Reliable water supply system design under uncertainty

G. Chung a, K. Lansey a,*, G. Bayraksan b

a Department of Civil Engineering and Engineering Mechanics, The University of Arizona, Tucson, AZ 85721, USA
b Department of Systems and Industrial Engineering, The University of Arizona, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 25 March 2008
Received in revised form 1 August 2008
Accepted 12 August 2008
Available online 26 November 2008

Keywords:
Water supply system
Data uncertainty
Spatially correlated data
Robust optimization

ABSTRACT

Given the natural variability and uncertainties in long-term predictions, reliability is a critical design factor for water supply systems. However, the large scale of the problem and the correlated nature of the involved uncertainties result in models that are often intractable. In this paper, we consider a municipal water supply system over a 15-year planning period with initial infrastructure and possibility of construction and expansion during the first and sixth year on the planning horizon. Correlated uncertainties in water demand and supply are applied on the form of the robust optimization approach of Bertsimas and Sim to design a reliable water supply system. Robust optimization aims to find a solution that remains feasible under data uncertainty. Such a system can be too conservative and costly. In the Bertsimas and Sim approach, it is possible to vary the degree of conservatism to allow for a decision maker to understand the tradeoff between system reliability and economic feasibility/cost. The degree of conservatism is incorporated in the probability bound for constraint violation. As a result, the total cost increases as the degree of conservatism (and reliability) is increased. In the water supply system application, a tradeoff exists between the level of conservatism and imported water purchase. It was found that the robust optimization approach addresses parameter uncertainty without excessively affecting the system. While we applied our methodology to hypothetical conditions, extensions to real-world systems with similar structure are straightforward. Therefore, our study shows that this approach is a useful tool in water supply system design that prevents system failure at a certain level of risk.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

A water supply system typically includes multiple sources and demand centers (agricultural, domestic, industrial and commercial users). System components are designed to treat relatively good quality source waters from an aquifer or various surface supplies and deliver it to users in a water distribution system that is sized to provide fire flows. Water system design represents a tradeoff between treatment plant size (economy of scale) and the number of plants, pipe and pump sizes and energy consumption, and component sizes, travel time and water quality. To accommodate growth, affected communities have found it necessary to shift from reliance on traditional water supplies—ground water or relatively meager surface flows—toward a combination of large, engineered water projects, water reuse and conservation measures. All of these decisions are made subject to uncertainties introduced by future growth rates and locations, water resource availability, and changing social and institutional conditions.

Much research has been conducted on the simulation of the water supply system (Cai, 2008; Chung et al., 2008; Makropoulos et al., 2008; Mitchell et al., 2008). A municipal water supply system is defined as the physical infrastructure to treat, deliver water to and collect water from users. The design of the capacities of alternative components in a water supply system is usually based upon predictions of future population and climatic conditions. Uncertainty in predicting these conditions is inherent in all water supply systems. Thus, a decision made with a deterministic model that is based on satisfying demand/supply conditions without consideration of uncertainty may result in two consequences: (i) lower net benefits than expected (i.e., it is more costly to provide the desired water) or, (ii) some probability of system failure, where failure is defined as not meeting a given demand or other system constraint (Watkins and McKinney, 1997). These consequences may be rectified in real-time operations at some cost but flexibility must be built into the system during the design process to allow for those adjustments. Deterministic optimization removes this flexibility, thus, a reliability-based design tool is needed that can assist decision makers plan a long-term water supply scheme to cope with the future changes in water demands and supplies.

The complexity of a water system and the correlated uncertainties make incorporating uncertainty a challenging exercise.
A number of stochastic optimization approaches have been applied to water supply system design and operation. Most works have adopted two-stage or multi-stage linear or nonlinear stochastic programming with recourse. The main objectives of these studies were to minimize expected total cost for water transfer to spot-markets (Lund and Israel, 1995); to develop long- as well as short-term water supply management strategies (Wilchfort and Lund, 1997); to manage water supply capacity under water shortage conditions (Elshorbagy et al., 1997). On the other hand, markets (Lund and Israel, 1995); to develop long- as well as short-term water supply system design and operation. Most works have focused on the design of water supply systems, including the use of stochastic programming approaches to handle uncertainty. The main objectives of these studies were to minimize expected total cost for water supply management under uncertainty. This type of robust optimization was first introduced by Soyster (1973) for linear programming problems. However, the classical assumption in deterministic mathematical programming is that all parameters (input data) are known precisely. This is rarely the case in real applications since many parameters contain uncertainties such as future predictions or measurement errors. One way to deal with uncertainty is to design a system that is “robust” to parameter changes. That is, the system remains feasible and operates in a near-optimal fashion for a variety of values that the uncertain parameters can take. Soyster (1973) formulated the following linear programming model to find

\[ \min \{ c^T x + \epsilon^T y : A x + b \leq \epsilon_0 + \delta, \quad x, y \geq 0 \} \]

where \( c, A, b, \) and \( \epsilon_0, \delta \) are known parameters, and \( \epsilon \) is a vector of uncertain parameters. This model significantly constrains the objective function to assure robustness; thus conservative solutions are found that may be practically unrealistic. Ben-Tal and Nemirovski (1999, 2000), El-Ghaoi and Lebret (1997), and El-Ghaoi et al. (1998) extended the Soyster model. These extensions, however, introduced a higher degree of non-linearity. Since real systems themselves are likely to be nonlinear, these approaches make problem more complicated and difficult to find a solution. The approach of Bertsimas and Sim controls the degree of conservatism for the system reliability without increasing the difficulty in solving the original problem.

In this paper, the robust optimization framework of Bertsimas and Sim (2004) is used to develop a reliable water supply system design. A robust solution can be defined as one that remains feasible under uncertainty. The type of robust optimization was first introduced by Soyster (1973) for linear programming problems. Soyster’s model significantly constrains the objective function to assure robustness; thus conservative solutions are found that may be practically unrealistic. Ben-Tal and Nemirovski (1999, 2000), El-Ghaoi and Lebret (1997), and El-Ghaoi et al. (1998) extended the Soyster model. These extensions, however, introduced a higher degree of non-linearity. Since real systems themselves are likely to be nonlinear, these approaches make problem more complicated and difficult to find a solution. The approach of Bertsimas and Sim controls the degree of conservatism for the system reliability without increasing the difficulty in solving the original problem.
Uncertainty is modeled in the same way as Ben-Tal and Nemirovski. That is, for the $i$th constraint, $f_i$ represents the set of indices that correspond to uncertain $\tilde{a}_{ij}$. It is assumed that $\tilde{a}_{ij}, j \in f_i$, are independent, symmetric and bounded random variables as given in Eq. (2).

To control the degree of conservatism, Berti"{s}mas and Sim introduce an additional parameter, $\Gamma$, that can take any real value within the range of $[0, |f_i|]$, in a manner that the most significant coefficients up to the $|f_i|$th order is fully allowed to vary within their uncertainty intervals and the $(|f_i| + 1)$th order significant coefficient, $a_{ij}$, is bounded by $(|f_i| - 1)a_{ij}$, while the remaining coefficients are fixed at their nominal values. Then, Eq. (4) is reformulated in a robust form to improve the system reliability as:

\[
\text{maximize } cx \\
\text{subject to } \sum_{j=1}^{n} \tilde{a}_{ij}x_j \leq b_i, \quad \forall i, j \in f_i \quad (4)
\]

\[
I \leq x \leq u.
\]

At optimality, $y_j = |x_j'|$ for all $j$.

The nominal (deterministic) problem would have constraint $\sum \tilde{a}_{ij}x_j \leq b_i$ instead of Eq. (5). In the above robust formulation, however, an additional max-term keeps the system feasible as allowed by the degree of conservatism (or the degree of uncertainty) represented by $\Gamma$. Note that when $\Gamma = 0$, the constraint is equivalent to that of the nominal problem, $\sum \tilde{a}_{ij}x_j \leq b_i$, and when $\Gamma = |f_i|$, the constraint is completely protected against uncertainty. As the max-term in the constraint increases, the nominal-term must be reduced to satisfy the constraint bound, $b_i$. In other words, system conservatism will have an increasing effect on the max-term as $\Gamma$ value is increased and accordingly the capacity of system component, $x_i$, will be reduced to satisfy the constraint. In this manner, the tradeoff between the degree of conservatism and corresponding system capacity, i.e., system economic feasibility, can be evaluated.

The above formulation assumes that the uncertain coefficients are independent. This assumption is unlikely to be valid in many systems including our water supply system so, an extended uncertainty model for correlated random variables was proposed by Bertsimas and Sim (2004). Suppose a number of different sources of uncertainty affect the system, the randomness in the $i$th constraint can then be represented as:

\[
\tilde{a}_{ij} = \tilde{a}_{ij} + \sum_{k \in K_i} \eta_{kij} \tilde{a}_{kij}, \quad \forall j \in f_i
\]

where $\eta_{kij}$ are independently and symmetrically distributed random variables in the range $[-1, 1]$, $|K_i|$ is the number of uncertainty sources that affect the coefficients in the $i$th constraint, as before $\tilde{a}_{ij}$ is the nominal value of $\tilde{a}_{ij}$, and $f_i$ is the set of indices of

Subject to

\[
\begin{align*}
\sum_{k \in K_i} \eta_{kij} \tilde{a}_{kij} & \leq b_i, \quad \forall i, j \in f_i \\
-\sum_{k \in K_i} \eta_{kij} \tilde{a}_{kij} & \leq b_i, \quad \forall i, j \in f_i \\
I & \leq x \leq u, \\
y & \geq 0.
\end{align*}
\]

Subject to

\[
\begin{align*}
\sum_{k \in K_i} \eta_{kij} \tilde{a}_{kij} & \leq b_i, \quad \forall i, j \in f_i \\
-\sum_{k \in K_i} \eta_{kij} \tilde{a}_{kij} & \leq b_i, \quad \forall i, j \in f_i \\
I & \leq x \leq u, \\
y & \geq 0.
\end{align*}
\]
uncertain parameters in the ith constraint that is subject to uncertainty. \( \tilde{a}_{ij} \) represents a maximum level of correlation effect on uncertain parameter \( a_{ij} \) from source \( k \in K_i \).

The robust model with correlated uncertain coefficients can then be rewritten as:

\[
\text{maximize } \mathbf{c} \mathbf{x} \\
\text{subject to } \sum_{j \in J} a_{ij} x_j + \max_{\{S_i \cup \{t_i\} : S_i \subseteq K_i, |S_i| - |I_i|, t_i \in K \setminus S_i\} \sum_{j \in J} g_{ij} x_j \leq b_i, \quad \forall i \\
I \leq \mathbf{x} \leq \mathbf{u}.
\]

(7)

The problem solution is affected in a similar manner as Eq. (5) with the objective function value decreasing with increasing \( \Gamma_i \). With this formulation and also in Eq. (5), the max-term can be represented as a linear program and thus, the robust solutions of nonlinear constraint functions but all the constraints with uncertainties are linear and, therefore, we use a similar linear representation of the resulting max-terms.

The level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show the level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show the level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show the level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show the level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show the level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show the level of conservatism, \( \Gamma_i \), is a useful tool to investigate system robustness against failure. If system failure can be presented as a probability, it would give better understanding of system safety. It is possible to relate \( \Gamma_i \) to a probability level and show.

\[
B(n, \Gamma_i) \leq \sum_{i=1}^{n} C(n, I_i),
\]

(9)

where \( n = |K_i|, \mu = (\Gamma_i + n)/2, \mu = \nu - |\nu| \) and

\[
C(n, I_i) = \left\{ \begin{array}{ll}
\frac{1}{\sqrt{2\pi} \sqrt{n}} \exp \left( \frac{n}{2} \log \left( \frac{n}{2(\pi - 1)} \right) + \log \left( \frac{n}{2(\pi - 1)} \right)^2 \right) & \text{if } I_i = 0 \text{ or } I_i = n \\
\end{array} \right.
\]

(10)

For the model with correlated data (Eq. (7)) and \( n = |K_i| \), the bound is calculated in a similar fashion.

To compute \( \Gamma_i \) values, a desired probability level in Eq. (8) and \( B(n, \Gamma_i) \) is specified for each of \( i \) uncertain constraints. Assuming Eq. (9) is tight, \( \Gamma_i \) can then be directly computed using Eq. (10). Each uncertain constraint is considered independently to determine its corresponding \( \Gamma_i \). With the set of \( \Gamma_i \), the optimal solution for the desired probability level is determined by solving problem in Eq. (5) or Eq. (7). A range of probability levels can be evaluated to provide the decision maker the tradeoff between robustness and cost.

3. Water supply system

3.1. Problem statement and notation

The robust optimization methodology is applied to a realistic hypothetical water supply system but is capable of considering a general system. As schematic of the water system including subsurface (aquifer), surface, and imported water sources, domestic and agricultural irrigation users, and water and wastewater treatment plants is shown in Fig. 1. It is possible that an external water source can be imported at a cost per unit volume plus the cost of the conveyance system to transport the water to the community.

In some semi-arid regions, wastewater effluent is discharged to a normally dry or low flow channel. Over time, a downstream riparian habitat develops that is sustained by the effluent flow. If communities move to using reclaimed wastewater effluent for non-potable and, potentially, potable uses, this water would no longer

![Fig. 1. Water supply system network schematic. Solid arcs represent conveyance structures to be sized and dashed arcs represent precipitation or infiltration from users and sources to the aquifer.](image-url)
be released to the riparian area. Thus, communities face depletion of both surface and subsurface water sources and the decision to maintain environmental flows. Therefore, as part of their planning processes for a sustainable water supply, communities will need to develop new water supply structures and sources while preserving environmental flows in the river stream and storage in an aquifer.

The objective function of the problem is to minimize the cost of construction and operation/maintenance of system components as well as the cost of purchasing water from outside sources to meet water demand. Design decisions are pipe sizes, pump design flow, pump design head, canal depth, and water and wastewater treatment plant capacities. Flow allocations over the water supply network are operational variables.

To model the water supply system, we have a graph $G = (N, A)$, where $N$ is the set of nodes and $A$ is the set of arcs (Fig. 1). $N$ consists of eight nodes: sources ($N_S$ – imported water, aquifer, upstream river, and downstream river), users ($N_U$ – domestic and irrigation), and water and wastewater treatment plants ($N_W$ and $N_WT$, respectively). A river is represented by upstream ($N_U$) and downstream ($N_{DU}$) nodes and a connecting arc. Sources are divided into two subsets: storage sources ($N_S$) and non-storage sources ($N_{WS}$) depending on their capability of storing water from the previous time step. For instance, a groundwater aquifer is a storage source while imported water ($N_U$), upstream ($N_U$) and downstream ($N_{DU}$) rivers are non-storage sources.

In the application system, $A$, the set of arcs consists of 18 arcs $(i, j)$ that connects nodes $i$ and $j$. Arcs represent canals ($A_C$ – arcs 1, 2, 3, 4 and 5 in Fig. 1), pipe lines ($A_P$ – arcs 8, 9, 10 and 11), pump stations ($A_U$ – arcs 6 and 7), rainfall or mountain front recharge ($A_R$ – arcs 12, 13, 14 and 15), and seepage or infiltration ($A_I$ – arcs 16, 17, and 18). Pump stations are located in arcs to overcome friction losses and elevation differences through pipe connections. To permit water banking, arc 4 represents a canal to carry imported water to recharge basins. Annual seepage and infiltration are constant values.

A 15-year planning period is evaluated with existing infrastructure in place at year 1, and new facilities can be constructed in year 1 and 6. $T$ and $O$ denote the set of construction/expansion times $t$ and operational times $o$, respectively. The system can be constructed or expanded in year $t$. Operational variables ($d_{ij}$) over arcs $(i,j)$ during each operation period, $o$, are determined each year ($\Delta o = 1$) for the first 5 years (first design period) and every other year ($\Delta o = 2$) for the last 10 years (second design period).

The system constraints are:

1. meet water demand for water users,
2. satisfy conservation of mass through nodes,
3. meet required river discharge at downstream river node,
4. meet required groundwater storage to maintain a sustainable system,
5. restrict the amount of imported water by the external water availability,
6. do not exceed water and wastewater treatment plant capacity,
7. limit canal flow by maximum canal capacity,
8. maintain pump operating efficiency,
9. meet minimum pressure requirement at the end of pipe and pump arcs.

Constraints (3) and (4) are environmental constraints to sustain the natural water supplies. The next three set of constraints (5)–(7) are the capacity constraints and the last two constraints (8) and (9) are dictated by the hydraulic design. The objective is to build and operate the system with minimum cost. Table 1 shows the particular values of data used in the application.

The remainder of this section presents the objective function and constraints of the water supply problem. In Section 3.2, the uncertain parameters are identified and the constraint conversion to the robust formulation is described. In Section 5, the water supply system application and the results from the robust optimization are presented.

### 3.2. Objective function

As noted, the objective is to minimize the total cost for construction of the system components (pipes, canals, pumps, and water and wastewater treatment facilities), for operation and maintenance of the system and for purchasing imported water:

$$
\text{minimize } z = f_1(x_{ij}^0) + f_2(d_{ij}) + f_3(x_{ij}^0H_{ij}^0) + f_4(w_{ij}^0)
+ f_5(q_{ij}^0, w_{ij}^0) + f_6(q_{ij}^0).
$$

(11)

The first term, the pipe construction cost, is a function of pipe diameters that connect nodes $i$ and $j$ at year $t$ ($x_{ij}^0$) and computed following Clark et al. (2002).

$$
f_1(x_{ij}^0) = \sum_{t \in T} \left( \frac{1}{1 + t} \cdot \sum_{(i,j) \in A_C} x_{ij}^0 \left( 57.198 + 0.35 x_{ij}^0 + 0.62 c_{ij}^{1/24} 
+ 0.0018 c_{ij}^{1/9} + 0.0062 c_{ij}^{1/8} - 0.062 c_{ij}^{1/7} - 0.02 c_{ij}^{1/6}
+ 0.23 c_{ij}^{1/5} + 0.0022 c_{ij}^{1/4} \right) \right).
$$

(11a)

The product of the pipe length, $l_{ij}$, and the constant term gives a positive cost even when the pipe diameter, $\kappa$, is 0 (i.e., no connection is desired). Therefore, a binary variable ($x_{ij}^0$) is added to the model to define the existence of a pipe from node $i$ to $j$ at time $t$.

Canal flows are driven by gravity and canal construction cost is a function of the channel depth ($d_{ij}^0$) (US. Army Corps of Engineers, 1980):

$$
f_2(d_{ij}^0) = \sum_{t \in T} \sum_{(i,j) \in A_C} c_{ij}d_{ij}^0
= \sum_{t \in T} \left( \frac{1}{1 + t} \cdot \sum_{(i,j) \in A_C} \left( 1.45 \cdot \left( 55.30d_{ij}^0 \cdot \text{ENR} \cdot \text{CITY} \right) / 2877 \right) \right).
$$

(11b)

\text{CITY} and \text{ENR} are parameters that account for local cost variations and the inflation rate, respectively. The \text{ENR} factor for year $t$ is computed by:

### Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Darcy–Weisbach coefficient, $f$</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Manning’s coefficient, $n$</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Canal side slope, $z$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hydraulic conductivity, COND</td>
<td>9.14</td>
<td>m/yr</td>
</tr>
<tr>
<td>Imported water availability, $W$</td>
<td>19.6</td>
<td>m$^3$/s</td>
</tr>
<tr>
<td>Interest rate, $r$</td>
<td>2.0</td>
<td>%/yr</td>
</tr>
<tr>
<td>City multiplier, CITY</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Annual precipitation, $P$</td>
<td>33.4</td>
<td>mm/yr</td>
</tr>
<tr>
<td>Correlation coefficients, $p$</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Basin area, $A_B$</td>
<td>12,645</td>
<td>km$^2$</td>
</tr>
<tr>
<td>Basin area contributing to imported water, $A_I$</td>
<td>13,909</td>
<td>km$^2$</td>
</tr>
<tr>
<td>Required groundwater storage, $RS$</td>
<td>9.93</td>
<td>km$^3$</td>
</tr>
<tr>
<td>Required downstream river flow, $RO_{13}$</td>
<td>11.4</td>
<td>m$^3$/s</td>
</tr>
<tr>
<td>Unit cost of purchasing imported water, $C_{IP}$</td>
<td>0.81</td>
<td>$/m^3$</td>
</tr>
<tr>
<td>Agricultural consumptive use (1–5 periods), $I_{AC}$</td>
<td>12.5</td>
<td>m$^3$/s</td>
</tr>
<tr>
<td>Agricultural consumptive use (6–10 periods), $I_{AC}$</td>
<td>11.3</td>
<td>m$^3$/s</td>
</tr>
<tr>
<td>Initial domestic demand</td>
<td>9.78</td>
<td>m$^3$/s</td>
</tr>
<tr>
<td>Initial population</td>
<td>1,200,000</td>
<td></td>
</tr>
<tr>
<td>Population growth rate</td>
<td>2.7</td>
<td>%/yr</td>
</tr>
</tbody>
</table>
\( ENR^t = -7.7 \times 10^8 + 15.7t - 12.0 \times 10^3 t^2 + 4.1t^3 - 0.5 \times 10^{-3} t^4 \)

The rated pump head \((h_{ij}^r)\) and discharge \((x_{ij}^r)\), which define the most efficient pump operation point, are the pump design variables. A binary variable \( \mu_{ij}^f \) indicates the existence of a pump. The pump construction costs are computed by Walski et al. (1987) and summed over the set of pumps \((A_P \cup A_U)\) to determine the total cost:

\[
f_3(x_{ij}^r, h_{ij}^r) = \sum_{t \in T} \left( \frac{1}{1 + f^t} \sum_{(i,j) \in A_P \cup A_U} \mu_{ij}^f \left( 500x_{ij}^{\mu^t \times h_{ij}^{\mu^t}} \right) \right).
\]

Water and wastewater treatment facility construction and expansion costs are computed from their capacities \((w_{ij}^f)\) (Tang et al., 1987):

\[
f_4(w_{ij}^f) = \sum_{t \in T} \left( \frac{1}{1 + f^t} \sum_{k \in \text{NWT}} (2897.13w_{ij}^f + 35.987) \right)
+ \sum_{t \in T} \left( \frac{1}{1 + f^t} \sum_{j \in \text{NWT}} (10,811.92w_{ij}^f + 5,454,228) \right).
\]

All costs are converted to their equivalent present values at year 1 by applying the present worth factor of \(1/(1 + I)^{10t} \) where \(I\) is the interest rate. Operations and maintenance (O&M) costs are calculated for each operation period, \(o\), by (Clark et al., 2002; US, Army Corps of Engineers, 1980; Walski et al., 1987; Tang et al., 1987):

\[
f_5(q_{ij}^0, w_{ij}^o) = \sum_{o=1}^{10} \left( \frac{1}{1 + f^t} \sum_{(i,j) \in A_P} x_{ij}^o (27.7 + 0.3 q_{ij}^o) L_{ij} \right)
+ \sum_{(i,j) \in A_C} \left( 0.02544 L_{ij} q_{ij}^{0.572} \right)
+ \left( 0.078 + 0.0135 q_{ij}^o \right) L_{ij} \frac{\text{ENR}^o}{1850}
+ \sum_{(i,j) \in A_P \cup A_U} \left( 79.47 L_{ij} q_{ij}^o + 4560 q_{ij}^{0.55} \right)
+ 320q_{ij}^{0.55} \frac{\text{ENR}^o \times \text{CTY}^o}{2877}
+ \sum_{j \in \text{NWT}} (28.97 w_{ij}^o + 360)
+ \sum_{j \in \text{NWT}} (108.12 w_{ij}^o + 54,542) \right) \Delta o^o.
\]

The terms in Eq. (11e) are related to pipes, canals, pumps, water treatment and wastewater treatment plants in terms of facility capacities and/or operation flows rate \((q_{ij}^o)\), respectively. The parameters \(A_{ij}\) in the pump term are the elevation differences between the pipe endpoints.

Water \((\text{IW}^o)\) can be purchased and brought to the supply system at a unit cost of \(C_{\text{IW}}\). A time step factor \((\Delta o^0)\) accounts for the variable decision period durations. The imported water cost is:

\[
f_6(q_{ij}^o) = \sum_{o=1}^{10} \left( \frac{1}{1 + f^t} \sum_{j \in \text{N}} q_{ij}^o \text{IW}_{ij} \right) \Delta o^o.
\]

### 3.3. Flow constraints through nodes

Water demands for each operation period \(o\) \((D_{ij}^o)\), must be satisfied for each demand center \(j\) (agriculture and domestic areas) by supplying from upstream sources, \(i \in \text{NS}^o\):

\[
\sum_{i \in \text{NS}} q_{ij}^o \geq D_{ij}^o, \quad \forall j \in \text{NS}, \quad \forall o \in \text{O}.
\]

Similarly, the total flow through a water or wastewater treatment plant cannot exceed the plant capacity, \(w_{ij}^f\) or:

\[
\sum_{i \in \text{N}} q_{ij}^o \leq w_{ij}^f, \quad \forall j \in \text{NWT} \cup \text{NWWT}, \quad t \leq o, \quad \forall o \in \text{O}, \quad \forall t \in \text{T}.
\]

The amount of imported water inflow to the system is limited by external water availability \((\text{IW}^o)\) and is computed as the sum of the outflows from the imported water node \((\text{IW})\) or:

\[
\sum_{j \in \text{N} \cup \text{N}^c} q_{ij}^o \text{IW}_{ij} \leq \text{IW}^o, \quad \forall o \in \text{O}.
\]

Natural runoff resulting from precipitation \((\tilde{P}^o)\) on the upstream watershed that has area, \(A_L\), contributes to the river flow and aquifer. Assuming 60% of the precipitation is abstracted to interception and depression storage and evaporates, 30% of the rainfall is assumed to be an inflow to the upstream river node \((\text{NRU})\) and 10% of rainfall recharges the aquifer \((\text{NS}^o)\). The volumes of streamflow and aquifer recharge are the product of the precipitation and the contributing area or:

\[
q_{ij}^o = 0.3 \tilde{P}^o A_L, \quad i = \text{precipitation}, \quad \forall j \in \text{NRU}, \quad \forall o \in \text{O}.
\]

By conservation of mass, inflows and outflows at some non-storage nodes \((\text{NSS})\) must balance. For node \(i\), conservation of mass constraint at period \(o\) is:

\[
\sum_{j \in \text{N}} \sum_{o} q_{ij}^o - \sum_{j \in \text{N}} q_{ij}^0 = 0, \quad \forall i \in \text{NSS} \setminus (\text{NRU} \cup \text{NRD}), \quad \forall o \in \text{O}.
\]

To account for changing stream conditions and the location of inflows and outflows along a river, rivers are modeled with an upstream and downstream node. Both are non-storage nodes. Each node must supply a minimum downstream flow, \(RQ\), to satisfy environmental requirements or:

\[
\sum_{j \in \text{N}} q_{ij}^0 - \sum_{j \in \text{N}} q_{ij}^o \geq RQ_i, \quad \forall i \in \text{NRU} \cup \text{NRD}, \quad \forall o \in \text{O}.
\]

Storage nodes - groundwater aquifers and surface reservoir (surface reservoir is not shown in the following application) - retain water over time. To maintain a sustainable system, the water storage must exceed a required volume, \(R_S\), for all storage sources \(i\):

\[
\text{WS}_i^t \geq R_S^i, \quad \forall i \in \text{NSS}, \quad \forall o \in \text{O}.
\]

where \(\text{WS}^t\) is computed for every operation time step \((\Delta o^0)\) for the storage node by conservation of mass:

\[
\text{WS}_i^t = \text{WS}_i^{t-1} + \left( \sum_{j \in \text{N}} q_{ij}^0 - \sum_{j \in \text{N}} q_{ij}^o \right) \Delta o^0, \quad \forall i \in \text{NSS}, \forall o \in \text{O}.
\]

By modifying the precipitation coefficient on the inflows (Eq. (16)), this constraint can be also written for a surface storage reservoir.
3.4. Flow constraints through arcs

Arc flows are based on hydraulic relationships for the components and introduce arc flow constraints. A canal’s capacity is estimated using Manning’s open channel flow equation by the defined channel hydraulic characteristics (slope \(S\), roughness \(n\), canal side slope \(z\) and geometry). The channel depth \(d_{ij}\) is a decision variable. Flow in each canal \(A_C\) during each time period must be less than its capacity or:

\[
q_{ij}^o \leq \frac{1.49}{n_{ij}} \sqrt{2 \left( \frac{1 + z_{ij}^2}{2 z_{ij}} \right) d_{ij}^{3.5} n_{ij}^{1/2}}, \quad \forall (i,j) \in A_C.
\]

\[
t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]  

(21)

To maintain reasonable pump operating efficiencies, pump discharge flow rates must be maintained between 50% and 150% of the pump design capacity \(q_{ij}^o\) for all pump arcs and all operational times or:

\[
q_{ij}^o \geq 0.5 x_{ij}^o, \quad \forall (i,j) \in A_U, \quad t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]  

(22)

\[
q_{ij}^o \leq 1.5 x_{ij}^o, \quad \forall (i,j) \in A_U, \quad t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]  

(23)

Pipelines carry all potable supplies. A complete distribution system is too complex to model in this formulation so the arc to domestic nodes is intended to be representative of that system. Pipeline arcs must provide water to the downstream node with a minimum energy. Here, a minimum pressure head \(H_{min,j}\) of 14.0 m (equivalent to 137.9 kPa or 20 psi) was required. If the upstream elevation head is insufficient to overcome friction and provide that downstream pressure, a pump station may be installed.

Conservation of energy is written for each pipeline arc as:

\[
\left( \frac{4 H_{ij}^t - H_{ij}^o q_{ij}^o}{3 x_{ij}^o} - \frac{8 f_{ij} q_{ij}^o d_{ij}^t}{g \pi^2 x_{ij}^o q_{ij}^o - D_{ij}^o} \right) \geq H_{min,j} - 10,000 \left( 1 - \mu_{ij}^o \right), \quad \forall (i,j) \in A_P, \quad t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]

(24)

The upstream node is assumed to be a free surface (no pressure or velocity head) and flow is carried to the downstream location at an elevation difference of \(D_{ij}^t\). Pipe friction losses are computed using the Darcy–Weisbach equation assuming fully turbulent flow and a constant friction factor \(f\). Finally, the energy added by the pump is given by the first two terms on the left hand side of Eq. (24). This representative pump curve equation (Walski et al., 1987) is a function of the design variables; pump head and discharge \((H_{ij}^t, x_{ij}^o)\) or:

\[
\left( \frac{4 H_{ij}^t - H_{ij}^o q_{ij}^o}{3 x_{ij}^o} - \frac{8 f_{ij} q_{ij}^o d_{ij}^t}{g \pi^2 x_{ij}^o q_{ij}^o - D_{ij}^o} \right) \geq H_{min,j} - 10,000 \left( 1 - \mu_{ij}^o \right), \quad \forall (i,j) \in A_U, \quad t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]

(25)

3.5. Simple decision bounds

Simple decision variable bounds and the pipe length constraints are:

\[
q_{ij}^o, H_{ij}^t, d_{ij}^t, w_{ij}^t \geq 0, \quad \forall (i,j) \in A, \quad \forall \in N_{WT} \cup N_{WWT}, \quad \forall t \in T, \quad \forall o \in O.
\]

(26)

\[
x_{ij}^o, \chi_{ij}^o \geq \epsilon, \quad \forall (i,j) \in A_P \cup A_U, \quad \forall t \in T.
\]

(27)

\[
q_{ij}^o \leq M_{ij}^o \bar{x}_{ij}^o, \quad \forall (i,j) \in A_P, \quad t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]  

(28)

\[
q_{ij}^o \leq M_{ij}^o \bar{\mu}_{ij}^o, \quad \forall (i,j) \in A_U, \quad t \leq 0, \quad \forall o \in O, \quad \forall t \in T.
\]  

(29)

Eq. (26) requires that, arc flows, pump design head, canal depth, pump design capacity, water and wastewater treatment plant capacities must be non-negative. The lower bound on pipe diameters and pump capacities (Eq. (27)) is assigned a small value \((10^{-5})\) instead of 0 to avoid numerical error because hydraulic relationships given in Eqs. (24) and (25) cannot be divided by 0. Binary variables are included to denote the construction of a new pipe \((x_{ij}^o)\) or pump station \(\left(\mu_{ij}^o\right)\). The terms, \(M_{ij}^o\), in Eqs. (28) and (29) are assigned large values to define upper bounds on the corresponding flows. If the binary variables are set to 0, the flow rate variable for the corresponding arc must also be 0. Otherwise, flow can be allocated to those arcs.

4. Data uncertainty and robust formulation

4.1. Correlated model of data uncertainty

Predictions of future conditions inherently involve uncertainty. The most significant uncertainties for a water supply system are the water demands and supplies that arise from the predictions of future population and precipitation, respectively. The uncertainties are complicated by correlations between the variables. Consumptive use and imported water are dependent on the amount of water. It is likely that during drought conditions, water demand, particularly consumptive use, will increase while less water will be available. Precipitation appears directly in the relationships for estimating streamflow and groundwater storage (Eqs. (15) and (16)). In this study, uncertainties in the parameters of precipitation \(\left(\bar{P}^o\right)\), water demand \(\left(D_L^t\right)\) and imported water availability \(\left(W^o\right)\), and the relationship between precipitation and water demand and availability are considered.

All random variables are assumed to have bounded, symmetric distributions. For instance, water demand for agricultural area has a lower and an upper bound and is a random variable within this range. A non-symmetric distribution could also be modeled. As an independent random variable, the precipitation is expressed as:

\[
\bar{P}^o = \bar{P}^o + \bar{\eta}_1^o \bar{P}^o, \quad \forall o \in O.
\]

(30)

where \(\bar{P}^o\) is the nominal precipitation in operational period \(o\), \(\bar{P}^o\) is half of the precipitation interval that is assumed to be 10% of the nominal precipitation, and \(\eta_1^o\) is a random variable in the interval \([-1, 1]\). Therefore, the range of precipitation is \(\left[\bar{P}^o - \bar{\eta}_1^o \bar{P}^o, \bar{P}^o + \bar{\eta}_1^o \bar{P}^o\right]\).

User demands can be expressed in a similar manner. Domestic and agricultural area demands take values according to a bounded, symmetric distribution with mean equal to the nominal value of \(D_L^t\), its half interval, \(D_L^t\), and its correlation with precipitation, \(-\rho D_L^t \bar{P}^o\).

Assuming that water demands are random variables and are correlated with precipitation, we express demand as:

\[
\bar{D}_L^t = \bar{D}_L^t + \bar{\eta}_1 \bar{D}_L^t - \bar{\eta}_1 \bar{P}^o \bar{D}_L^t A_{DO,OUT}, \quad \forall o \in N_U, \quad \forall o \in O.
\]

(31)

In particular, for our application,

\[
\bar{D}_L^o = \bar{D}_L^o + \bar{\eta}_2 \bar{D}_L^o - \bar{\eta}_2 \bar{P}^o \bar{D}_L^o A_{DO,OUT} \quad \text{(domestic area)}
\]

(31a)

\[
\bar{D}_L^o = \bar{D}_L^o + \bar{\eta}_3 \bar{D}_L^o - \bar{\eta}_3 \bar{P}^o \bar{D}_L^o A_{AG} \quad \text{(agricultural area)}
\]

(31b)

where \(A_{DO,OUT}\) and \(A_{AG}\) are the outdoor land areas in domestic and agricultural irrigation. \(\bar{D}_L^o\) and \(\bar{D}_L^o\) are total water demand in domestic...
and agricultural area, respectively. Water demand in a domestic area is calculated based on the population at period \(0\). Since population is assumed to increase over time, domestic water demand rises and the range of its random parameter increases because the half interval of random parameter is set as 10% of nominal value. This is a realistic assumption because long-term future predictions have more uncertainty than short-term predictions. Relationship between precipitation and user demands is negatively correlated with precipitation while short-term predictions have more uncertainty than short-term predictions. That is, during periods of high precipitation within the basin, user demand decreases. Similarly with high precipitation on the basin contributing to the external water source, the available imported water decreases. Furthermore, the uncertainty in user water demand including its correlation to other user demands is negatively correlated with precipitation while short-term predictions have more uncertainty than short-term predictions. Thus, the degree of conservatism, \(\Gamma\), is introduced by Bertsimas and Sim (2004) controls the level of system reliability as described previously in Section 2.

In the robust formulation, the constraints that contain the uncertain water demand, imported water availability and precipitation (Eqs. (12), (14)-(16)) are rewritten in the form of Eq. (7). To do this, we first rewrite the constraints explicitly, using the random parameters as modeled in Section 4.1. Then, we introduce \(\Gamma\) into the constraint and explicitly write the max-term as in Eq. (7). We note that while our formulation is nonlinear and involves discrete variables, we are able to use robust framework for these constraints as they are linear, involve continuous variables and we are able to calculate the max-terms explicitly without requiring linear programming dual information. We start with robust formulation of Eq. (12).

The uncertainty in user water demand including its correlation with precipitation can be included in a general form in Eq. (12) as:

\[
\sum_j q_{ji}^o \geq Di_j^o + \eta_i D_i^o + \eta_j \rho_j \hat{P}^o A_i j, \quad \forall i \in N_i, \forall o \in O.
\]  

Writing Eq. (33) specifically for domestic and agricultural users in our application (Fig. 1) gives:

\[
q_{45} + q_{25} + A_{DO_OUT} (P^o + \eta_4 \hat{P}^o) 
\geq D_5^o + \eta_2 D_2^o + \eta_1 \rho_1 \hat{P}^o A_{DO_OUT}.
\]

for domestic areas, and

\[
q_{16} + q_{36} + q_{46} + q_{56} + A_{AG} (P^o + \eta_6 \hat{P}^o) 
\geq D_6^o + \eta_3 D_3^o + \eta_2 \rho_2 \hat{P}^o A_{AG}.
\]

for agricultural areas.

Eq. (34) is rearranged to the form:

\[
-q_{45} + q_{25} + D_5^o - A_{DO_OUT} P^o + \eta_2 D_2^o + \eta_1 \rho_1
\left(-A_{DO_OUT} \hat{P}^o(1 + \rho_1)\right) \leq 0.
\]

To write the robust constraint for water demand of domestic users, we introduce the parameter \(\Gamma\) which can take values in the range \([0, 2]\) since only two random parameters appear in Eq. (36). Following the form of Eq. (7), the robust formulation can be written explicitly as

\[
-q_{45} - q_{25} - D_5^o - A_{DO_OUT} P^o + \max \left\{ D_5^0 + \Gamma (1 - 1) \right\} 
- A_{DO_OUT} \hat{P}^o(1 + \rho_1) (1 - \Gamma) + \left| -A_{DO_OUT} \hat{P}^o(1 + \rho_1) \right| \leq 0
\]

and when \(\Gamma < 1\), the robust form of Eq. (36) in the form of Eq. (7) can be written explicitly as:

\[
-q_{45} - q_{25} - D_5^o - A_{DO_OUT} P^o + \max \left\{ \Gamma D_5^0; \Gamma \right\} 
- A_{DO_OUT} \hat{P}^o(1 + \rho_1) \right| \leq 0.
\]

### Table 2

<table>
<thead>
<tr>
<th>(\Gamma)</th>
<th>Probability of constraint violation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>(f_1)</td>
<td>2.00</td>
</tr>
<tr>
<td>(f_2)</td>
<td>2.00</td>
</tr>
<tr>
<td>(f_3)</td>
<td>2.00</td>
</tr>
<tr>
<td>(f_4)</td>
<td>1.00</td>
</tr>
<tr>
<td>(f_5)</td>
<td>1.00</td>
</tr>
<tr>
<td>(f_6)</td>
<td>2.00</td>
</tr>
<tr>
<td>(f_7)</td>
<td>3.00</td>
</tr>
<tr>
<td>(f_8)</td>
<td>4.00</td>
</tr>
<tr>
<td>(f_9)</td>
<td>5.00</td>
</tr>
<tr>
<td>(f_{10})</td>
<td>6.00</td>
</tr>
<tr>
<td>(f_{11})</td>
<td>6.92</td>
</tr>
<tr>
<td>(f_{12})</td>
<td>7.64</td>
</tr>
<tr>
<td>(f_{13})</td>
<td>8.10</td>
</tr>
<tr>
<td>(f_{14})</td>
<td>8.20</td>
</tr>
</tbody>
</table>
We note that the max-terms above can be calculated offline, before optimization. We introduce \( I_2 \) for the agricultural area water demand constraint, and write it when \( I_2 \geq 1 \) in its robust form as:

\[
-q_{16}^o - q_{26}^o - \alpha q_{36}^o - q_{46}^o - D_{6}^o - A_{PG}^o(1 + \rho_2) + \max\left\{ \frac{\Delta}{2} - (I_2 - 1) \right. \\
\left. - A_{PG}^o(1 + \rho_2) + \left( (I_2 - 1) D_{6}^o + \max \left\{ \begin{array}{c} \frac{\Delta}{2} + (I_2 - 1) \frac{\Delta}{2} + \max \left\{ \begin{array}{c} \frac{\Delta}{2} \end{array} \right\} \end{array} \right\} \right) \right\} \leq 0 \tag{39}
\]

where once again, the max-term can be calculated offline. When \( I_2 < 1 \), the robust form is similar and hence omitted here.

Eq. (14), which requires that external water volume use must be less than imported water availability, can also be rewritten using the uncertainty model as:

\[
q_{12}^o + q_{14}^o + q_{16}^o \leq \Delta + \eta_4 \Delta + \eta_3 \Delta \geq A_{PG}^o + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \end{array} \right\} \end{array} \right\} \right\} \right\} \leq 0. \tag{40}
\]

The minimum river flow constraint (Eq. (18)) can be rewritten by substituting the random terms of the inflow due to precipitation and the assumption that 30% of the precipitation reaches the river or:

\[
0.3A_{0}\left( \frac{\Delta}{2} + \eta_4 \Delta + \eta_3 \Delta \right) + q_{28}^o - q_{34}^o - q_{36}^o - q_{32}^o \geq \Delta + \eta_4 \Delta + \eta_3 \Delta \geq A_{PG}^o + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \end{array} \right\} \end{array} \right\} \right\} \leq 0. \tag{42}
\]

Introducing \( I_4 \) and rearranging the terms of Eq. (42) gives the robust form of Eq. (15) as:

\[
-q_{28}^o + q_{34}^o + q_{36}^o - q_{32}^o - 0.3A_{0} \Delta + I_4^o + 0.3A_{0} \Delta \geq \Delta + \eta_4 \Delta + \eta_3 \Delta \geq A_{PG}^o + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \end{array} \right\} \end{array} \right\} \leq 0. \tag{43}
\]

Finally, precipitation within the basin causes uncertainty in flows to the groundwater systems and in the upstream river inflows. Thus, the minimum groundwater storage constraint (Eq. (19)) is reformulated as a robust constraint by accounting for the uncertainty in input over time as:

\[
IGS + \sum_{k=1}^{0} \left( q_{12}^o + q_{32}^o + q_{52}^o + q_{62}^o + q_{32}^o + q_{25}^o + q_{26}^o \right) + 0.1A_{0} \sum_{k=1}^{0} \left( \eta_4 \Delta + \eta_3 \Delta \right) \geq \Delta + \eta_4 \Delta + \eta_3 \Delta \geq A_{PG}^o + \max \left\{ \begin{array}{c} \Delta \\
\left( I_3 - 1 \right) \Delta + \max \left\{ \begin{array}{c} \Delta \end{array} \right\} \end{array} \right\} \leq 0. \tag{44}
\]

where IGS is initial groundwater storage. Note that the 0.1 coefficient relates to the assumption that the defined 10% of precipitation infiltrates to the aquifer. Eq. (44) contains o random parameters,
So, we introduce $\eta_1, \eta_2, \ldots, \eta_o$. So, we introduce $o$ parameters, $\Gamma_{4+1}, \Gamma_{4+2}, \ldots, \Gamma_{4+o}$ and formulate the robust counterpart as:

$$-\text{IGS} - 0.1A_b \sum_{k=1}^{o} \overline{p}^k - \sum_{k=1}^{o} \left( q_{12}^k + q_{13}^k + q_{14}^k - q_{25}^k - q_{26}^k \right)$$

$$+ RS_2 + \left| -0.1A_b \sum_{k=1}^{o} \Gamma_{k+4}^k \right| \leq 0,$$

(45)

where $\Gamma_{k+4}^k \in [0, o], k \in \{1, 2, \ldots, o\}$. Since there are 10 operational periods in our application, this set of constraints results in 10 parameters, $\Gamma_5, \Gamma_6, \ldots, \Gamma_{14}$.

4.3. Probability bounds

In the robust formulation, Eqs. (37), (39), (41), (43), and (45) replace their deterministic forms, Eqs. (12), (14)–(16). Fourteen additional constants, $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_{14})$, are included to control system conservatism. The values of $|\Gamma_k|$ (i.e., number of uncertain parameters in constraint $k$) are the upper bounds of the parameter $\Gamma_k$. The first three $\Gamma_k$ are domestic and agricultural demand satisfaction and imported water availability with ranges equal to $[0, 2]$. Inflow from precipitation to river system is controlled by $\Gamma_4$ that has a range of $[0, 1]$. The last 10 constants,
Table 3
Flow allocations (m³/s) along operational periods for nominal problem.

<table>
<thead>
<tr>
<th>Operational Period</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRIuTAG</td>
<td>0.84</td>
<td>0.84</td>
<td>0.81</td>
<td>0.61</td>
<td>0.41</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>FWWTWT</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FWWTGW</td>
<td>7.17</td>
<td>7.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FWWTAG</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FGVTDW</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1.61</td>
<td>0.93</td>
</tr>
<tr>
<td>FGVTAG</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FWWTGD</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FWWTDO</td>
<td>6.05</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
</tr>
<tr>
<td>FGWTAG</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FGWTGD</td>
<td>2.19</td>
<td>2.27</td>
<td>2.48</td>
<td>2.71</td>
<td>2.94</td>
<td>3.47</td>
<td>3.86</td>
<td>4.39</td>
<td>4.95</td>
<td>5.54</td>
</tr>
<tr>
<td>FGWTDO</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FWDOTW</td>
<td>6.05</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
<td>6.16</td>
</tr>
<tr>
<td>FWWTDO</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

FRIuTWT – flow allocation from upstream river to water treatment plant, FRIuTAG – flow allocation from upstream river to agricultural area, FWWTWT – flow allocation from imported water to water treatment plant, FWWTGD – flow allocation from imported water to groundwater, FWWTAG – flow allocation from imported water to agricultural area, FWWTDO – flow allocation from water treatment plant to downstream river, FWDOTW – flow allocation from downstream river to water treatment plant, FWWTAG – flow allocation from wastewater treatment plant to agricultural area, FWWTGD – flow allocation from wastewater treatment plant to downstream river.

5. Results and discussion

The mixed-integer nonlinear problem was solved using the GAMS/BARON global optimization solver with the relative termination tolerance of 0.05 (Sahinidis and Tawarmalani, 2005). The Branch-And-Reduce Optimization Navigator (BARON) is a GAMS solver for the global solution of nonlinear (NLP) and mixed-integer nonlinear programs (MINLP). The BARON implements algorithms of the branch-and-bound type enhanced with a variety of constraint propagation and dual properties for reducing ranges of variables in the course of the algorithm. The relative termination tolerance means that the optimizer will stop with a solution whose objective function value is within this tolerance of the objective function value of the best possible solution.

To demonstrate the effect of robustness on the model results, the system is optimized for violation probabilities ranging from 0.1 (the most conservative) to 1.0 (nominal). Figs. 4 and 5 show the optimal value of the flow and design variables at year 1 in nominal (i.e., all $\Gamma = 0$) and robust problems ($P = 0.1$). When the constraint violation probability equals 0.1, the solution ensures that all of the $\Gamma_S, \Gamma_D, ..., \Gamma_{14}$, are related to the inflow to the aquifer due to precipitation during period $o \in \{1, 2, ..., 10\}$ and $\Gamma_{i,o}$ has the range of $[0, o]$. For a set of values of $\Gamma$, the robust form of the problem is formulated for the system shown in Fig. 1. To examine the effect of uncertainties and begin to answer the question of how reliable the system should be designed for, the problem is solved for alternative values of the conservatism parameters and the decision maker can then judge the tradeoff between the conservatism and total cost. For instance, with the $\Gamma$ values equal to 0, the max-term equals 0, i.e. the mean parameter values are used in the optimization model and no uncertainty is considered. As described in Section 2, the values of $\Gamma$ can be calculated using Eq. (9) for a given allowable constraint violation probability and are listed in Table 2. For the same probability level, the $\Gamma$ values vary due to the different number of random variables that appear in problem constraints. Fig. 3 shows the probability bounds of constraint violation vs. the corresponding value of $\Gamma$ for $n = |K| = 10$ and $n = |K| = 2$. Generally, probability bounds have log-function shape similar to Fig. 2a. However, small number of uncertain parameters in a constraint $i, n = |K| = 2$, causes a different shape as shown in Fig. 2b.
constraints will remain feasible at least 90% of the time. The values of the $14 \Gamma_k$ in this case ($P = 0.1$) are {2.00, 2.00, 2.00, 1.00, 1.00, 2.00, 3.00, 3.72, 4.20, 4.34, 4.61, 4.91, 4.93, 5.29} (Table 2).

Groundwater storage requirements (Eq. (45)) have more uncertain parameters ($|K_k|$) in later operation periods ($\Gamma_5 \sim \Gamma_{14}$). Uncertainty in yearly precipitation is generated independently and the total uncertainty increases over time. Domestic demands increase with larger populations.

Optimal arc flows for the nominal and robust problems are listed in Tables 3 and 4, respectively. Most noticeably, to meet the water demands, imported water at year 1 in the nominal problem (7.17 m$^3$/s) is increased to 10.02 m$^3$/s when $P = 0.1$. Due to the expense of imported water ($0.81$/m$^3$), both alternatives replenish the aquifer with those waters in year 1 and no additional imported water is purchased. Both cases use reclaimed water from the wastewater treatment plant as the primary agricultural source. The robust solution requires other sources while the nominal solution only uses a small amount beyond reclaimed water.

Table 5 and 6 list the optimal design decisions for the two problems. Domestic area demand increases over time as the population grows while agricultural demand decreases after the fifth operation year. Therefore, few components are expanded in the second design epoch (year 6). Capacities and heads of pumps that provide water to and from domestic areas are expanded under the nominal condition to meet the higher demand. Increased domestic demand is supplied from the upstream river through the wastewater treatment plant that, in turn, reduces water supply to the agricultural area from the upstream river. Reclaimed water from wastewater treatment plant replaces this flow after year 5 and requires the pump capacity to be expanded (Table 5).

When the constraint violation probability is 0.1, only the pump capacity from wastewater treatment plant to agricultural area is expanded; from 5.78 m$^3$/s to 7.46 m$^3$/s because of increasing uncertainty in precipitation (Table 6). Water and wastewater treatment plants in nominal condition have the capacity of 7.13 m$^3$/s at design period 1 and expanded 8.12 m$^3$/s and 10.06 m$^3$/s at design period 2, respectively. When the constraint violation probability is 0.1, the optimal water treatment plant capacity is 6.16 m$^3$/s and no expansion is necessary. On the other hand, the initial wastewater treatment plant capacity is expanded to 8.67 m$^3$/s in design period 1 and further enlarged to 11.20 m$^3$/s in design period 2. A smaller water treatment plant is constructed in the robust solution because supplying water to the domestic area from the aquifer is more economical than expanding the water treatment plant. In this case, the aquifer is recharged with imported water. The wastewater treatment plant capacity, however, increases when the constraint violation probability is 0.1 because effluent from domestic area increases as demand rises.

As seen in Fig. 6, the optimal total cost depends upon the degree of conservatism. Note that the shape of the optimal cost is similar to the shape in Fig. 2b that shows the relationship between constraint violation and the values of $\Gamma_1$, $\Gamma_2$, and $\Gamma_3$. In particular, the optimal cost and the $I$ values increase sharply between constraint violation probabilities of 0.7 and 0.5. These $I$ are related to domestic and agricultural demands and imported water availability. Thus, it appears that uncertainties in these terms dominate the total system cost.

As conservatism is raised (and constraint violation probability is decreased), system reliability is insured by expanding component sizes or purchasing more imported water. Both options increase the system cost. At high constraint violation probabilities, the cost is relatively flat. Above about $P = 0.7$, an expanded water treatment plant supplies water to domestic users. The flatness of the curve demonstrates the economies of scale in meeting needs through the single treatment plant.

As the robustness requirement is increased, the strategy to meet demands changes. The water treatment plant is not expanded rather its capacity remains at its initial size. Given the demand, developing or expanding the conveyance system (canal and pumping system) to import water and deliver it to domestic users becomes viable. As seen in Fig. 6, the amount of imported water parallels the increase in cost. Below the constraint violation probability level of 0.5, economies of scale again dominate and the incremental cost of supply results in a relatively flat response.

Lastly, Monte Carlo simulation was implemented to analyze the probability of the system failure. One hundred thousand realizations were evaluated for the nominal (i.e., all $I = 0$) and robust ($P = 0.1$) problem solutions. Defining system failure as when the demand of one or more users is not satisfied, the probabilities of system failure in the nominal and robust condition are found to be
51% and 7.1%, respectively by the Monte Carlo simulation. The nominal solution’s failure probability is expected to near 50% given the symmetric probability distributions. The low failure probability of the $P = 0.1$ solution appears reasonable given the defined robustness level. We have repeated the Monte Carlo simulation for other values of $P$ to estimate the actual probability of system failure $(\hat{P})$. Fig. 7 shows the comparison of theoretical and simulated probability of constraint violation of the water supply system. The dotted lines indicate the upper and lower bounds estimated from the Monte Carlo simulation for 90% of confidence interval $\frac{\hat{P} - 1.645 \sqrt{\hat{P}(1 - \hat{P})/100,000}}{\hat{P} + 1.645 \sqrt{\hat{P}(1 - \hat{P})/100,000}}$.

Since the discrepancies are within the interval, the solutions obtained from the robust optimization are acceptable. Especially, there is a good fit between theoretical and simulated probability of constraint violation around the mean which is most important for the water supply. Therefore, the robust optimization applied in this study is suitable for the water supply system.

6. Conclusions

In this study, a robust optimization approach was applied in a hypothetical water supply system to minimize the total system cost. It is shown that the Bertsimas and Sim approach can be a useful tool in water supply system design without introducing additional complexity into the optimization problem to prevent system failure at a certain level of risk. This approach can also be applied to a general water supply system without adding complexity to the optimization problem.

Considering that data in real systems inherently involve uncertainties, it is important to consider these uncertainties during the design process to improve system reliability and robustness. Uncertainties in precipitation, water demand, water availability, and the relationship between precipitation and water demand and water availability were considered and resulted in 14 additional model parameters. The probability of violating a constraint is related to those parameters. The robust problem formulation is virtually identical in structure as the original problem; however, it remains feasible under the worst case.

In the application system, the problem was solved for a range of values of $\tau$ using the GAMS software using BARON global optimization solver (Sahinidis and Tawarmalani, 2005). As system robustness requirements are increased, the optimal solution structure changes in terms of imported water and facility expansions. Because of the high cost to purchase imported water, this water source is only purchased in year 1 and is used to recharge the aquifer. Compared to the nominal solution, conservative solutions import more water to maintain system reliability and preserve environmental flows. As a result, the total construction and operation costs increase with higher robustness levels. The solutions were analyzed by Monte Carlo simulation. The discrepancies between theoretical and simulated probability of constraint violation are within 90% of confidence interval.

In terms of facility expansion, at low robustness levels, expansion of the existing water treatment plant is more economical than building a canal or a pump/pipe to transport flow to the domestic area. As the constraint violation increases to about 0.7 the cost is relatively flat but rises quickly to a second plateau for constraint violation probabilities below 0.5. To achieve increased robustness, a larger canal and pumping station are
constructed to transport additional imported water and meet larger domestic demands.

Future work includes better representations of the system and its decisions including adding additional uncertain parameters to account for the temporal correlation of precipitation and using integer optimization variables for pipe diameters to represent commercially available sizes. Also, distributed water and wastewater treatment plants can be examined in more complex systems to investigate the cost tradeoff between treatment plant construction and transporting water.

Acknowledgements

This material is based upon work supported by SAHRA (Sustainability of semi-Arid Hydrology and Riparian Areas) under the STC Program of the National Science Foundation, Agreement No. EAR-9876800 and the Technology and Research Initiative Fund (TRIF) of the Water Sustainability Program (WSP) at the University of Arizona. The authors appreciate the reviewers’ comments that improved the clarity of this manuscript.

References


