Proceedings of the 2013 Industrial and Systems Engineering Research Conference
A. Krishnamurthy and W.K.V. Chan, eds.

A Likelihood Robust Method for Water Allocation under Uncertainty

David Love
University of Arizona
Tucson, AZ

Güzin Bayraksan
The Ohio State University
Columbus, OH

Abstract

We adapt and extend the likelihood robust optimization method recently proposed by Wang, Glynn, and Ye to examine water allocation under the ambiguous distribution of future available supply and demand in a two-stage setting. We examine the value of collecting additional data and the cost of finding a solution robust to an ambiguous probability distribution. A decomposition-based solution algorithm to solve the resulting model is given. Computational results are presented on a long-term water allocation problem in southeast area of Tucson.

Keywords
Optimization under uncertainty, water resources management, ambiguous stochastic programming, robust optimization, environmental sustainability

1. Introduction and Motivation

More than 25 million people in the southwestern United States depend on the water supplied by the Lower Colorado River Basin for their livelihood. More than half of Tucson's water, for instance, comes from this source. The Colorado River Basin has experienced a sustained period of drought in recent years, which has led to questions about the adequacy of the Colorado to meet future demands, especially as the population of Arizona (and of other states that depend on this water source) increases. Thus, the problem of allocating Colorado water is of critical importance.

In this paper, we present a generalized network model of Colorado River water allocation in Tucson, Arizona motivated by the CALVIN (CALifornia Value Integrated Network) optimal water allocation model of California created by Draper et al. [4]. More general models of Colorado River water distribution have also been studied, such as the Colorado River Reservoir Model [2] and the Colorado River Budget Model [1]. Our model is modified to incorporate future uncertainty by using the Likelihood Robust Optimization (LRO) approach of Wang, Glynn and Ye [7]. LRO is a data-driven method that uses observations of random or unknown parameters to account not only for inherent stochasticity of a problem, but the uncertainty of the probabilistic model itself. This is especially important in our application as we are looking 40 years into the future and there is considerable ambiguity in the uncertainties.

The LRO is especially attractive because only those scenarios of interest, obtained either through observation or simulation, are used directly in the calculations. Thus the size of the problem is polynomial in the sample size, making it computationally tractable. Furthermore, the samples used in the LRO can represent select scenarios that are of interest to the authorities—for instance, generated by the scenario planning process [3, 6]. This fits with practice well; the scenario planning process results in several scenarios that the agencies would especially like to be robust against.

This paper is organized as follows: in Section 2 we state the generalized network model in deterministic and stochastic forms; in Section 3 we extend the Likelihood Robust Optimization (LRO) model for a two-stage stochastic program with recourse; in Section 4 we discuss a method of measuring the value of gathering additional data for input to the LRO; in Section 5 we present a decomposition method for solving the LRO model; and in Section 6 we present some computational results for our application. Finally, we end in Section 7 with conclusions and future work.

2. Generalized Network Water Model

We begin with a multi-period generalized network flow model of Colorado River water allocation in Tucson, defined by a set of nodes and directed arcs (N, A). The nodes represent available water supply from the Colorado River, water treatment plants, reservoirs, and water demand sites. The arcs represent the conveyance system (pipes, etc.) that carry water between the nodes. Water can be stored in between time periods in reservoirs to meet future demands. The model aims to find the minimal cost water flows considering energy, treatment, storage, and transportation costs over the planning period. Generalized network water allocation models have been used to find water allocations and delivery reliabilities and to assess values of different water use operations; see, e.g., [4].

Water flows on arc $(i, j) \in A$ during time period $t = 1, \ldots, P$ are represented by decisions x_{ijt}. Each arc $(i, j) \in A$ and time period t has a unit cost c^x_{ijt}, loss coefficient $0 \le a_{ijt} \le 1$ to account for evaporation, leakage from the pipes, etc., and bounds on the flow $l^x_{ijt} \le x_{ijt} \le u^x_{ijt}$. Each node $j \in N$ has a supply/demand for time period t, b_{jt}. Nodes representing reservoirs are able to store water between time periods. Stored water available at node j at the beginning of time period t is s_{jt}, with associated cost c^s_{jt} and bounds $l^s_{jt} \le s_{jt} \le u^s_{jt}$. Finally, water released into the environment from node j in period t is given by r_{jt}, with bounds $l^r_{jt} \le r_{jt} \le u^r_{jt}$. The deterministic model is a multi-period generalized network flow model of the form

$$
\begin{aligned}
\min_{x,s,r} \ & \sum_{(i,j)\in A} \sum_{t=1}^{P} c^x_{ijt} x_{ijt} + \sum_{j\in N} \sum_{t=1}^{P} c^s_{jt} s_{jt} \\
\text{s.t.} \ & \sum_{i\in N} x_{jit} + s_{j,t+1} + r_{jt} = \sum_{i\in N} a_{ijt} x_{ijt} + s_{jt} + b_{jt}, \quad \forall j,t \\
& l^x_{ijt} \le x_{ijt} \le u^x_{ijt}, \quad \forall i,j,t \\
& l^s_{jt} \le s_{jt} \le u^s_{sjt}, \quad \forall j,t \\
& l^r_{jt} \le r_{jt} \le u^r_{sjt}, \quad \forall j,t.
\end{aligned}
$$

In practice, many parameters are unknown, especially as we further look into the future. This is particularly true of the water supply and demands, denoted by b_{jt} in the above model. To capture this uncertainty, we make this model stochastic by rewriting it as a two-stage linear program. The model has a total of P time periods, which is split into two stages of P_1 and $P_2 = P - P_1$ years each. The water supplies and demands b_{jt} are assumed to be known during the first stage of P_1 years and in the remaining P_2 years, they are uncertain and modeled using a discrete set of scenarios with a specific probability distribution. Note that we also assume some lower and upper bounds to be stochastic in the second stage. The model thus becomes

$$
\begin{aligned}
\min_{(x,s,r)\in L^1} \ & \sum_{(i,j)\in A} \sum_{t=1}^{P_1} c^x_{ijt} x_{ijt} + \sum_{j\in N} \sum_{t=1}^{P_1} c^s_{jt} s_{jt} + \sum_{\omega=1}^{n} p_\omega h^\dagger_\omega(s) \qquad (1) \\
\text{s.t.} \ & \sum_{i\in N} x_{jit} + s_{j,t+1} + r_{jt} = \sum_{i\in N} a_{ijt} x_{ijt} + s_{jt} + b_{jt}, \quad \forall j, 1 \le t \le P_1,
\end{aligned}
$$

where the ω index indicates the n second-stage scenarios with probabilities p_ω,

$$
\begin{aligned}
h^\dagger_\omega(s) = \min_{(x,s,r)\in L^2_\omega} \ & \sum_{(i,j)\in A} \sum_{t=P_1+1}^{P} c^x_{ijt} x_{ijt} + \sum_{j\in N} \sum_{t=P_1+1}^{P} c^s_{jt} s_{jt} \qquad (2) \\
\text{s.t.} \ & \sum_{i\in N} x_{jit} + s_{j,t+1} + r_{jt} = \sum_{i\in N} a_{ijt} x_{ijt} + s_{jt} + b^\omega_{jt}, \quad \forall j, P_1 + 1 \le t \le P,
\end{aligned}
$$

and L^1 and L^2_ω represent the feasible regions defined by the lower and upper variable bounds.

In the rest of the paper, we simplify the notation for the first-stage (1) and second-stage (2) problems as follows. In the first stage, decision variables $\{x_{ijt}\}$, $\{s_{jt}\}$ and $\{r_{jt}\}$ become the vector $\mathbf{x}$, costs $\{c^x_{ijt}\}$ and $\{c^s_{jt}\}$ are written as the row vector $\mathbf{c}$, the supply/demand parameters b_{jt} become the vector $\mathbf{b}$ and the constraint matrix is written as A. In the second stage, we denote the decisions as $\mathbf{y}^\omega$, the costs as $\mathbf{q}^\omega$, the supply/demands as $\mathbf{d}^\omega$, and the constraint matrices multiplying $\mathbf{y}^\omega$ and $\mathbf{x}$ as D^ω and B^ω, respectively. This notation puts the generalized network model in the form of a two-stage stochastic linear program with recourse (SLP-2), formulated as

$$
\begin{aligned}
\min_{\mathbf{x}} \ & \mathbf{c}\mathbf{x} + \sum_{\omega=1}^{n} p_\omega h^\dagger(\mathbf{x}) \qquad (3) \\
\text{s.t.} \ & A\mathbf{x} = \mathbf{b} \\
& \mathbf{x} \ge 0
\end{aligned}
$$

where

$$h^{\dagger}_{\omega}(\mathbf{x}) = \min_{y^{\omega}} \mathbf{q}^{\omega} y^{\omega} \quad (4)$$
$$\text{s.t. } D^{\omega} y^{\omega} = B^{\omega}\mathbf{x} + \mathbf{d}^{\omega}$$
$$y^{\omega} \geq 0.$$

For simplicity we assume relatively complete recourse; i.e., the second-stage problems $h^{\dagger}_{\omega}(\mathbf{x})$ are feasible for every feasible solution $\mathbf{x}$ of the first-stage problem. In our application, there are penalty costs when demand is not met; therefore, our model has relatively complete recourse.

3. LRLP-2 Formulation

The SLP-2 formulation assumes that the distribution $\{p_{\omega}\}$ is known. However, in many applications, including our water planning, the distribution is itself unknown. One technique to deal with this is to replace the known distribution with an *ambiguity set* of distributions; i.e., a set of distributions which is believed to contain the true distribution. In the likelihood robust formulation, we assume scenario ω has been observed N_{ω} times, with $N = \sum_{\omega=1}^{n} N_{\omega}$ total observations. In SLP-2, this would correspond to probability of scenario ω to be $p_{\omega} = N_{\omega}/N$, which is the maximum likelihood distribution. The *ambiguity set*, however, is defined by the set of distributions with sufficiently high empirical likelihood $\prod_{\omega=1}^{n} p_{\omega}^{N_{\omega}}$. By replacing the specific distribution in SLP-2 with a set of distributions with high empirical likelihood, we create a model that we refer to as two-stage likelihood robust linear program with recourse (LRLP-2).

To derive the LRLP-2, we begin by writing SLP-2 given in (3)–(4) in extensive form

$$\begin{array}{lrll}
\min_{\mathbf{x},y^{\omega}} & \mathbf{cx} & +\sum_{\omega} p_{\omega}\mathbf{q}^{\omega}y^{\omega} & \\
\text{s.t.} & A\mathbf{x} & & = \mathbf{b} \\
 & -B^{\omega}\mathbf{x} & +D^{\omega}y^{\omega} & = \mathbf{d}^{\omega},\ \forall\omega \\
 & \mathbf{x} & & \geq 0 \\
 & & y^{\omega} & \geq 0,\ \forall\omega.
\end{array}$$

The SLP-2 formulation is then augmented by the set of distributions with sufficiently high likelihood. To be robust against all these possible distributions, the distribution that results in the maximum expected cost is considered. Then, the objective function is minimized with respect to this worst-case distribution selected from the ambiguity set of distributions. The resulting minimax formulation of LRLP-2 is

$$\min_{\mathbf{x},y^{\omega}} \max_{p} \ \mathbf{cx} + \sum_{\omega} p_{\omega}\mathbf{q}^{\omega}y^{\omega} \quad (5)$$
$$\text{s.t. } A\mathbf{x} = \mathbf{b}$$
$$-B^{\omega}\mathbf{x} + D^{\omega}y^{\omega} = \mathbf{d}^{\omega},\ \forall\omega$$
$$\sum_{\omega} N_{\omega}\log p_{\omega} \geq \gamma \qquad (\lambda) \quad (6)$$
$$\sum_{\omega} p_{\omega} = 1 \qquad (\mu) \quad (7)$$
$$\mathbf{x} \geq 0$$
$$y^{\omega}, p_{\omega} \geq 0,\ \forall\omega.$$

Following Wang et al. [7], we have introduced the likelihood parameter γ, and used it to construct the ambiguity set of distributions $\{p_{\omega}\}$ satisfying constraints (6) and (7). Note the likelihood constraint (6) is equivalent to $\prod_{\omega=1}^{n} p_{\omega}^{N_{\omega}} \geq e^{\gamma}$, which explicitly states that the empirical likelihood should be above a certain desired level dictated by γ. Constraint (7), along with nonnegativity constraints on p_{ω}, simply ensures that $\{p_{\omega}\}$ constitutes a probability distribution. Let $0 \leq \gamma' \leq 1$ be the *relative likelihood parameter* that expresses γ as a *proportion* of the maximum likelihood; i.e., $\gamma = \log(\gamma' \prod_{\omega}(\frac{N_{\omega}}{N})^{N_{\omega}})$. In the rest of the paper, we will be referring to γ'.

Taking the dual of the inner maximization problem, with dual variables λ and μ, of constraints (6) and (7), respectively, yields

$$\min_{\lambda,\mu} \ \mu + \bar{N}\lambda + \sum_{\omega} N_{\omega}\lambda(\log\lambda - \log(\mu - \mathbf{q}^{\omega}y^{\omega}))$$
$$\text{s.t. } \lambda \geq 0$$
$$\mu \geq \mathbf{q}^{\omega}y^{\omega},\ \forall\omega,$$

where $\bar{N} = N(\log N - 1) - \log\gamma$. Combining the two minimizations gives LRLP-2 in extensive form

$$\begin{aligned}\min_{\mathbf{x},\lambda,\mu,y^\omega} \ & \mathbf{cx} + \mu + \bar{N}\lambda + \sum_\omega N_\omega \lambda(\log\lambda - \log(\mu - \mathbf{q}^\omega y^\omega)) \\ \text{s.t. } & A\mathbf{x} = \mathbf{b} \\ & -B^\omega \mathbf{x} + D^\omega y^\omega = \mathbf{d}^\omega, \ \forall\omega \\ & \mu \geq \mathbf{q}^\omega y^\omega, \ \forall\omega \\ & \mathbf{x}, \lambda, y^\omega \geq 0, \ \forall\omega. \end{aligned} \tag{8}$$

Finally, we wish to return the LRLP-2 to two-stage formulation. All terms inside the sum over ω will be put into the second stage. To make the formulation as similar to SLP-2 as possible, we choose to express the second stage as an expected value using the maximum likelihood distribution $\frac{N_\omega}{N}$. The formulation becomes

$$\begin{aligned}\min_{\mathbf{x},\lambda,\mu} \ & \mathbf{cx} + \mu + \bar{N}\lambda + \sum_\omega \frac{N_\omega}{N} h_\omega(\mathbf{x},\lambda,\mu) \\ \text{s.t. } & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x}, \lambda \geq 0, \end{aligned} \tag{9}$$

where

$$h_\omega(\mathbf{x},\lambda,\mu) = \min_{y^\omega} N\lambda(\log\lambda - \log(\mu - \mathbf{q}^\omega y^\omega)) \tag{10}$$

$$\begin{aligned}\text{s.t. } & -B^\omega \mathbf{x} + D^\omega y^\omega = \mathbf{d}^\omega \\ & \mu - \mathbf{q}^\omega y^\omega \geq 0 \\ & y^\omega \geq 0. \end{aligned} \tag{11}$$

Since log is uniformly increasing, we can rewrite the second stage problem as $h_\omega(\mathbf{x},\lambda,\mu) = (-N\lambda)\log(\mu - \min_{y^\omega \in Y^\omega} \mathbf{q}^\omega y^\omega)$ with $Y^\omega = \{y^\omega | -B^\omega \mathbf{x} + D^\omega y^\omega = \mathbf{d}^\omega, \mu - \mathbf{q}^\omega y^\omega \geq 0, y^\omega \geq 0\}$. Thus we can state the second stage of LRLP-2 in terms of the second stage of SLP-2, $h_\omega(\mathbf{x},\lambda,\mu) = N\lambda\left[\log\lambda - \log(\mu - h^\dagger_\omega(\mathbf{x}))\right]$.

As noted in [7], the KKT conditions for (8) give the relation between primal and dual variables

$$p_\omega = \frac{\lambda N_\omega}{\mu - h^\dagger_\omega(\mathbf{x})}. \tag{12}$$

4. The Value of Data

With a data driven formulation such as LRLP-2, it is natural to ask how the behavior changes as more data is gathered. In particular, for robust formulations like LRLP-2 one might be concerned about being overly conservative in the problem formulation and thus missing the opportunity to find a better solution to the true distribution. For instance, for our water application, initial estimates might show that water demand would be high or low with approximately equal probabilities. Based on new data, later studies could give low demand an increased weight. For LRLP-2, this means that the initial model is likely to be more conservative in an effort to be robust, while the new information will make the model less conservative because low water demand results in lower costs and our belief is increased that the future demand will be low. In this section, we present a simple method of estimating the probability that taking an additional sample will eliminate the old worst-case distribution and allow for better optimization; i.e., a lower-cost solution.

Consider again the deterministic equivalent formulation of LRLP-2 in (8). Let $f_N(\mathbf{x},\mu,\lambda) = \mathbf{cx} + \mu + \bar{N}\lambda + \sum_\omega N_\omega \lambda$ $(\log\lambda - \log(\mu - h^\dagger(\mathbf{x})))$ be the objective function, and $z_N = \min_{\mathbf{x},\mu,\lambda} f_N(\mathbf{x},\mu,\lambda)$. We wish to find a simple estimate of the decrease in the optimal cost associated with taking an additional sample, $z_N - z_{N+1}$, looking in particular for a condition under which $z_N - z_{N+1} > 0$.

Let $\mathbf{x}^*_N, \mu^*_N, \lambda^*_N \in \operatorname{argmin} f_N(\mathbf{x},\mu,\lambda)$ be optimal solutions to the N-sample problem. Then $z_N - f_{N+1}(\mathbf{x}^*_N, \mu^*_N, \lambda^*_N)$ is a lower bound on the decrease in optimal cost $z_N - z_{N+1}$. Let $\hat{\omega}$ be the scenario that is selected with the additional sample, then

$$z_N - f_{N+1}(\mathbf{x}^*_N, \mu^*_N, \lambda^*_N) = \left[\overline{N} - \overline{N+1} - \log\lambda^*_N + \log(\mu^*_N - h^\dagger_{\hat{\omega}}(\mathbf{x}^*_N))\right]\lambda^*_N.$$

We can bound $\overline{N} - \overline{N+1} = N\log N - (N+1)\log(N+1) + 1$ by using the tangent lines

$$\log\mathbf{x} + 1 \leq (\mathbf{x}+1)\log(\mathbf{x}+1) - \mathbf{x}\log\mathbf{x} \leq \log(\mathbf{x}+1) + 1$$

to get $\overline{N} - \overline{N+1} \geq -\log(N+1)$.

Combining these results gives the condition

$$z_N - f_{N+1}(\mathbf{x}_N^*, \mu_N^*, \lambda_N^*) \geq \left[-\log(N+1) - \log\lambda_N^* + \log(\mu_N^* - h^\dagger(\mathbf{x}_N^*))\right]\lambda_N^* > 0.$$

Note that $\lambda_N^* > 0$, so to guarantee a drop in optimal cost we must show that the first term is positive. This then simplifies to

$$\frac{\mu_N^* - h^\dagger(\mathbf{x}_N^*)}{\lambda_N^*(N+1)} > 1.$$

Using the KKT condition (12), this can be rewritten as

$$\frac{N_{\hat{\omega}}}{N} > \left(\frac{N+1}{N}\right) p_{\hat{\omega}}. \tag{13}$$

The left-hand side of (13) is the empirical probability of scenario $\hat{\omega}$ given by the observations used to solve (9). The right-hand side contains $p_{\hat{\omega}}$, the worst-case probability of scenario $\hat{\omega}$ computed by solving (9), using a total of N observations.

We can interpret (13) as follows. If an additional sample is taken from the unknown distribution and the resulting observed scenario $\hat{\omega}$ satisfies (13), then the $(N+1)$-sample problem will have a lower cost than the N-sample problem that was already solved. This is equivalent to saying that an additional observation of $\hat{\omega}$ will rule out the computed worst-case distribution given by $\{p_\omega\}$ given in (12).

Finally, we would like a lower bound on the probability that the next sample will decrease the optimal cost. Let $L = \left\{\omega : \frac{N_\omega}{N} > \left(\frac{N+1}{N}\right) p_\omega\right\}$, where p_ω is the worst-case distribution discussed above. That is, L gives the set of scenarios that, if sampled one more observation, would result in a decrease in the optimal cost in LRLP-2. We can estimate a lower bound on the probability of sampling a scenario in L by using the same likelihood ambiguity set that was used to formulate LRLP-2 given in (6) to solve the minimization problem

$$\begin{aligned} \min_{\omega\in L} \; & \sum_{\omega\in L} q_\omega \\ \text{s.t.} & \sum_\omega N_\omega \log q_\omega \geq \gamma \\ & \sum_\omega q_\omega = 1 \\ & q_\omega \geq 0, \; \forall\omega, \end{aligned} \tag{14}$$

where we have introduced the dummy variables q_ω to distinguish the minimization in (14) from the worst-case distribution $\{p_\omega\}$ calculated in (12). Solving (14) yields an estimated lower bound on the probability that an additional sample will result a likelihood ambiguity set that does not contain the current worst-case distribution $\{p_\omega\}$ using the current set of observations. Let $N_L = \sum_{\omega\in L} N_\omega$ be the number of observations in set L. Note that $\min_{\omega\in L}\sum_{\omega\in L} q_\omega \leq \frac{N_L}{N}$, because the maximum likelihood distribution is always within the likelihood ambiguity set. We will use $\frac{N_L}{N}$ as a benchmark in Figures 1, 2a and 4.

We solve (14) by taking its dual, which results in the two dimensional nonlinear program

$$-\min_{\mu,\lambda} \mu + \bar{N}\lambda + N\lambda\log\lambda - \lambda N_L \log(\mu - 1) - \lambda(N - N_L)\log\mu, \tag{15}$$

where $\bar{N} = N(\log N - 1) - \log\gamma'$.

We can view the optimal value of (15) as a function of three parameters: the total sample size N, the relative likelihood parameter γ', and the number of observations in the set L, N_L. The behavior of this lower bound estimate is studied in Figure 1 as a function of the ratio $\frac{N_L}{N}$, and in Figures 2a and 2b as a function of the relative likelihood parameter γ'. Figure 1 shows that the estimated lower bound on the probability of optimal cost decrease stays relatively close to the identity line for most values of γ', getting closer to the identity line as γ' is increased; i.e., the ambiguity set (or robustness) is decreased. Figures 2a and 2b give a closer look at how (15) differs from $\frac{N_L}{N}$ as γ' is changed.

Notice, however, that the three parameters discussed, γ', N, and N_L do not play the same role in the LRLP-2. The first two of these parameters are also parameters of the LRLP-2 problem (9). The third, N_L, is computed from the optimal solution of (9) via (12) and (13). As such, N_L should be viewed as changing with γ'. In general, N_L will decrease as γ' increases. To see this, recall that values of γ' close to 1 consider increasingly limited set of distributions, the ones closest to the maximum likelihood distribution $p_\omega = \frac{N_\omega}{N}$, $\forall\omega$. So, the condition given in (13) is satisfied for a smaller number of scenarios. As N_L changes, the estimated bound on the probability of cost decrease will have one or more jump discontinuities, moving from one line to another line below it in Figure 2b, as seen in Figure 4. This behavior is studied in Section 6.

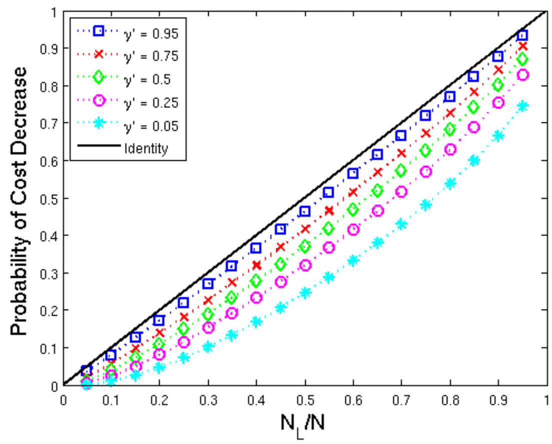
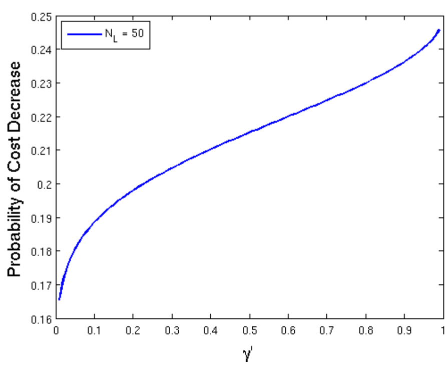
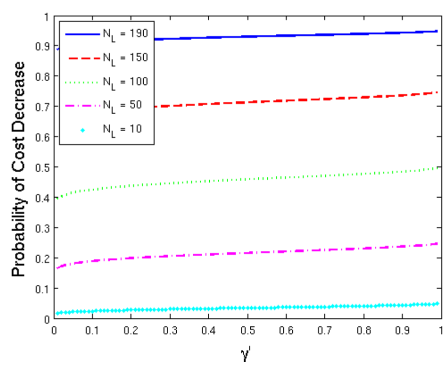


Figure 1: The probability that an additional sample decreases the optimal cost of the LRLP-2 as a function of the ratio $\frac{N_L}{N}$ for total sample size $N = 20$.



(a) $N_L = 50$ $\left(\frac{N_L}{N} = 0.25\right)$



(b) $N_L = 10, 50, 100, 150$ and 190

Figure 2: The probability that an additional sample decreases the optimal cost of the LRLP-2 as a function of γ' for total sample size $N = 200$, (a) for a single value of $N_L = 50$ and (b) for multiple values of N_L.

5. Decomposition-Based Solution Method

As the model gets larger, as in our water application presented in Section 6, a direct solution of LRLP-2 becomes computationally expensive. Decomposition-based methods could significantly reduce the solution time and allow for larger problems to be solved efficiently. In this section, we propose a Bender's decomposition-based method for solving LRLP-2. Our algorithm uses the LRLP-2 second-stage problems $h_\omega(\mathbf{x},\lambda,\mu)$ (10) to form the cuts, ensuring a linear master problem. The algorithm removes constraint (11) from the second-stage problem (10) and exchanges it with a series of feasibility constraints (or cuts) in the first-stage problem. Making this change ensures that the second-stage problems are solved using the formulation of $h^\dagger_\omega(\mathbf{x})$ for SLP-2 given in (4), and is more efficient. The master problem is given by

$$\begin{aligned}\min_{\mathbf{x},\lambda,\mu}\ & \mathbf{cx}+\mu+\bar{N}\lambda+\theta \\ \text{s.t. } & A\mathbf{x}=\mathbf{b} \\ & \theta \geq T_j(\mathbf{x},\lambda,\mu)^T+t_j,\ j\in J \\ & \mu \geq M_k\mathbf{x}+m_k,\ k\in K \\ & \mathbf{x},\lambda \geq 0,\end{aligned} \tag{16}$$

where $T_j(\mathbf{x},\lambda,\mu)^T+t_j$ are the objective cuts, $M_k\mathbf{x}+m_k$ are the feasibility cuts on constraint (11), and J and K are the sets of objective and feasibility cuts, respectively. We next discuss these cuts in more detail.

5.1 Objective Cuts

Let $(\hat{\mathbf{x}},\hat{\lambda},\hat{\mu})$ be the candidate solution from the master problem (16). An objective cut can be computed by solving the SLP-2 subproblems $h^\dagger_\omega(\hat{\mathbf{x}})$ along with optimal dual solutions $\pi^{*,\omega}$ to each second-stage problem, and using these to compute the partial (sub)derivatives of the LRLP-2 subproblems as

$$\begin{aligned}\frac{\partial h_\omega}{\partial \mathbf{x}}(\hat{\mathbf{x}},\hat{\lambda},\hat{\mu}) &= \left(\frac{N\hat{\lambda}}{\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})}\right)\pi^{*,\omega}B^\omega \\ \frac{\partial h_\omega}{\partial \lambda}(\hat{\mathbf{x}},\hat{\lambda},\hat{\mu}) &= N+N\log\hat{\lambda}-N\log(\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})) \\ \frac{\partial h_\omega}{\partial \mu}(\hat{\mathbf{x}},\hat{\lambda},\hat{\mu}) &= \frac{-N\hat{\lambda}}{\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})}\end{aligned}$$

Recall that $h^\dagger_\omega(\mathbf{x})=\min_{y^\omega\geq 0}\{\mathbf{q}^\omega y^\omega | D^\omega y^\omega=\mathbf{d}^\omega+B^\omega\mathbf{x}\}$. The cuts are then given by

$$\begin{aligned}T^\omega_j &= \left(\left(\frac{N\hat{\lambda}}{\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})}\right)\pi^{*,\omega}B^\omega,\quad N+N\log\hat{\lambda}-N\log(\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})),\quad -\frac{N\hat{\lambda}}{\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})} \right) \\ t^\omega_j &= \frac{N\hat{\lambda}h^\dagger_\omega(\hat{\mathbf{x}})-N\hat{\lambda}\pi^{*,\omega}B\hat{\mathbf{x}}}{\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})}.\end{aligned}$$

For the single-cut master problem proposed, $T_j=\sum_\omega \frac{N_\omega}{N}T^\omega_j$ and $t_j=\sum_\omega \frac{N_\omega}{N}t^\omega_j$.

5.2 Feasibility Cuts

After the subproblems $h^\dagger_\omega(\hat{\mathbf{x}})$ are solved, it may be the case that $\hat{\mu}-h^\dagger_\omega(\hat{\mathbf{x}})<0$ for some ω, rendering $\hat{\mu}$ infeasible. This is corrected using the feasibility problem

$$\begin{aligned}U_\omega(\mathbf{x},\mu) = \min_{y_\omega,z\geq 0}\ & z \\ \text{s.t. } & z+\mu-\mathbf{q}^\omega y^\omega \geq 0 \\ & D^\omega y^\omega=\mathbf{d}^\omega+B^\omega\mathbf{x},\end{aligned}$$

which is solved by $z^*=h^\dagger_\omega(\mathbf{x})-\mu$. The subdifferentials can be easily found as $\frac{\partial z^*}{\partial \mathbf{x}}=\pi^{*,\omega}B^\omega$ and $\frac{\partial z^*}{\partial \mu}=-1$. Then for infeasible candidate solution $(\hat{\mathbf{x}},\hat{\lambda},\hat{\mu})$ we get the inequality

$$U_\omega(\mathbf{x},\mu)\geq \pi^{*,\omega}B^\omega(\mathbf{x}-\hat{\mathbf{x}})-(\mu-\hat{\mu})+h^\dagger_\omega(\hat{\mathbf{x}})-\hat{\mu},$$

and setting $U_\omega(\mathbf{x},\mu)=0$ to find a feasible solution gives the feasibility cut

$$\mu\geq \pi^{*,\omega}B^\omega\mathbf{x}+(h^\dagger_\omega(\hat{\mathbf{x}})-\pi^{*,\omega}B^\omega\hat{\mathbf{x}}).$$

Once the feasibility cut is generated, we may need to find a feasible (and reasonable) value of μ to generate an objective cut, or to initialize the next iteration of the master problem. This can be done quickly by minimizing the objective function of (8) with respect to μ while keeping $\hat{\mathbf{x}}$ and $\hat{\lambda}$ constant, which is equivalent to minimizing $\mu - \sum_\omega N_\omega \hat{\lambda} \log(\mu - h^\dagger_\omega(\hat{\mathbf{x}}))$. We do this by solving the equation $\sum_\omega \frac{N_\omega \hat{\lambda}}{\mu - h^\dagger_\omega(\hat{\mathbf{x}})} = 1$ with Newton's method.

5.3 Decomposition Algorithm

We solve the LRLP-2 with the following decomposition algorithm with tolerance level `TOL`

```
Initialize z_l = −∞, z_u = ∞
Solve first stage (16) with θ = 0 to generate x
Solve all second stage scenarios h†_ω(x) (4)
Initialize λ ← 1, μ that minimizes μ − Σ_ω N_ω λ̂ log(μ − h†_ω(x̂))
Generate initial objective cut
while z_u − z_l ≥ TOL min{|z_u|, |z_l|} do
    Solve master problem (16), get x, λ, μ, θ_M
    Solve sub-problems h†_ω(x) (4)
    θ_true ← Σ_{ω=1}^n (N_ω/N) h_ω(x, λ, μ)
    if μ < max_ω h†_ω(x) then
        Generate feasibility cut
        Find μ that minimizes μ − Σ_ω N_ω λ̂ log(μ − h†_ω(x̂))
    else
        z_l ← master optimal cost cx + μ + N̄λ + θ_true
    end if
    Generate objective cut
    if cx + μ + N̄λ + θ_true < z_u then
        z_u ← cx + μ + N̄λ + θ_true
        x_best ← x, λ_best ← λ, μ_best ← μ
        p_ω ← λ_best N_ω / (μ_best − h†_ω(x_best)) for i = 1, ..., n
    end if
end while
```

5.4 Computational Enhancements

In order to enhance the performance of the above decomposition-based algorithm, we made some adjustments. First, we included an L_∞-norm trust region which is scaled up (by a factor of 3) or down (by a factor of $\frac{1}{4}$) when the trust region inhibits finding the optimal solution or when the polyhedral lower approximation is far from the second-stage expected cost, respectively. The trust region is an implementation of Algorithm 4.1 in [5].

Because we are also interested in the worst-case LRO probabilities given in the primal variables and not computed directly, we include a second tolerance as a stopping condition, ensuring that $|1 - \sum_{i=1}^n p_\omega| < \texttt{TOL2}$ when the algorithm is completed. This must be satisfied in addition to the original condition $z_u - z_l < \texttt{TOL}\min\{|z_u|, |z_l|\}$.

6. Application and Computational Results

Using the generalized network model (1) of Colorado River water allocation in the southeastern portion of Tucson, we created a likelihood robust water allocation problem of the form LRLP-2. The southeastern portion of Tucson is a newly developing area and is expected to grow considerably. The LRLP-2 water allocation model would help authorities with future water plans in this area while being robust to uncertainties in water supplies and demands.

The model has a total of $P = 41$ time periods, representing years 2010–2050. For each time period, the network has 62 nodes representing demand for potable and nonpotable (reclaimed) water, pumps, water treatment plants, and the available water supply from the Colorado River. The network in each time period has 102 arcs, representing the pipe network carrying the water between the nodes physically and connecting the network to the five reservoirs that connect the time stages in the model. We use $P_1 = 10$ time periods for the first stage. Uncertainty in the second stage takes the form of uncertain population (thus, demand for water) and supply of water. There are a total of 4 scenarios considered in this test instance: (i) high population, high supply, (ii) high population, low supply, (iii) low population, high supply, and (iv) low population, low supply. Each scenario is assumed to have five observations. The high population scenarios are more costly as the system needs to meet demand or pay for unmet demand. The low population scenarios, on the

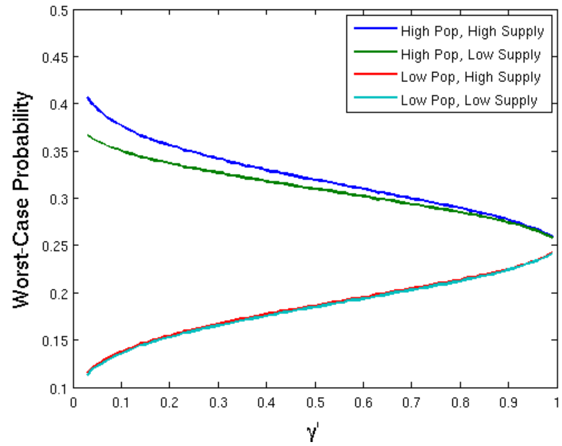


Figure 3: Worst-case distribution for the likelihood robust water allocation problem.

other hand, are not as costly. We applied the decomposition-based solution algorithm presented in Section 5 to solve this model and selected tolerances $\texttt{TOL} = 10^{-5}$ and $\texttt{TOL2} = 10^{-3}$ for our computational experiments.

Figure 3 shows how the worst-case distribution changes with γ'. When γ' is close to 1, we use the maximum likelihood distribution, which has equal $\frac{1}{4}$ probabilities on each of the four scenarios. As γ' is decreased, the ambiguity set increases, and the worst-case distribution used by LRLP-2 changes. It gives higher than $\frac{1}{4}$ probability to the two high-population scenarios and lower than $\frac{1}{4}$ probability to the two low-population scenarios, making the solution more robust to costly scenarios. Note that the scenarios fall into two similar pairs because the cost of each scenario depends strongly on the projected demand but only weakly on the projected supply of Colorado River water. A closer look at the optimal solutions reveals that as γ' is decreased, or as robustness is increased, the solution uses more and more reclaimed water (treated wastewater that is reused for nonpotable purposes such as irrigation) in an effort to meet demands in a least-costly way.

The results of the water model were then analyzed with the value of data techniques from Section 4. Figure 4 shows the estimated probability that an additional sample will remove the worst-case distribution from the likelihood region, resulting in a lower-cost solution. The dashed line in Figure 4 depicts the computed values of $\frac{N_L}{N}$, which provide an upper bound on the estimated probabilities. Because the low-population scenarios have lower costs, an additional sample of either low-population scenario will result in a lower expected cost. This is what we see through most of the computed values of γ', with $\frac{N_L}{N} = 0.5$, indicating that the sufficient condition (13) was satisfied for both low-population scenarios. For extremely large values of γ'—above 0.97—we see the ratio $\frac{N_L}{N}$ quickly drops to zero. This occurs because (13) only compares scenario probabilities in the empirical (N_ω/N) and worst-case (p_ω) distributions and does not use the computed cost of the scenarios. As γ' increases and the ambiguity set shrinks, the worst-case probabilities become so close the empirical probabilities that (13) can no longer be satisfied, resulting in $\frac{N_L}{N}$ decreasing to zero.

7. Conclusion and Future Work

In this paper, we proposed an extension of the Likelihood Robust Optimization (LRO) method of Wang et al. [7] to general two-stage stochastic programs with recourse, creating a two-stage likelihood robust program with recourse, denoted LRLP-2. The LRO models use the empirical likelihood function to define an ambiguity set of probability distributions using observed data and optimize the worst-case expected cost with respect to this likelihood ambiguity set. We provided a computationally simple method to estimate the probability that an additional sample will produce a likelihood ambiguity set that does not contain the current worst-case distribution and will result in a lower-cost solution. We have also provided a Bender's decomposition-based solution algorithm for the LRLP-2 and applied this method to planning future water distribution in Tucson, Arizona.

Our future work includes the following. We plan to augment the existing model first with a richer set of second-stage scenarios. In addition to more varied estimates for future population, we will integrate climate change predictions into

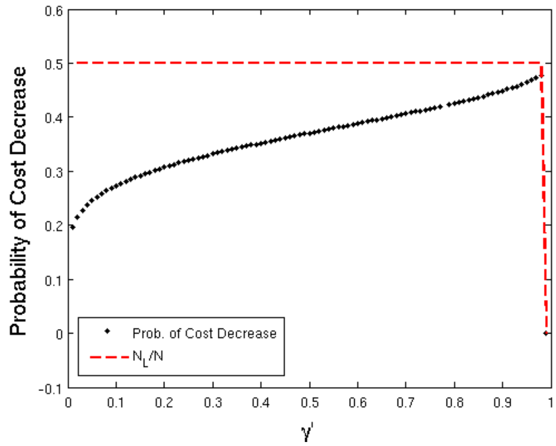


Figure 4: Probability that an additional sample causes a decrease in worst-case expected cost for the likelihood robust water allocation problem. The red line shows the upper bound probability $\frac{N_L}{N}$.

the model to generate scenarios for future water supply from the Colorado River. This model is intended to include a facility location problem to determine the best places for an additional waste water treatment plants to increase the use of reclaimed water in the most cost-efficient manner. On the methodological side, we plan to provide guidelines on selecting the robustness parameter γ with respect to the sample size and investigate the asymptotic behavior of the model as the sample size increases.

Acknowledgements

The authors would like to thank the Water Sustainability Program at the University of Arizona for providing funding for this research. This work has also been partially supported by the National Science Foundation through grant CMMI-1151226.

References

[1] T.P. Barnett and D.W. Pierce. Sustainable water deliveries from the Colorado River in a changing climate. *Proceedings of the National Academy of Sciences*, 106(18):7334–7338, 2009.

[2] N.S. Christensen, A.W. Wood, N. Voisin, D.P. Lettenmaier, and R.N. Palmer. The effects of climate change on the hydrology and water resources of the Colorado River Basin. *Climatic Change*, 62(1):337–363, 2004.

[3] City of Tucson, Water Department. Water plan: 2000-2050. http://cms3.tucsonaz.gov/water/waterplan, 2004.

[4] A.J. Draper, M.W. Jenkins, K.W. Kirby, J.R. Lund, and R.E. Howitt. Economic-engineering optimization for California water management. *Journal of Water Resources Planning and Management*, 129:155–164, 2003.

[5] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Verlag, 1999.

[6] U.S. Department of the Interior, Bureau of Reclamation. Colorado River Basin water supply and demand study: Technical report A – Scenario development. Technical report, 2011.

[7] Z. Wang, P.W. Glynn, and Y. Ye. Likelihood robust optimization for data-driven newsvendor problems. Technical report, Department of Management Science and Engineering, Stanford University, USA, 2010.