Optimal design of decentralized constructed wetland treatment system under uncertainties

Abstract

Keywords:

1. Introduction

2. Problem statement

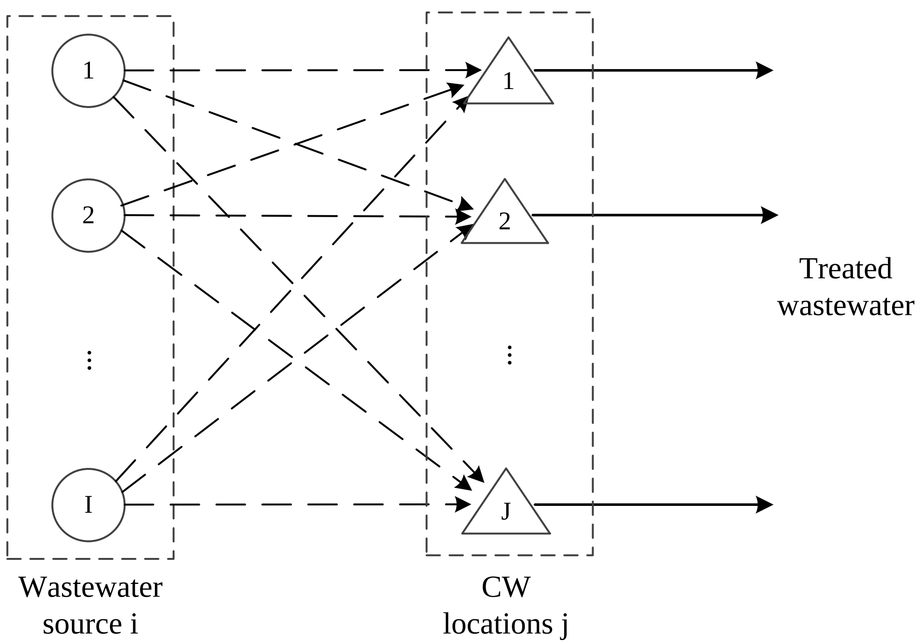

Figure 1: Superstructure of a decentralized CW treatment system network

We consider a general decentralized CW treatment system as shown in Fig.1. Given a set of wastewater sources I and a set of potential CW locations J, wastewater is generated from the sources that contain certain pollutants, then treated in CWs, and finally discharged or reused on site. To guarantee the quality of treated wastewater, we impose treatment targets on a set of pollutants M. That is, we require the pollutant concentrations in the effluent to be below certain guidelines. We also allow different treatment targets to be set for different CW sites based on the local conditions and requirements. For example, if a location aims to reuse the treated wastewater for the purpose of landscape irrigation, regulations on irrigation should be adopted. The goal of this problem is to obtain a feasible planning solution that determines the location and design of CWs, the interconnections between wastewater sources and CWs, and the corresponding allocation amounts. A planning solution is defined to be "feasible" if it could collectively satisfy all constraints, including the treatment targets. Unfortunately, due to the ubiquitous uncertainties in the environment itself as well as the pollutant concentration of influent wastewater, the performance of CWs removing pollutants could be greatly affected. On the other hand, in reality, it is often unrealistic to assure that treated wastewater is certain to achieve all the treatment targets collectively out of the budget concern. Hence, the goal of our problem is to find a planning solution such that the treatment targets could be satisfied as well as possible. Specifically, we make the following assumptions:

- wastewater from each source $i = 1, ..., I$ can be allocated to multiple CW sites and each CW can also treat wastewater from multiple wastewater sources. Sewer lines are needed to be constructed between wastewater sources and CWs.
- for any location $j = 1, ..., J$, if it is determined to construct a CW, K design options $k = 1, ..., K$ could be selected. Otherwise, we use $k = 0$ to indicate that this location is not chosen to construct any CWs.
- a list of M pollutants $m = 1, ..., M$ are evaluated. Treatment target τ_j^m is set for each pollutant m and each CW site j.

A basic list of model parameters and decision variables is provided in Table 1. Other notations would be introduced and defined as per required in the rest of the paper. Before we address the uncertainty issues, let us begin with a deterministic problem that solves a feasible planning solution. The following constraints formulate the CW construction option, wastewater allocation between sources and CWs, pollutant removal performance, treatment target fulfillment and budget restriction:

CW construction option: Let y_{jk} be a binary variable that describes whether construction option k is selected for CW site j. Particularly, $k = 0$ indicates that site j is determined not to construct any CWs. Otherwise, one of the design options $k = 1, ..., K$ would be chosen. Different design options have different CW scales and pollutant removal capacities, and induce different construction fees. Generally, if a CW has a better treatment performance, the corresponding construction cost would be higher. The constraints

$$\sum_{k=0}^{K} y_{jk} = 1, \quad \forall j$$
$$y_{jk} \in \{0, 1\}, \quad \forall j, k$$

ensure that only one specific construction option is chosen for each site.

Wastewater allocation: A binary decision variable x_{ij} is used to describe whether sewer lines are constructed between wastewater source i and location j. If source i and location j are connected, x_{ij} then takes value of 1; otherwise, x_{ij} equals 0. Let z_{ij} denote the wastewater amount assigned from source i to CW site j, we have the following constraints:

Table 1: Notations of model parameters and decision variables

Indices	
i	index of wastewater sources, $i \in \{1, 2, ..., I\}$
j	index of potential CW locations, $j \in \{1, 2, ..., J\}$
m	index of evaluated water pollutants, $m \in \{1, 2, ..., M\}$
k	index of CW construction options, $k \in \{0, 1, 2, ..., K\}$
Model parameters	
ε_i^m	concentration of pollutant m in the wastewater source i (mg/m^3)
$\varepsilon_{in,j}^m$	concentration of pollutant m in the influent of CW site j (mg/m^3)
$\varepsilon_{out,j}^m$	concentration of pollutant m in the effluent of CW site j (mg/m^3)
τ_j^m	treatment target for pollutant m in CW site j (mg/m^3)
F_i	total wastewater flow generated by source i (m^3/d)
Q_{jk}	flow capacity of CW in option k for site j (m^3/d)
A_{jk}	area of CW in option k for site j (m^2)
$c_{cw,jk}$	construction cost of CW in design option k for site j (\$)
d_{ij}	distance between wastewater source i and site j (m)
c_s	unit construction cost of sewer lines per distance (\$/m)
Decision variables	
x_{ij}	binary variable, $x_{ij} = 1$ if sewer lines are constructed from wastewater source i to CW site j and 0 otherwise
y_{jk}	binary variable, $y_{jk} = 1$ if construction option k is chosen for site j and 0 otherwise. In particular, y_{j0} denotes the choice of not constructing any CWs in site j
z_{ij}	wastewater flow assigned from wastewater source i to the CW in site j

$$
\begin{aligned}
&\sum_{i=1}^{I} z_{ij} \le \sum_{k=0}^{K} Q_{jk} y_{jk}, && \forall j \\
&\sum_{j=1}^{J} z_{ij} = F_i, && \forall i \\
&0 \le z_{ij} \le F_i x_{ij}, && \forall i, j \\
&x_{ij} \in \{0, 1\}, && \forall i, j,
\end{aligned}
$$

where Q_{jk} is the flow capacity of CW in design option k and F_i is the wastewater flow rate generated from source i.

These constraints ensure that

- a CW site can treat wastewater from any sources only if the CW is open and its maximum flow capacity is not exceeded. In particular, we set Q_{j0} to be 0, which allows no wastewater to be treated in site j when no CWs are constructed there.
- wastewater generated from each source must be fully treated.
- wastewater could be assigned from source i to CW site j only if they are connected.

Pollutant removal performance: Let ε_i^m be the concentration of pollutant m in the wastewater source i. We firstly formulate the problem by assuming that ε_i^m is deterministic. Later we would extend ε_i^m to be random under the uncertainty assumptions. Similarly, we denote the pollutant concentration in the influent and effluent of CW in site j by $\varepsilon_{in,j}^m$ and $\varepsilon_{out,j}^m$, respectively. Given an allocation strategy $\boldsymbol{z}$, $\varepsilon_{in,j}^m$ could be calculated as the average concentration in the mixed wastewater that flows into CW j:

$$
\varepsilon_{in,j}^m \sum_{i=1}^{I} z_{ij} = \sum_{i=1}^{I} z_{ij} \varepsilon_i^m, \quad \forall j, m. \tag{1}
$$

To account for the relationship between pollutant concentrations in the influent $\varepsilon_{in,j}^m$ and effluent $\varepsilon_{out,j}^m$, we introduce the following general expression to describe the pollutant removal performance of CWs:

$$
\varepsilon_{out,j}^m = \sum_{k=0}^{K} (a_{jk}^m \varepsilon_{in,j}^m + b_{jk}^m) y_{jk}, \quad \forall j, m, \tag{2}
$$

where a_{jk}^m and b_{jk}^m are design parameters that depend on the option k and the pollutant type m. We can observe that once the design choice k is determined, the effluent concentration $\varepsilon_{out,j}^m$ is affinely dependent on the influent concentration $\varepsilon_{in,j}^m$. In particular, a_{jk}^m is nonnegative, hence $\varepsilon_{out,j}^m$ is non-decreasing in $\varepsilon_{in,j}^m$. As reviewed in **?**, regression equations are widely assumed in the CW design models and demonstrated to be effective. Furthermore, even for the more complex physical models of CW performances (e.g., first-order k-C^* model), they also follow this affine relationship once the design parameters are fixed. To substitute $\varepsilon_{in,j}^m$ in (2) with (1), we can then obtain the following expressions of effluent pollutant concentrations:

$$\varepsilon_{out,j}^{m}\sum_{i=1}^{I}z_{ij}=\sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^{m}\varepsilon_{i}^{m}+b_{jk}^{m})y_{jk}z_{ij},\quad\forall j,m.$$

Treatment targets: The constraints

$$\varepsilon_{out,j}^{m}\leq\tau_{j}^{m},\quad\forall j,m$$

assure that each pollutant concentration in the effluent $\varepsilon_{out,j}^{m}$ achieves the corresponding treatment target τ_{j}^{m}.

Budget restriction: We only consider the capital cost at the design phase and restrict it within a budget, denoted by B:

$$\sum_{i=1}^{I}\sum_{j=1}^{J}c_{s}d_{ij}x_{ij}+\sum_{j=1}^{J}\sum_{k=0}^{K}c_{cw,jk}y_{jk}\leq B,$$

where c_s is the unit construction cost of sewer lines, $c_{cw,jk}$ is the construction cost of CW in design option k for site j, and d_{ij} is the distance between source i and site j. Hence, the left hand side accounts for the total capital cost of the treatment network, which is restricted within budget B.

Consolidating the above, the resulting feasibility problem (Model F) under deterministic environment can be formulated as follows:

$$\begin{aligned}\text{Model F}:\quad\text{s.t.}\quad&\sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^{m}\varepsilon_{i}^{m}+b_{jk}^{m})y_{jk}z_{ij}\leq\tau_{j}^{m}\sum_{i=1}^{I}z_{ij},\quad\forall j,m\\&\sum_{i=1}^{I}\sum_{j=1}^{J}c_{s}d_{ij}x_{ij}+\sum_{j=1}^{J}\sum_{k=0}^{K}c_{cw,jk}y_{jk}\leq B\\&\sum_{i=1}^{I}z_{ij}\leq\sum_{k=0}^{K}Q_{jk}y_{jk},\quad\forall j\\&z_{ij}\leq F_{i}x_{ij},\quad\forall i,j\\&\sum_{j=1}^{J}z_{ij}=F_{i},\quad\forall i\\&\sum_{k=0}^{K}y_{jk}=1,\quad\forall j\\&x_{ij},y_{jk}\in\{0,1\},z_{ij}\geq 0,\quad\forall i,j,k.\end{aligned}\tag{3}$$

As can be seen, Model F in (3) is bilinear which involves the product of binary variable y_{jk} and continuous variable z_{ij}. We can also observe that z_{ij} is bounded in $[0,F_i]$. Hence, the bilinear constraints could be explicitly linearized with *McCormick's inequalities* (**?**). Replace the bilinear term $y_{jk}z_{ij}$ with an added variable θ_{ijk}, whilst adjoining a sequence of linear inequalities, we can obtain

$$\sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m\varepsilon_i^m + b_{jk}^m)\theta_{ijk} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j, m$$

$$\begin{pmatrix} \theta_{ijk} \geq 0 \\ \theta_{ijk} \geq z_{ij} + z_{ij}^U y_{jk} - z_{ij}^U \\ \theta_{ijk} \leq z_{ij}^U y_{jk} \\ \theta_{ijk} \leq z_{ij} \end{pmatrix}, \quad \forall i, j, k,$$

where z_{ij}^U denotes the upper bound of z_{ij}. Integrating the above inequalities, the feasibility problem in (3) could then be reformulated as the following *Mixed-Integer Linear Programming* (MILP) problem (Model F-1):

$$\begin{aligned}
\text{Model F-1 :} \quad \text{s.t.} \quad & \sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m\varepsilon_i^m + b_{jk}^m)\theta_{ijk} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j, m \\
& \sum_{i=1}^{I}\sum_{j=1}^{J} c_s d_{ij} x_{ij} + \sum_{j=1}^{J}\sum_{k=0}^{K} c_{cw,jk} y_{jk} \leq B \\
& \sum_{i=1}^{I} z_{ij} \leq \sum_{k=0}^{K} Q_{jk} y_{jk}, \quad \forall j \\
& z_{ij} \leq F_i x_{ij}, \qquad \forall i, j \\
& \theta_{ijk} \geq z_{ij} + z_{ij}^U y_{jk} - z_{ij}^U, \quad \forall i, j, k \\
& \theta_{ijk} \leq z_{ij}^U y_{jk}, \qquad \forall i, j, k \\
& \theta_{ijk} \leq z_{ij}, \qquad \forall i, j, k \\
& \sum_{j=1}^{J} z_{ij} = F_i, \quad \forall i \\
& \sum_{k=0}^{K} y_{jk} = 1, \quad \forall j \\
& x_{ij}, y_{jk} \in \{0, 1\},\ z_{ij}, \theta_{ijk} \geq 0, \quad \forall i, j, k.
\end{aligned} \tag{4}$$

In Model F-1, all the parameters are assumed to be deterministic and we aim to find a feasible planning solution $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ that satisfies all our requirements. However, one real-world concern in the environmental decision making is the ubiquitous uncertainty. On the one hand, the solution obtained under deterministic environment may become infeasible when uncertain parameters are revealed. On the other hand, it may not be realistic for decision makers to implement a scheme that remains feasible under all realizations of uncertainties due to the economic concern. For the CWs design, studies show that pollutant concentrations in the wastewater sources (i.e., ε_i^m) present a high level of variety in practice such that the treatment targets would probably be violated. Hence, one interesting question is how can we develop a planning strategy to ensure that the quality of treated wastewater could best fulfill the treatment targets under different realizations.

In view of this, our work proposes a novel feasibility criterion to measure the level of solutions satisfying treatment targets. Optimization models for CW system planning problems embedded with the feasibility criterion are developed and tractable models are also provided.

3. Feasibility criteria and optimization models

Let pollutant concentrations in the wastewater sources be random, denoted by $\tilde{\boldsymbol{\varepsilon}} := [\tilde{\boldsymbol{\varepsilon}}_1, ..., \tilde{\boldsymbol{\varepsilon}}_I] \in \Re^{M\times I}$, where $\tilde{\boldsymbol{\varepsilon}}_i = [\tilde{\varepsilon}_i^1, ..., \tilde{\varepsilon}_i^M]'$. We introduce $\mathcal{U}(\boldsymbol{\alpha})$ to denote the *feasible uncertainty set*, from which any realization of $\tilde{\boldsymbol{\varepsilon}}$ would not violate the feasibility of a given solution $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$. In other words, $\mathcal{U}(\boldsymbol{\alpha})$ is the set of all possible realizations that retain a solution to be feasible. In particular, $\boldsymbol{\alpha} \in \Re^{M\times I}$, and $\mathcal{U}(\cdot)$ is designed such that $\mathcal{U}(\boldsymbol{\alpha}_1) \subseteq \mathcal{U}(\boldsymbol{\alpha}_2) \subseteq \mathcal{W}$ for all $\mathbf{0} \leq \boldsymbol{\alpha}_1 \leq \boldsymbol{\alpha}_2$, where $\mathcal{W}$ denotes the support of random variable $\tilde{\boldsymbol{\varepsilon}}$. Obviously, if a feasible solution $(\boldsymbol{x}_2, \boldsymbol{y}_2, \boldsymbol{z}_2)$ results in a strictly larger $\mathcal{U}(\cdot)$ compared with $(\boldsymbol{x}_1, \boldsymbol{y}_1, \boldsymbol{z}_1)$, $(\boldsymbol{x}_2, \boldsymbol{y}_2, \boldsymbol{z}_2)$ is then definitely preferable since it could remain feasible under an exactly broader range of uncertainty realizations. However, in many cases, we cannot straightforwardly obtain the comparison result $\mathcal{U}(\boldsymbol{\alpha}_1) \subseteq \mathcal{U}(\boldsymbol{\alpha}_2)$. Therefore, we propose a feasibility criterion to measure the feasibility level of solutions.

3.1. Case of discrete probability distribution

Let us first consider the scenario where $\tilde{\boldsymbol{\varepsilon}}$ follows discrete probbaility distributions. Particularly, we model $\tilde{\varepsilon}_i^m$ on the discrete support $\mathcal{W}_i^m = \{\varepsilon_i^{m1}, \varepsilon_i^{m2}, ..., \varepsilon_i^{mL(i,m)}\}$ and define $\mathcal{W} := \mathcal{W}_1 \times \mathcal{W}_2 ... \times \mathcal{W}_I$, $\mathcal{W}_i := \mathcal{W}_i^1 \times \mathcal{W}_i^2 \times ... \times W_i^M$. Denote the underlying distribution of $\tilde{\boldsymbol{\varepsilon}}$ by $\mathbb{P}$ and the realization of $\tilde{\boldsymbol{\varepsilon}}$ by $\boldsymbol{\varepsilon}$, we assume that $\tilde{\boldsymbol{\varepsilon}}$ has strictly positive mass functions $\Pr(\tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}^n) = p^n$, $\forall \boldsymbol{\varepsilon}^n \in \mathcal{W}$. We model the feasible uncertainty set as a "box": $\mathcal{U}(\boldsymbol{\alpha}) = \{\boldsymbol{\varepsilon} \in \mathcal{W} : \boldsymbol{\varepsilon} \in [\mathbf{0}, \boldsymbol{\alpha}]\}$ and propose the following *probabilistic feasibility criterion* $\rho(\boldsymbol{\alpha})$:

$$\rho(\boldsymbol{\alpha}) = \Pr\left(\tilde{\boldsymbol{\varepsilon}} \in \mathcal{U}(\boldsymbol{\alpha})\right) = \Pr(\mathbf{0} \leq \tilde{\boldsymbol{\varepsilon}} \leq \boldsymbol{\alpha}),$$

which represents the joint probability of $\tilde{\boldsymbol{\varepsilon}}$ being in the set $\mathcal{U}(\boldsymbol{\alpha})$ evaluated on the distribution $\mathbb{P}$. In other words, $\rho(\boldsymbol{\alpha})$ measures how the random outcomes of uncertainties are collectively met by the feasible uncertainty set. Intuitively, a solution resulting in a larger value of $\rho(\boldsymbol{\alpha})$ should be preferable since the probability of the solution being feasible under any random realization is higher. Hence, by allowing $\boldsymbol{\alpha}$ to be adjustable, we formulate the optimization model for the decentralized CW system planning problem based on $\rho(\boldsymbol{\alpha})$ as follows:

$$\begin{aligned} \text{Model P}: \quad & \max \Pr(\mathbf{0} \leq \tilde{\boldsymbol{\varepsilon}} \leq \boldsymbol{\alpha}) \\ \text{s.t.} \quad & \sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m \varepsilon_i^m + b_{jk}^m) y_{jk} z_{ij}(\boldsymbol{\varepsilon}) \leq \tau_j^m \sum_{i=1}^{I} z_{ij}(\boldsymbol{\varepsilon}), \quad \forall \boldsymbol{\varepsilon} \in \mathcal{U}(\boldsymbol{\alpha}),\ j, m \\ & \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y},\ \boldsymbol{z}(\boldsymbol{\varepsilon}) \in \mathcal{Z}, \quad \forall \boldsymbol{\varepsilon} \in \mathcal{W} \\ & x_{ij}, y_{jk} \in \{0,1\},\ z_{ij}(\boldsymbol{\varepsilon}) \geq 0, \quad \forall \boldsymbol{\varepsilon} \in \mathcal{W}, i, j, k, m. \end{aligned} \tag{5}$$

Note that Model P in (5) is a two-stage problem where the allocation amount $z_{ij}(\boldsymbol{\varepsilon})$ is a second-stage decision variable. The value of $z_{ij}(\boldsymbol{\varepsilon})$ depends on the specific realization $\boldsymbol{\varepsilon}$, that is, allocation decisions could be made when pollutant concentrations in each source are observed. The other two variables x_{ij} and y_{jk} regarding the construction of sewer lines and CWs are long-term decisions since they involve a huge one-off investment. For the conciseness of model formulation and transformation, the other linear constraints in (4) have been modeled into decision spaces $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$. We can observe that among the other difficulties, one main challenge in solving Model P is the evaluation of joint probability. Hence, in Section 4, we propose different ways to make Model P computationally tractable and could be solved efficiently by any commercial MILP solvers.

4. Tractable models for feasibility criteria

4.1. Tractable models for probabilistic feasibility criterion

4.1.1. A concave function based approximation and tractable models

In this section, we propose an efficient approximation on the joint probability $\Pr(0 \leq \tilde{\boldsymbol{\varepsilon}} \leq \boldsymbol{\alpha})$ and simplify the two-stage problem Model P to be a one-stage tractable MILP problem. The drawbacks of probabilistic feasibility criterion $\rho(\boldsymbol{\alpha})$ are twofold. First, joint probability is nonconcave and the consequent optimization model is typically very difficult to solve. Secondly, probability measures are insensitive to the degree of shortfalls (i.e., $\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\varepsilon}}$). It does not distinguish the level of treated wastewater that falls short of treatment targets. On the other hand, these issues could be addressed by taking expectations over nondecreasing concave functions. Here, we propose an approximation of $\rho(\boldsymbol{\alpha})$ based on a nondecreasing concave function as below:

$$\rho'(\boldsymbol{\alpha}) = \max_{\boldsymbol{v} \geq \mathbf{0}} \mathbb{E}_{\mathbb{P}} \left[\min_{i,m} \left\{ v_i^m(\alpha_i^m - \tilde{\varepsilon}_i^m), 1 \right\} \right],$$

where $\min_{i,m} \left\{ v_i^m(\alpha_i^m - \tilde{\varepsilon}_i^m), 1 \right\}$ is a nondecreasing concave function of $\tilde{\boldsymbol{\varepsilon}}$: $\Re^{M \times I} \mapsto \Re$. We can observe that $\rho'(\boldsymbol{\alpha})$ is sensitive to the magnitude of targets' shortfalls because the multipliers v_i^m are nonnegative whenever the random variable is not in the feasible set, i.e., $\tilde{\varepsilon}_i^m > \alpha_i^m$, reflecting the penalization against the shortfall of treatment targets.

The following result shows that $\rho'(\boldsymbol{\alpha})$ is a lower bound on the probabilistic feasibility criterion and reflects the uncertainty aversion.

Proposition 1.

$$\Pr(\mathbf{0} \leq \tilde{\boldsymbol{\varepsilon}} \leq \boldsymbol{\alpha}) \geq \max_{\boldsymbol{v} \geq \mathbf{0}} \mathbb{E}_{\mathbb{P}} \left[\min_{i,m} \left\{ v_i^m(\alpha_i^m - \tilde{\varepsilon}_i^m), 1 \right\} \right].$$

Proof.

Observe that the joint probability can be expressed as an expectation over a step function:

$$\Pr(\mathbf{0} \leq \tilde{\boldsymbol{\varepsilon}} \leq \boldsymbol{\alpha}) = \mathbb{E}_{\mathbb{P}}(h(\tilde{\boldsymbol{\varepsilon}})),$$

where the step function $h(\tilde{\boldsymbol{\varepsilon}})$ is given by

$$h(\tilde{\boldsymbol{\varepsilon}}) = \begin{cases} 1 & \text{if } \tilde{\boldsymbol{\varepsilon}} \leq \boldsymbol{\alpha} \\ 0 & \text{otherwise.} \end{cases}$$

The inequality then follows from the observation that the step function dominates the concave function, i.e., $h(\tilde{\boldsymbol{\varepsilon}}) \geq \min_{i,m} \{v_i^m(\alpha_i^m - \tilde{\varepsilon}_i^m), 1\}$, $\forall \boldsymbol{v} \geq \mathbf{0}$. Hence, we can have for any $\boldsymbol{v} \geq \mathbf{0}$,

$$\mathbb{E}_{\mathbb{P}}(h(\tilde{\boldsymbol{\varepsilon}})) \geq \mathbb{E}_{\mathbb{P}}\left[\min_{i,m} \{v_i^m(\alpha_i^m - \tilde{\varepsilon}_i^m), 1\}\right].$$

Taking the maximum of the right hand side, the desired result follows. $\square$

Approximating the probabilistic feasibility criterion via its lower bound $\rho'(\boldsymbol{\alpha})$, we can reformulate Model P as follows:

$$\begin{aligned}
&\max \mathbb{E}_{\mathbb{P}}\left[\min_{i,m} \{v_i^m(\alpha_i^m - \tilde{\varepsilon}_i^m), 1\}\right] \\
\text{s.t.}\quad &\sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m \varepsilon_i^m + b_{jk}^m) y_{jk} z_{ij}(\boldsymbol{\varepsilon}) \leq \tau_j^m \sum_{i=1}^{I} z_{ij}(\boldsymbol{\varepsilon}), \quad \forall \boldsymbol{\varepsilon} \in \mathcal{U}(\boldsymbol{\alpha}),\ j, m \\
&\boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y},\ \boldsymbol{z}(\boldsymbol{\varepsilon}) \in \mathcal{Z}, \quad \forall \boldsymbol{\varepsilon} \in \mathcal{W} \\
&x_{ij}, y_{jk} \in \{0,1\},\ z_{ij}(\boldsymbol{\varepsilon}), v_i^m, \alpha_i^m \geq 0, \quad \forall \boldsymbol{\varepsilon} \in \mathcal{W}, i, j, k, m.
\end{aligned} \tag{6}$$

However, solving (6) remains challenging because we are optimizing over arbitrary recourse functions $z_{ij}(\boldsymbol{\varepsilon})$. In each pass of searching $\boldsymbol{\alpha}$, all realizations $\boldsymbol{\varepsilon}$ from $\mathcal{U}(\boldsymbol{\alpha})$ are modeled to solve the optimal solution. Consequently, (6) turns out to be a large-scale problem. Our next result shows that instead of optimizing over a sequence of recourse functions, solving (6) could be simplified by regarding all decision variables as first-stage decision variables and solving its robust counterpart.

Proposition 2. *Solving* (6) *is equivalent to solving the following model:*

$$\begin{aligned}
&\max \mathbb{E}_{\mathbb{P}}\left[\min_{i,m} \{v_i^m\left(\alpha_i^m - \tilde{\varepsilon}_i^m\right), 1\}\right] \\
s.t.\quad &\sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m \alpha_i^m + b_{jk}^m) y_{jk} z_{ij} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j, m \\
&\boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y},\ \boldsymbol{z} \in \mathcal{Z} \\
&x_{ij}, y_{jk} \in \{0,1\},\ z_{ij}, v_i^m, \alpha_i^m \geq 0, \quad \forall i, j, k, m.
\end{aligned} \tag{7}$$

Proof.

To demonstrate the optimality equivalence of (6) and (7), we prove from the following two aspects:

(i) Given an optimal solution $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ to (7), we show that it is feasible in (6).
(ii) Given an optimal solution $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ to (7), we show that it is also optimal in (6).

Since the objective functions in both models are the same, we focus on analyzing the different treatment target constraints. Rewrite the treatment target constraints compactly in both models as below:

$$\boldsymbol{y}_j A_j^m(\boldsymbol{\varepsilon})\boldsymbol{z}_j'(\boldsymbol{\varepsilon}) \leq \tau_j^m \boldsymbol{z}_j'(\boldsymbol{\varepsilon})\mathbf{1}, \quad \forall \boldsymbol{\varepsilon} \in \mathcal{U}(\boldsymbol{\alpha}), j, m$$
$$\boldsymbol{y}_j A_j^m(\boldsymbol{\alpha})\boldsymbol{z}_j' \leq \tau_j^m \boldsymbol{z}_j'\mathbf{1}, \quad \forall j, m,$$

where $\boldsymbol{y}_j = [y_{j1}, y_{j2}, ..., y_{jK}]$, $\boldsymbol{z}_j = [z_{1j}, z_{2j}, ..., z_{Ij}]$, and $A_j^m(\boldsymbol{\varepsilon})$ is a matrix in $\Re^{K\times I}$. The entries of $A_j^m(\boldsymbol{\varepsilon})$, denoted by $a_{ijk}^m(\varepsilon_i^m)$, are affine functions of the random variable realizations ε_i^m, i.e., $a_{ijk}^m(\varepsilon_i^m) = a_{jk}^m \varepsilon_i^m + b_{jk}^m$.

Result (i) could be easily shown as we can observe that

$$\boldsymbol{y}_j A_j^m(\boldsymbol{\alpha})\boldsymbol{z}_j' - \tau_j^m \boldsymbol{z}_j'\mathbf{1} \geq \boldsymbol{y}_j A_j^m(\boldsymbol{\varepsilon})\boldsymbol{z}_j' - \tau_j^m \boldsymbol{z}_j'\mathbf{1}, \quad \forall \boldsymbol{\varepsilon} \in \mathcal{U}(\boldsymbol{\alpha}), j, m$$

hold for all $\boldsymbol{y} \in \mathcal{Y}$, $\boldsymbol{z} \in \mathcal{Z}$ since $A_j^m(\boldsymbol{\varepsilon})$ is non-decreasing in $\boldsymbol{\varepsilon}$.

It then suffices to show (ii) that $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ is also optimal in (6). Consider a subset of $\mathcal{U}(\boldsymbol{\alpha})$, denoted by $\mathcal{U}'(\boldsymbol{\alpha})$, such that $\mathcal{U}'(\boldsymbol{\alpha}) \subset \mathcal{U}(\boldsymbol{\alpha})$. We relax the optimization model (6) by narrowing down $\mathcal{U}(\boldsymbol{\alpha})$ to $\mathcal{U}'(\boldsymbol{\alpha})$. Denote the optimal objective values of (6) subject to $\mathcal{U}'(\boldsymbol{\alpha})$ and $\mathcal{U}(\boldsymbol{\alpha})$ by $O(\mathcal{U}'(\boldsymbol{\alpha}))$ and $O(\mathcal{U}(\boldsymbol{\alpha}))$ respectively, we can obtain that $O(\mathcal{U}'(\boldsymbol{\alpha})) \geq O(\mathcal{U}(\boldsymbol{\alpha}))$ hold since the constraints are relaxed when we consider a smaller uncertainty set. In the extreme case when $\mathcal{U}'(\boldsymbol{\alpha}) = \{\boldsymbol{\alpha}\}$, the problem (6) subject to $\mathcal{U}'(\boldsymbol{\alpha})$ would exactly turn to be (7). Hence, we can easily conclude that the objective value calculated by $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ in (6) is at least as large as $O(\mathcal{U}(\boldsymbol{\alpha}))$. The optimality of $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ in (6) then follows as $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ is a feasible solution to (6).

□

Intuitively, Proposition 2 simplifies (6) by considering the worst-case scenario where maximum realizations $\boldsymbol{\alpha}$ are revealed. This is because, increasing pollutant concentrations ε_i^m will always have a negative impact on the CW's treatment ability to fulfill targets τ_j^m collectively. Hence, if we aim to find a solution which keeps feasible under all realizations from the uncertainty set, we must guarantee its feasibility in the worst-case scenario. Apart from its remarkable computational benefits, Proposition 2 provides important practical implications. The optimal one-stage solution obtained by (7) could result in a feasibility level at least as high as that of the two-stage problem in (6). That means we could avoid troubles in frequently adjusting allocation amounts during operation since a high feasibility level could always be guaranteed. Now the only difficulty in solving (7) is the multi-linear terms. However, under the assumption of discrete probability distributions, this issue could be easily addressed.

First, we let $\alpha_i^m \in \mathcal{W}_i^m$ and model them as $\alpha_i^m = \sum_{l=1}^{L(i,m)} \gamma_i^{ml}\varepsilon_i^{ml}$, $\sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1$. γ_i^{ml} is a binary variable indicating whether the support variable ε_i^{ml} is selected as the upper bound of $\mathcal{U}(\cdot)$. The feasible uncertainty set could then be written as a function of $\boldsymbol{\gamma}$:

$$\mathcal{U}(\boldsymbol{\gamma}) = \left\{ \varepsilon_i^m \in \mathcal{W}_i^m \middle| \, 0 \leq \varepsilon_i^m \leq \sum_{l=1}^{L(i,m)} \gamma_i^{ml}\varepsilon_i^{ml}, \quad \forall i, m \right\}.$$

Following which, problem (7) can be written as Model P-1:

$$\text{Model P-1}: \max \mathbb{E}_{\mathbb{P}} \left[\min_{i,m} \left\{ \sum_{l=1}^{L(i,m)} \left(v_i^m \gamma_i^{ml} \varepsilon_i^{ml} - v_i^m \tilde{\varepsilon}_i^m \right), 1 \right\} \right]$$
$$\begin{aligned} \text{s.t.} \quad & \sum_{i=1}^{I} \sum_{k=0}^{K} \sum_{l=1}^{L(i,m)} \left(a_{jk}^m \gamma_i^{ml} \varepsilon_i^{ml} + b_{jk}^m \right) y_{jk} z_{ij} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j, m \\ & \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \quad \forall i, m \\ & \boldsymbol{x} \in \mathcal{X}, \ \boldsymbol{y} \in \mathcal{Y}, \ \boldsymbol{z} \in \mathcal{Z} \\ & x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0,1\}, \ z_{ij}, v_i^m \geq 0, \quad \forall i, j, k, m, l. \end{aligned} \tag{8}$$

Next, we linearize Model P-1 with McCormick's inequalities. Specifically, we replace bilinear terms $v_i^m \gamma_i^{ml}$, $y_{jk} z_{ij}$, $\gamma_i^{ml} \theta_{ijk}$ with introduced variables u_i^{ml}, θ_{ijk}, λ_{ijk}^{ml} respectively, and add a set of linear constraints:

$$\max \mathbb{E}_{\mathbb{P}} \left[\min_{i,m} \left\{ \sum_{l=1}^{L(i,m)} \left(u_i^{ml} \varepsilon_i^{ml} - v_i^m \tilde{\varepsilon}_i^m \right), 1 \right\} \right]$$
$$\begin{aligned} \text{s.t.} \quad & \sum_{i=1}^{I} \sum_{k=0}^{K} \sum_{l=1}^{L(i,m)} a_{jk}^m \varepsilon_i^{ml} \lambda_{ijk}^{ml} + \sum_{i=1}^{I} \sum_{k=0}^{K} b_{jk}^m \theta_{ijk} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \ \forall \ j, m \\ & u_i^{ml} \geq v_i^m + v_i^{mU} \gamma_i^{ml} - v_i^{mU}, \quad \forall i, m, l \\ & u_i^{ml} \leq v_i^{mU} \gamma_i^{ml}, \quad \forall i, m, l \\ & u_i^{ml} \leq v_i^m, \quad \forall i, m, l \\ & \theta_{ijk} \geq z_{ij} + z_{ij}^U y_{jk} - z_{ij}^U, \quad \forall i, j, k \\ & \theta_{ijk} \leq z_{ij}^U y_{jk}, \quad \forall i, j, k \\ & \theta_{ijk} \leq z_{ij}, \quad \forall i, j, k \\ & \lambda_{ijk}^{ml} \geq \theta_{ijk} + \theta_{ijk}^U \gamma_i^{ml} - \theta_{ijk}^U, \quad \forall i, j, k, m, l \\ & \lambda_{ijk}^{ml} \leq \theta_{ijk}^U \gamma_i^{ml}, \quad \forall i, j, k, m, l \\ & \lambda_{ijk}^{ml} \leq \theta_{ijk}, \quad \forall i, j, k, m, l \\ & \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \quad \forall i, m \\ & \boldsymbol{x} \in \mathcal{X}, \ \boldsymbol{y} \in \mathcal{Y}, \ \boldsymbol{z} \in \mathcal{Z} \\ & x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0,1\}, \ z_{ij}, v_i^m, u_i^{ml}, \theta_{ijk}, \lambda_{ijk}^{ml} \geq 0, \quad \forall i, j, k, m, l \end{aligned} \tag{9}$$

Assume that N realizations of $\tilde{\boldsymbol{\varepsilon}}$ are observed, denoted by $\boldsymbol{\varepsilon}^n, n = 1, 2, ..., N$, we propose *sample average approximation method* (SAA) to solve Model P-1 and consequently obtain the following result:

$$
\begin{aligned}
\text{Model P-2}: \quad & \max \frac{1}{N}\sum_{n=1}^{N}\eta^n \\
\text{s.t.} \quad & \sum_{l=1}^{L(i,m)} u_i^{ml}\varepsilon_i^{ml} - v_i^m\varepsilon_i^{mn} \geq \eta^n, \quad \forall i,m,n \\
& 1 \geq \eta^n, \qquad \forall n \\
& \sum_{i=1}^{I}\sum_{k=0}^{K}\sum_{l=1}^{L(i,m)} a_{jk}^m\varepsilon_i^{ml}\lambda_{ijk}^{ml} + \sum_{i=1}^{I}\sum_{k=0}^{K} b_{jk}^m\theta_{ijk} \leq \tau_j^m\sum_{i=1}^{I} z_{ij}, \ \forall\, j,m \\
& u_i^{ml} \geq v_i^m + v_i^{mU}\gamma_i^{ml} - v_i^{mU}, \quad \forall i,m,l \\
& u_i^{ml} \leq v_i^{mU}\gamma_i^{ml}, \qquad \forall i,m,l \\
& u_i^{ml} \leq v_i^m, \qquad \forall i,m,l \\
& \theta_{ijk} \geq z_{ij} + z_{ij}^U y_{jk} - z_{ij}^U, \qquad \forall i,j,k \\
& \theta_{ijk} \leq z_{ij}^U y_{jk}, \qquad \forall i,j,k \\
& \theta_{ijk} \leq z_{ij}, \qquad \forall i,j,k \\
& \lambda_{ijk}^{ml} \geq \theta_{ijk} + \theta_{ijk}^U\gamma_i^{ml} - \theta_{ijk}^U, \quad \forall i,j,k,m,l \\
& \lambda_{ijk}^{ml} \leq \theta_{ijk}^U\gamma_i^{ml}, \qquad \forall i,j,k,m,l \\
& \lambda_{ijk}^{ml} \leq \theta_{ijk}, \qquad \forall i,j,k,m,l \\
& \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \quad \forall i,m \\
& \boldsymbol{x}\in\mathcal{X},\ \boldsymbol{y}\in\mathcal{Y},\ \boldsymbol{z}\in\mathcal{Z} \\
& x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0,1\},\ z_{ij}, v_i^m, u_i^{ml}, \theta_{ijk}, \lambda_{ijk}^{ml} \geq 0, \quad \forall i,j,k,m,l
\end{aligned}
\tag{10}
$$

We can observe that Model P-2 has $IML+IJ+JK$ binary decision variables. Compared with the benchmark method in Section 4.2, (10) could be less computationally demanding since it has significantly reduced the number of binary variables. However, one possible limitation of (10) is the large number of constraints, which may impose additional computational efforts. Hence, we are now trying to reformulate the problem by applying *Big-M* modeling techniques.

First, we reformulate the treatment target constraints in the initial feasibility model (3):

$$a_j^m + b_j^m \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j,m \tag{11}$$

$$a_j^m \geq a_{jk}^m \sum_{i=1}^{I} \varepsilon_i^m z_{ij} - \Theta_{jk}(1-y_{jk}), \quad \forall j,m,k \tag{12}$$

$$b_j^m \geq b_{jk}^m \sum_{i=1}^{I} z_{ij} - \Theta_{jk}(1-y_{jk}), \quad \forall j,m,k, \tag{13}$$

where Θ_{jk} is assumed to be a sufficiently large real number. In particular, when design option k' is selected for CW site j, constraints in (12) and (13) enforce lower bounds on a_j^m and b_j^m with design parameters in k' choice, i.e., $a_j^m \geq a_{jk'}^m \sum_{i=1}^{I} \varepsilon_i^m z_{ij}$, $b_j^m \geq b_{jk'}^m \sum_{i=1}^{I} z_{ij}, \forall j, m$. The other options are not effective in that case. Hence, the above reformulations are equivalent to the original treatment target constraints in (3). In comparison, the advantage of (11)-(13) is that a smaller number of constraints are involved as bilinear terms $z_{ij}y_{jk}$ are not included in the reformulations.

Based on the reformulated constraints (11)-(13), we can rewrite Model P-1 as follows:

$$\text{Model P}'\text{-1: } \max \mathbb{E}_{\mathbb{P}} \left[\min_{i,m} \left\{ \sum_{l=1}^{L(i,m)} \left(u_i^{ml} \varepsilon_i^{ml} - v_i^m \tilde{\varepsilon}_i^m \right), 1 \right\} \right]$$

$$\begin{aligned}
\text{s.t. } & a_j^m + b_j^m \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j, m \\
& a_j^m \geq a_{jk}^m \sum_{i=1}^{I} \sum_{l=1}^{L(i,m)} \lambda_{ij}^{ml} \varepsilon_i^{ml} - \Theta_{jk}(1 - y_{jk}), \quad \forall, j, k, m \\
& b_j^m \geq b_{jk}^m \sum_{i=1}^{I} z_{ij} - \Theta_{jk}(1 - y_{jk}), \quad \forall j, k, m \\
& u_i^{ml} \geq v_i^m + v_i^{mU} \gamma_i^{ml} - v_i^{mU}, \qquad \forall i, m, l \\
& u_i^{ml} \leq v_i^{mU} \gamma_i^{ml}, \qquad \forall i, m, l \\
& u_i^{ml} \leq v_i^m, \qquad \forall i, m, l \\
& \lambda_{ij}^{ml} \geq z_{ij} + z_{ij}^U \gamma_i^{ml} - z_{ij}^U, \qquad \forall i, j, m, l \\
& \lambda_{ij}^{ml} \leq z_{ij}^U \gamma_i^{ml}, \qquad \forall i, j, m, l \\
& \lambda_{ij}^{ml} \leq z_{ij}, \qquad \forall i, j, m, l \\
& \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \quad \forall i, m \\
& \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y},\ \boldsymbol{z} \in \mathcal{Z} \\
& x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0, 1\},\ z_{ij}, v_i^m \geq 0, \quad \forall i, j, k, m, l.
\end{aligned}$$

Similarly, we can solve Model P′-1 by SAA method and obtain the following model:

$$
\begin{aligned}
\text{Model P}'\text{-2: } & \max \frac{1}{N}\sum_{n=1}^{N}\eta^n \\
\text{s.t.} \quad & \sum_{l=1}^{L(i,m)} u_i^{ml}\varepsilon_i^{ml} - v_i^m\varepsilon_i^{mn} \ge \eta^n, \quad \forall i,m,n \\
& 1 \ge \eta^n, \quad \forall n \\
& a_j^m + b_j^m \le \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j,m \\
& a_j^m \ge a_{jk}^m \sum_{i=1}^{I}\sum_{l=1}^{L(i,m)} \lambda_{ij}^{ml}\varepsilon_i^{ml} - \Theta_{jk}(1-y_{jk}), \quad \forall j,m,k \\
& b_j^m \ge b_{jk}^m \sum_{i=1}^{I} z_{ij} - \Theta_{jk}(1-y_{jk}), \quad \forall j,m,k \\
& u_i^{ml} \ge v_i^m + v_i^{mU}\gamma_i^{ml} - v_i^{mU}, \qquad \forall i,m,l \\
& u_i^{ml} \le v_i^{mU}\gamma_i^{ml}, \qquad \forall i,m,l \\
& u_i^{ml} \le v_i^m, \qquad \forall i,m,l \\
& \lambda_{ij}^{ml} \ge z_{ij} + z_{ij}^U\gamma_i^{ml} - z_{ij}^U, \qquad \forall i,j,m,l \\
& \lambda_{ij}^{ml} \le z_{ij}^U\gamma_i^{ml}, \qquad \forall i,j,m,l \\
& \lambda_{ij}^{ml} \le z_{ij}, \qquad \forall i,j,m,l \\
& \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \quad \forall i,m \\
& \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y},\ \boldsymbol{z} \in \mathcal{Z} \\
& x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0,1\},\ z_{ij}, v_i^m, u_i^{ml}, \lambda_{ij}^{ml} \ge 0,\ \forall i,j,k,m,l.
\end{aligned} \tag{14}
$$

Compared with Model P-2, we can observe that although the binary decision variables in both models are the same, the number of constraints and continuous decision variables in Model P′-2 has been successfully cut down.

4.1.2. Sample based probabilistic feasibility criterion and tractable models

In the above section 4.1.1 , we have mainly developed two tractable MILP models based on the nondecreasing concave approximation function $\rho'(\boldsymbol{\alpha})$. In this section, we provide another way to model probabilistic feasibility criterion $\rho(\boldsymbol{\alpha}) = \Pr(0 \le \tilde{\boldsymbol{\varepsilon}} \le \boldsymbol{\alpha})$. Following the assumption of discrete probability distribution, we can express $\rho(\boldsymbol{\alpha})$ explicitly over a set of concave functions:

$$
\Pr(\mathbf{0} \le \tilde{\boldsymbol{\varepsilon}} \le \boldsymbol{\alpha}) = \lim_{\delta \to 0^+} \max_{\boldsymbol{v} \ge \mathbf{0}} \mathbb{E}_{\mathbb{P}}\left[\min_{i,m}\left\{v^n(\alpha_i^m + \delta - \varepsilon_i^{mn}), 1\right\}\right] \tag{15}
$$

where $\boldsymbol{v}$ is an N dimensional vector: $\boldsymbol{v} = \{v^1, v^2, ..., v^N\}$, α_i^m takes value from the support $\mathcal{W}_i^m$, and

δ is assumed to be a sufficiently small positive number. We can observe that for any realization n, if all random outcomes $\varepsilon_i^{mn}, \forall i, m$ are strictly smaller than $\alpha_i^m + \delta$, the right hand side of (15) takes the value of 1. In this case, we allow parameters v^n to be dependent on the specific realization ε_i^{mn}. Hence, no punishment would be imposed on the target shortfalls.

With the expression of $\rho(\boldsymbol{\alpha})$ in (15), we can write the CW system planning problem as follows:

$$
\begin{aligned}
& \max \lim_{\delta \to 0^+} \mathbb{E}_{\mathbb{P}} \left[\min_{i,m} \left\{ v^n \left(\sum_{l=1}^{L(i,m)} \gamma_i^{ml} (\varepsilon_i^{ml} + \delta) - \varepsilon_i^{mn} \right), 1 \right\} \right] \\
\text{s.t.} \quad & \lim_{\delta \to 0^+} \sum_{i=1}^{I} \sum_{k=0}^{K} \left[a_{jk}^m \sum_{l=1}^{L(i,m)} \gamma_i^{ml} \left(\varepsilon_i^{ml} + \delta \right) + b_{jk}^m \right] y_{jk} z_{ij} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \quad \forall j, m \\
& \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \quad \forall i, m \\
& \boldsymbol{x} \in \mathcal{X}, \ \boldsymbol{y} \in \mathcal{Y}, \ \boldsymbol{z} \in \mathcal{Z} \\
& x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0, 1\}, \ v^n, z_{ij} \geq 0, \quad \forall i, j, k, m, l, n.
\end{aligned}
\tag{16}
$$

Since the constraints in (16) are almost the same with Model P-1, we still follow the previous proposed methods to simplify (16) including with and without using Big-M techniques.

One simplification way is to directly apply McCormick's inequalities on those bilinear terms in (16). We can obtain the following model:

$$
\begin{aligned}
& \max \frac{1}{N} \sum_{n=1}^{N} \eta^n \\
\text{s.t.} \quad & \lim_{\delta \to 0^+} \sum_{l=1}^{L(i,m)} u_i^{mln}(\varepsilon_i^{ml} + \delta) - v^n \varepsilon_i^{mn} \geq \eta^n, \quad \forall i, m, n \\
& 1 \geq \eta^n, \qquad \forall n \\
& \lim_{\delta \to 0^+} \sum_{i=1}^{I} \sum_{k=0}^{K} \sum_{l=1}^{L(i,m)} a_{jk}^m (\varepsilon_i^{ml} + \delta) \lambda_{ijk}^{ml} + \sum_{i=1}^{I} \sum_{k=0}^{K} b_{jk}^m \theta_{ijk} \leq \tau_j^m \sum_{i=1}^{I} z_{ij}, \ \forall\, j, m \\
& u_i^{mln} \geq v^n + v^{nU} \gamma_i^{ml} - v^{nU}, \qquad \forall i, m, l, n \\
& u_i^{mln} \leq v^{nU} \gamma_i^{ml}, \qquad \forall i, m, l, n \\
& u_i^{mln} \leq v^n, \qquad \forall i, m, l, n \\
& \theta_{ijk} \geq z_{ij} + z_{ij}^U y_{jk} - z_{ij}^U, \qquad \forall i, j, k \\
& \theta_{ijk} \leq z_{ij}^U y_{jk}, \qquad \forall i, j, k \\
& \theta_{ijk} \leq z_{ij}, \qquad \forall i, j, k \\
& \lambda_{ijk}^{ml} \geq \theta_{ijk} + \theta_{ijk}^U \gamma_i^{ml} - \theta_{ijk}^U, \qquad \forall i, j, k, m, l \\
& \lambda_{ijk}^{ml} \leq \theta_{ijk}^U \gamma_i^{ml}, \qquad \forall i, j, k, m, l \\
& \lambda_{ijk}^{ml} \leq \theta_{ijk}, \qquad \forall i, j, k, m, l \\
& \sum_{l=1}^{L(i,m)} \gamma_i^{ml} = 1, \qquad \forall i, m \\
& \boldsymbol{x} \in \mathcal{X}, \ \boldsymbol{y} \in \mathcal{Y}, \ \boldsymbol{z} \in \mathcal{Z} \\
& x_{ij}, y_{jk}, \gamma_i^{ml} \in \{0,1\}, \ z_{ij}, v^n, u_i^{mln}, \lambda_{ijk}^{ml} \geq 0, \ \forall i, j, k, m, l, n.
\end{aligned}
\tag{17}
$$

Another way to simplify (16) is to first apply Big-M modeling techniques for the objective function as well as the treatment target constraint, just as the way we have done in Model P′-1, and then linearize the model using Mccormick's inequalities. The following model shows the result.

$$
\begin{aligned}
&\max \frac{1}{N}\sum_{n=1}^{N}\eta^n \\
\text{s.t.}\quad &\lim_{\delta\to 0^+} v^n(\varepsilon_i^{ml}+\delta-\varepsilon_i^{mn})+\Theta_i^{ml}(1-\gamma_i^{ml})\geq \eta^n, \quad \forall i,m,l,n \\
&1\geq \eta^n, \qquad \forall n \\
&a_j^m+b_j^m\leq \tau_j^m\sum_{i=1}^{I} z_{ij}, \quad \forall j,m \\
&a_j^m\geq \lim_{\delta\to 0^+} a_{jk}^m\sum_{i=1}^{I}\sum_{l=1}^{L(i,m)}\lambda_{ij}^{ml}(\varepsilon_i^{ml}+\delta)-\Theta_{jk}(1-y_{jk}), \quad \forall j,m,k \\
&b_j^m\geq b_{jk}^m\sum_{i=1}^{I} z_{ij}-\Theta_{jk}(1-y_{jk}), \quad \forall j,m,k \\
&\lambda_{ij}^{ml}\geq z_{ij}+z_{ij}^U\gamma_i^{ml}-z_{ij}^U, \qquad \forall i,j,m,l \\
&\lambda_{ij}^{ml}\leq z_{ij}^U\gamma_i^{ml}, \qquad \forall i,j,m,l \\
&\lambda_{ij}^{ml}\leq z_{ij}, \qquad \forall i,j,m,l \\
&\sum_{l=1}^{L(i,m)}\gamma_i^{ml}=1, \quad \forall i,m \\
&\boldsymbol{x}\in\mathcal{X},\ \boldsymbol{y}\in\mathcal{Y},\ \boldsymbol{z}\in\mathcal{Z} \\
&x_{ij},y_{jk},\gamma_i^{ml}\in\{0,1\},\ z_{ij},v^n,\lambda_{ij}^{ml}\geq 0,\ \forall i,j,k,m,l,n
\end{aligned}
\tag{18}
$$

where Θ_i^{ml} and Θ_{jk} are assumed to be sufficiently large real numbers. Similar with the constraints (12)-(13), the first set of constraints in (18) are effective only when the specific variable ε_i^{ml} is selected to be the upper bound. Compare (17) with (18), we can observe that the advantage of applying Big-M techniques is that the number of constraints and continuous variables could be significantly reduced. However, one possible drawback is that the decision space obtained from Big-M constraints might be large due to large numbers Θ_i^{ml}.

4.2. Benchmark method: joint success probability on treatment targets

In this section, we provide one benchmark method by directly maximizing the joint success probability of the treated wastewater collectively achieving multiple treatment targets, denoted as $\max \Pr(\tilde{\boldsymbol{\varepsilon}}_{out}\leq\boldsymbol{\tau})$, where $\tilde{\boldsymbol{\varepsilon}}_{out} =: [\tilde{\boldsymbol{\varepsilon}}_{out,1},..,\tilde{\boldsymbol{\varepsilon}}_{out,J}]\in\Re^{M\times J}$ and $\tilde{\boldsymbol{\varepsilon}}_{out,j}=[\tilde{\varepsilon}_j^1,...,\tilde{\varepsilon}_j^M]'$. The *joint success probability based model* (Model S) is then formulated as follows:

$$
\begin{aligned}
\text{Model S:}\quad &\max \Pr(\tilde{\boldsymbol{\varepsilon}}_{out}\leq\boldsymbol{\tau}) \\
\text{s.t.}\quad &\tilde{\varepsilon}_{out,j}^m\sum_{i=1}^{I} z_{ij}(\boldsymbol{\varepsilon})=\sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m\tilde{\varepsilon}_i^m+b_{jk}^m)y_{jk}z_{ij}(\boldsymbol{\varepsilon}), \quad \forall\boldsymbol{\varepsilon}\in\mathcal{W},\ j,m \\
&\boldsymbol{x}\in\mathcal{X},\ \boldsymbol{y}\in\mathcal{Y},\ \boldsymbol{z}(\boldsymbol{\varepsilon})\in\mathcal{Z},\ \ \forall\boldsymbol{\varepsilon}\in\mathcal{W}
\end{aligned}
\tag{19}
$$

Evaluating the joint success probability is typically intractable. Unfortunately, the approximation techniques with nondecreasing concave functions on joint probability cannot be efficiently applied here to make Model S tractable. It is because $\tilde{\boldsymbol{\varepsilon}}_{out}$ are not the primitive uncertainty parameters, but nonlinear functions of $\tilde{\boldsymbol{\varepsilon}}$. Consequently, the approximated objective function would remain intractable as a result of high-dimensional multi-linear terms and products of continuous decision variables. Here we propose to solve Model S using SAA method. Assume that N number of random uncertainty outcomes are observed, denoted by $\varepsilon_i^{mn}, n = 1, ..., N$. For a given set of realizations, we introduce binary variable γ^n to indicate whether all the pollutant concentrations in the treated wastewater reach the treatment targets. γ^n takes value of 1 if all the effluent pollutant concentrations reach the targets collectively, and 0 otherwise. The linearized sample average model (Model S-1) for (19) can then be stated as:

$$
\begin{aligned}
\text{Model S-1:} \max\ & \frac{1}{N}\sum_{n=1}^{N}\gamma^n \\
\text{s.t.}\ \ & \tau_j^m \sum_{i=1}^{I} z_{ij}^n - \sum_{i=1}^{I}\sum_{k=0}^{K}(a_{jk}^m \varepsilon_i^{mn} + b_{jk}^m)\theta_{ijk}^n \geq \Theta^n(\gamma^n - 1), \quad \forall j, m, n \\
& \theta_{ijk}^n \geq z_{ij}^n + z_{ij}^{nU} y_{jk} - z_{ij}^{nU}, \quad \forall i, j, k, n \\
& \theta_{ijk}^n \leq z_{ij}^{nU} y_{jk}, \qquad \forall i, j, k, n \\
& \theta_{ijk}^n \leq z_{ij}^n, \qquad \forall i, j, k, n \\
& \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y},\ \boldsymbol{z} \in \mathcal{Z} \\
& x_{ij}, y_{jk}, \gamma^n \in \{0, 1\},\ z_{ij}^n, \theta_{ijk}^n \geq 0 \quad \forall i, j, k, n.
\end{aligned}
\tag{20}
$$

where Θ^n is a real number assumed to be sufficiently large. In (20), there are totally $N + IJ + JK$ binary decision variables. Conceivably, due to the additional binary variables γ^n, (20) will be remarkably more difficult to solve than our proposed models in the previous section when the number of scenarios N is large.

5. Case study: a decentralized CWs system for municipal wastewater treatment

5.1. Introduction

5.1.1. Mobile, Alabama

With 412,992 people, Mobile County is the second most populated county within the state Alabama in the United States. The population of Mobile County has been steadily increasing (United States Census Bureau (2002)). As populations grow, several fringe communities are created. To manage the wastewater produced by such communities, the traditional approach is to link these fringe communities up with long length large diameter pipes to transport all the wastewater to a single municipal wastewater treatment plant to be processed and then released into nearby water sources. The Mobile Area Water & Sewer System (MAWSS) serves approximately 530 square kilometres in Mobile County (Mobile Area Water & Sewer System (????)), mostly through centralised wastewater treatment facilities. However, annual operations and maintenance costs for such centralised wastewater management systems have been shown to be costlier than decentralised ones, where the wastewater is processed and released within the community (White (????)). Thus, implementation of decentralised wastewater management systems such as constructed wetlands are being considered to manage costs in the long term.

5.1.2. Decentralised constructed wetlands wastewater management

Cost

As urbanised areas grow in size and population, the amount of wastewater produced per day in an urban area increases. A centralised wastewater management system becomes unwieldy and costly to maintain on a large scale. In order to tackle this, decentralised wastewater management systems have been proposed to reduce the distance from the wastewater source to the release point, cutting down on the cost of transporting wastewater to a dedicated facility. Various case studies have conducted a cost-benefit analysis on the implementation of a centralised wastewater management system against a decentralised one and have concluded that the decentralised wastewater management is generally cheaper to maintain (White (????); Prihandrijanti et al. (2008)).

Maintenance

One approach to decentralised wastewater management is the constructed wetlands concept. Essentially, constructed wetlands aim to simulate real wetlands where water flows through and has its nutrients removed via biological processes. In the constructed version, wastewater flows through the wetland to provide the pollutants within as a nutrient source for plant absorption. The end product of wastewater going through these processes is water that meets the standards for release. The use of plants to treat the water allows the procedure of treating wastewater to be more hands off, thus requiring less maintenance.

Factors for success

However, the cost effectiveness of constructed wetlands is based on the assumption that the decentralised sites are in optimal locations and are of suitable sizes to balance the cost of transporting wastewater and building the site. The size of the constructed wetland has to be big enough to serve the amount of wastewater produced, but not too big as the system may collapse due to a lack of nutrients for the plants towards the end of the wetlands. Additionally, the location of the constructed wetlands has to be taken into consideration to determine the shortest distance of pipes required to link the wastewater sources to the treatment area. In order to determine the ideal location and size of constructed wetlands in a defined area, a mathematical model has been developed to represent the problem. Solving this model will give the ideal location and size of constructed wetlands to be designed.

5.2. Objective

This paper aims to determine the optimal configuration (location and size) of constructed wetlands in a specified area in Mobile, Alabama as a case study area.

6. Conclusions

Appendix

References

Mobile Area Water & Sewer System (????), “What We Do Section,” .

Prihandrijanti, M., Malisie, A., and Otterpohl, R. (2008), *Cost–Benefit Analysis for Centralized and Decentralized Wastewater Treatment System (Case Study in Surabaya-Indonesia)*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 259–268.

United States Census Bureau (2002), “Measuring America: The Decennial Censuses From 1790 to 2000,” .

White, K. D. (????), “Decentralized Wastewater Management - An Overview of Experiences in Mobile, AL,” .