Robust stochastic optimization for reservoir operation. *Water Resources Research*, *51*(1), 409–429. https://doi.org/10.1002/2014WR015380

Pankaj, B. S., Jaykrishnan, G., & Ostfeld, A. (2022). Optimizing water quality treatment levels for water distribution systems under mixing uncertainty at junctions. *Journal of Water Resources Planning and Management*, *148*(5), 04022013. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001544

Perelman, G. (2023). GalPerelman/wds-aro: WDS ARO - v1.0.0. https://doi.org/10.5281/ZENODO.8378673

Perelman, G., Ostfeld, A., & Fishbain, B. (2023). Robust optimal operation of water distribution systems. *Water*, *15*(5), 963. https://doi.org/10.3390/W15050963

Perelman, L., Housh, M., & Ostfeld, A. (2013). Robust optimization for water distribution systems least cost design. *Water Resources Research*, *49*(10), 6795–6809. https://doi.org/10.1002/wrcr.20539

Postek, K., den Hertog, D., Kind, J., & Pustjens, C. (2019). Adjustable robust strategies for flood protection. *Omega*, *82*, 142–154. https://doi.org/10.1016/J.OMEGA.2017.12.009

10.1029/2023WR035508

Quintiliani, C., & Creaco, E. (2019). Using additional time slots for improving pump control optimization based on trigger levels. *Water Resources Management*, *33*(9), 3175–3186. https://doi.org/10.1007/S11269-019-02297-6/FIGURES/6

Rao, Z., & Salomons, E. (2007). Development of a real-time, near-optimal control process for water-distribution networks. *Journal of Hydroinformatics*, *9*(1), 25–37. https://doi.org/10.2166/hydro.2006.015

Salomons, E., & Housh, M. (2020). Practical real-time optimization for energy efficient water distribution systems operation. *Journal of Cleaner Production*, *275*, 124148. https://doi.org/10.1016/j.jclepro.2020.124148

Schwartz, R., Housh, M., & Ostfeld, A. (2016). Least-cost robust design optimization of water distribution systems under multiple loading. *Journal of Water Resources Planning and Management*, *142*(9), 04016031. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000670

See, C. T., & Sim, M. (2009). Robust approximation to multiperiod inventory management. *Operations Research*, *58*(3), 583–594. https://doi.org/10.1287/OPRE.1090.0746

Selek, I., Bene, J. G., & Hs, C. (2012). Optimal (short-term) pump schedule detection for water distribution systems by neutral evolutionary search. *Applied Soft Computing*, *12*(8), 2336–2351. https://doi.org/10.1016/J.ASOC.2012.03.045

Sharif, M. N., Haider, H., Farahat, A., Hewage, K., & Sadiq, R. (2019). Water–energy nexus for water distribution systems: A literature review. *Environmental Reviews*, *27*(4), 519–544. https://doi.org/10.1139/ER-2018-0106

Yanıkoğlu, İ., Gorissen, B. L., & den Hertog, D. (2019). A survey of adjustable robust optimization. *European Journal of Operational Research*, *277*(3), 799–813. https://doi.org/10.1016/J.EJOR.2018.08.031