University of Tennessee, Knoxville
TRACE: Tennessee Research and Creative Exchange

Doctoral Dissertations | Graduate School

5-2019

Optimizing Urban Infrastructure Resilience Under Precipitation and Population Growth Uncertainties

Masoud Barah
University of Tennessee, mbarah@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

Recommended Citation

Barah, Masoud, "Optimizing Urban Infrastructure Resilience Under Precipitation and Population Growth Uncertainties. " PhD diss., University of Tennessee, 2019.
https://trace.tennessee.edu/utk_graddiss/5358

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a dissertation written by Masoud Barah entitled "Optimizing Urban Infrastructure Resilience Under Precipitation and Population Growth Uncertainties." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

Anahita Khojandi, Major Professor

We have read this dissertation and recommend its acceptance:

Xueping Li, Jon Mitchel Hathaway, Oleg Shylo

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Optimizing Urban Infrastructure Resilience Under Precipitation and Population Growth Uncertainties

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Masoud Barah

May 2019

© by Masoud Barah, 2019
All Rights Reserved.

To my wife, Narges

Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Anahita Khojandi for the continuous support of my Ph.D study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis.

This research is based upon work supported by the National Science Foundation under Grant No. CMMI-1634975.

Abstract

Increased urbanization, infrastructure degradation, and climate change threaten to overwhelm stromwater systems across the nation, rendering them ineffective. Green Infrastructure (GI) practices are low cost, low regret strategies that can contribute to urban runoff management. However, questions remain as to how to best distribute GI practices through urban watersheds given the precipitation uncertainty and the hydrological responses to them.

First, we develop a two-stage stochastic robust programming model to determine the optimal placement of GI practices across a set of candidate locations in a watershed to minimize the total expected runoff under medium-term precipitation uncertainties. We develop a systemic approach to downscale the existing daily precipitation projections into hourly units and efficiently estimate the corresponding hydrological responses. We conduct a case study for an urban watershed in a mid-sized city in the U.S., perform sensitivity analyses and provide insights.

Second, we develop a mathematical model to optimally place GI practices when (re-)designing an urban area, subject to uncertainties in population growth and future precipitation. Specifically, we develop a finite-horizon Markov decision process model to determine the extent to which GI practices need to be incorporated in different parts of a given urban area to maximize their benefits, considering the dynamic changes in population density and precipitation. We conduct a case study, perform sensitivity analyses and provide insights.

Finally, we consider a problem of scheduling maintenance crew following a storm event to efficiently maintain GI practices across a watershed to mitigate surface runoff due to future events. Specifically, we investigate a condition for which the polyhedron of the flow shop scheduling problem is integer-optimal. This condition is used to construct a column generation algorithm to solve the problem to optimality. The solution approach is boosted with a heuristic that sequentially solves a series of linear programming models to generate a quality initial solution. The solution approach is also integrated with a commercial solver, which results in significant computational savings. Computational experiments show that the developed algorithm can efficiently solve test problems to near-optimality.

Table of Contents

List of Tables

List of Figures

Chapter 1

Optimizing Green Infrastructure Placement Under Precipitation Uncertainty

One of the most important factors threatening infrastructure in the U.S. is climate change. Climate change affects the frequency, intensity, spatial extent, duration as well as timing of extreme events [48]. Over the past decade, we have observed more frequent, intense and untimely events damaging infrastructure and impacting people and businesses (e.g., Hurricane Katrina, Superstorm Sandy). Thus, there are major concerns as to whether cities are protected against these projected increasing number of extreme weather events. To mitigate these effects, municipalities are beginning to seek opportunities to improve the resiliency of infrastructure through better urban planning and taking advantage of innovative solutions. This is extremely timely, as by the end of next decade, 60% of the world population will live in cities [103].

Provision of scientifically-based methodologies for understanding and evaluating climate impacts will be critical to the development of adaptation strategies designed to avoid the increasing socioeconomic costs of severe weather-related damages to urban landscapes [82]. Despite this understanding, city managers are forced to make infrastructure decisions complicated by massive amounts of data and uncertainty. In a time when multiple, sometimes conflicting, climate projections exist, tools to distill these data into a usable format for such individuals are critical. Hence, city managers need a tool which addresses the complexity and uncertainty of climate projections to allow optimized choices for building resiliency into urban systems.

In the 2013 "Report Card" for American infrastructure, the nation's stormwater systems (in combination with wastewater) were awarded a D^+, indicating the poor state of these critical components of the urban landscape. Exacerbating this need is the specter of climate change, leading us to the age of non-stationarity, where past trends of precipitation may

no longer be relevant as a basis of design for civil infrastructure. Stormwater systems are particularly susceptible, as the size of pipes is selected based on how much stormwater needs to be conveyed for a given storm of interest or design storm applied to the watershed. As design storms are determined based on historical rainfall data, climate change threatens to overwhelm pipes that are in poor condition and undersized relative to changing weather patterns. Thus climate change and the associated overwhelming of stormwater pipe systems is likely to cause increased flooding in urban watersheds, escalating the already present trend of flooding and flash flooding as (on average) the leading cause of weather-related fatalities in the U.S., beyond even hurricanes and tornadoes [63]. For instance, in April 2016, a 17-inch rainfall resulted in significant flooding in Houston Texas, with the estimated total cost of 1.2 billion dollars.

Replacing existing stormwater sewers with pipes of larger capacity would be prohibitively expensive and time consuming in many urban environments due to surrounding infrastructure and social conflicts. However, building resiliency into urban stormwater systems through the use of green infrastructure (GI) is an increasing trend nationwide. The 2014 Intergovernmental Panel on Climate Change (IPCC) has identified changes to urban drainage systems as a key adaption issue for North America and recommends consideration of low–regret strategies such as GI to reduce flooding while also providing co-benefits to freshwater provision, ecological processes, and freshwater fish populations [5, 73]. The U.S. Environmental Protection Agency (EPA) is promoting GI as a means to enable communities to avoid costly water infrastructure replacement and repair by using vegetation and soil to manage rainwater where it falls, thereby reducing the burden on aging sewer pipes [20]. These systems act as localized storage centers, where stormwater can enter, be detained, then leave the system as evaporation, infiltration, or as runoff with diminished energy and volume. As such, GI has been deemed as a way to build better infrastructure as part of the National Academy of Engineering's Grand Challenge to restore and improve urban infrastructure.

In recent years, researchers have considered the impact of GI on urban flooding at the watershed scale [34, 67]. Kim et al. [65] studied the impact of urban green spaces on reducing urban flood risk. As their case study, they considered a flooded area in Seoul, South Korea. They divided the case study area into four regions based on topographic and physical characteristics, and used logistic regression to determine how flooding probabilities change with respect to green space area. Based on their results, the probability of flooding could be reduced by over 50% depending on the location of green spaces and their types. In a related study, Liu et al. [68] developed a simulation model to determine the reduction of peak flow rate in flooding for an urban community in Beijing, China. They reported that an integrated GI configuration can reduce peak flow by 92.8-100%. Liu et al. [69] also investigated the impact of GI practice types and sizes on reducing urban flooding. They reported that expanding green spaces, concave green space, storage pond, and porous brick pavements are effective in reducing urban flooding. Using different sizes of these GI,

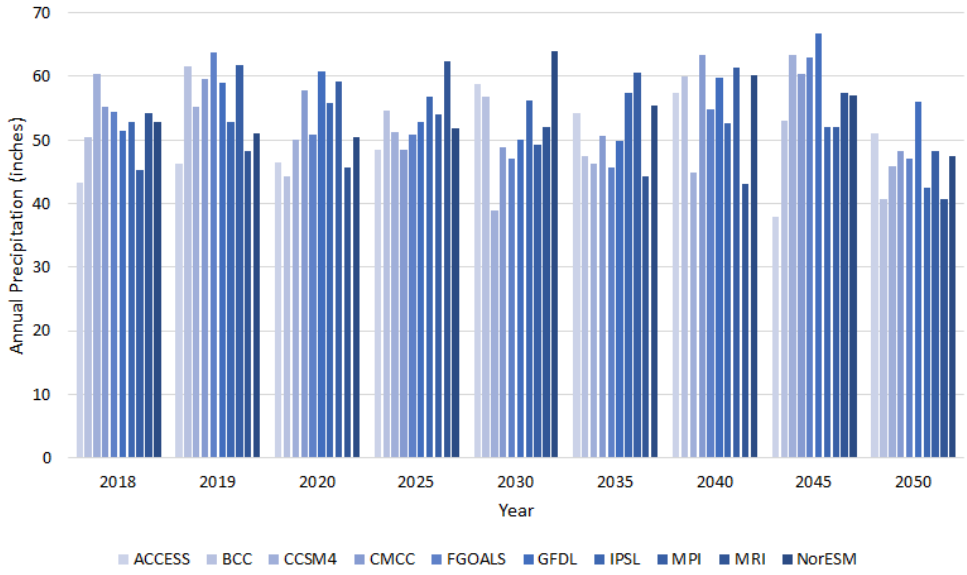

Figure 1.1: Projected annual precipitation over the city of Knoxville, Tennessee, in inches, between 2018 and 2050 under 10 popular climate models.

they studied runoff reduction in 5-year recurrence storm and concluded that the proper GI combination together with appropriate GI sizing is necessary for urban stormwater runoff management. Thus, the properties of the GI and how it is configured in a given watershed have shown to be an important factor in literature for determining the effectiveness of these interventions [45, 44].

Although the current body of work provides invaluable insights, to improve the resiliency of infrastructure, we need to modify our approach to infrastructure planning to account for future changes in climate. Accounting for extreme events does not necessarily translate into planning for the worst-case scenario; instead, it requires policymakers to allocate the budget and effort for future urban planning and maintenance actions by accounting for a wide range of factors *under uncertainty*. In our context, climate parameters, specifically future precipitation, are the main uncertainty. One important factor to consider when trying to optimize a measure of interest under uncertainty, is that not only is knowledge about climate patterns limited and inherently stochastic, but there are multiple climate models that at times make inconsistent predictions. For example, Figure 1.1 gives the projected annual precipitation, in inches, between 2018 and 2050 in the City of Knoxville, Tennessee, using 10 coupled general circulation models (CGCMs) [56]. As seen in the figure, there is significant difference between these 10 models in terms of annual precipitation levels, e.g., in year 2021, standard deviation of precipitation is 7.27 inches. Hence, if placing GI practices in an urban watershed is performed under one projected scenario, it may fall extremely short of addressing the true stormwater management needs if another scenario is realized.

Stochastic and robust programming have been used extensively for decision making under uncertainty, e.g., power systems [105], finance [66, 88], and many engineering applications [47, 52]. Specifically, these approaches has been extensively used in modeling facility location under uncertainty [90]. To the best of our knowledge, the use of these two important methodologies in environmental engineering applications has been limited, especially when it comes to placing GI practices in an urban environment under various uncertainties. Ramshani et al. [84] is perhaps one the few of such studies, and uses a stochastic programming model to optimally place PV panels and green roofs in a mid-sized city under climate change uncertainty to maximize the overall profit from energy generated and saved.

In this paper, we use stochastic and robust programming to account for the uncertainty in future precipitation when placing GI practices in an urban watershed. Specifically, we first develop a two-stage stochastic programming model to determine the optimal placement of GI practices across a set of candidate locations in an urban watershed to minimize the total expected surface runoff under medium-term precipitation uncertainty. Using statistical analysis on the performance of GI practices, we then develop a robust two-stage stochastic programming to produce alternative solutions to the problem of placing GI practices in an urban watershed. We conduct a case study for a watershed in the City of Knoxville, Tennessee, in which we calibrate the model using literature, historical precipitation data, future precipitation projections, watershed hydrological responses to precipitation and GI installations, and expert opinion. We provide the results obtained from the two modeling approaches under various levels of available budget, investigate their differences, conduct extensive sensitivity analyses, and provide insights.

No work has been identified in literature that addresses GI placement in an urban watershed under precipitation uncertainties. Perhaps the closet work is Loáiciga et al. [70]. Their objective was to minimize total construction cost such that volumetric water balance, stormwater volumes, and water-quality characteristics fell within an allowable range. However, this work does not account for the uncertainty in future precipitation projections.

The rest of the paper is organized as follows. First, we formulate the model in Section 2.1. Next, in Section 2.2, we calibrate our model for a watershed in a mid-size city in the U.S. In Section 2.3, we provide the computational results for our case study and draw insights. Finally, we provide a summary and additional insights in Section 1.4.

1.1 Model Formulation

In this study, our goal is to minimize the expected total runoff volume over a medium-term planning horizon under future precipitation uncertainty, given an available budget for investment. This is consistent with challenges currently facing city planners throughout the world. Various types of GI differ in their expense, requirements for advanced planning, necessary land allocation, and their efficiency in reducing surface runoff following

precipitation. Accordingly, in this study, we consider two groups of GI and two stages for placing them. The first set of decisions are made to install large-scale GI practices during the planning horizon to prevent excess runoff before a precipitation scenario is realized. The second set of decisions, which involve placing small-scale GI practices to further improve the overall performance and increase runoff volume reduction, are made after realizing a precipitation scenario.

1.1.1 Two-Stage Stochastic Programming Model

In this section, we develop a two-stage stochastic programming model. The goal is determine the extent to which each sub-catchment must be covered by each of the available types of GI, in the two stages, to minimize the expected total runoff over the planning horizon under precipitation uncertainty. For brevity, we refer to this model as 'stochastic model' in the remainder of the manuscript.

Let $V = \{1, 2, , \ldots, |V|\}$ denote the set of sub-catchments within a watershed whose impervious areas are candidates for placing GI practices. For any given sub-catchment, let $G = \{1, 2, \ldots, |G|\}$ denote the set of all available types of GI practices. We assume that each GI practice may be installed in various levels within a given sub-catchment, e.g., to cover 5%, 7.5%, and 10% of the impervious area within any given sub-catchment with GI. Let $L = \{1, 2, \ldots, |L|\}$ denote the set of available levels of installation of GI practices within a given sub-catchment.

As discussed in Section 2, although CGCMs may be used to project future precipitation in a given region, the resulting projections from different models do not necessarily agree. Hence, the variability across these precipitation projections are the source of uncertainty in our model. Let T denote the length of the planning horizon in years and Ψ denote the finite set of projected precipitation time series for the watershed over the planning horizon T. We let $\psi \in \Psi$ denote a projected precipitation time series, corresponding to scenarios in the model, and π^{ψ} denote the realization probability of scenario $\psi \in \Psi$.

As discussed earlier in Section 2.1, we consider two groups of large- and small-scale GI practices in this study, where practices from the former and latter groups can be placed before and after a CGCM is realized. Let $\overline{T} \leq T$ denote the year in which a precipitation scenario is realized. Also, let G^I and G^{II}, where $G^I \cup G^{II} = G, G^I \cap G^{II} = \emptyset$, denote the set of possible types of GI practices available for placement at $t \leq \overline{T} - 1$ and $\overline{T} \leq t \leq T$, respectively. Consequently, let $x^t_{i,j,l}$ denote the first stage binary decision variable indicating whether or not a GI practice of type $j \in G^I \subset G$ in level l is placed within sub-catchment i in year $t \leq \overline{T}-1$. Similarly, let $y^{\psi,t}_{i,j,l}$ denote the second stage binary decision variables indicating whether or not a GI practice of type $j \in G^{II} \subset G$ in level l is placed within sub-catchment i in $\overline{T} \leq t \leq T$. The decision variables assume the value 1 if the corresponding practice is installed, and the value 0, otherwise. Lastly, we let $\delta_{i,j,l}$ denote the corresponding area (in

square feet) of GI practice type $j \in G$ installed in level $l \in L$, within sub-catchment $i \in V$. In this study, we assume only one type of GI can be placed in each sub-catchment, mainly due to the sizes of GI practices considered, compared to the sizes of the sub-catchments.

Precipitation that is not infiltrated into the soil becomes surface runoff. We incorporate precipitation scenarios into our model by quantifying their impact on each sub-catchments' surface runoff reduction. Let $Q_i^{\psi,t}$ denote the total *baseline* surface runoff under scenario $\psi \in \Psi$ over sub-catchment $i \in V$ in year t when no GI practice is placed. Similarly, let $\hat{Q}_{i,j,l}^{\psi,t}$ denote the surface runoff captured by GI practice of type $j \in G$ installed in level $l \in L$ within sub-catchment $i \in V$ under scenario $\psi \in \Psi$ in year t. Hence, clearly for any given $i \in V$, the difference between $Q_i^{\psi,t}$ and $\hat{Q}_{i,j,l}^{\psi,t}$ gives the total surface runoff in sub-catchment i over year t under scenario $\psi \in \Psi$ as a result of installing GI practices of type $j \in G$ in level $l \in L$ within the sub-catchment.

In this study, we assume that once a GI practice is constructed, it must be maintained annually to preserve its runoff reduction properties. Let $C_{i,j}^t$ denote the per square feet present total cost of placing GI practice of type j within sub-catchment i in year t. Also, let B denote the total available budget at the beginning of the planning horizon for placing GI practices.

A key goal to achieve in planning GI is connectivity as it provides additional resilience against storms and flooding [36, 40, 51, 62, 75]. For instance, all else held constant, a series of connected GI practices is more effective in managing water quantity and quality than a set of disjoint GI practices that are surrounded by urban development [87]. This is mainly because runoff that flows from a sub-catchment to a downstream sub-catchment can be slowed or captured by GI practices before reaching downstream [51, 62]. This impact is particularly pronounced in adjacent/neighboring sub-catchments with respect to watershed hydrology as the connected GI practices can further mitigate runoff resulting from 'directly connected impervious areas,' reducing runoff volumes, peak discharge, and baseflow effects [79].

Figure 1.2 illustrates a subset of a watershed consisting of seven sub-catchments and its main stream. Placing GI practices in any of the sub-catchments reduces the surface runoff in that sub-catchment. Additionally, dependent on sub-catchments characteristics [35], placing a GI practice in an upstream sub-catchment, may further reduce the surface runoff in a downstream sub-catchment. Lastly, simultaneous placement of GI practices has the potential to further mitigate the surface runoff, if the sub-catchments are 'hydrologically connected.' For instance, because sub-catchments 3 and 5 are hydrologically connected, placing GI practices in sub-catchment 3 can potentially also reduce the amount of run-off over sub-catchment 5, even if no GI is placed on the latter sub-catchment. Furthermore, placing GI practices in both sub-catchments 3 and 5 can potentially result in a larger reduction in surface runoff, compared to that obtained from placing the same type/level of GI practices in the two sub-catchments if they were not hydrologically connected.

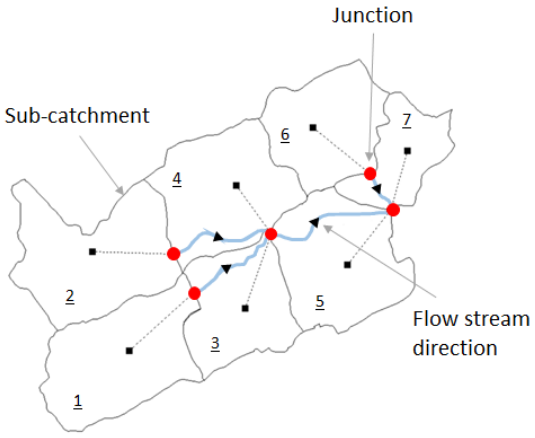

Figure 1.2: A small portion of a watershed consisting of seven sub-catchments and its main stream.

We capture sub-catchment connectivity in a watershed using a directed acyclic graph. Specifically, let the directed acyclic graph $\mathbb{G}(V, A)$ denote the system of sub-catchments where V is the set of nodes in the graph, corresponding to the sub-catchments in the watershed, and A denotes the set of sub-catchment connectivity arcs, where there exists an arc $a_{i',i} \in A$ if and only if sub-catchments $i', i \in V$ are connected.

Consider a given pair of connected sub-catchments i' and i, $a_{i',i} \in A$. When accounting for surface run-off reduction over sub-catchment i due to a GI practice placed upstream, assuming large-scale GI practice installations only, three distinct cases must be considered: (a) a GI practice is placed within upstream sub-catchment i' in year t' after a GI is placed within downstream sub-catchment i in year t such that $0 \leq t \leq t' \leq \overline{T} - 1$; (b) a GI is placed within downstream sub-catchment i in year t after a GI is placed within upstream sub-catchment i' in year t' such that $0 \leq t' \leq t \leq \overline{T} - 1$; and (c) a GI is placed within upstream sub-catchment i' in year t' and no GI placed in downstream sub-catchment i by the beginning of year $\bar{T}$, i.e., $0 \leq t' \leq \overline{T} - 1$.

To be able to account for the adjustment in surface runoff reduction due to GI installations in connected sub-catchments as described in cases (a)-(c), we introduce the runoff 'adjustment factor' $\beta_{i,j,l}^{i',j',l'}$ and the variable $z_{t,i,j,l}^{t',i',j',l'}$. Specifically, for any given pair of connected sub-catchments i' and i, where $a_{i',i} \in A$, we let $0 \leq \beta_{i,j,l}^{i',j',l'} \leq 1$ denote the runoff 'adjustment factor' over the downstream sub-catchment $i \in V$, when a GI practice of type $j' \in G^I$ in level $l' \in L$ is installed within upstream sub-catchment $i' \in V$ and no GI practice or a GI practice of type $j \in G^I$ in level $l \in L$ is installed within the downstream sub-catchment $i \in V$. We use $j = 0$ to indicate that no GI is installed within a sub-catchment. In addition, we let $z_{t,i,j,l}^{t',i',j',l'}$ denote the binary variable indicating whether or not GI practices of types $j', j \in G^I$ in levels $l', l \in L$ are installed within sub-catchment $i', i \in V$ in years $t', t \leq \bar{T} - 1$, respectively. The variable assumes the value 1 if the corresponding practices are installed and

equals 0, otherwise. In addition, we define $z_{t,i,0,l}^{t',i',j',l'} = 0$ for all t', i', j', l', t, i, and l, to account for cases when only the upstream sub-catchment is selected for installing large-scale GI practices. Note that a downstream sub-catchment can be hydrologically connected to more than one upstream sub-catchment. We assume that the 'adjustments' over downstream sub-catchments are additive. Lastly, for completeness, we define $\hat{Q}_{i,0,l}^{\psi,t} = 0$ for all ψ, t, i, and l to account for the case where no GI is installed in sub-catchment $i \in V$.

In addition to accounting for potential adjustments in runoff reduction as a result of hydrological connectivity, we require first stage decision variables to fulfill a certain connectivity constraint to ensure that the model provides at least a minimum desired level of connectivity among large-scale GI practices by the beginning of year $\overline{T}$. Specifically, we define GI connectivity as a 1-neighbor constraint on first stage decision variables, which prescribe large-scale GI practice installations. That is, a first stage GI practice can be installed in sub-catchment i if there exists at least one placed first stage GI practice in one of the sub-catchments that are hydrologically connected to sub-catchment i. For simplicity of notation, in the remainder we use $x = [x_{i,j,l}^t]$, $z = [z_{t,i,j,l}^{t',i',j',l'}]$, $y = [y_{i,j,l}^{\psi,t}]$ to refer to the vectors of the corresponding variables. The notation is summarized in Appendix A.

We let $\phi_S(x, z, y)$ denote the total expected surface runoff across the watershed $\mathbb{G}(V, A)$ over the planning horizon, T, under the decision variables x, z, and y for the stochastic model. Therefore, given the total available budget, B, the following model minimizes $\phi_S(x, z, y)$, i.e.,

$$
\begin{aligned}
\min_{x,z,y} \phi_S(x,z,y) = \min_{x,z,y} \sum_{\psi\in\Psi} \pi^{\psi} \cdot \Bigg[& \sum_{i\in V} \sum_{\{t|0\le t\le T\}} Q_i^{\psi,t} \\
& - \sum_{i\in V} \sum_{j\in G^I} \sum_{l\in L} \sum_{\{t|0\le t\le \overline{T}-1\}} \sum_{\{t'|t\le t'\le T\}} \hat{Q}_{i,j,l}^{\psi,t'} \cdot x_{i,j,l}^{t} \\
& - \sum_{a_{i',i}\in A} \sum_{j\in G^I\cup\{0\}} \sum_{j'\in G^I} \sum_{l\in L} \sum_{l'\in L} \sum_{\{t|0\le t\le \overline{T}-1\}} \sum_{\{t'|0\le t'\le \overline{T}-1\}} \sum_{\{t''|\max\{t\cdot\mathbb{1}_{\{j\neq 0\}},t'\}\le t''\le T\}} \\
& \qquad \beta_{i,j,l}^{i',j',l'} \left(Q_i^{\psi,t''} \left(x_{i',j',l'}^{t'} - z_{t,i,j,l}^{t',i',j',l'} \right) + \hat{Q}_{i,j,l}^{\psi,t''} \cdot z_{t,i,j,l}^{t',i',j',l'} \right) \\
& - \sum_{a_{i',i}\in A} \sum_{j\in G^I\cup\{0\}} \sum_{j'\in G^I} \sum_{l\in L} \sum_{l'\in L} \sum_{\{t|0\le t\le \overline{T}-1\}} \sum_{\{t'|0\le t'\le t-1\}} \sum_{\{t''|t'\le t''\le t-1\}} \beta_{i,j,l}^{i',j',l'} \cdot Q_i^{\psi,t''} \cdot z_{t,i,j,l}^{t',i',j',l'} \\
& - \sum_{i\in V} \sum_{j\in G^{II}} \sum_{l\in L} \sum_{\{t|\overline{T}\le t\le T\}} \sum_{\{t'|t\le t'\le T\}} \hat{Q}_{i,j,l}^{\psi,t'} \cdot y_{i,j,l}^{\psi,t} \Bigg],
\end{aligned}
\tag{1.1}
$$

$$\text{s.t.} \sum_{i\in V}\sum_{j\in G^I}\sum_{l\in L}\sum_{\{t|0\le t\le \overline{T}-1\}} C^t_{i,j}\cdot\delta_{i,j,l}\cdot x^t_{i,j,l} + \sum_{i\in V}\sum_{j\in G^{II}}\sum_{l\in L}\sum_{\{t|\overline{T}\le t\le T\}} C^t_{i,j}\cdot\delta_{i,j,l}\cdot y^{\psi,t}_{i,j,l} \le B, \quad \forall\psi\in\Psi, \tag{1.2}$$

$$x^{t'}_{i',j',l'} + x^t_{i,j,l} \le z^{t',i',j',l'}_{t,i,j,l} + 1, \quad \forall i', i\in V, a_{i',i}\in A, \forall j', j\in G^I, \forall l', l\in L, 0\le t', t\le \overline{T}-1, \tag{1.3}$$

$$x^t_{i,j,l} \ge \sum_{j'\in G^I}\sum_{l'\in L}\sum_{\{t'|0\le t'\le\overline{T}-1\}} z^{t',i',j',l'}_{t,i,j,l}, \quad \forall i, i'\in V, a_{i',i}\in A, \forall j\in G^I, \forall l\in L, 0\le t\le\overline{T}-1, \tag{1.4}$$

$$x^{t'}_{i',j',l'} \ge \sum_{j\in G^I}\sum_{l\in L}\sum_{\{t|0\le t\le\overline{T}-1\}} z^{t',i',j',l'}_{t,i,j,l}, \quad \forall i, i'\in V, a_{i',i}\in A, \forall j'\in G^I, \forall l'\in L, 0\le t'\le\overline{T}-1, \tag{1.5}$$

$$z^{t',i',j',l'}_{t,i,j,l} = 0, \quad \forall i', i\in V, a_{i',i}\in A, j=0, \forall j'\in G^I, \forall l, l'\in L, 0\le t, t'\le\overline{T}-1, \tag{1.6}$$

$$\sum_{j\in G^I}\sum_{l\in L}\sum_{\{t|0\le t\le\overline{T}-1\}} x^t_{i,j,l} \le \sum_{a_{i',i}\in A}\sum_{j\in G^I}\sum_{l\in L}\sum_{\{t|0\le t\le\overline{T}-1\}} x^t_{i',j,l}, \quad \forall i\in V, \tag{1.7}$$

$$\sum_{j\in G^I}\sum_{l\in L}\sum_{\{t|0\le t\le\overline{T}-1\}} x^t_{i,j,l} + \sum_{j\in G^{II}}\sum_{l\in L}\sum_{\{t|\overline{T}\le t\le T\}} y^{\psi,t}_{i,j,l} \le 1, \quad \forall i\in V, \psi\in\Psi, \tag{1.8}$$

$$x^t_{i,j,l}, y^{\psi,t''}_{i,j'',l}, z^{t',i',j',l'}_{t,i,j,l} \in \{0,1\}, \quad \forall i', i\in V, a_{i',i}\in A, \forall j', j\in G^I, \forall j''\in G^{II}, \forall l', l\in L, \forall\psi\in\Psi, 0\le t, t'\le\overline{T}-1, \overline{T}\le t''\le T. \tag{1.9}$$

The objective function (1.1) minimizes the total expected surface runoff across the sub-catchments within the watershed over the planning horizon. The first term in (1.1) captures

the total baseline runoff. The second term in (1.1) presents the reduction in surface runoff over the sub-catchments as a result of first stage GI installations within the sub-catchments. The third and forth terms in (1.1) address the adjustment in surface runoff reduction due to GI installations in connected sub-catchments (see Appendix B for more details). Finally, the last term in (1.1) presents the reduction in surface runoff over the sub-catchments as a result of second-stage GI installations within the sub-catchments.

Constraint (1.2) enforces budget limitations for placing GI practices. Constraints (1.3)-(1.6) establish the relationship between variables x and z and enforces the latter to assume the value one when large-scale GI practices are installed within two connected sub-catchments, and to assume the value zero, otherwise. Constraint (1.7) ensures the 1-neighbor connectivity among first-stage GI practices. Constraint (1.8) assures that at most one GI practice is installed in any given sub-catchment throughout the planning horizon. Finally, constraint (1.9) enforces binary restrictions on the decision variables. Let Ω denote the feasible set of the problem, i.e., $\Omega = \{\chi = (x, z, y) | (1.2) - (1.9)\}$. Accordingly, we let $\chi_S^\star \in \Omega$ denote the optimal solution to the stochastic model, i.e., $\phi_S(\chi_S^\star) \leq \phi_S(\chi)$ for all $\chi \in \Omega$.

1.1.2 Two-Stage Stochastic Robust Programming Model

Here, we reformulate the problem as a two-stage stochastic robust model, using the same notation introduced in Section 1.1.1. Similar to the previous formulation, the model prescribes the extent to which each sub-catchment must be covered by each type of GI practice in the two stages. Different from the previous formulation in which the baseline surface runoff volume, $Q_i^{\psi,t}$, and surface runoff volume captured by a GI practice, $\hat{Q}_{i,j,l}^{\psi,t}$, were assumed to be readily known, in this formulation we assume there is uncertainty in calculating these runoff volumes. For brevity, we refer to this model as 'robust model' in the remainder of the manuscript.

Specifically, we redefine $Q_i^{\psi,t}$ to denote the *average* baseline surface runoff volume within sub-catchment $i \in V$ under scenario $\psi \in \Psi$ in year t, and let $2q_i^{\psi,t}(\alpha)$ denote the width of the $100(1-\alpha)\%$ confidence interval (CI) for the corresponding average baseline surface runoff volume. Similarly, we redefine $\hat{Q}_{i,j,l}^{\psi,t}$ to denote the *average* surface runoff volume captured by GI practice of type $j \in G$ installed in level $l \in L$ within sub-catchment $i \in V$ under scenario $\psi \in \Psi$ in year t, and let $2\hat{q}_{i,j,l}^{\psi,t}(\alpha)$ denote the width of the $100(1-\alpha)\%$ CI for the corresponding average surface runoff volume captured by the GI practice. Consequently, $[Q_i^{\psi,t} - q_i^{\psi,t}(\alpha), Q_i^{\psi,t} + q_i^{\psi,t}(\alpha)]$ and $[\hat{Q}_{i,j,l}^{\psi,t} - \hat{q}_{i,j,l}^{\psi,t}(\alpha), \hat{Q}_{i,l,j}^{\psi,t} + \hat{q}_{i,j,l}^{\psi,t}(\alpha)]$ give the corresponding $100(1-\alpha)\%$ CI for the average baseline surface runoff volume and runoff volume captured, respectively. Accordingly, the average baseline runoff volume and runoff volume captured by the given GI practice within sub-catchment $i \in V$ under scenario $\psi \in \Psi$ in year t are no worse than the CI upper bound $Q_i^{\psi,t} + q_i^{\psi,t}(\alpha)$ and the CI lower bound $\hat{Q}_{i,j,l}^{\psi,t} - \hat{q}_{i,j,l}^{\psi,t}(\alpha)$, respectively, $100(1-\alpha)\%$ of the time.

Finally, consistent with Soyster's method [91], in our objective function of the robust model, compared with that of the stochastic model in equation (1.1), we use the $100(1-\alpha)\%$ CI upper bound and lower bounds of the estimated values for $Q_i^{\psi,t}$ and $\hat{Q}_{i,j,l}^{\psi,t}$ to take a conservative view. Accordingly, we let $\phi_{R_\alpha}(x,z,y)$ denote the total expected surface runoff volume across the watershed $\mathbb{G}(V,A)$ over the planning horizon, T, under the decision variables x, z, and y for the robust model. Therefore, given the total available budget, B, the following model minimizes $\phi_{R_\alpha(x,z,y)}$. Note that analogous to the stochastic model, we let $\chi_R^\star \in \Omega$ denote the optimal solution to the robust model.

$$
\begin{aligned}
\min_{x,z,y} \phi_{R_\alpha}(x,z,y) = \min_{x,z,y} \sum_{\psi\in\Psi} \pi^\psi \cdot \Bigg[& \sum_{i\in V} \sum_{\{t|0\le t\le T\}} Q_i^{\psi,t} + q_i^{\psi,t}(\alpha) \\
& - \sum_{i\in V} \sum_{j\in G^I} \sum_{l\in L} \sum_{\{t|0\le t\le \overline{T}-1\}} \sum_{\{t'|t\le t'\le T\}} \left(\hat{Q}_{i,j,l}^{\psi,t'} - \hat{q}_{i,j,l}^{\psi,t'}(\alpha)\right) \cdot x_{i,j,l}^{t} \\
& - \sum_{a_{i',i}\in A} \sum_{j\in G^I\cup\{0\}} \sum_{j'\in G^I} \sum_{l\in L} \sum_{l'\in L} \sum_{\{t|0\le t\le \overline{T}-1\}} \sum_{\{t'|0\le t'\le \overline{T}-1\}} \sum_{\{t''|\max\{t\cdot\mathbb{1}_{\{j\neq 0\}},t'\}\le t''\le T\}} \\
& \qquad \beta_{i,j,l}^{i',j',l'} \left(\left(Q_i^{\psi,t''} - q_i^{\psi,t''}(\alpha)\right) \left(x_{i',j',l'}^{t'} - z_{t,i,j,l}^{t',i',j',l'}\right) + \left(\hat{Q}_{i,j,l}^{\psi,t''} - \hat{q}_{i,j,l}^{\psi,t''}(\alpha)\right) \cdot z_{t,i,j,l}^{t',i',j',l'} \right) \\
& - \sum_{a_{i',i}\in A} \sum_{j\in G^I\cup\{0\}} \sum_{j'\in G^I} \sum_{l\in L} \sum_{l'\in L} \sum_{\{t|0\le t\le \overline{T}-1\}} \sum_{\{t'|0\le t'\le t-1\}} \sum_{\{t''|t'\le t''\le t-1\}} \\
& \qquad \beta_{i,j,l}^{i',j',l'} \cdot \left(Q_i^{\psi,t''} - q_i^{\psi,t''}(\alpha)\right) \cdot z_{t,i,j,l}^{t',i',j',l'} \\
& - \sum_{i\in V} \sum_{j\in G^{II}} \sum_{l\in L} \sum_{\{t|\overline{T}\le t\le T\}} \sum_{\{t'|t\le t'\le T\}} \left(\hat{Q}_{i,j,l}^{\psi,t'} - \hat{q}_{i,j,l}^{\psi,t'}(\alpha)\right) \cdot y_{i,j,l}^{\psi,t} \Bigg],
\end{aligned}
\tag{1.10}
$$

$$\text{s.t. } \chi = (x,z,y) \in \Omega.$$

Although the objective function in the robust model may seem overly conservative compared with the one in the stochastic model, that only accounts for average volumes, we believe such a model is practical in our context. Note that the intensity of precipitation, i.e., the amount of precipitation in a period of time (especially for short periods, e.g., 24 hours) is an important predictor of, and is negatively correlated with, GI practice performance [33]. Assuming that the precipitation intensity is relatively similar across all sub-catchments in a relatively small watershed, when intense precipitation occurs, the performance of all GI practices are expected to get worse. This means that the resulting runoff across all sub-catchments would increase accordingly and, in turn, in equation (1.1) all coefficients pertaining to baseline surface runoff volume, $Q_i^{\psi,t}$, and surface runoff volume captured by GI practices, $\hat{Q}_{i,j,l}^{\psi,t}$, must be adjusted.

1.2 Model Calibration

In this section, we calibrate the mathematical models presented in Section 2.1 using literature, historical data, precipitation projections, and expert opinion for an urban watershed of a mid-sized city in the U.S. First, in Section 2.2.1 we calibrate the parameters associated with GI practices. Next, in Section 1.2.2 we discuss the preprocessing performed on precipitation projections to convert them into the requisite format. Finally, in Section 1.2.3 we describe the hydrological simulations performed to characterize the surface runoff resulting from precipitation projections in the sub-catchments, under potential GI placements.

1.2.1 GI Practices

The performance of a GI practice can be described as the volume of surface runoff that the practice can infiltrate on an hourly basis [89]. The performance of GI practices depend on an array of factors including design specifications (such as surface storage volume, media storage, and media composition and depth, etc.) and climate patterns (such as precipitation event intensity and duration, etc.) [37, 93]. In addition, maintenance activities must be performed for GI practices to continue their performance [6].

GI types, G, and GI installation levels, L. In this study, we consider two common types of GI practices, namely, bioretention and rain garden [41], hence $|G|= 2$. The former is typically installed in relatively large, commercial scales and is held to a higher design standard, whereas the latter is a smaller system with lower design standards and is placed in residential lots [9]. Accordingly, we let bioretentions and rain gardens be available for placement in the first and second stages, respectively. Hence, in this study we use $|G^I|= 1$ and $|G^{II}|= 1$.

The amount of surface runoff reduction by GI practices in any given sub-catchment of a watershed is closely related to the surface area that they cover from the corresponding sub-catchment [86]. We account for three levels of installation for each of the two GI practice types considered, i.e., $|L|= 3$. National Association of City Transportation Officials (NACTO) [13] recommends using the effective impervious surface area in the drainage region (sub-catchments) as a key design factor when sizing bioretentions [14]. To that end, and due to the larger scale of bioretention installation, in this study we allow bioretentions to cover 5%, 7.5%, and 10% of the impervious area of each sub-catchment. Given the size of a sub-catchment, these ratios can be translated into square feet to obtain the corresponding values of $\delta_{i,j,l}$. For rain gardens, due to their residential-scale implementation, we allocate the total areas of 2,500, 5,000, and 7,500 square feet for placing the GI practices within each sub-catchment. Finally, note that for the general attributes of the two types of GI practices considered, e.g., minimum media depth, ponding depth, media permeability, we

use the stormwater training manuals from State of Tennessee Department of Environment & Conservation [17].

GI costs, $C^t_{i,j}$. The total cost of placing GI practices includes construction and maintenance costs. Let $c^t_{i,j}$ and $\mathfrak{c}^t_{i,j}$ denote the per square feet construction and annual maintenance costs of a GI practice of type j in sub-catchment i in year $t \leq T$, respectively. We assume the maintenance cost incurs annually starting from the year of construction and is subject to an annual increase with the average annual inflation rate r. Hence, the present value at time zero of the total per square feet cost of placing GI practice of type j in sub-catchment i at time t is given by

$$C^t_{i,j} = \frac{1}{(1+r)^t} \cdot \left(c^t_{i,j} + \frac{1 - r^{T-t}}{1 - r} \cdot \mathfrak{c}^t_{i,j} \right).$$

We use the inflation-adjusted EPA Opti-Tool [43] and the University of Texas A&M's AGRILIFE Report [46] to obtain the per square feet construction cost of bioretentions and rain gardens, respectively. For instance, the reported per square feet construction cost of bioretentions was \$15.46 in 2016, and that of rain gardens was \$6.00 in 2012. To estimate the corresponding costs during the planning horizon, we adjust the values using the U.S. Labor Department's Consumer Price Index (CPI) inflation calculator [96] . We do not consider land cost in this study as we assume all GI practices are placed on public land or on land parcels offered by private property owners. Based on published reports [92], the annual GI maintenance cost ranges between 3%-6% of its construction cost. Let ρ denote the ratio of maintenance cost to construction cost, i.e., $\rho = \mathfrak{c}^t_{i,j}/c^t_{i,j}$. In this study, we set $\rho = 3\%$. Lastly, we use the average annual inflation rate $r = 1.86\%$, which equals the average annual U.S. inflation rate over the period 2007-2017 [77].

1.2.2 Pre-Processing of Precipitation Projections

As discussed in Section 2, CGCMs project future precipitation, which are next fed to hydrological simulators to calculate the resulting surface runoff, at various degrees of GI installation. In this study, we use precipitation projections for the City of Knoxville produced by ten CGCMs (see Table C.1 in Appendix C for more detail.) Note that using CGCMs to produce projections are computationally expensive and hence, the projections are usually only produced in daily units. Let $\hat{\psi} \in \hat{\Psi}$ denote a daily precipitation projection time series produced by one of the ten CGCMs. To accurately capture the GI response to precipitation, more granular data, i.e., hourly precipitation projections, are required due to quick transport of runoff in urban watersheds. Therefore, the daily precipitation projections must be converted into hourly precipitation projections, denoted by $\psi \in \Psi$, before they can be fed into hydrological simulators to calculate corresponding amounts of surface runoff.

Note that hourly precipitation projections can be uniquely aggregated to produce daily precipitation projections; however, the reverse is not true. In this section, we present an approach for pre-processing daily precipitation projections to generate one of the many likely hourly precipitation projections. Specifically, we use quartile-based and seasonal-based temporal distributions [27] to convert a daily precipitation time series into an hourly precipitation time series. Temporal distributions of precipitation summarize the historical cumulative percentages of precipitation up to any point during a precipitation event and provide the proportion of time that the pattern was observed.

Figure 1.3 presents the quartile-based distributions of 24-hour precipitation of Ohio river basin (including the City of Knoxville), adopted from Precipitation-Frequency Atlas of the United States [27]. Specifically, Figures 1.3(a)-1.3(d) present the cumulative probability plots of temporal distributions, where the highest percentage of precipitation during the 24-hour period occurred in the first-fourth quarters of the day, respectively. For instance, Figure 1.3(a) presents the temporal distributions, where the highest amount of daily precipitation occurred during the first quarter of the day. The nine cumulative distributions in each panel present the nine general patterns according to which the corresponding amount of precipitation was accumulated. The shades of the distributions present the percentage of time that the particular pattern was observed. For instance, given that the highest percentage of daily precipitation occurred in the first quantile (Figure 1.3(a)), in 10% of the cases, 55.1% of total daily precipitation occurred during the first 8.3% of the time-period, i.e., the first 2 hours of the day.

Note that the quartile-based temporal distributions in Figure 1.3 are generated under the assumption of the homogeneity of monthly precipitation. In the absence of monthly precipitation homogeneity, seasonality must be considered [81]. Hence, we follow the procedure described in Huff (1967) to generate seasonal-based temporal distributions. As the input, we use 20 years of precipitation data (i.e., year 1997–year 2016) in the City of Knoxville, obtained from National Center for Environmental Information (NCEI) [15], stratified across the four seasons. Figure 1.4 presents the resulting seasonal temporal distributions of daily precipitation.

Lastly, to analyze the homogeneity of monthly precipitation to determine whether quartile-based or seasonal-based temporal distributions can be applied to convert the daily projections into hourly projections, we use coefficient of variability [38, 81]. Let p_k denote the accumulated precipitation in month k and let η denote the precipitation coefficient of variability. The value of η is given by

$$\eta = \frac{\sum_{k=1}^{12} p_k^2}{(\sum_{k=1}^{12} p_k)^2}.$$

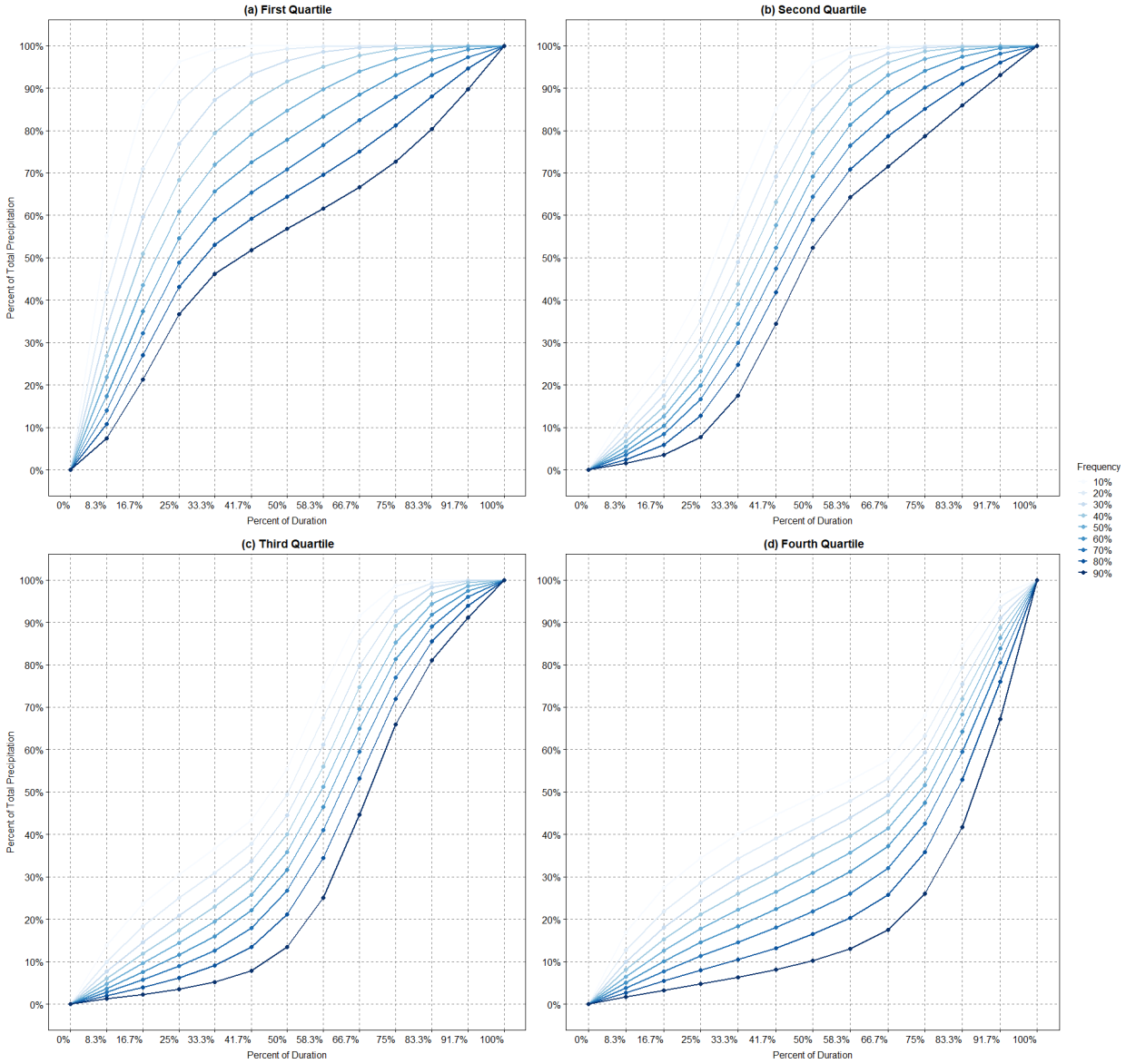

Figure 1.3: Quartile-based Temporal distributions of daily precipitation, adapted from [27].

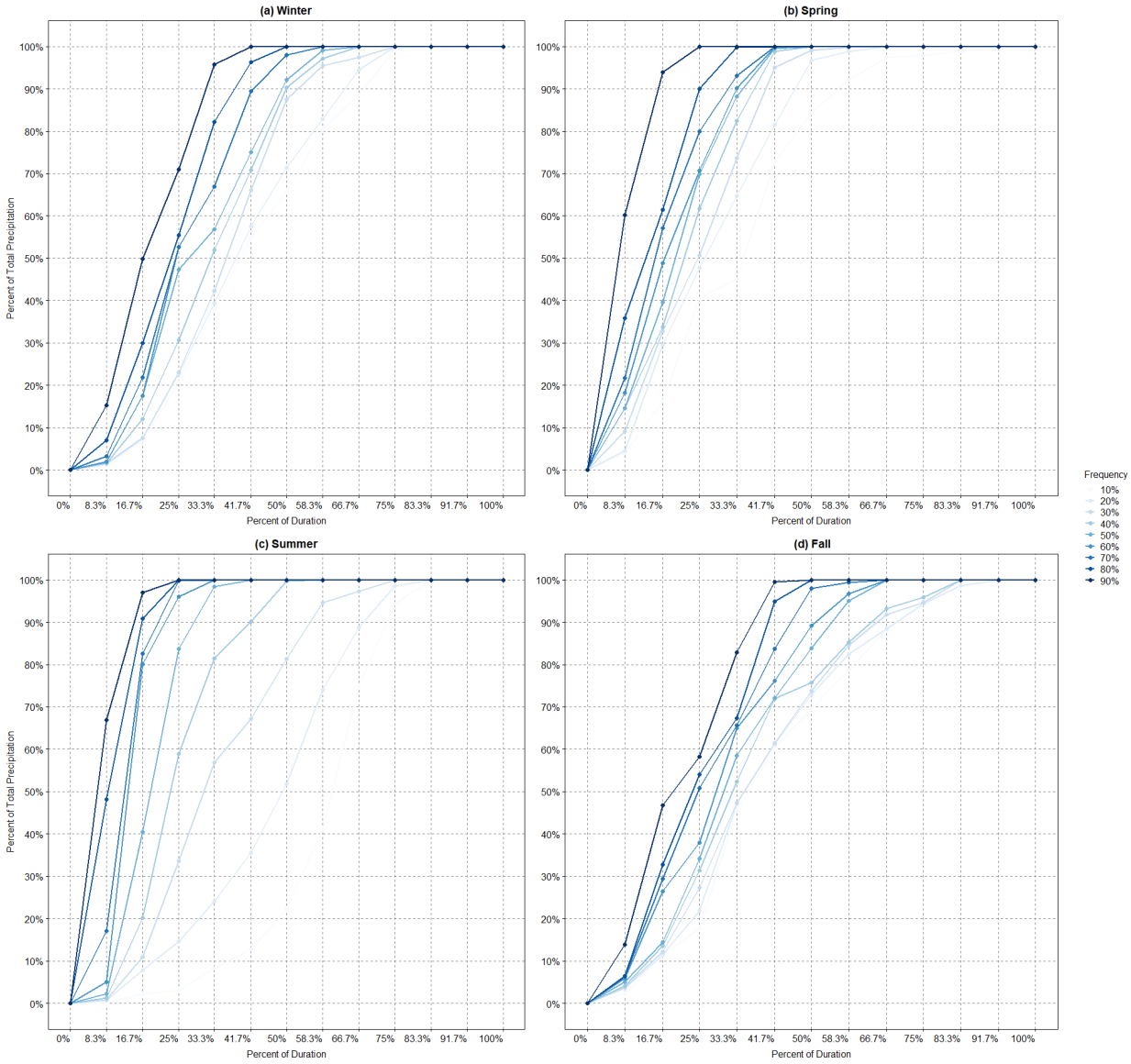

Figure 1.4: Seasonal-based temporal distributions of daily precipitation.

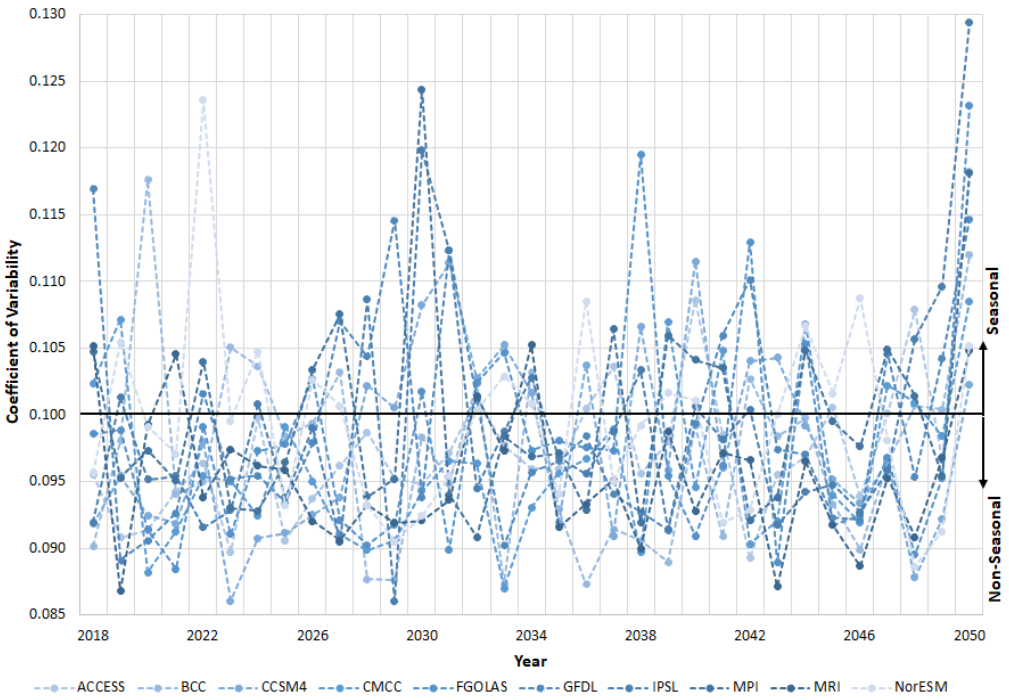

Figure 1.5: Coefficient of variability for the ten precipitation projections for the City of Knoxville, for years 2018 to 2050.

If the value of η ranges between 0 and 0.1, it suggests that precipitation is relatively uniformly distributed across the months, i.e., homogeneity of monthly precipitation. In contrast, if the value of η ranges between 0.1 and 0.2, it indicates seasonal patterns for precipitation. Note that if the value of η is greater than 0.2, it indicates that there are distinct monthly precipitations and thus, monthly precipitations are fully heterogeneous. Figure 1.5 presents the coefficient of variability, η, for the 10 precipitation projections during the planning time horizon. As seen in the figure, η ranges from 0.085 to 0.125, and is always less than 0.2.

Finally, we use the following procedure to pre-process any given daily precipitation projection (from any of the ten CGCMs) to generate an hourly projection. First, we break down the daily precipitation projection by year. For any given year, we first examine the value of coefficient of variability to determine whether quartile-based or seasonal-based temporal distributions apply. If quartile-based distributions apply, we first determine the proportion of time that the highest percentage of precipitation occurred in the first-fourth quarters of the day using the historical precipitation data collected in the Ohio River Basin, which includes the City of Knoxville [27]. Next, for any given day of the year in the daily precipitation projections, we generate a weighted random number according to these proportions to determine which quartile to use. Next, we generate a weighted random number according to the probability of observing each of the cumulative distributions in the

corresponding quantile. Finally, once a cumulative distribution is chosen, we use it to project the amount of precipitation in that day into an hourly time series. Similarly, if seasonal-based temporal distributions apply, for any given day in any given season, we generate a weighted random number according to the probability of observing each of the cumulative distributions for that season. We then use the selected cumulative distribution to project the amount of precipitation in that day into an hourly time series.

1.2.3 Hydrological Simulations And Estimating Surface Runoff

As discussed in Section 1.2.2, precipitation projections need to be fed into hydrological simulators to calculate the surface runoff during any given precipitation event, at various degrees of GI installation. In this study, to perform hydrological simulations we use EPA SWMM [86], a widely used software in literature [31, 39, 61, 72, 76, 85, 94, 95]. Note that conducting brute-force SWMM simulations can be time-consuming. Hence, in this section, we first discuss the computational difficulties of executing such simulations and then provide an approach for sampling events to estimate the total baseline surface runoff, $Q_i^{\psi,t}$, and the surface runoff captured by GI practices, $\hat{Q}_{i,j,l}^{\psi,t}$, under various projected precipitation scenarios, ψ, for the stochastic model. Next, we describe the approach used for calculating the $100(1-\alpha)\%$ CIs for the amount of surface runoff captured by GI practices that gives the estimated value of $\hat{q}_{i,j,l}^{\psi,t}(\alpha)$ for the robust model. Finally, we describe the approach used for calculating the runoff adjustment factor $\beta_{i,j,l}^{i',j',l'}$ over any given downstream sub-catchment.

SWMM partitions rainfall to runoff and routes it through the watershed and the potential GI practices, while accounting for several adjustments such as (i) rainfall interception from depression storage, (ii) infiltration of rainfall into unsaturated soil layers, and (iii) percolation of infiltrated water into groundwater layers [86]. Note that SWMM simulation can be extremely computationally expensive, given a large watershed and a long time horizon for the input precipitation. For instance, based on our experiments, each SWMM simulation performed on a 2.4 GHz CPU (single core) to obtain the surface runoff after placing GI practices within a single sub-catchment can take on the order of approximately 25 minute to execute for a time series that spans only one year, expressed in hourly units. Note that increasing the planning horizon proportionally increases the simulation time. In addition, given the total number of the sub-catchments, $|V|$, the number of GI types to place in each sub-catchment, $|G|$, and the number of possible installation levels, $|L|$, a total of $(|G|\cdot|L|+1)^{|V|}$ SWMM simulations must be executed to calculate the surface runoff for all possible combinations of GI placements if all sub-catchments are hydrologically connected. Hence, using a brute-force simulation approach is computationally intractable even for a medium-sized watershed, with approximately 100 sub-catchments.

Therefore, in this study, we exploit three approaches to mitigate the prohibitively long simulation time to estimate the surface runoff. First, we use a sampling method

to approximate surface runoff resulting from precipitation scenarios using only a series of sampled events instead of the entire precipitation time series spanning the planning horizon. Note that we use the sampling method along with bootstrapping to also calculate CIs for the estimated surface runoff vlolumes. Second, we run SWMM simulations for all sub-catchments simultaneously, i.e., we execute one simulation under no GI practice placement to calculate the baseline surface runoff over all sub-catchments, and a total of $|G|\cdot|L|$ simulations where the same GI practice of type j in the same level l is placed within all sub-catchments to calculate the corresponding runoff after placement. Clearly, if hydrological connectivity among sub-catchments are not captured in the watershed model, the estimated runoff volumes give the parameters $Q_i^{\psi,t}$ and $\hat{Q}_{i,l,j}^{\psi,t}$, respectively. However, if hydrological connectivity among sub-catchments are captured in the watershed model, the former and latter groups of estimated runoff volumes need to be adjusted back by the adjustment factor $\beta_{i,j,l}^{i',j',l'}$ to estimate the parameters $Q_i^{\psi,t}$ and $\hat{Q}_{i,l,j}^{\psi,t}$, respectively. In our main SWMM simulation model for the watershed, the hydrological connectivity among sub-catchments is not entirely captured, which simplifies the estimation of $Q_i^{\psi,t}$ and $\hat{Q}_{i,l,j}^{\psi,t}$. However, at the same time, it complicates the process of estimating the adjustment factors $\beta_{i,j,l}^{i',j',l'}$. To be able to estimate the adjustment factors, we develop a complementary SWMM model, which we calibrate based on the characteristics of the sub-catchments and their hydrological connectivity in the watershed. Lastly, we stratify sub-catchments based on their characteristics and only calculate the adjustment factors $\beta_{i,j,l}^{i',j',l'}$ for a reduced number of sub-catchment type pairs using the set of sampled events.

Sampling events and calculating surface runoff parameters, $Q_i^{\psi,t}$ and $\hat{Q}_{i,l,j}^{\psi,t}$, for the stochastic model. As discussed in Section 2.1, precipitation intensity, i.e., the amount of precipitation in a period of time, is an important predictor of, and is negatively correlated with, GI practice performance. Depending on the intensity of precipitation events, a GI practice may present different performance levels. For instance, GI practices generally exhibit a lower performance under a series of short but intense events, but a higher performance under long but mild events. Hence, we use precipitation intensity as a basis for sampling events.

Recall that $\hat{\psi} \in \hat{\Psi}$ denotes daily precipitation projection time series produced by the CGCMs, and $\psi \in \Psi$ denotes hourly precipitation projection time series, corresponding to scenarios in the stochastic model. As discussed in Section 1.2.2, the hourly precipitation projection time series, ψ, resulting from a daily precipitation projection time series, $\hat{\psi}$, is not unique. In fact, each of the resulting hourly precipitation projection time series can have very different daily precipitation intensities in any given day. Therefore, we rely on repeated sampling to produce a large set of hourly precipitation projection time series, ψ, and then aggregate them based on precipitation intensities to estimate runoff volumes. The detailed steps are as follows.

- **Initialization**: For each of the ten daily precipitation projections, $\hat{\psi} \in \hat{\Psi}$, use precipitation coefficient of variability, η, to identify the relevant set of temporal distributions for any given rainy day, i.e., when precipitation volume is greater than zero, over the span of 32 years, i.e., 2018-2050. Next, for each of these rainy days, randomly select from the corresponding set of temporal distributions to project daily precipitation into hourly basis. Repeat the procedure to generate 100 time series of hourly precipitation projections for each of the 10 daily precipitation projections, $\hat{\psi} \in \hat{\Psi}$. This results in 1,000 time series of hourly precipitation projections, $\psi \in \Psi$, each of which consist of a series of hourly precipitation events with various intensities.

- **Aggregation**: Use all $\psi \in \Psi$ to calculate the histogram of hourly event intensities, using Sturge's rule to break the intensity range into categories.

- **SWMM Simulation**: For any given 100 hourly precipitation projections corresponding to daily precipitation projection $\hat{\psi}$, calculate the histogram of hourly event intensities using the previously defined categories. Randomly select a set of 10 events from the category to use in SWMM simulations. If a category has fewer than 10 events, use all in the simulation. For any chosen event, execute SWMM simulation when no GI practice is placed in any of the watershed sub-catchments, $i \in V$. For any given sub-catchment, calculate the category's corresponding baseline average 'volume-based runoff coefficient,' i.e., the ratio of runoff volume to the precipitation volume [53], using all selected events in the category. Next, for any selected event, execute SWMM simulation when identical GI practice j in level l is placed across all sub-catchments. For any given sub-catchment i, calculate the category's corresponding average 'runoff coefficient' with respect to the placed GI practice of type j with level l, using all selected events in the category. Follow the procedure for all $\hat{\psi} \in \hat{\Psi}$ and calculate all runoff coefficients.

- **Estimating Runoff**: Given an hourly precipitation projection $\psi \in \Psi$ for sub-catchment $i \in V$, use the *expanded rational method* [53] to calculate the baseline runoff using the corresponding baseline runoff coefficients of the corresponding daily precipitation projection $\hat{\psi} \in \hat{\Psi}$. That is, for any given rainy day in the projection $\psi \in \Psi$, calculate the total daily runoff by multiplying the runoff volume by the runoff coefficient that corresponds to the precipitation intensity in that day, obtained from the corresponding $\hat{\psi} \in \hat{\Psi}$. The overall yearly baseline runoff for the hourly precipitation projection ψ over sub-catchment i, i.e., $Q_i^{\psi,t}$ for all $0 \leq t \leq T$, is the summation of calculated total daily runoff volumes in that year. Use the same method to calculate the overall runoff for sub-catchment $i \in V$ with respect to placed GI practice $j \in G$ in level $l \in L$ in year $t \leq T$. Let $\tilde{Q}_{i,l,j}^{\psi,t}$ denote the surface runoff over sub-catchment $i \in V$ given that GI practice of type $j \in G$ is installed in level $l \in L$ within the

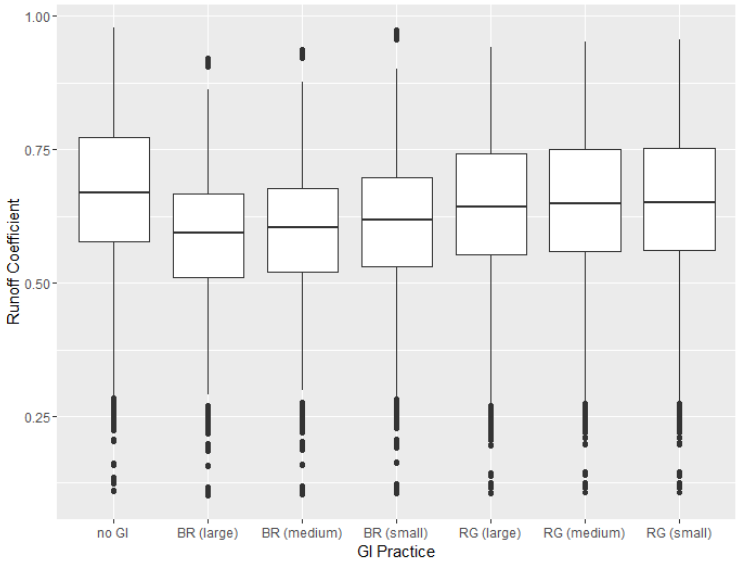

Figure 1.6: Boxplots of the estimated volume-based runoff coefficients for all sub-catchments in the watershed of interest, for all given hourly precipitation projections corresponding to CGCM ACCESS over the years 2018–2050, under various GI practice installation. BR and RG stand for bioretention and rain garden, respectively. The three levels of installation are described in Section 2.2.1.

sub-catchment under the hourly precipitation projection $\psi \in \Psi$ in year t. Hence, under hourly precipitation projection ψ, the corresponding surface runoff captured by the GI practice, $\hat{Q}^{\psi,t}_{i,j,l}$, for all $i \in V, j \in G, l \in L, 0 \leq t \leq T$ is obtained as follows: $\hat{Q}^{\psi,t}_{i,j,l} = Q^{\psi,t}_{i} - \tilde{Q}^{\psi,t}_{i,l,j}$. Repeat this process for all 1,000 hourly precipitation projections to estimate the corresponding surface runoff volumes $Q^{\psi,t}_{i}$ and $\hat{Q}^{\psi,t}_{i,j,l}$ for all sub-catchments in the watershed.

Figure 1.6 presents the variation in the estimated volume-based runoff coefficient across all sub-catchments in our watershed of interest under various GI practice installation. BR and RG stand for bioretention and rain garden, respectively, and the three levels of installation are described in Section 2.2.1. As seen in the figure, the runoff coefficient is generally lower after installing GI practices, compared with the the baseline (i.e., no treatment). In addition, bioretention generally have a lower runoff coefficient, hence present a better performance in reducing runoff compared with rain gardens. Lastly, the larger the GI practice, especially in bioretentions, the higher the performance.

Calculating confidence intervals for runoff volumes for the robust model. Recall that the robust model requires the $100(1-\alpha)\%$ CIs for surface runoff for any given GI practice in any given sub-catchment. We use bootstrapping to generate these intervals [42]. In contrast to stochastic model in which we use a total of 1,000 hourly precipitation projections as scenarios to estimate the corresponding surface runoff volumes, in the robust model we redefine scenarios to be the aggregate measure of 100 hourly precipitation projections

produced from any given CGCM. We then use these scenarios to estimate the runoff volumes as follows.

First, we follow the first three steps in the procedure used for calibrating the stochastic model, i.e., Initialization, Aggregation, and SWMM Simulation, from which we obtain volume-based runoff coefficients for all identified categories for any given CGCM. Next, for each CGCM, we group all volume-based runoff coefficients regardless of the categories and use bootstrapping to replicate large enough bootstrap samples to calculate the corresponding $100(1-\alpha)\%$ CI for runoff coefficients for each sub-catchment, under all GI practice placement (and no treatment). Finally, we use the center and half of the width of each CI in the rational method to estimate the corresponding $\hat{Q}_{i,l,j}^{\psi,t}$ and $\hat{q}_{i,l,j}^{\psi,t}(\alpha)$.

Calculating runoff adjustment factor, $\beta_{i,j,l}^{i',j',l'}$. As discussed in Section 2.1, the surface runoff volume over a downstream sub-catchment is not only a function of the amount of precipitation, the sub-catchment's hydorlogical characteristics, and the placed GI practices within the sub-catchment, but also it is affected by (large-scale) GI practices placed within upstream sub-catchment(s) that are hydrologically connected to this downstream sub-catchment. Also recall that we assume the adjustments over downstream sub-catchments are additive when large-sale GI practices are placed within more than one of its upstream sub-catchments. As discussed earlier in this section, our main SWMM simulation model for the watershed does not capture the entire hydrological connectivity among sub-catchments. Hence, to be able to estimate the adjustment factors, we develop a complementary SWMM model, which we calibrate based on the characteristics of the sub-catchments and their hydrological connectivity in the watershed.

Specifically, we develop a SWMM model that consists of two hydrologically connected sub-catchments, where the residual runoff from the upstream sub-catchment flows onto the downstream sub-catchment. We run the simulation for any given pairs of sub-catchment characteristics to estimate the adjustment factor, $\beta_{i,j,l}^{i',j',l'}$, under various GI practice placements as well as no treatment. To further reduce the computation time, we only use the most important sub-catchment characteristics related to runoff reduction, as identified in the literature [35], and stratify sub-catchments accordingly (see Appendix D for details).

Figure 1.7 presents a subset of the estimated runoff adjustment factors over the downstream sub-catchment, where a large bioretention is placed in upstream and a small bioretention is placed downstream, for all observed combinations of sub-catchment characteristics as described in Appendix D. Sub-catchment characteristics are shown as tuples, where the three elements correspond to percent of imperviousness, percent of slope, and Manning's n for overland flow over the pervious portion of the sub-catchment, each of which are categorized into three levels of 0-2, encoding low, medium, and high, respectively. As seen in the figure, the adjustment factor varies based on the characteristics of the pair of sub-catchments, ranging between 0.2% and 0.75%. In general, a higher level of

imperviousness results in a larger amount of runoff. Hence, when the upstream sub-catchment has a higher level of imperviousness, it contributes a larger amount of flow onto hydrologically connected downstream sub-catchments. As a result, and as seen in the figure, placing a large bioretention in a highly impervious upstream sub-catchment contributes to a larger adjustment in runoff over the downstream sub-catchment. In contrast, the impact is less pronounced when the upstream sub-catchment is relatively pervious.

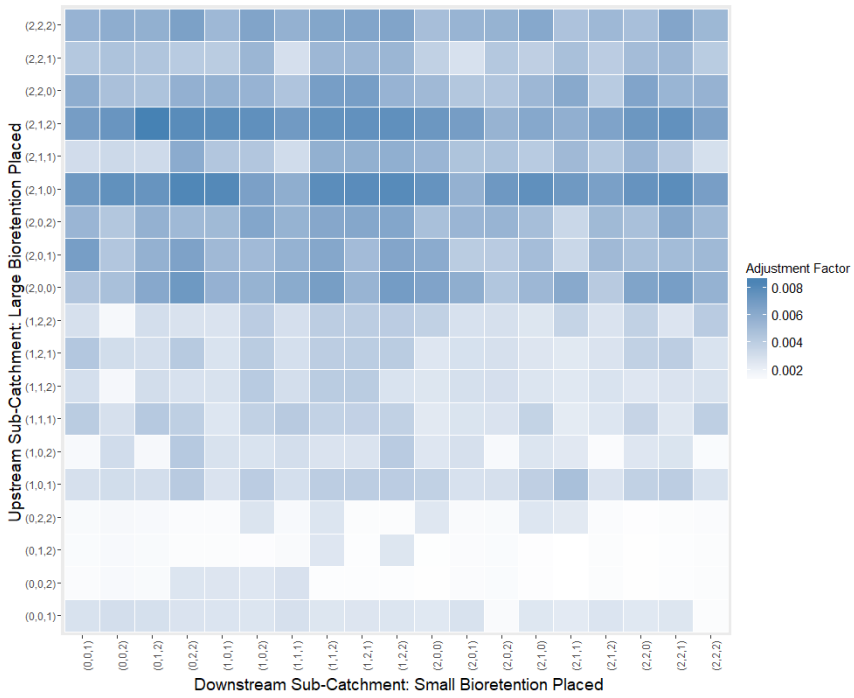

Figure 1.7: Heat map of the runoff adjustment factors over the downstream sub-catchment, where a large bioretention is placed in upstream and a small bioretention is placed downstream, for all observed combinations of sub-catchment characteristics' categories as described in Appendix D. Sub-catchment characteristics are shown as tuples, where the three elements correspond to percent of imperviousness, percent of slope, and Manning's n for overland flow over the pervious portion of the sub-catchment, each of which are categorized into three levels of 0-2, encoding low, medium, and high, respectively.

1.3 Case Study

In this section, we first conduct a case study for a watershed in a mid-sized city in the U.S. We then conduct sensitivity analysis, investigate the relationship between 1-neighbor constraint and the runoff adjustment factor, discuss the findings and provide insights on the implications of our modeling approaches.

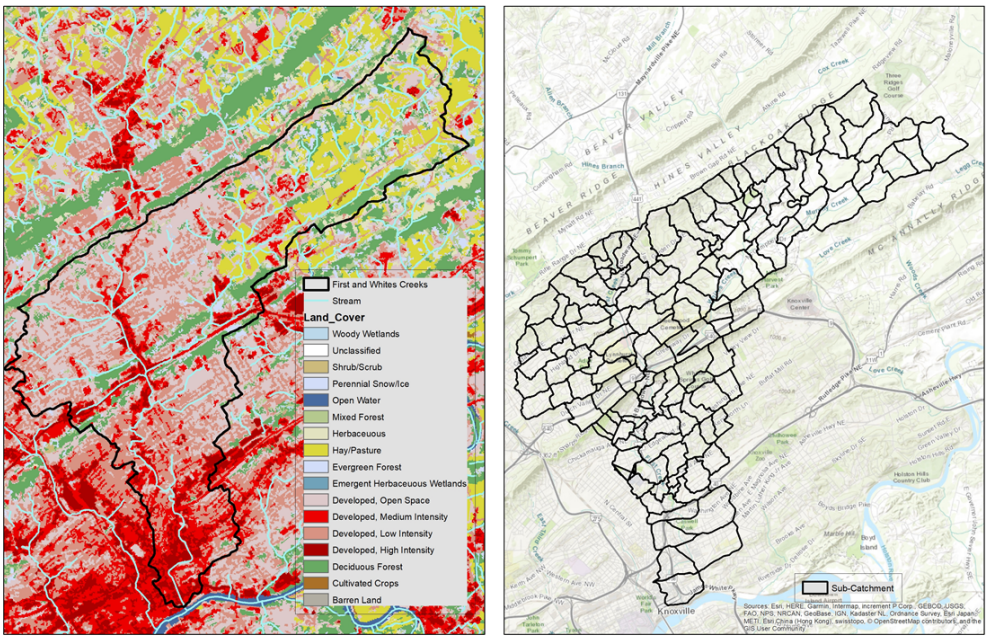

Figure 1.8: Map of land cover [57] (left panel) and hydrological sub-catchments (right panel) of the First Creek, Knoxville, Tennessee.

1.3.1 Case Study Specifications

As a case study, we consider the First Creek in the City of Knoxville, Tennessee. The creek is located entirely within the City of Knoxville and have been identified as the principal sources of flooding in Knox County, Tennessee [21]. The watershed's combined area is 14,805 acres and encompasses parts of the most densely populated regions of the city, including Downtown Knoxville. The First and Whites Creek hydrological model was provided to us by the Stormwater Engineering Division of the City [80]. The hydrological model divides the creek into 140 sub-catchments, all of which are associated with one rain gauge. This model also includes 365 junction nodes and 439 conduit links that direct the flow into the Tennessee River.

Figure 1.8 illustrates the map of land cover (left panel) and hydrological sub-catchments (right panel) of the First Creek. The red shades on the left panel represent level of development, from low (mostly meadow and forest land cover) to high. As seen in the figure, the southern region of the watershed, which is where Downtown Knoxville is located, is highly developed. Subsequently, this dense region has larger amount of impervious area, compared with other regions in the watershed (see Appendix E for more details).

In this case study, we use a planning horizon of length 32 years ($T = 32$), i.e., for years 2018–2050, for which the precipitation projections are available. We let the first and second stage decision variables be taken in the beginning of the planning horizon, i.e., in year 2018, and 10 years into the planning horizon, i.e., in year 2027 ($\overline{T} = 10$).

1.3.2 Computational Results of the Case Study

In this section, we present the computational results of the case study for both stochastic and robust models. In all computational experiments, we use the IBM ILOG CPLEX 12.8 (64-bit edition) on a PC running Microsoft Windows 7 (64-bit edition) with a Core i7, 4 GHz processor and 32 Gigabyte of RAM. Overall, solving the stochastic model is much harder than the robust model. On average, the computational time of the stochastic model is on the order of 3-4 minutes, which is approximately two orders of magnitude larger than that of the robust model, which takes on the order of 1 second to solve.

Given the 1,000 generated hourly precipitation time projections discussed in Section 1.2.3, we let Φ_S denote the projected total expected runoff volume, i.e., $\Phi_S = \sum_{\psi\in\Psi} \pi^\psi \cdot \sum_{i\in V} \sum_{0\leq t\leq T} Q_i^{\psi,t}$, over the First Creek equals 4.57×10^{11} gallons. This volume is used as total baseline surface runoff under no treatment (i.e., no GI practice placed) in the stochastic model. For the robust model, given the same projections that are aggregated regardless of their daily precipitation intensity categories, we let Φ_{R_α} denote the sample average baseline surface runoff under $100(1-\alpha)\%$ confidence level. Accordingly, for the given scenarios, $\Phi_{R_{0.05}} = 4.56\times10^{11}$ and the estimated 95% CI for the expected baseline surface runoff equals $4.56\times10^{11}\pm1.75\times10^{9}$ Note that these runoff volumes correspond to no treatment (i.e., no GI practice placed) in the robust model. In our computational results, we report the percentage reduction in total expected runoff volume under the optimal GI practice placement across the watershed, i.e., $(\Phi_S - \phi_S(\chi_S^\star))/\Phi_S$ and $(\Phi_{R_{0.05}} - \phi_{R_{0.05}}(\chi_R^\star))/\Phi_{R_{0.05}}$ for the stochastic and robust models, respectively.

First, we solve the models under the available budgets of 10, 20, and 50 million dollars and compare the corresponding optimal GI practice placements. Figure 1.9 presents the first stage decision variables under the optimal solution for all cases considered. That is, it presents the sub-catchments in which bioretentions are placed and their level of installation. In addition, Table 1.1 summarizes the second stage decision variables under the optimal solution for all cases considered. That is, it presents the average percentages of sub-catchments in which rain gardens are placed, along with the distribution of their level of installation, given that the scenarios from one of the ten CGCMs are realized. As expected, and seen in the figure and table, as the available budget increases, a larger number of sub-catchments are selected for bioretention installation in the first stage and the sizes of placed rain gardens stochastically increases in the second stage, under all scenarios generated from the ten CGCMs.

Recall that the unit construction cost of bioretentions (and rain gardens) are equal across all sub-catchments. However, as discussed in Section 2.2.1, the area used for bioretention installation corresponds to the level of imperviousness in that sub-catchment, i.e., a fixed percentage of the impervious area of the sub-catchment is treated with bioretentions.

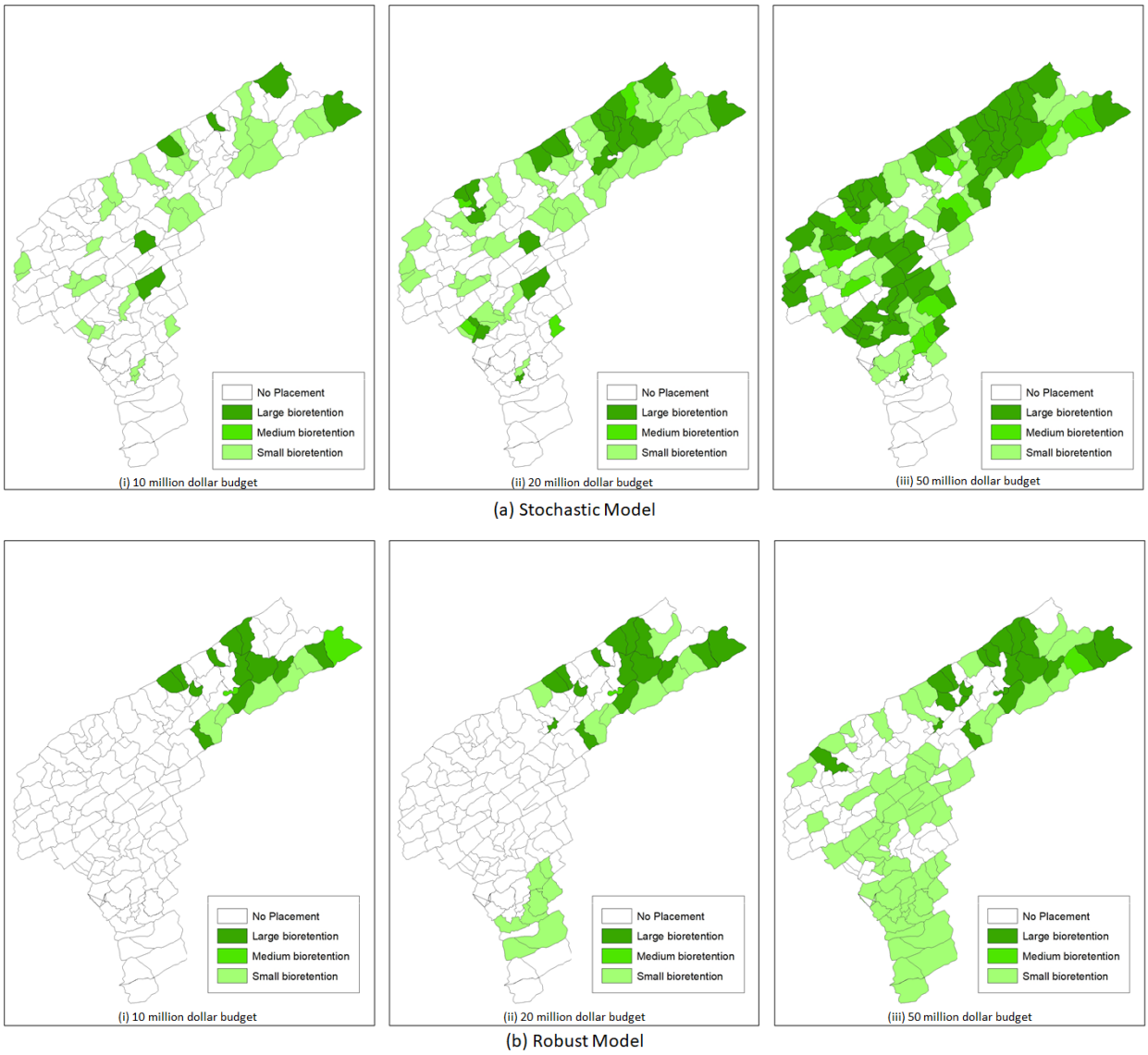

Figure 1.9: Map of placed bioretentions and their level of installation (first stage decision variables) under the optimal solution for various levels of available budget.

As a result, it is much more expensive to place bioretentions in highly developed sub-catchments as the level of development largely correlates with the level of imperviousness. Therefore, in Figure 1.9, sub-catchments in the southern region of the watershed, where the highly developed Downtown Knoxville is located, are not usually selected for bioretention installation. Indeed, as seen in the figure, given a limited budget, e.g., 10 million dollars, solutions to both stochastic and robust models consist of placing bioretention in sub-catchments with a low level of development, where the construction cost is generally lower. Recall that in both models, the 1-neighbor constraint on first stage decision variables ensures connectivity among large-scale GI practice placements. Therefore, if a highly developed sub-catchment is selected as part of the first stage decisions, the available budget should be enough to cover the costs of placing bioretentions not only in that sub-catchment, but also in at least one of its hydrologically connected neighbors. This, in turn, makes placing bioretentions in general very costly within highly developed regions, e.g., sub-catchments in the southern region of the watershed, where the highly developed Downtown Knoxville is located. Accordingly, only as the amount of available budget increases, it becomes optimal to place bioretentions in some of the more developed sub-catchments. It is interesting to note that some of the placed bioretentions in Figure 1.9 are stand-alone. Note that this does not violate the 1-neighbor constraint as these sub-catchments are not downstream to any of their neighboring sub-catchments, i.e, they have no upstream hydrologically connected sub-catchments and hence, 1-neighbor connectivity constraint does not apply to them.

As seen in Figure 1.9, the solutions to the stochastic and robust models are not necessarily identical under the given available budget; however, comparing the results shows similar reduction in total expected surface runoff over the planning horizon for the two models. For instance, under 50 million dollars available budget, the optimal GI placement contributes to 12.01% and 11.30% reduction in total expected runoff for the stochastic and robust models, respectively. It is interesting to note that these reductions are achieved under different allocations of budget in the first and second stages under the two models. Specifically, in the stochastic model, the percentages of budget spent in the first stage are 81%, 88%, and 97% under 10, 20, and 50 million dollars available budget, respectively. Compare these percentages, respectively, with 64%, 82%, and 96% spent in the first stage in the robust model. This suggests a slightly more conservative allocation of budget in the first stage under the robust model, compared with the stochastic model, especially when the available budget is relatively low. As seen in Table 1.1, this relatively conservative allocation of budget is compensated for in the second stage, where the average percentages of sub-catchments in which GI practices are placed is generally larger under the robust model, compared with the stochastic model, across the scenarios generated from the ten CGCMs.

Next, we more extensively compare the percentage reduction in total expected runoff under the stochastic and robust models. Figure 1.10 presents a comparison of percentage reduction in total expected runoff under stochastic and robust models, where the available

Table 1.1: Average percentages of sub-catchments in which rain gardens are placed, along with the distribution of their level of installation, given that the scenarios from one of the ten CGCMs are realized (second stage decision variables) for various levels of available budget.

		Budget											
		10 million dollars				20 million dollars				50 million dollars			
Model	CGCM	% of Sub-Cat.	Large	Med.	Small	% of Sub-Cat.	Large	Med.	Small	% of Sub-Cat.	Large	Med.	Small
Stochastic	ACCESS	64.1%	11.1%	3.8%	85.1%	47.6%	54.4%	4.2%	41.4%	23.6%	67.3%	10.8%	21.9%
	BCC	63.7%	11.0%	4.8%	84.3%	49.7%	45.3%	13.4%	41.3%	23.4%	70.4%	6.9%	22.7%
	CCSM4	67.1%	9.6%	1.0%	89.4%	53.1%	40.0%	10.9%	49.1%	23.4%	69.6%	7.6%	22.8%
	CMCC	66.4%	5.5%	10.5%	84.0%	52.9%	39.1%	13.7%	47.2%	23.6%	72.7%	0.0%	27.3%
	FGOALS	67.0%	9.6%	1.3%	89.1%	51.0%	46.1%	6.4%	47.4%	23.6%	69.7%	6.1%	24.2%
	GFDL	57.2%	16.1%	8.9%	75.0%	52.8%	39.3%	13.4%	47.3%	23.6%	70.2%	5.1%	24.7%
	IPSL	65.4%	8.2%	7.0%	84.8%	49.2%	49.0%	8.1%	42.9%	22.9%	74.4%	4.3%	21.4%
	MPI	66.3%	5.2%	11.3%	83.5%	47.0%	53.5%	8.9%	37.6%	22.9%	74.9%	3.3%	21.8%
	MRI	54.4%	20.6%	7.3%	72.2%	45.7%	56.3%	9.4%	34.4%	22.6%	68.6%	19.3%	12.1%
	NorESM	68.2%	7.3%	3.6%	89.1%	52.3%	42.3%	9.3%	48.4%	23.1%	71.2%	8.3%	20.6%
Robust	ACCESS	80.0%	43.8%	9.8%	46.4%	72.9%	51.0%	7.8%	41.2%	40.0%	60.7%	8.9%	30.4%
	BCC	82.1%	43.5%	5.2%	51.3%	74.3%	51.0%	3.8%	45.2%	40.7%	61.4%	3.5%	35.1%
	CCSM4	81.4%	42.1%	9.6%	48.2%	73.6%	49.5%	8.7%	41.7%	38.6%	66.7%	5.6%	27.8%
	CMCC	84.3%	39.8%	7.6%	52.5%	75.7%	47.2%	7.5%	45.3%	40.7%	59.6%	7.0%	33.3%
	FGOALS	84.3%	41.5%	4.2%	54.2%	77.1%	46.3%	5.6%	48.1%	40.7%	63.2%	0.0%	36.8%
	GFDL	82.9%	41.4%	7.8%	50.9%	74.3%	49.0%	7.7%	43.3%	39.3%	63.6%	7.3%	29.1%
	IPSL	83.6%	41.0%	6.8%	52.1%	75.7%	48.1%	5.7%	46.2%	40.7%	61.4%	3.5%	35.1%
	MPI	82.9%	38.8%	12.9%	48.3%	74.3%	47.1%	11.5%	41.3%	39.3%	63.6%	7.3%	29.1%
	MRI	83.6%	40.2%	8.5%	51.3%	75.0%	47.6%	8.6%	43.8%	40.0%	62.5%	5.4%	32.1%
	NorESM	82.9%	40.5%	9.5%	50.0%	74.3%	47.1%	11.5%	41.3%	40.0%	60.7%	8.9%	30.4%

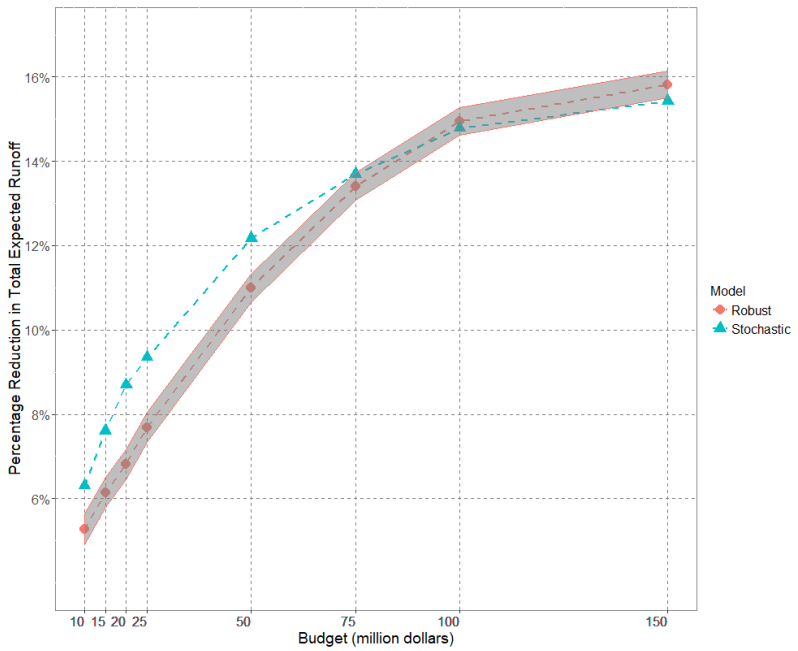

Figure 1.10: Comparison of percentage reduction in total expected runoff under stochastic and robust models, where the available budget ranges between 10 and 150 million dollars. The shaded area represents the 95% CI for reduction in total expected runoff for the robust model.

budget ranges between 10 and 150 million dollars. The shaded area represents the 95% CI for reduction in total expected runoff for the robust model. Note that in reporting the percentage reduction in total expected runoff under the robust model, we use the center of the 95% CI, as depicted in the figure. As seen in Figure 1.10, the stochastic model performs relatively better under lower budgets, i.e., 75 million dollars or less. For instance, under 20 million dollars available budget, the stochastic model outperforms the robust model by 1.89% reduction in total expected runoff. However, this difference between the objective values decreases in the amount of available budget and at higher budgets, the robust model performs relatively better than the stochastic model. For instance, under 150 million dollars available budget, the robust model outperforms the stochastic model by 0.39% reduction in total expected runoff.

1.3.3 Sensitivity Analyses

In this section, we examine the sensitivity of the solutions with respect to some of the important calibrated parameters, including the years to realize a scenario, $\overline{T}$, the ratio of maintenance cost to construction cost, ρ, and the runoff adjustment factor, $\beta_{i,j,l}^{i',j',l'}$. In all cases, we conduct the sensitivity analysis under a wide range of available budgets.

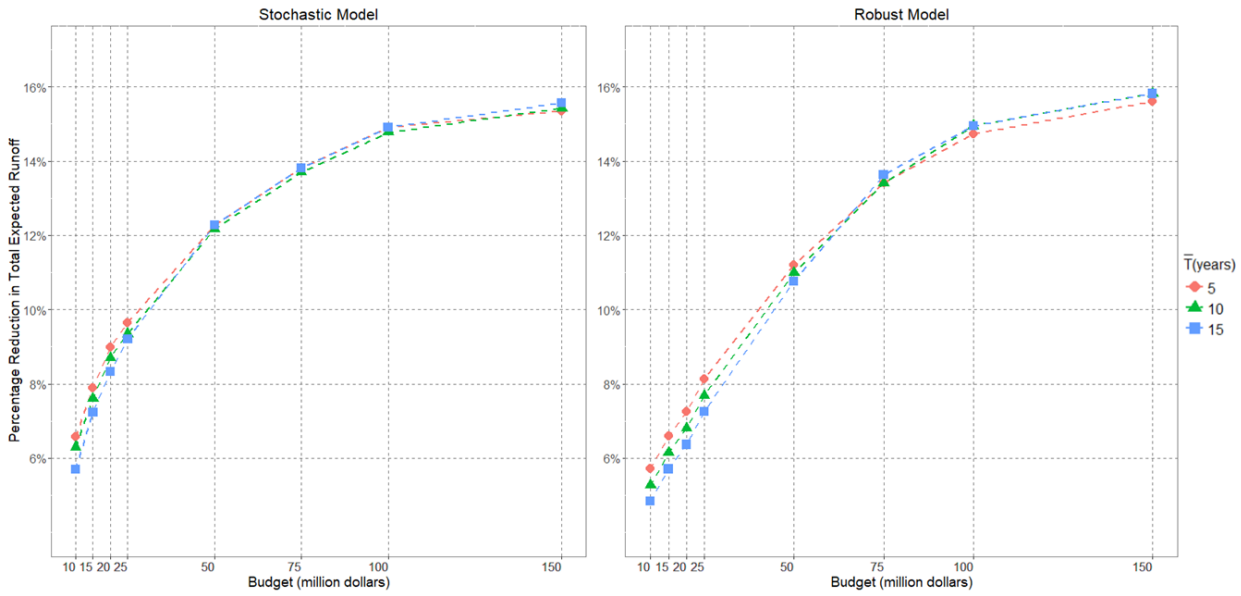

Figure 1.11: Percentage reduction in total expected runoff for $\overline{T} = 5$, 10, and 15 years, where the available budget ranges between 10 and 150 million dollars.

First, we perform sensitivity analysis on the years to realize a scenario, $\overline{T}$, under different budget limitations. Figure 1.11 shows the percentage reduction in total expected runoff for $\overline{T} = 5$, 10, and 15, with the total available budget ranging between 10 and 150 million dollars. As seen in the figure, realizing a scenario sooner, i.e., smaller values of $\overline{T}$, results in a larger reduction in total expected runoff. However, the differences among the percentage runoff reductions for the three cases is low, and decreases in the amount of available budget. For instance, given 10 million dollars available budget, the maximum difference among the percentage runoff reductions equals 1.00% and 0.85% for the stochastic and robust models, respectively. This maximum difference decreases to almost zero for budgets larger than 50 million dollars for both models. This is mainly because under a large enough available budget, large-scale bioretentions are placed within almost all sub-catchments in the first stage. Therefore, because at most one type of GI practice can be placed within any given sub-catchment, there would be few vacant sub-catchments in which rain gardens can be placed after realizing a scenario at time $\overline{T}$ in the second stage. This, in turn, decreases the impact of second stage decisions, resulting in almost no significant difference between the three cases under larger amounts of budget.

Next, we conduct sensitivity analysis on the amount of maintenance cost. As discussed in Section 2.2.1, we set the annual annual GI maintenance cost equal to 3% of its construction cost, i.e., $\rho = 3\%$. Table 1.2 presents the percentage reduction in total expected runoff under different ratios of maintenance cost to construction cost, ρ, ranging between 1% and 10%, where the available budget ranges between 10 and 150 million dollars. In general, as expected, given any available budget, the percentage runoff reduction non-increases in ρ.

Table 1.2: Percentage reduction in total expected runoff for different ratios of maintenance cost to construction cost, ρ, where the available budget ranges between 10 and 150 million dollars.

Budget (million dollars)	Ratio of maintenance cost to construction cost, ρ							
	Stochastic Model				Robust Model			
	1%	**3%**	6%	10%	1%	**3%**	6%	10%
10	6.26%	**6.22%**	6.16%	6.09%	5.32%	**5.30%**	5.27%	5.23%
15	7.59%	**7.55%**	7.50%	7.43%	6.12%	**6.09%**	6.05%	6.00%
20	8.61%	**8.56%**	8.51%	8.44%	6.90%	**6.87%**	6.81%	6.75%
25	9.36%	**9.33%**	9.26%	9.20%	7.68%	**7.64%**	7.58%	7.49%
50	12.18%	**12.14%**	12.07%	11.98%	11.04%	**10.97%**	10.88%	10.76%
75	13.75%	**13.71%**	13.65%	13.58%	13.52%	**13.47%**	13.37%	13.25%
100	14.78%	**14.74%**	14.69%	14.62%	14.90%	**14.86%**	14.79%	14.70%
150	15.37%	**15.37%**	15.37%	15.37%	15.77%	**15.77%**	15.77%	15.77%

This is because as ρ increases, a larger portion of the budget must be allocated to maintain the GI practices to be placed. For instance, under 10 million dollars available budget, in the stochastic model, the runoff reduction decreases by 0.13% when ρ increases from 3% to 10%. Similarly, in the robust model, the corresponding runoff reduction decrease equals to 0.07%. Note that for large amounts of available budget, i.e., under 150 million dollars available budget, changing ρ no longer impacts the solution as the available budget is high enough that covers all construction and maintenance costs.

Finally, we conduct sensitivity analysis with respect to the value of runoff adjustment factor, $\beta_{i,j,l}^{i',j',l'}$. Table 1.3 presents the percentage reduction in total expected runoff under the estimated adjustment factors, no adjustment, and where the estimated adjustment is modified by 50%, where the available budget ranges between 10 and 150 million dollars. As seen in the table, connectivity, captured through adjustment factors, contributes to up to 0.16% and 0.27% reduction in total runoff under various available budgets for the stochastic and robust models, respectively. Also, note that connectivity contributes to a higher percentage of runoff reduction under higher levels of available budget. This is mainly because in such cases, a larger number of bioretentions are placed across the watershed, which potentially results in a higher number of pairs of hydrologically connected sub-catchments.

1.3.4 Investigating the Relationship Between 1-Neighbor Constraint and the Runoff Adjustment Factor

In this section, we evaluate the importance of accounting for the adjustment in surface runoff reduction due to GI placements in connected sub-catchments using the runoff adjustment factor, $\beta_{i,j,l}^{i',j',l'}$. In this section, for simplicity of notation, we let $\beta = [\beta_{i,j,l}^{i',j',l'}]$ denote the vector of all adjustment factors. Specifically, we evaluate the expected opportunity loss due to installing a potentially sub-optimal solution as a result of not accounting for runoff

Table 1.3: Percentage reduction in total expected runoff under different levels of runoff adjustment factors, where the available budget ranges between 10 and 150 million dollars.

Budget (million dollars)	Levels of Runoff Adjustment Factor							
	Stochastic Model				Robust Model			
	No Adj.	-50%	**Estimated Adj.**	50%	No Adj.	-50%	**Estimated Adj.**	50%
10	6.16%	6.19%	**6.22%**	6.25%	5.25%	5.28%	**5.30%**	5.32%
15	7.46%	7.50%	**7.55%**	7.60%	6.03%	6.06%	**6.09%**	6.12%
20	8.46%	8.51%	**8.56%**	8.62%	6.78%	6.83%	**6.87%**	6.92%
25	9.21%	9.26%	**9.33%**	9.39%	7.53%	7.59%	**7.64%**	7.70%
50	11.99%	12.07%	**12.14%**	12.22%	10.78%	10.88%	**10.97%**	11.07%
75	13.57%	13.64%	**13.71%**	13.78%	13.20%	13.33%	**13.47%**	13.60%
100	14.58%	14.66%	**14.74%**	14.83%	14.64%	14.74%	**14.86%**	14.97%
150	15.22%	15.30%	**15.37%**	15.45%	15.60%	15.69%	**15.77%**	15.85%

adjustment factors. We conduct the analyses with and without considering the 1-neighbor constraint to draw insights.

First, let $\tilde{\chi}_S \in \Omega$ denote the optimal solution of the stochastic model, where all adjustment factors are set to zero in the stochastic objective function (1.1), i.e., $\tilde{\chi}_S = _{\chi\in\Omega} \phi_S(\chi; \beta = 0)$. Similarly, let $\tilde{\chi}_R \in \Omega$ denote the optimal solution of the robust model, where all adjustment factors are set to zero in the robust objective function (1.10), i.e., $\tilde{\chi}_R = _{\chi\in\Omega} \phi_R(\chi; \beta = 0)$. Recall that $\chi^\star_S$ and $\chi^\star_R$ denote the optimal solutions of the stochastic and robust models under the estimated values for the adjustment factors, respectively. Clearly, $\phi_S(\chi^\star_S) \leq \phi_S(\tilde{\chi}_S)$ and $\phi_R(\chi^\star_R) \leq \phi_R(\tilde{\chi}_R)$, where the equalities respectively hold when $\tilde{\chi}_S$ and $\tilde{\chi}_R$ are optimal solutionsto the calibrated models with $\beta \geq 0$.

It is interesting to note that per our numerical experiments, both $\tilde{\chi}_S$ and $\tilde{\chi}_R$ are indeed optimal solutions to their corresponding models, i.e., $\phi_S(\tilde{\chi}_S) = \phi_S(\chi^\star_S)$ and $\phi_{R_{0.05}}(\tilde{\chi}_R) = \phi_{R_{0.05}}(\chi^\star_R)$, where $\chi^\star_S, \tilde{\chi}_R \in \Omega$, specifically when the 1-neighbor constraint is included in the models. Our intuition is that because 1-neighbor connectivity constraint (1.7) enforces placing large-scale GI practices (first-stage decisions) in hydrologically connected sub-catchments, it protects the solution to remain optimal, regardless of accounting for adjustment factors. Note that in our watershed of interest, there are only five sub-catchments (out of a total of 140 sub-catchments) that have more than one upstream sub-catchments; these five sub-catchments each have exactly two upstream sub-catchments. Suprisingly, for all these five sub-catchments, the adjustment factors of the two upstream sub-catchments are rather identical. This further reduces the importance of including the exact adjustment factors in the model. Hence, we conclude that for our watershed of interest, given the structure of the corresponding graph $\mathbb{G}(V, A)$, enforcing the 1-neighbor constraint (1.7) is enough to obtain the optimal solution, contributing to a dramatic reduction in calibration efforts.

To further verify this hypothesis, we replicate the analysis without accounting for the 1-neighbor constraint (1.7). Specifically, we let χ'_S and $\bar{\chi}_S$ denote the optimal solutions to

the stochastic model under the estimated values for the adjustment factors and where all adjustment factors are zero, respectively, when relaxing the 1-neighbor constraint (1.7), i.e.,

$$\chi'_S =_{\chi\in\Omega\setminus\{(1.7)\}} \phi_S(\chi),\ \bar{\chi}_S =_{\chi\in\Omega\setminus\{(1.7)\}} \phi_S(\chi;\beta=0).$$

Clearly, $\phi_S(\chi'_S) \leq \phi_S(\bar{\chi}_S)$, where the equality holds when $\bar{\chi}_S$ is an optimal solution to the calibrated model with $\beta \geq 0$ when relaxing the 1-neighbor constraint (1.7). Analogously, we let χ'_R and $\bar{\chi}_R$ denote the optimal solutions to the robust model under the estimated values for the adjustment factors and where all adjustment factors are set to zero, respectively, when relaxing the 1-neighbor constraint (1.7). Hence, similar to the stochastic model, for the robust model we have $\phi_{R_{0.05}}(\chi'_R) \leq \phi_{R_{0.05}}(\bar{\chi}_R)$, where the equality holds when $\bar{\chi}_R$ is an optimal solution to the calibrated model with $\beta \geq 0$ when relaxing the 1-neighbor constraint (1.7). Consistent with our intuition, our numerical experiments show that $\bar{\chi}_S$ and $\bar{\chi}_R$ are indeed sub-optimal solutions to their corresponding stochastic and robust problems, respectively, when relaxing the 1-neighbor constraint, i.e., $\phi_S(\chi'_S) < \phi_S(\bar{\chi}_S)$ and $\phi_{R_{0.05}}(\chi'_R) < \phi_{R_{0.05}}(\bar{\chi}_R)$.

Table 1.4 summarizes the numerical analyses on characterizing the relationship between 1-neighbor constraint (1.7) and the runoff adjustment factor, β, at various levels of available budget. The second through fourth columns show the percentage reduction in total expected runoff volume under $\chi^\star_S$, χ'_S, and $\bar{\chi}_S$, respectively. The fifth through seventh columns show the percentage reduction in total expected runoff volume under $\chi^\star_R$, χ'_R, and $\bar{\chi}_R$, respectively. First note that, as discussed, the percentage reduction in total expected runoff volume under $\tilde{\chi}_S$ and $\tilde{\chi}_R$ are the same as those under $\chi^\star_S$ and $\chi^\star_R$, respectively; hence, they are not included in the table. As seen in the table, the values under χ'_S are larger than those obtained under $\chi^\star_S$. Similarly, the values under χ'_R are larger than those obtained under $\chi^\star_R$. This suggests that, as expected, the 1-neighbor constraint (1.7) is binding under the optimal solutions to both stochastic and robust models with the original feasible set Ω. In addition, as discussed, any difference between the values under χ'_S and $\bar{\chi}_S$, and those under χ'_R and $\bar{\chi}_R$ indicates that $\bar{\chi}_S$ and $\bar{\chi}_R$ are respectively sub-optimal solutions to the stochastic and robust models with the feasible set $\Omega \setminus \{(1.7)\}$. Accordingly, as seen in the table, $\bar{\chi}_S$ is sub-optimal at almost all budget levels, except 20 and 150 million dollars, in the stochastic model, and $\bar{\chi}_R$ is sub-optimal under available budgets of 25 and 50 million dollars in the robust model.

In summary, this analysis show that, given the structure of the underlying graph of sub-catchments $\mathbb{G}(V, A)$ in our study, 1-neighbor (1.7) constraint guarantees the optimality of a solution, regardless of accounting for adjustment factors. This has the potential to dramatically reduce the calibration efforts. However, note that using a set of well-estimated adjustment factors in models result in more accurate estimated values for the corresponding objective functions.

Table 1.4: Percentage reduction in total expected runoff volume under the solutions $\chi^\star_M$, χ'_M, and $\bar{\chi}_M$, $M \in \{S, R\}$, where the budget available ranges between 10 and 150 million dollars. The corresponding values under $\tilde{\chi}_M$ and $\chi^\star_M$, $M \in \{S, R\}$, are the same; hence the former are not included in the table.

	Stochastic			Robust		
Budget (million dollars)	$\frac{\Phi_S - \phi_S(\chi^\star_S)}{\Phi_S}$	$\frac{\Phi_S - \phi_S(\chi'_S)}{\Phi_S}$	$\frac{\Phi_S - \phi_S(\bar{\chi}_S)}{\Phi_S}$	$\frac{\Phi_{R_{0.05}} - \phi_{R_{0.05}}(\chi^\star_R)}{\Phi_{R_{0.05}}}$	$\frac{\Phi_{R_{0.05}} - \phi_{R_{0.05}}(\chi'_R)}{\Phi_{R_{0.05}}}$	$\frac{\Phi_{R_{0.05}} - \phi_{R_{0.05}}(\bar{\chi}_R)}{\Phi_{R_{0.05}}}$
10	6.22%	7.71%	7.16%	5.30%	5.34%	5.34%
15	7.55%	8.41%	8.40%	6.09%	6.16%	6.16%
20	8.56%	9.25%	9.25%	6.87%	6.95%	6.95%
25	9.33%	9.96%	9.95%	7.64%	7.72%	7.70%
50	12.14%	12.41%	12.40%	10.97%	11.01%	10.99%
75	13.71%	13.90%	13.90%	13.47%	13.47%	13.47%
100	14.74%	14.84%	14.83%	14.86%	14.86%	14.86%
150	15.37%	15.37%	15.37%	15.77%	15.77%	15.77%

1.4 Summary and Insights

Climate change threatens to overwhelm stromwater systems across the nation, rendering them ineffective. Green Infrastructure (GI) practices are low cost, low regret strategies that can contribute to urban runoff management. However, questions remain as to how to best distribute GI practices through urban watersheds given precipitation uncertainty and hydrological responses to their installation. In this work, we showcase an approach that can enable city managers to incorporate the complexity and uncertainty of climate projections to make optimized choices for building resiliency into urban systems.

In this study, we developed two-stage stochastic programming and robust programming to determine the optimal placement of GI practices across a set of candidate locations in a watershed to minimize the total expected surface runoff under medium-term precipitation uncertainties, given an available budget. We proposed a novel scenario generation process that allowed us to efficiently evaluate the impact of precipitation on the entire watershed system under various combinations of GI practice placements. We calibrated the model using literature, historical precipitation data, future precipitation projections, and expert opinion and conducted a case study for an urban watershed in the City of Knoxville. We provided computational results and conducted extensive sensitivity analyses. Our results show that the optimal placement of GI practices within our watershed of interest can contribute to up to approximately 9.5% reduction in total expected runoff over the planning horizon, with a limited budget of 25 million dollars. The reduction in total expected runoff obtained by the two modeling approaches are comparable. The two models, however, are quite different with respect to the computational time. That is, the computational time of the stochastic model is approximately two orders of magnitude larger than that of the robust model. This is mainly because of the lower number of scenarios used in the latter approach due to pre-processing

of the precipitation projections, i.e., using CIs for the baseline runoff volume and surface runoff captured by a GI practice in any given sub-catchment, instead of all 100 scenarios per CGCM.

In our models, we accounted for hydrological connectivity in the watershed using an underlying acyclic connectivity graph of sub-catchments. Specifically, we introduced a 1-neighbor connectivity constraint over the graph to ensure that a large-scale GI practice can be placed in a given sub-catchment if there exists at least one large-scale GI practice in one of the sub-catchments that are hydrologically connected to it. In addition, we carefully calibrated the runoff adjustments over pairs of hydrologically connected sub-catchments to more accurately estimate the impact of large-scale GI practices on runoff reduction not only within the sub-catchments in which they are placed, but also in their downstream sub-catchments. Our analysis shows that the 1-neighbor constraint protects the optimality of a solution in our watershed of interest, regardless of accounting for adjustment factors. This is mainly because of the particular structure of the connectivity graph of sub-catchments. More in-depth analysis is needed to establish sufficient conditions under which calibration of runoff adjustment factors is completely unnecessary.

In this study, we only accounted for two relatively similar types of GI practices, i.e., bioretentions and rain gardens. The selected types of GI practices are considered to be very efficient not only in reducing runoff volume, but also in treating stormwater quality. Note that the model developed is very versatile and allows for including more than two types of GI practices. Hence, accordingly, city planners can use the model using a wide array of GI practices to determine the best course of GI practice planning.

In this study, we accounted for future precipitation uncertainty using an array of CGCMs. This enabled us to account for climate change uncertainty when planning GI practices. Although we accounted for precipitation uncertainty, in this study we did not account for population growth and future urban development that can give rise to an increase in impervious area. Additional studies are needed to account for a close-loop system where a more livable city leads to urban population growth, which in turn leads to more runoff.

Lastly, in this study, we only accounted for runoff capturing properties of GI practices. As thoroughly discussed in the literature, GI practices provide a wide array of benefits, e.g., improving water and air quality, contributing to urban aesthetics, etc. Future multi-objective mathematical programs need to be developed to account for all benefits of GI practices when optimizing GI practice placement within an urban watershed.

Chapter 2

Optimizing Dynamic Green Infrastructure Placement in an Urban Watershed Precipitation and Population Growth Uncertainties

Green infrastructure (GI) systems provide a variety of social, economic and ecological benefits [26], well exemplified by several real life implementations. New York City avoided the need to spend $6-$8 billion on new water filtration and treatment plants by instead purchasing and protecting watershed land in the Catskill Mountains for $1.5 billion [19]. Likewise Arnold, Missouri dramatically reduced the cost to taxpayers of disaster relief by purchasing threatened properties and creating a greenway in the flood plain [71, 18]. Portland, Oregon initiated the Grey to Green (G2G) Initiative to better handle approximately ten billion gallons of stormwater runoff annually by expanding the city's green infrastructure. Nagoya, Japan is working to reduce temperatures by increasing the percentage of vegetated from 25 to 40 percent by 2050, installing green roofs, trees, recreational greenways and planting more trees [64, 55]. It is our contention that most GI projects tend to be *reactive* i.e. implemented to counter an existing problem and a more proactive approach in integrating large scale GIs in urban planning could be far more beneficial. To assist policy makers and urban planners decide how many GIs to place, which ones to select from a given set of choices and where to place them, we propose an approach based on three important considerations: stormwater reduction, population proximity and GI connectivity.

GI can be used to mitigate stormwater runoffs in conjunction with conventional stormwater management systems such as sewers, pipes etc. Gray infrastructure used to handle stormwater has has marginal benefits at best - they are costly to build and maintain, reduce on-site absorption of precipitation thereby creating more runoffs and climate change

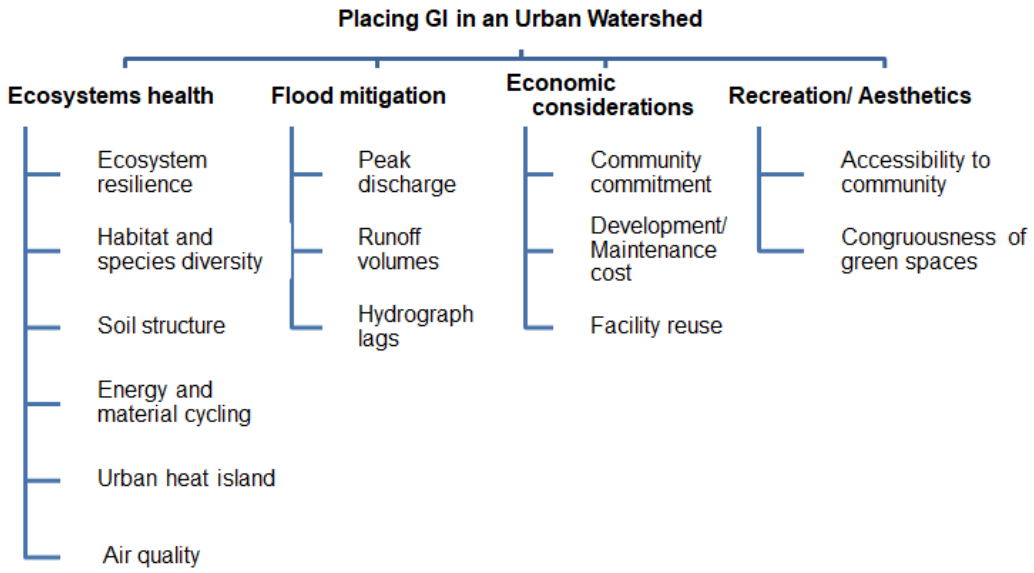

Figure 2.1: Summary of factors that are influenced by/ contribute to placing GI practices in an urban watershed.

considerations only serve to further reduce the actual benefits from these systems. A large drainage and stormwater system built for the current levels of precipitation would lose some of its utility if the precipitation in the region decreases over time and vice-versa. Compared to such single utility of conventional stormwater management systems, GI is a more attractive alternative because in addition to handling runoffs just as well, they can have several other benefits. Our model uses projections of precipitation from several climate models to ascertain the level of GI required to achieve set targets of runoff reduction. To maximize the benefits of GI, it is imperative to place them close to centers of population. They provide cleaner air and water, increased recreational opportunities; improved health and better connection to nature and sense of place [26]. Well placed green space has also been shown to increase property values [97, 98, 99, 78]. They have been shown to reduce surrounding temperatures which is important because an important consequence of urbanization besides stormwater runoff is the creation of 'urban heat islands', i.e., an increase in the local temperatures. The reasons for this are manifold: emissions (from generators, buildings, HVAC units, traffic), reduction in evotranspiration i.e. evaporation of water from the surface, reduction in fauna (trees in specific) and their cooling affect, entrapment of heat from sunlight by buildings, etc. In short, the important factors in placing a GI practice can be categorized into four subjects: (i) flood mitigation, (ii) ecosystem health, (iii) economic consideration, and (iv) recreation/ aesthetics. Figure 2.1 summarizes the factors and sub-factors that are influenced by/ contributes to placing GI in an urban watershed.

To the best of our knowledge, there is no study that encompasses the dynamic nature of the GI placement given uncertainties in population and precipitation projections. As such,

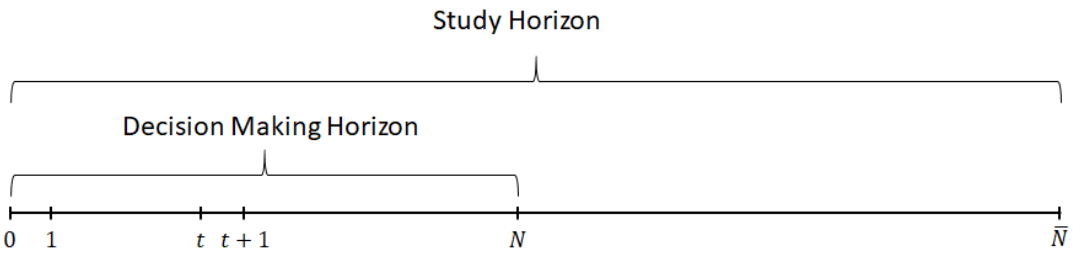

Figure 2.2: A schematic representation of the study horizon. The GI placement intervention occurs in decision making horizon, i.e., first N years of the study horizon.

in this study, our concentration is on developing a mathematical model through which we address the placement of GI practices in an urban watershed, upon future population and precipitation projections, so that the dynamic nature of the problem is efficiently served.

The rest of the chapter is organized as follows. First, we formulate the model in Section 2.1. Next, in Section 2.2, we calibrate our model for a watershed in a mid-size city in the U.S. In Section 2.3, we provide the computational results for our case study and draw insights. Finally, we conclude in Section 3.5.

2.1 Model Formulation

In this section, we develop a finite horizon Markov Decision Process (MDP) model to optimally place GI in an urban watershed. Specifically, we develop a deterministic policy solution in response to uncertainty associated with the population and precipitation projections so that the total expected (discounted) reward placed GI practices across urban watershed is maximized throughout the study horizon. For any given decision making time period and any given location, our deterministic policy prescribes set of actions to be taken with respect to the type and the level of GI practice.

A watershed consist of sub-catchments whose impervious areas candidate locations to place GI practices during the study horizon. Let $\bar{N}$ denote the study horizon where the first N time periods is the decision making horizon. Figure 2.2 shows an schematic concept of the study and decision making horizons. At the beginning of any given decision making time period, we make a decision on placing GI practices. Specifically, we decide on the type, level (size), and the location (sub-catchment) of a GI practice with respect to the available budget and realized population projection. Due to our made decision in a time period, the population and the GI level is subject to uncertainty in the next time period. Also, our decision contributes to runoff volume reduction over the period. Figure 2.3b shows the timeline of events in time period t.

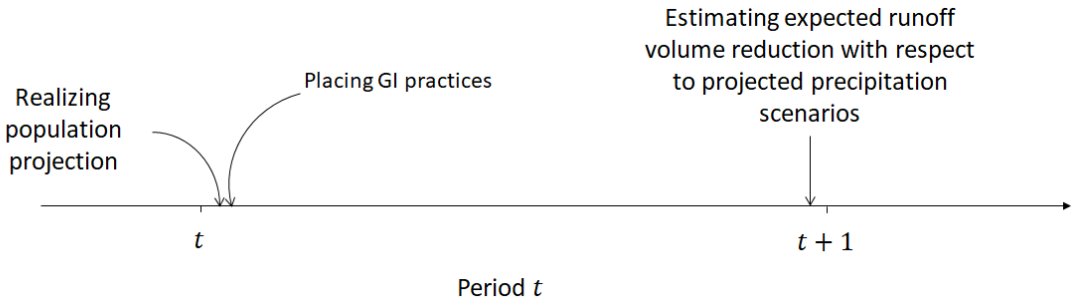

Figure 2.3: Timeline of events in time period t

Let $\Delta = \{1, 2, \ldots, |\Delta|\}$ denote the set of sub-catchments of the watershed of interest. Let $M = \{1, 2, \cdots, |M|\}$ denote the set of projected population levels in any given sub-catchment. Regarding the GI practices, let $G = \{1, 2, \ldots, |G|\}$ denote the set of types of GI practices. Also, let $L = \{1, 2, \ldots, |L|\}$ denote the set of levels associates with the GI types. Accordingly, we let pair $\psi = (g, l) \in G \times L$ denote the GI characteristics within any given sub-catchment. Finally, let $T = \{1, 2, \ldots, N\}$ denote the the decision making horizon.

The state of the process over sub-catchment i at time t is represented by $s_t(i, \mu, \psi)$ where μ and ψ denote the population level and characteristics of GI within the sub-catchment i. We start with s_t at time t. For sub-catchment i in time period t, available actions are defined by finite set $A_{s_t(i,\mu,\psi)} = \{0, 1, 2, .., |A_{s_t(i,\mu,\psi)}|\}$ in which 0 denote the *do nothing* (maintain current GI characteristic), 1 denotes *one-level-upgrade* (upgrade GI characteristics by one level), and so on. Also, we let $a_i \in A_{s_t(i,\mu,\psi)}$ denote the an action available to take in sub-catchment i in time t, to which $c_t((i, \mu, \psi), a_i)$ denotes the corresponding cost.

For any given sub-catchment, transitioning between states from period t to the next period is impacted by actions taken for the sub-catchment. Accordingly, let $p_{a_i}(\psi'|\psi)$ denote the probability of transitioning from GI level ψ to ψ' under action a_i. Also, let $p(\mu'|\mu, \psi)$ the probability that the population level transitions from level μ to μ', given the GI characteristic ψ.

The revenue of placing a GI practice in any sub-catchment in any time period is deemed as a factor of surface runoff volume reduction . Clearly, the main source of the runoff in any time period is the amount of precipitation during the period which is subject to uncertainty. Let the random variable ζ_t denote the amount of precipitation in period t, with the probability mass function p_{ζ_t} and cumulative density function F_{ζ_t}. Accordingly, we let $b_t(i, \mu, \psi, x)$ denote the immediate expected runoff volume reduction for sub-catchment i in period t as a function of population density level, μ, GI characteristics, ψ, and amount of precipitation during the period, ζ.Let $r_t((i, \mu, \psi), a_i)$ denote the immediate expected reward starting from state s_t under action a_i for sub-catchment$i$ which is given by

$$r_t((i,\mu,\psi),a_i) = \sum_{\psi' \in \Psi} \sum_{\zeta \in \zeta_t} p_{a_i}(\psi'|\psi) b_t(i,\mu,\psi',\zeta). \tag{2.1}$$

Let $V_t(s)$ denote the expected discounted reward for GI practice placement in any sub-catchment starting from state s. Also, let $R(i,\mu,\psi)$ denote the lump-sum reward at time T. Note that we assume in period N onward, i.e., the next $\overline{T} - T$ periods, the only action available for the system is to maintain the current level of GI practice across all tracts. Therefore, the Bellman equations for the expected discounted reward of GI practice placement for sub-catchment i, starting from state s across the planning horizon is given by

$$V_t(i,\mu,\psi) = \begin{cases} \max\limits_{a_i} \left\{ r_t((i,\mu,\psi),a_i) + \lambda \sum_{\mu' \in M} \sum_{\psi' \in \Psi} p_{a_i}(\psi'|\psi) p(\mu'|\mu,\psi') V_{t+1}(i,\mu',\psi') \right\}, & t < N, \\ R(i,\mu,\psi) & t = N, \end{cases} \tag{2.2}$$

where λ denotes the discount rate.

We are interested in adding the budget constraint to our policy for placing GI practices within sub-catchments of study across the watershed. Moreover, we are interested in deterministic policy rather than randomized policy. To do so, we opt to work with the dual formulation of our proposed Bellman equations (2.2). Let B denote the overall available budget.Let $\alpha(i,\mu,\psi)$, for all $i \in \Delta, \mu \in M, \psi \in \Psi, t \in T$, denote the initial state for sub-catchment i. Let $x_t((i,\mu,\psi),a_i)$ be a non negative variable denoting the the average number of times we observe the sub-catchment i with state $s_t(i,\mu,\psi)$ and take action a_i, $a_i \in A_{s_t(i,\mu,\psi)}$ at time period t, $t < T$. Analogously, let $y_N(i,\mu,\psi)$ denote the number of times we observe sub-catchment i with state $s_N(i,\mu,\psi)$ and take the default action, e.g, 'do nothing', at time period $t = N$. To enforce deterministic policy, we let $z_t((i,\mu,\psi),a_i)$ be a binary variable corresponding to $x_t((i,\mu,\psi),a_i)$, where it takes positive value if its corresponding variable, $x_t((i,\mu,\psi),a_i)$, is positive. The dual linear programming formulation of Bellman equations (2.2) with budget constraint and enforced deterministic policy solution is given in the following.

$$\begin{aligned} \max \sum_{i \in \Delta} \sum_{\mu \in M} \sum_{\psi \in \Psi} \sum_{t \in T \setminus \{N\}} \sum_{a_i \in A_{s_t(i,\mu,\psi)}} & r_t((i,\mu,\psi),a_i) x_t((i,\mu,\psi),a_i) \\ & + \sum_{i \in \Delta} \sum_{\mu \in M} \sum_{\psi \in \Psi} R(i,\mu,\psi) y_N(i,\mu,\psi) \end{aligned} \tag{2.3a}$$

$$\text{s.t.} \sum_{a_i \in A_{s_t(i,\mu,\psi)}} x_t((i,\mu,\psi),a_i) = \alpha(i,\mu,\psi), \qquad \begin{array}{l}\forall i \in \Delta, \forall \mu \in M,\\ \forall \psi \in \Psi, t = 1,\end{array} \tag{2.3b}$$

$$\begin{array}{l}\displaystyle\sum_{a_i \in A_{s_t(i,\mu,\psi)}} x_t((i,\mu,\psi),a_i)\\ \quad - \lambda \displaystyle\sum_{\mu' \in M} \sum_{\psi' \in \Psi} \sum_{a'_i \in A_{s_t(i,\mu,\psi)}} p_{a'_i}(\psi|\psi') p(\mu|\mu',\psi') x_{t-1}((i,\mu',\psi'),a'_i)\\ \quad = 0,\end{array} \qquad \begin{array}{l}\forall i \in \Delta, \forall \mu \in M,\\ \forall \psi \in \Psi,\\ \forall t \in T \setminus \{1, N\},\end{array} \tag{2.3c}$$

$$\begin{array}{l} y_N(i,\mu,\psi)\\ \quad - \lambda \displaystyle\sum_{\mu' \in M} \sum_{\psi' \in \Psi} \sum_{a'_i \in A_{s_t(i,\mu,\psi)}} p_{a'_i}(\psi|\psi') p(\mu|\mu',\psi') x_{t-1}((i,\mu',\psi'),a'_i)\\ \quad = 0,\end{array} \qquad \begin{array}{l}\forall i \in \Delta, \forall \mu \in M\\ \forall \psi \in \Psi, \forall t = N,\end{array} \tag{2.3d}$$

$$\sum_{i \in \Delta} \sum_{\mu \in M} \sum_{\psi \in \Psi} \sum_{t \in T \setminus \{N\}} \sum_{a_i \in A_{s_t(i,\mu,\psi)}} c_t((i,\mu,\psi),a_i) x_t((i,\mu,\psi),a_i) \leq B \tag{2.3e}$$

$$x_t((i,\mu,\psi),a_i) \leq z_t((i,\mu,\psi),a_i), \qquad \begin{array}{l}\forall i \in \Delta, \forall \mu \in M,\\ \forall \psi \in \Psi,\\ \forall a_i \in A_{s_t(i,\mu,\psi)},\\ \forall t \in T \setminus \{N\},\end{array} \tag{2.3f}$$

$$\sum_{a_i \in A_{s_t(i,\mu,\psi)}} z_t((i,\mu,\psi),a_i) = 1, \qquad \begin{array}{l}\forall i \in \Delta, \forall \mu \in M,\\ \forall \psi \in \Psi,\\ \forall t \in T \setminus \{N\},\end{array} \tag{2.3g}$$

$$z_t((i,\mu,\psi),a_i) \in \{0,1\}, x_t((i,\mu,\psi),a_i), y_N(i,\mu,\psi) \geq 0, \qquad \begin{array}{l}\forall i \in \Delta, \forall \mu \in M,\\ \forall \psi \in \Psi,\\ \forall a_i \in A_{s_t(i,\mu,\psi)},\\ \forall t \in T \setminus \{N\}.\end{array} \tag{2.3h}$$

In dual model (2.3), the constraint set (2.3b)-(2.3d) includes well-know flow conservation constraints, enforcing the balance between the expected flow-in and flow-out for any state and any time period. Constraint (2.3e) enforces the total expected budget not exceed the

allocated budget, B. Constraint (2.3f) enforces enforces the binary variable $z_t((i,\mu,\psi),a_i)$ take the value of one if its corresponding variable $x_t((i,\mu,\psi),a_i)$ is positive. The deterministic policy is guaranteed by constraint (2.3g). Finally, constraint (2.3h) maintains the domains of variables.

2.2 Model Calibration

In this section, we calibrate our mathematical model presented in Section 2.1 using the literature, historical data, precipitation projections, and expert opinion for a given urban watershed of a mid-sized city in the U.S. First, in Section 2.2.1 we calibrate the parameters associated with characteristics of GI practices. Next, we discuss our approach in calculating the population transition probabilities in Section 2.2.2. Next, in Section 2.2.3 we discuss the preprocessing performed on calculating the immediate and lump-sum terminal rewards.

2.2.1 GI Practices

The GI practices are, generally, implemented in an urban watershed to capture surface runoff volumes. Also, their performance in improving the water quality during infiltration is another metric of interest. Hence, depending on the implementation criterion, the GI type and its other design characteristics (such as surface vegetation, surface storage volume, media storage composition, underdrain pipe size, etc.) are calibrated accordingly.

GI type, G ,and implementation levels, L. In this study, we consider urban bioretention areas, $|G|= 1$, implemented in three levels, $|L|= 3$, corresponding to a sub-catchment impervious area, namely low, medium and high levels that respectively treat 5%, 7.5%, and 10% of the sub-catchment impervious area. Hence, overall GI characteristic $|\Psi|= (|G|\times|L|) = 3$. Note that, regardless of the bioretention levels of implementation and their underlying sub-catchments, the remaining design characteristics of the bioretention areas are identical. These design characteristics are adapted from stormwater training manuals from State of Tennessee Department of Environment & Conservation [17].

GI cost, $c_t((i,\mu,\psi),a_i)$. The total cost of placing GI practices includes construction and maintenance costs per square feet which are adjusted based on the level of implementation in any given sub-catchment. For any given time period, depending on the action taken, we accrue a cost. That is, for instance if a low level bioretention area is already implemented in a given sub-catchment, and the policy recommends maintain action, we just accrue the maintenance cost associated with low level bioretention area in the sub-catchment. Otherwise, if the the action is upgrade bu one level to medium bioretention, we accordingly accrue medium level construction cost and maintenance cost associated with that level during the period. Note that, depending on the period at which we, adjust the cost with respect to the present value of money. Hence, $c_t((i,\mu,\psi),a_i)$ represent the value of money at $t = 0$.

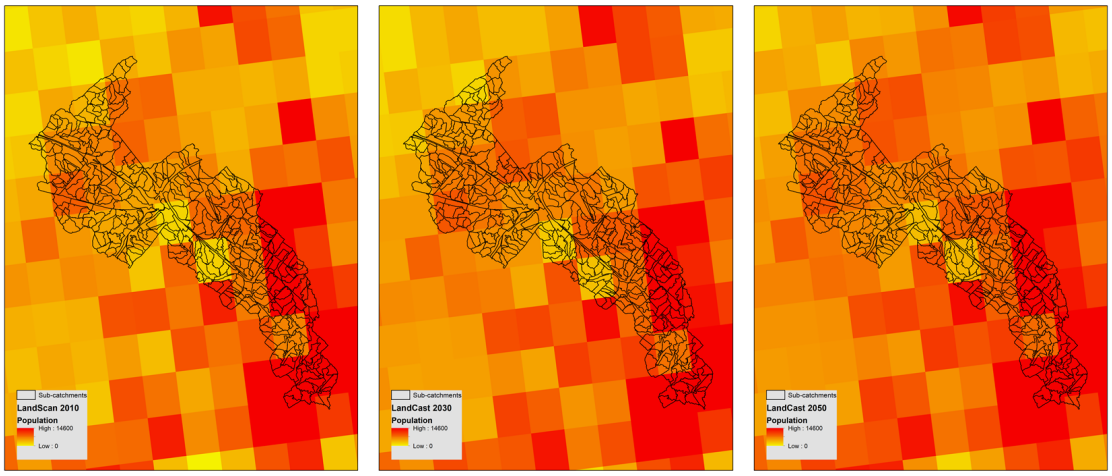

Figure 2.4: Baseline population in year 2010 (left panel) and projection in years 2030 (middle panel) and 2050 (right panel) over Second Creek, Knoxville, Tennessee [74].

We use the inflation-adjusted EPA Opti-Tool [43] to estimate the per square feet cost of implementation for bioretention areas; the value was \$15.46 in 2016. We use the U.S. Labor Department's Consumer Price Index (CPI) inflation calculator [43] to adjust the value. Regarding the annual maintenance cost a bioretention area, we use published reports [92] that recommend 3%-6% of its construction cost. We let ρ denote the ratio of maintenance cost to the construction cost. we use $\rho = 3\%$. Lastly, we use the average annual inflation rate of 1.86%, which equals the average annual U.S. inflation rate over the period 2007-2017 [77].

2.2.2 Population Transition Probabilities

Preparing the population data. We use the land scan data available for the years 2010, 2030, and 2050. Figure 2.4(a) shows the land scan raster data for the mentioned years in Knox County, TN. Note that the raster data resolution is 1:1524, whereas sub-catchments' areas are more granular. Hence, the a population raster cell data may cover more that one sub-catchment (See Figure 2.4(b)). To project raster cell data to a sub-catchment, we break down the raster cell population cells data proportional to the cell area which intersects with a certain sub-catchment. Note that we assume that the raster cell population is equally spread across the cell area.

Simulating stochastic population transition in sub-catchment level, $p(\mu|\mu', \psi')$. We first use the interpolating spline method [49] to interpolate the yearly population data between 2018 and 2050 for any given sub-catchment. For any given year, the resulting interpolated points is then used as a average population in each sub-catchment. Also for the given year t and the given sub-catchment i, we assume that the population has normal distribution with the average equal to the interpolated point, $\hat{d}_t^i$ and a standard deviation σ^i. To estimate the σ^i, we use the standard deviation calculated from three different sources that has projected

Table 2.1: Population in 2010 and projected population in 2030 and 2050 from three different sources for Knox County, Tennessee.

Source	Population 2010	Projected Population 2030	Projected Population 2050
LandScan/LandCast [74]	458,877	518,921	643,226
Boyd Center Tennessee Population Projections: 2016-2070 [29]	-	509,363	578,740
ProximityOne [83]	433,056	501277	564081

Knox County population for the years 2030 and 2050. That is, for any given year, we calculate the standard deviation as a proportion of the average population in that year. Then, we take the average of the three proportions and set it as σ^i, for all $i \in \Delta$.

Next, for any given sub-catchment in any given interpolated year, We generate 1,000 trajectories from these normal distributions. Then, we calculate the population densities out of calculated population trajectories across all sub-catchments. For any given sub-catchment, we stratify population across all sub-catchments into quartiles (extracting first and third quartiles). We then categorize the population below first quartile as 'low', between first and third quartiles as 'medium' and, greater than equal to the third quartile as 'high' population levels. Finally, we calculate the population transitions for any trajectory, for two consecutive years in the planning horizon, by counting the number of times that we observe a transition from one population level to (another) level in the next consecutive year, e.g., ('low','low'), ('low', 'medium'), etc. Figure 2.5 shows our simulation algorithm for estimating stochastic population transition matrix across all sub-catchments.

Stochastic dominance for population with respect to placed GI practices. We assume that for any given sub-catchment, our actions in placing GI practices, impacts the population transition probabilities. To establish our assumption, we generate new transition matrices corresponding to our action set. That is, for maintain action we use the original simulated population transition matrix. For one-level upgrade to three-level upgrade we use matrices that are 5%, 10%, and 15% shifted to the right, i.e., they are stochastically greater than the original simulated matrix in terms of transition from low level to high level of population.

2.2.3 Preprocessing Immediate and lump-sum Rewards

As discussed in Section 2.2.1, our focus is on runoff capturing performance of placed bioretention areas within the watershed of study. Moreover, we are interested in treating sub-catchments that are more populated. To calculate the performance of placed bioretention areas within any given sub-catchment with respect to projected precipitation in the study horizon.

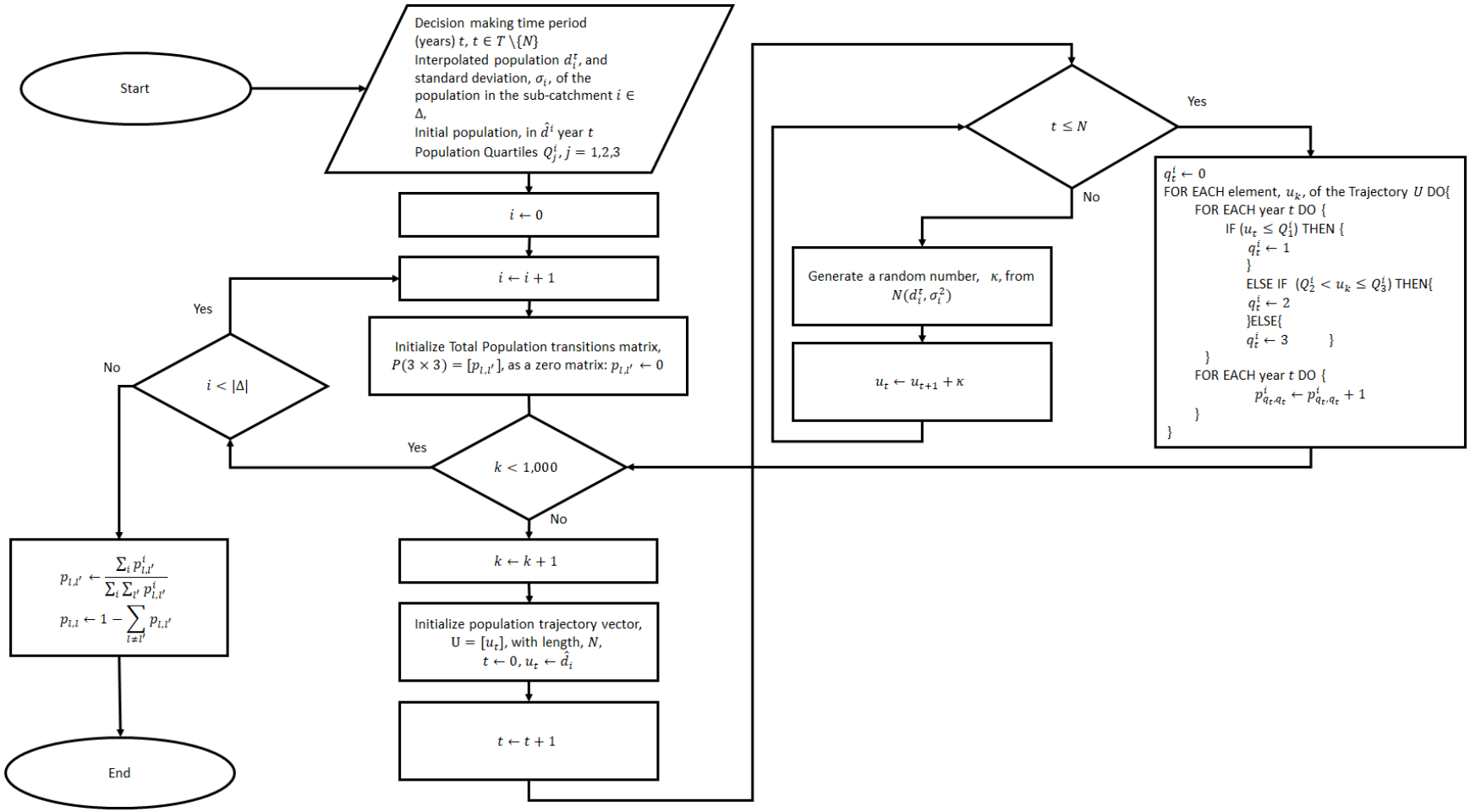

Figure 2.5: Simulation algorithm for calculating the population transition probabilities.

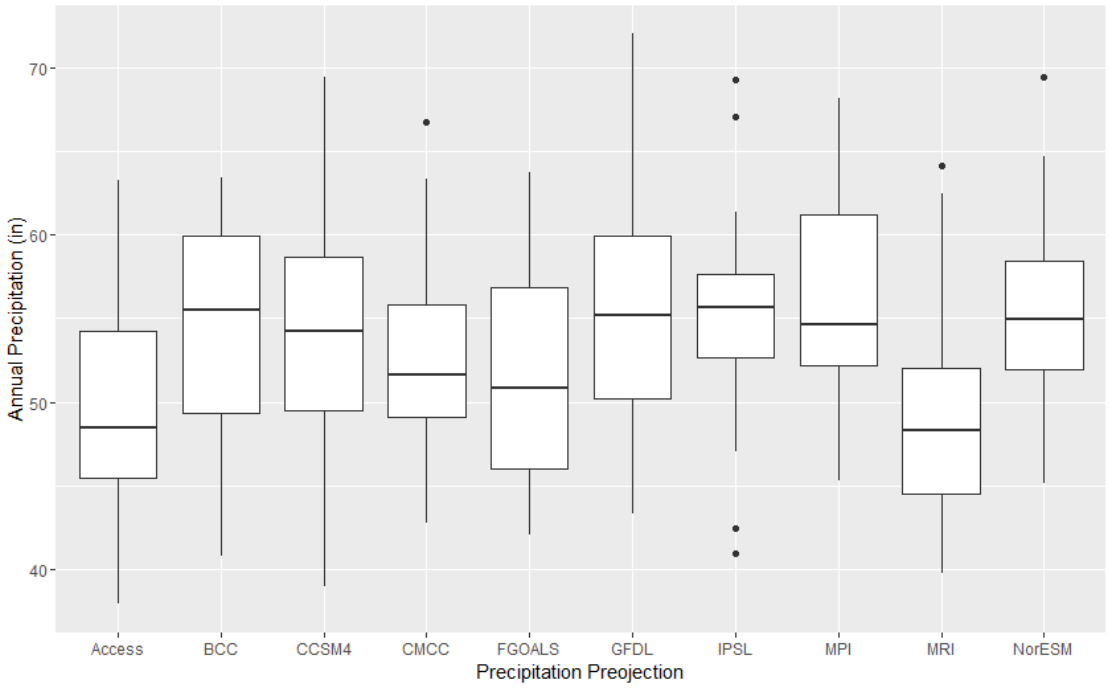

Figure 2.6: Box plots of annual precipitation projections over Knox County, Tennessee between 2018 to 2050.

In this study, we use 10 precipitation projections for the City of Knoxville produced by ten well known coupled global circulation models (CGCMs). Figure 2.6 shows the box-plots of the annual precipitation (in inches) using 10 CGCMs. As the figure shows, the annual precipitation is varying from one CGCM to another, indicating possible discrepancy in estimated runoff volumes. We first interpolate these precipitation projections into hourly projections and then run the EPA SWMM simulation engine [86] to estimate the baseline runoff volume as well as runoff captured by any given bioretention area within our watershed of study. Our preprocessing to estimate the surface runoff volumes is described in detail in [23].

Calculating the immediate reward, $r_t((i, \mu, \psi), a_i)$. The immediate reward of any action in time period t is the expected runoff volume captured by bioretention area multiplied by the median of the population category. Hence, with this setting, placing bioretention areas within more populated sub-catchments are rewarded more.

Calculating lump-sum reward, $R((i, \mu, \psi))$. We set up a Markov chain to calculate the lump-sum reward. That is, from year N onward, we set up a special MDP in which the only decision is 'maintain' regardless of the population transitions in those years. Note that for this special case MDP, the immediate reward calculation is identical to that of our original MDP. Also, for any given sub-catchment, the initial state is defined by state of the sub-catchment in year N. Lastly, for this special MDP, the lump-sum reward is zero.

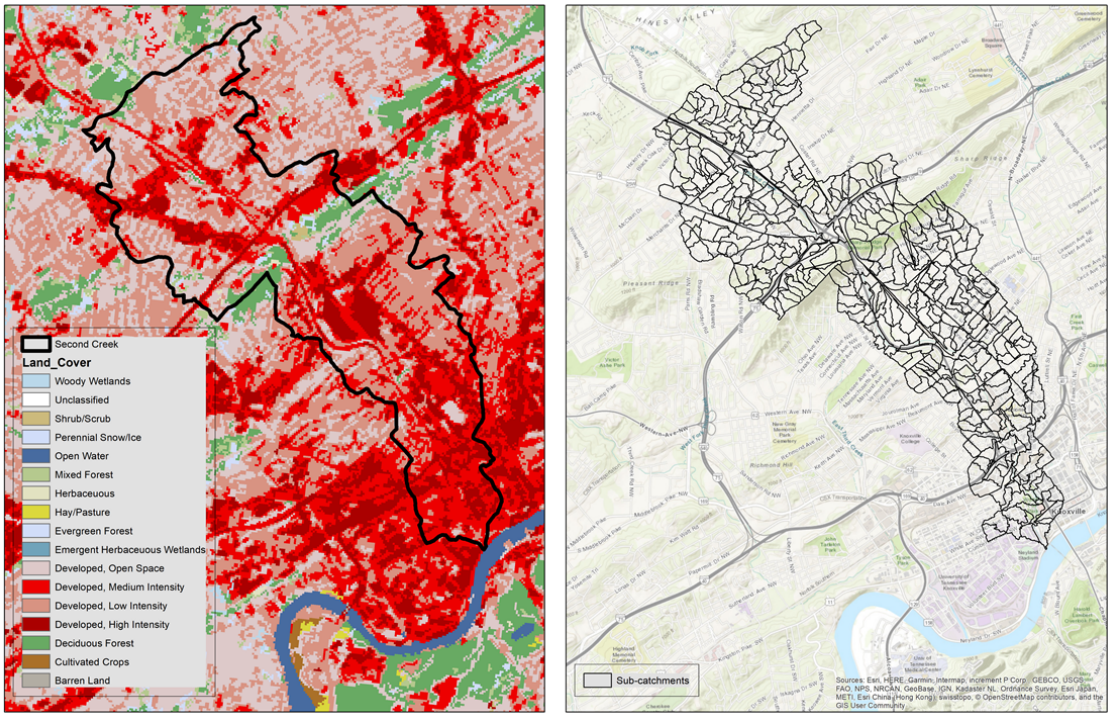

Figure 2.7: Map of land cover [57] (left panel) and hydrological sub-catchments (right panel) of the Second Creek, Knoxville, Tennessee

2.3 Case Study

As a case study, we consider the Second Creek in the City of Knoxville, Tennessee. The creek are located entirely within the City of Knoxville and have been identified as one of the principal sources of flooding in Knox County, Tennessee [21]. The creek is one of the most dense areas within the City of Knoxville. The creek area is 4913 acres. Our SWMM model of the creek contains 419 sub-catchments.

Figure 2.7 shows the map of land cover and land cover [57] (left panel) and hydrological sub-catchments (right panel) of the Second Creek. The red shades on the left panel denotes level of development ranges from low (e.g., forest land cover) to high. AS the panel shows, the southern region of the creek near to Tennessee river is highly developed. Subsequently, this highly developed region has the densest population throughout the City of Knoxville among years 2010, 2030, and 2050.

In solving the MDP model, we let the study horizon be 32 years (2018-2050), in which the decision making horizon consist of are years 2018-2029 ($N = 12$). Recall each time period is associated with a year. In the beginning of each year of decision making horizon, we decide on placing bioretention practices. Through years 2030-2050 we run a Markov chain model to estimate the lump-sum reward. beyond year 2051 the reward of placed bioretentions is zero.

Table 2.2: Summarizing total expected reward with respect to available budget.

Budget (million dollars)	Total Expected Reward	Percent Increase	Solution Time (s)
0.5	3.89E+11	0.00%	23
1	5.19E+11	33.42%	26
1.5	6.00E+11	54.24%	26
2	6.54E+11	68.12%	25
3	7.17E+11	84.32%	29
4	7.40E+11	90.23%	25
5	7.41E+11	90.49%	26

2.3.1 Computational Study

In this section, we present the computational results of the case study for both stochastic and robust models. In all computational experiments, we use the IBM ILOG CPLEX 12.8 (64-bit edition) on a PC running Microsoft Windows 7 (64-bit edition) with a Core i7, 2.8 GHz processor and 32 Gigabyte of RAM. Overall, solving the dual MDP model is easy; the average solution time with respect to different available budgets is less than one minute.

The average expected runoff volume across the watershed during 2018-2050 is estimated as 3.72E+10 gallons. Placing low, medium and high levels of bioretention across the entire sub-catchment result is 1.88E+9, 2.81E+9, and 3.75E+9 gallons of runoff volume captured, respectively. Also, in terms of percentage reduction in total expected runoff volume, low, medium, and high levels of bioretention contribute to 5.04%, 7.55% and 10.07%, respectively. The cost associated to these decisions are $2,310,000, $3,462,000, and $4,615,000.

We solve the dual MDP model under available budgets of 0.5, 1, 1.5, 2, 3, 4, and 5 million dollars and compare the total expected rewards. Table 2.2 summarizes the results. The total expected reward is increasing in available budget, however, the increase slows down in higher budget levels. For instance, increasing available budget from half million dollar to one million contributes to 33.42% improvement in total expected reward, whereas increasing budget from four million dollar to five million only results in 0.29% total expected reward improvement.

Next, we explore the distribution of recommended actions across sub-catchments while budget is limited, i.e., two million dollars. Table 2.3 shows the annual distribution of recommended actions across all sub-catchments while other factors of the state of the system is given. As the Table shows, action 'do nothing' (maintain) is dominate action for any given year, however, the level of dominance is different with respect to the decision making time period (year). For instance, in year 2018, the action takes over approximately 38% of recommended actions given that the population level is low; whereas, the action is 99% in year 2029 given that the population is low. Also note that for any given population level,

the recommend action of improving GI level is non increasing; this indicates that as system evolves the need of intervention, i.e., placing more GI, fades away gradually.

2.3.2 Sensitivity Analyses

In this section, we examine the sensitivity of the solution with respect to some important calibrated parameters, including the fidelity of decision making horizon and the dynamic nature of decision making. In former case we use a wide range of available budget, whereas for the later case we only consider a certain value for available budget .

First, we perform sensitivity analysis on the the fidelity of decision making horizon, i.e., number of time periods in the horizon. To that extend, we compare the default annual decision making time periods with bi-annual and six-years decision making periods. Figure 2.8 shows the total expected reward under different available budget and annual, bi-annual, and six-years decision making time periods. Consistent with our expectation, we observe that annual decision making time period provides the best results in terms of total expected rewards. Also, as the budget increases, the gap in total expected rewards widens. Note that under limited budget, e.g., half million dollar, the dual MDP model does not have enough space to take advantage of the dynamic GI placement, hence, higher fidelity decision making time periods contributes to less significant reward, comparing to the that of lower fidelity time periods. In contrast, under higher budget, the dual MDP model does have enough space to take advantage of the dynamic GI placement, resulting in widened gap in total expected results. For instance, the gap in total expected reward under half million dollar and five million dollars are 2.40E+10 and 5.10E+10, respectively.

Next, we compare the solution to the dual MDP model with two other competitive solution approaches to illustrate the impact of dynamic decision making on solution quality. The first competitive solution approach is to divide the available budget into equal values and then solve the dual MDP model with respect to the annual available budget; the second approach is to allocate all available budget in the first year. For this analysis, we consider two million dollar budget available under which we have summarized the distribution of optimal (deterministic) recommended actions. Table 2.4 compares the total expected rewards associated with each of the three approaches. As the table shows, the dual MDP model results is higher quality solutions comparing to the two other approaches. As the table shows the total expected reward for the dual MDP model is 14.52% better than the the approach that we allocate equally-divided budget each year. This stems from the fact that dual MDP model is less restrictive than the later approach as the solution time is approximately four order of magnitudes less than the that of that approach. Without no surprise, the dual MDP approach is still producing higher quality solution comparing the approach in which we allocate all budget in the first year. However, the difference in total expected reward is less magnified, indicating that it is better to allocate budget in the first year rather than

Table 2.3: Distribution of optimal (deterministic) recommended actions for each year, population, and bioretention area levels across all sub-catchments (bioretention area level 3 is skipped since the only action for that case is 'do nothing'). The budget is set to 2 million dollars.

		Bioretention Level								
		0				1			2	
Year	Population Level	Recommended Action								
		0	1	2	3	0	1	2	0	1
0	0	38%	37%	19%	7%	100%	0%	0%	100%	0%
	1	100%	0%	0%	0%	100%	0%	0%	100%	0%
	2	100%	0%	0%	0%	100%	0%	0%	100%	0%
1	0	97%	1%	2%	0%	100%	0%	0%	100%	0%
	1	86%	9%	3%	2%	92%	5%	2%	98%	2%
	2	86%	10%	2%	2%	91%	6%	3%	96%	4%
2	0	97%	1%	1%	1%	98%	2%	0%	99%	1%
	1	88%	8%	3%	1%	90%	6%	3%	97%	3%
	2	89%	8%	1%	2%	90%	6%	3%	95%	5%
3	0	96%	1%	1%	1%	96%	2%	2%	99%	1%
	1	89%	7%	3%	1%	90%	7%	3%	97%	3%
	2	90%	7%	2%	1%	89%	7%	3%	94%	6%
4	0	96%	1%	1%	2%	95%	3%	2%	98%	2%
	1	89%	6%	3%	1%	89%	8%	3%	97%	3%
	2	90%	7%	1%	2%	89%	8%	3%	93%	7%
5	0	97%	1%	1%	1%	94%	3%	3%	98%	2%
	1	90%	5%	3%	1%	90%	7%	3%	97%	3%
	2	90%	7%	1%	1%	89%	8%	4%	93%	7%
6	0	97%	1%	1%	1%	93%	4%	3%	97%	3%
	1	89%	6%	3%	1%	88%	8%	4%	97%	3%
	2	91%	6%	1%	1%	88%	9%	4%	93%	7%
7	0	97%	1%	1%	1%	93%	4%	3%	96%	4%
	1	90%	5%	3%	1%	88%	7%	5%	96%	4%
	2	91%	6%	2%	1%	89%	8%	3%	92%	8%
8	0	97%	1%	1%	1%	91%	5%	4%	95%	5%
	1	92%	5%	2%	0%	89%	6%	5%	95%	5%
	2	92%	5%	2%	1%	90%	8%	2%	92%	8%
9	0	97%	1%	1%	1%	91%	6%	4%	95%	5%
	1	94%	4%	2%	0%	89%	6%	5%	94%	6%
	2	94%	4%	1%	1%	89%	8%	3%	92%	8%
10	0	98%	0%	1%	1%	93%	4%	3%	94%	6%
	1	95%	3%	2%	0%	90%	6%	4%	93%	7%
	2	97%	2%	1%	0%	91%	5%	3%	92%	8%
11	0	99%	0%	0%	0%	95%	2%	3%	94%	6%
	1	96%	3%	1%	0%	94%	4%	2%	94%	6%
	2	99%	1%	0%	0%	94%	4%	2%	94%	6%

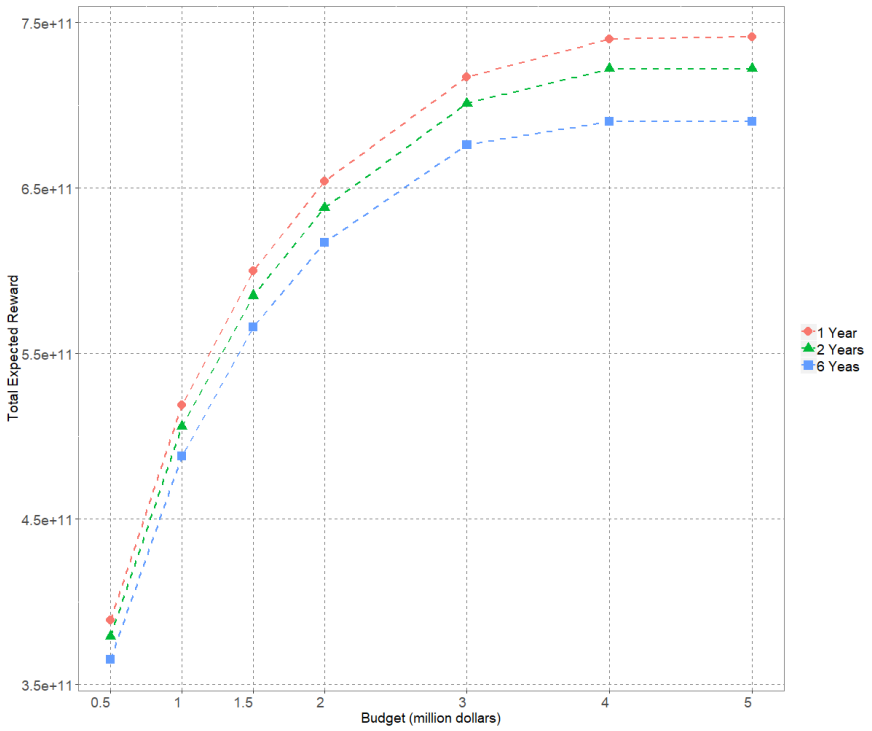

Figure 2.8: Optimal total expected reward with respect to the fidelity of decision making horizon under different budget available.

Table 2.4: Comparing objective of the dual MDP approach with other solution approaches under two million dollars budget

Solution Approaches	Total Expected Reward	Difference with MDP Reward (%)	Solution Time (s)
Dual MDP approach	6.54E+11	–	25
Equally allocating budget each year	5.59E+11	-14.52%	11,573
Allocating all budget in first year	6.03E+11	-7.86%	2

allocate equally-divided budget each year in each year of decision making horizon. Recall that distribution of optimal (deterministic) recommended actions indicates significant GI placement in the first year (see Table 2.3). Nevertheless, we loose total expected reward by 7.86% if we just allocate the all budget in the first year. In other words, the dynamic decision making can contribute to up to 7.86% better total expected reward.

2.4 Summary and Insights

In this study, we developed a mathematical model to optimally place GI practices when (re-)designing an urban area, subject to uncertainties in population growth and future precipitation. Specifically, we developed a MDP model that addresses the problem of

dynamic placement of GI practices across an urban watershed with respect to stochastic nature of population growth and precipitation projections. Furthermore, we enhanced our model in terms of producing a deterministic policy which is more comprehensible for city planners. We conducted our analysis on an congested urban watershed in the City of Knoxville. Our analysis shows that dynamic GI placement is significantly better than the conventional methods.

In calibrating our model, we included stochasticity of population growth into our model through population transition matrices, we introduced simulation algorithm that takes into account the rare information of the population and produces the population transition matrix. However, there is limited number of resources that are for population projection in fine scale. Also, note that we faced with the lack of resources in quantifying the population transition in response to placed GI in an urban neighborhood. Conducting a survey research that may help us with quantifying the population transition more accurately when we install a GI in an urban neighborhood.

Lastly, in our model calibration we incorporated the flood mitigation (runoff volume reduction) and GI recreation/ aesthetics of the GI placement in an urban watershed. However, there are other factors in placing GI practices that would be center of interest, e.g., ecosystem health and economic considerations.

Chapter 3

An Efficient Exact Solution Approach to the Time-Discretized Job Shop Scheduling: An Urban Storm Recovery Case Study

Job shop scheduling problem is a classical optimization problem which has been center of interest since 1950s. The problem is NP-hard in strong sense [50], indicating that solving even small job shop problem to optimallity is very hard. Approximation and optimization are two major techniques in solving the classical deterministic job shop problem [60]. Considering optimization algorithms, Bowman [28], Wagner [102], and Balas [22] first used (mixed) integer mathematical formulations to solve the problem. Since then, several formulations and exact solution algorithms has been introduced for the job-shop problem. We refer reader to the survey papers [59, 104] for a comprehensive literature review of exact solution methods for the job shop problem.

To the best of our knowledge, the approach of time-discretized network is not used to model the classical job shop problem, despite its extensive usage in a variant of the problem, e.g., train scheduling [30, 32]. The major barrier in using such approach is that the number of binary variables for even small job-shop problems is extremely high that results in an inefficient formulation. In a recent research, Barah et al. [25] investigated conditions for which a time-discredited train scheduling problem is integer-optimal. Their approach resulted in new extended formulation that has drastically less number of binary variables, comparing the original binary formulation of the time-discredited train scheduling problem. Their effort can, somehow, lift the major barrier of using time-discretized network in formulating the job-shop problem.

In this paper, we extend Barah et al. [25] research to propose a set of valid inequalities. Furthermore, we propose a column generation algorithm to solve the problem to optimality.

The solution approach is also integrated with a commercial solver, which results in significant computational savings. Computational experiments show that the developed algorithm can efficiently solve test problems to near-optimality. The algorithm is used in a case study to schedule maintenance crew following a storm event to efficiently maintain green infrastructure practices across a watershed to mitigate surface runoff due to future events.

The rest of the paper is organized as follows. First, introduce the model formulation in Section 3.1. Next, in Section 3.2 we develop our valid inequalities and introduce a column generation technique followed by the computational experiments in Section 3.3. Next, we introduce a case study of the maintenance crew scheduling in Section 3.4. Finally, we provide conclusion in Section 3.5.

3.1 Model Formulation

Let $L = \{1, 2, \ldots, |L|\}$ be the set of jobs, $M = \{1, 2, \ldots, |M|\}$ the set of machines (excluding artificial source, σ, and sink, τ, machines), $T = \{1, 2, \ldots, |T|\}$ be the set of discrete time-intervals.

The job shop scheduling problem can be visualized by in terms of flows along a digraph. We let $D(V, A)$ as an acyclic directed network. Let vertices $v^+_{m,t}$ and $v^-_{m,t}, v^+_{m,t}, v^-_{m,t} \in V$, respectively denote the jobs entering the machine m at time t and jobs that their processing is done at time t on machine m. To enforce a flow conservation over vertices pertaining to all machines, we let σ and τ, representing dummy source and sink nodes respectively.

Let $(v^+_{m,t}, v^-_{m,t'})$ denote the arc corresponding to job l that represents that processing the job on machine m, beginning at $v^+_{m,t}$ and end at $v^-_{m,t'}$. Also, let $(v^-_{m,t}, v^+_{m',t'})$ denote the arc the represents job l is transferred from machine m at time $v^-_{m,t}$ and reaches to machine m' at time $v^+_{m',t'}$. Lastly, let the $(v^+_{m,t}, v^+_{m,t'})$ denote job post completion waiting at machine m from time t' to time t'. Hence the buffer occurs after processing the job on any given machine. to establish the flow for any given job, we let dummy arcs $(\tau, v^+_{m,t})$ and $(v^-_{m,t}, \sigma)$ denote the dummy arcs that begin and end flow, respectively. We let the set $A^l = \{(v^+_{m,t}, v^-_{m,t'}), (v^-_{m,t}, v^+_{m',t'}), (v^+_{m,t}, v^+_{m,t'}), (\tau, v^+_{m,t}), (v^-_{m,t}, \sigma)\}$ denote all arcs corresponding to job l. Hence, the set $A = A^1 \cup A^2 \cup \cdots \cup A^{|L|}$.

For any given job, we define a time window for beginning the process on the any given machine. Let $[\underline{T}^l_m, \overline{T}^l_m]$ denote the time window of beginning the process of job j on machine m. Note that the job can stay in machine buffer so long as it can be transferred to the next machine at a time that is within the time window for beginning the job on next machine. Let δ^l_m be the processing time of job l on machine m. Also, let $\kappa^l_{m,m'}$ be the required minimum transitioning time of job l between ordered pair of machines (m, m').

The set $A = A^1 \cup A^2 \cup \cdots \cup A^{|L|}$ consists of the arcs obtained from the union of all possible arcs for each job in set L. As arcs are specific to each job, we can think of $a \in A^l$ as having the color l. Each A^l is made up of three types of arcs: arcs representing processing

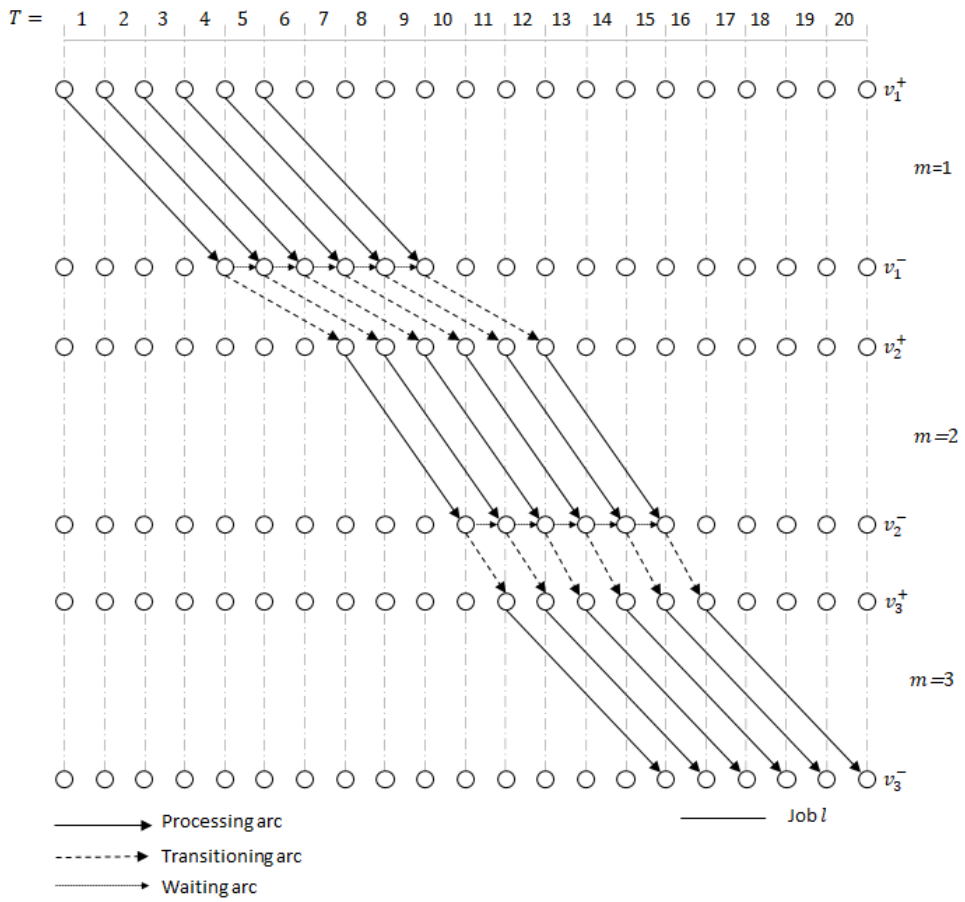

Figure 3.1: An example of the time-discretized job shop problem, adapted from [25]. For ease of exposition, we have omitted the artificial source and sink nodes and arcs corresponding to them.

on machines, arcs representing transitions from one machine to another machine, and arcs representing waiting at buffers. An example of a digraph D for job l is depicted in Figure 3.1. In this figure, Job l requires processing on machines 1, 2, and 3 sequentially. Machine 1 can begin processing job 1 sometime in $[\underline{T}^l_m = 1, \overline{T}^l_m = 6]$. There is a 3-minute transitioning time between machine 1 and 2 for job l, Note that the job gets the highest profit if it gets processed on machine 3 at time 15. Note that for ease of exposition, we have omitted the source and sink machines in this figure.

We consider two sets of constraints in our mathematical model. First, all jobs should physically traverse a sequentially from source to sink, referred to as flow conservation constraint in the paper. Second, no two jobs can be on the same machine at the same time interval, referred to as clique constraint in the paper.

In our mathematical model, we have a one-to-one correspondence between our variables and the arcs in the Digraph $D(V, A)$. Accordingly, we let $x^l_{m,t,t'}$, $y^l_{m,m',t,t'}$, and $w^l_{m,t,t'}$ be binary variables that are corresponding to arcs $(v^+_{m,t}, v^-_{m,t'})$, $(v^-_{m,t}, v^+_{m',t'})$, and $(v^+_{m,t}, v^+_{m,t'})$,

Table 3.1: The sets

Set	Definition
$L = \{1, 2, \ldots, \lvert L\rvert\}$	Set of jobs
$M = \{1, 2, 3, \lvert M\rvert\}$	Set of machines each job should be processed in its own sequence.
$M^l = \{1, 2, 3, \lvert M\rvert^l\}$	Ordered set of machines for job l
$T = \{1, 2, \ldots, \lvert L\rvert\}$	Set of discretized time periods
$V = \{v^+_{m,t}, v^-_{m,t'}\}$	Set of vertices corresponding to discretized time periods of all machine, where $v^+_{m,t}$ and $v^-_{m,t'}$ respectively denote the time discretized vertices of machine m corresponding to time period t at which a job enters the machine m for processing and the time a job processing on the machine is done
$A = \{A^l\}$	Set of all arcs, where, A^l denotes the arc set of job l

Table 3.2: The parameters

Parameter	Definition
σ	Artificial source machine for all jobs
τ	Artificial sink machine for all jobs
$v^+_{m,t}$	Time discretized vertex of machine m corresponding to time t at which a job can enter the machine for processing
$v^-_{m,t}$	Time discretized vertex of machine m corresponding to time t at which a job processing on machine m is done
δ^l_m	Required processing time of job l on machine m
$\kappa^l_{m,m'}$	Required minimum transferring time of job l between ordered pair of machines (m,m')
$\underline{T}^l_m$	The earliest time for job l that can be processed on machine m
$\overline{T}^l_m$	The latest time for job l that can be processed on machine m
$\pi^l_{m,t,t'}$	reward of arc $(v^+_{m,t}, v^-_{m,t'})$
$\mu^l_{m,t,t'}$	penalty of arc $(v^+_{m,t}, v^+_{m,t'})$

respectively. Also, let $y^l_{m,\sigma,t}$ and $y^l_{\tau,m,t}$ be binary variables corresponding to dummy arcs $(\tau, v^+_{m,t})$ and $(v^-_{m,t}, \sigma)$, respectively. For simplicity of notation, in the remainder we use $x = [x^l_{m,t,t'}]$, $y = [y^l_{m,m',t,t'}]$, $w = [w^l_{m,t,t'}]$ to refer to the vectors of the corresponding variables.

Lastly, our general objective is to schedule all jobs so that the makespan time is minimized. However, to prioritize particular jobs over other jobs, we assign a reward to the transitioning arcs. Also, we assign penalty to waiting arcs. Accordingly, let $\pi^l_{m,t,t'} \geq 0$ denote the reward of arc $(v^+_{m,t}, v^-_{m,t'})$ and let $\mu^l_{m,t,t'} \leq 0$ denote the penalty corresponding to arc $(v^+_{m,t}, v^+_{m,t'})$. The reward/penalty of other arcs is zero. Our notations are summarized in Tables 3.1, 3.2, and 3.3.

Table 3.3: The variables

Set	Definition
$x^l_{m,t,t'}$	Binary variable denoting if processing arc $(v^+_{m,t}, v^-_{m,t'})$ is chosen for job l
$y^l_{m,m',t,t'}$	Binary variable denoting if transitioning arc $(v^-_{m,t}, v^+_{m',t'})$ is chosen for job l
$w^l_{m,t,t'}$	Binary variable denoting if waiting arc $(v^+_{m,t}, v^+_{m,t'})$ is chosen for job l

Given the define sets, parameters and variable, our mathematical model for the time discretized job shop problem with processing time intervals is defined as following.

$$\max_{x,y,w} \sum_{m\in M^l} \sum_{t\in[\underline{T}^l_m,\overline{T}^l_m-1]} \pi^l_{m,t,t'} x^l_{m,t,t+\delta^l_m} + \sum_{m\in M^l\setminus\{|M^l|\}} \sum_{t\in[\underline{T}^l_m,\overline{T}^l_m-1]} \mu^l_{m,m',t,t+\kappa^l_m} y^l_{m,m',t,t+\kappa^l_m} \tag{3.1}$$

$$\text{s.t.} \sum_{t\in[\underline{T}^l_m{}',\overline{T}^l_m-1]} x^l_{m,t,t+\delta^l_m} = 1, \qquad \forall l\in L, \forall m\in M, \tag{3.2}$$

$$x^l_{m,t-\delta^l_m,t} = y^l_{m,m',t,t+\kappa^l_{m,m'}} + w^l_{m,t,t+1}, \qquad \begin{array}{l}\forall l\in L, \forall m,m'\in M^l,\setminus\{|M^l|\},\\ t=\underline{T}^l_m,\end{array} \tag{3.3}$$

$$x^l_{m,t-\delta^l_m,t} + w^l_{m,t-1,t} = y^l_{m,m',t,t+\kappa^l_{m,m'}} + w^l_{m,t,t+1}, \qquad \begin{array}{l}\forall l\in L, \forall m\in M^l\setminus\{|M^l|\},\\ \forall t\in[\underline{T}^l_m{}',\overline{T}^l_m-1]\end{array} \tag{3.4}$$

$$x^l_{m,t-\delta^l_m,t} + w^l_{m,t-1,t} = y^l_{m,m',t,t+\kappa^l_{m,m'}}, \qquad \begin{array}{l}\forall l\in L, \forall m\in M^l,\setminus\{|M^l|\},\\ t=\overline{T}^l_m\end{array} \tag{3.5}$$

$$\sum_{l\in L} \sum_{\{t'\in[\underline{T}^l_m,\overline{T}^l_m]|t\in[t',t'+\delta^l_m]\}} x^l_{m,t',t'+\delta^l_m} \le 1, \qquad \forall m\in M, \forall t\in T, \tag{3.6}$$

$$x,y,w\in\{0,1\}. \tag{3.7}$$

The objective function (3.1) maximizes the schedule profit. Constraint (3.2) ensures that exactly one flow is initiated for any given job. Constraints (3.3)-(3.5) are classical flow conservation constraints, enforcing flow-in flow out balance over vertex set, $V\setminus\{v^+_{m=1,t}, v^-_{m=|M|,t}, \forall t\in T\}$. Lastly, constraint (3.6) guarantees that at at time period, at most one job can be processed on a given machine.

Let $\Omega=\{(x,y,w)\in[0,1]^{|x|.|y|.|w|}|(3.3)-(3.5)\}$ and $\Pi=\{(x,y,w)\in[0,1]^{|x|.|y|.|w|}|(3.6)\}$ be polyhedrons, where their intersection, $\Omega\cap\Pi$, builds the the problem (3.2)-(3.6) polyhedron. It is easy to show that Ω and Π are block diagonal totally unimodular (TU) matrices. However, the intersect is not guaranteed to be TU. Barah et al. [25] show two conditions for which model (3.1)-(3.6) is integer-optimal. In short, for any given job l on any given machine m, (i) if the the length of processing arcs, δ^l_m, is less than or equal to the length of time window at which the job can be start processing on the machine, i,e, $\delta^l_m\le[\underline{T}^l_m,\overline{T}^l_m]$, and (ii) if all penalty coefficient of waiting arcs are non negative, $\mu^l_{m,t,t'}\le 0$, the resulting model is integer optimal. Utilizing their polyhedral study, they introduced a new set of binary variables z that maintains the condition (i); That is, for any given job on any given machine

m, they decompose the time windows into disjoint time windows, where each of which has a length at most as the length of processing arc δ_m^l on that machine, i.e.,

$$\sum_{t\in[\underline{T}_m^l+i\rho_m^l,\underline{T}_m^l+(i+1)\rho_m^l]} x_{m,t,t+\delta_m^l}^l - z_m^{l,i} \leq 0, \forall l \in L, \forall m \in M, \forall i \in \{0, \rho_m^l{}'\}, \rho_m^l = \left[\frac{\overline{T}_m^l - \underline{T}_m^l}{\delta_m^l}\right], \tag{3.8}$$

where, ρ_m^l denotes the maximum length at which the processing arcs corresponding to job l on machine m overlap at at list one time period. Furthermore, they add all new z variables along with equation (3.8) into their new model formulation that result in significant computational saving. Here, we further investigate the efficient ways in generating these variables, as well as introducing new valid inequalities so that we achieve more computational savings.

3.2 Developing Solution Algorithm

The cardinality of z variables is significantly less than the that of x variables, as a factor of processing times. Hence, lengthy processing times with short time windows makes the problem more simple to solve. Note that, adding z variables along with equation (3.8) to the model (3.1)-(3.7), with relaxed x, y, w variables, results in significant computational savings [25]. Next, we develop valid inequalities for these new binary variables in line with achieving more computational savings. Also, we introduce a column generation algorithm in which we add z variables more efficiently. First we introduce two sets of valid inequalities to the model (3.1)-(3.8). Next we introduce our column generation algorithm.

We re-define $z_m^{l,i}$ variables by adding one more index to it; let $z_{m,\hat{t}}^{l,i}$ be a binary variable where $\hat{t} = (\underline{T}_m^l + (i+1)\rho_m^l) - (\underline{T}_m^l + i\rho_m^l)/2$ denote the center of the time window that is covered by x variables that are corresponding to $z_m^{l,i}$ by (equation 3.8). Also let $\bar{\delta}_m$ be the minimum processing time of all jobs on machine m. Theorem 3.1 introduces two valid inequalities for the constraint set defined in (3.2)-(3.8).

Theorem 3.1. *For any given machine m, let $\bar{\delta}_m = \min\{\delta^l_m, \forall l \in L\}$. The following equations are valid to (3.2)-(3.8).*

$$\sum_{l\in L}\sum_{i\in\{0,\rho^l_m{}'\}}\sum_{\hat{t}\in[j\bar{\delta}_m,(j+1)\bar{\delta}_m]} z^{l,i}_{m,\hat{t}} \le 2, \qquad \forall m\in M, \forall j \in \left\{0, \left[\frac{|T|}{\bar{\delta}_m}\right]+1\right\} \tag{3.9}$$

$$\begin{aligned}&\sum_{l\in L}\sum_{i\in\{0,\rho^l_m{}'\}}\sum_{\{\hat{t}\in[\underline{T}^l_m+i\rho^l_m,\underline{T}^l_m+(i+1)\rho^l_m]|t\in[\hat{t}-\delta^l_m,\hat{t}+\delta^l_m-1]\}} z^{l,i}_{m,\hat{t}}\\ &+\sum_{l\in L}\sum_{\{t'\in[\underline{T}^l_m+i\rho^l_m,\underline{T}^l_m+(i+1)\rho^l_m]|t\in[t',t'+\delta^l_m]\}} x^l_{m,t',t'+\delta^l_m} \qquad \forall m\in M, \forall t\in T.\\ &-\sum_{l\in L}\sum_{\{t'\in[\underline{T}^l_m,\overline{T}^l_m]|t\notin[t',t'+\delta^l_m]\}} x^l_{m,t',t'+\delta^l_m} \le 2,\end{aligned} \tag{3.10}$$

Proof. Proof of validity of equation (3.9) by contradiction. For a given machine and given time window $[\bar{\delta}_m, 2\bar{\delta}_m]$, without loss of generality, we assume there would be three $z^{l,i}_{m,t}$ variables can take a positive value that is valid to (3.2)-(3.8). For simplicity of notation, we let z^1, z^2 and z^3, with corresponding centers t^1, t^2, t^3 (see definition of $\hat{t}$), be positive in a valid solution to (3.2)-(3.8). Since z^1, z^2 and z^3 are positive, their corresponding centers do not overlap and, hence, can be sorted in an strict ascending order. Let $\bar{\delta}_m \le t^1 < t^2 < t^3 \le 2\bar{\delta}_m$. On the other hand, let variable x^1, x^2 and x^3, corresponding to variables z^1, z^2 and z^3, respectively, be all positive in the solution. Recall that due to clique constraint (3.6), a valid solution requires not having overlap between the corresponding intervals to variables $x1$ and x^3. Hence, without loss of generality, we let $[t^1-\delta^1_m, t^1]$ and $[t^3, t^3+\delta^3_m]$. Hence, the maximum interval that is still vacant for variable x^2 is $[\bar{\delta}'_m, 2\bar{\delta}_m - 1]$ in which $\bar{\delta}'_m \le t^2 \le 2\bar{\delta}_m - 1$. Also, from the assumption of the theorem, we have $\bar{\delta}_m - 2 < \delta^3_m$. Therefore, there is no enough vacant time interval for variable x^2 for not having overlap with x^1 and x^3. That is, having positive x^1, x^2, and x^3, equivalently positive z^1, z^2, and z^3, violates our initial assumption of validity of the solution to (3.2)-(3.8).

The proof of the validity of the equation (3.10) is straightforward. Given machine m, suppose clique constraint (3.6) is binding. Hence, the time slot is occupied by at most one train, namely train l. In this case, the first and second terms in equation (3.10) associated with that train is one and the third term is zero. Note that there might be a positive z variable existing in the equation (3.10) that is associated with an other train which it is canceled out with the third term of the equation. □

Next, we describe our column generation algorithm that dynamically defines the z variables with their corresponding centers and adds them to the model. First, we solve the LP relaxation of the model (3.1)-(3.7). For any given job l on given machine m, let $\hat{x} = [x^l_{m,t,t+\delta^l_m}]$ denote vector of x variables that have fractional solution. Then, given the fractional solution variables corresponding arcs, we calculate the shortest time interval that

overlaps all the fractional solution variables. Let $[s, s']$ denote the interval. Next, beginning from time s we define $z^{l,i}_{m,\hat{t}}$ variable with corresponding $2\delta^l_m - 1$ time interval length. For example, beginning from time period s, we define $z^{l,1}_{m,s+\delta^l_m}$ and continue to define variables up to $\hat{t} = s' - \delta^l_m{}'$. Then, for a given job l and given machine m and corresponding centers $\hat{t}$, we solve the following set covering problem.

$$\min \sum_i z^{l,i}_{m,\hat{t}} \tag{3.11a}$$

$$\text{s.t.} \sum_i z^{l,i}_{m,\hat{t}} + \hat{x}^l_{m,t,t+\delta^l_m} \geq 1 \qquad \forall t \in [s, s'], \hat{t} \in [t, t + \delta^l_m], \tag{3.11b}$$

$$\sum_i \sum_{t \in [\hat{t}\delta^l_m, \hat{t}+\delta^l_m - 1]} z^{l,i}_{m,\hat{t}} \leq 1, \qquad \forall t \in [s, s'], \tag{3.11c}$$

$$z^{l,i}_{m,\hat{t}} \in \{0, 1\} \qquad \forall i. \tag{3.11d}$$

Given that vector $z^{l,i}_{m,\hat{t}}$ is the solution to set covering model (3.11), we add the variable to model (3.1)-(3.7) along with modified equation (3.8) that links the x variables to z variables. Then we solve the model (3.1)-(3.8) and capture the solution. If the solution is integral, we stop since we have reach to optimal solution. Otherwise, we continue our column generation by adding new z variables generated from model (3.11).

3.3 Computational Study

To evaluate our effort in enhancing our model formulation toward achieving more computational savings. We solved all computational experiments provided in [24]. We solved all instances using IBM ILOG Cplex 12.8 on a Microsoft windows 7 desktop PC with a 2.4 GHz 8 core CPU and 32 GB of RAM.

Table 3.4 summarizes efficiency of our proposed valid inequalities (3.9)-(3.10) in computational savings. In this table, we report the number of continues and binary variables in the third and forth columns, respectively. The MIP model is the one introduced in [25]. The Enhanced MIP is the MIP along with equations (3.9) and (3.10). In this table, we also show the number of processed nodes and number of nodes left. Note that we set the solution time limit to an hour for each instance.

As Table 3.4 shows, MIP is not able to solve all instances (instances number 17 and 20), whereas the enhanced MIP is capable to solve all instances to optimality. More insistingly, the enhanced MIP significantly outperforms the MIP in terms of average computational saving, i.e., 80%. Our promising results suggest that our valid inequalities are strong enough to significantly drop the computational time. The number of nodes left in the enhanced MIP suggest that equations (3.9) and (3.10) contribute to better lower bound which allows

the CPLEX to prone nodes of the branch and bound tree more quickly. In contrast, the MIP formulation explores more branch and bound tree nodes at the expense of 80% more computational times.

3.4 Case Study

Back-to-back storm events are, statistically rare, events that negatively impact the capability of GI practice runoff capturing volume [100]. Designing a system of GI practices that is resilient to back-to-back storms events is an ongoing research topic [100, 54, 101]. Although extreme back-to-back events statistically rare, due to climate change we may expect more frequent extreme events with large negative impact on GI practices in capturing runoff volumes. In case of (back-to-back) extreme events, an emergency maintenance operations maybe required to recover GI practices and restore them to their near normal runoff capturing volume.

As a case study, we consider the First and Whites Creeks in the City of Knoxville, Tennessee.These creeks are located entirely within the City of Knoxville and have been identified as the principal sources of flooding in Knox County, Tennessee [21]. to set up our case study, we divide the watershed into five regions (see Appendix B). We consider a three group of maintenance operators; the first response group that does rudimentary cleanups and asses the work load for the regular maintenance group, and finally the quality inspection group that performs post-maintenance audits to make sure the GI practise is recovered and restored properly. Our problem is to schedule the maintenance groups so that the overall maintenance operations is done as soon as possible and also those regions that are susceptible to destruction addressed immediately. By definition, our case study problem is an instance of flows shop problem in which jobs has a time window and prioritization. Specifically, we have flow shop with three machines each of which is corresponding to a maintenance group. The jobs are regions within the First and Whites Creek. Machines processing time can be interpreted as the average required workload to perform the fist response operations, regular maintenance operations, and quality inspections. Also the average travel time from one region to another is consist of our transitioning arcs. Finally, we assign a randomized rewards to regions to randomly prioritize regions for maintenance operations. Our analysis shows that the we can solve a model with up to 100 jobs and 3 machines to optimality in a reasonable time.

3.5 Conclusion

In this research, we addressed the time discretized job shop problem with specified time widows for jobs. Specifically, we extended the analysis of the underlying polyhedron of the

problem to propose a model formulation with valid inequalities. Furthermore, we utilized our analysis in developing a column generation algorithm. Our computational results showed that we achieved promising computational saving of 80%, comparing to our benchmark mathematical formulation, solving certain test problems.

Table 3.4: Computational comparison of the enhanced MIP with mixed integer programming model (MIP) published in [25].

No.	No. of Jobs (trains)	No. of machines (stations)	No. of vars	No. of binary vars	MIP				Enhanced MIP			
					Solution time	Best obj. value	Processed nodes	Nodes left	Sol. time	Best obj. value	Processed nodes	Nodes left
1	10	47	24928	1565	20.83	120653	103	0	21.15	120661.33	29	4
2	10	47	24177	1546	190.89	118037	1540	0	23.63	118046.5	153	2
3	10	47	24694	1575	15.09	125505	74	0	12.73	125505	23	0
4	10	47	23720	1516	13.78	119348	168	0	11.72	119348	29	0
5	15	47	34079	2176	69.81	161903	87	0	28.58	161903	11	0
6	15	47	30320	1955	45.81	154113	78	0	31.5	154113	41	0
7	15	47	35903	2292	601.5	171894	777	0	169.94	171898.49	129	4
8	15	47	31632	2047	8.8	155580	0	0	10.91	155580	0	0
9	15	47	30854	2001	40.99	151720	120	0	17.84	151720	9	0
10	20	47	40260	2562	2637.39	193649	7672	0	155.81	193662.92	255	4
11	20	47	42866	2728	1067.42	204182	1446	0	254.14	204197	93	2
12	20	47	40870	2623	377.56	200102	382	0	177.53	200102	61	0
13	20	47	39538	2511	344.27	195088	727	0	62.33	195088	38	0
14	20	47	38504	2453	165.08	186005	1578	0	31.5	186007	65	4
15	25	47	41372	2637	396.45	203448	690	0	56.52	203448	25	0
16	25	47	42059	2653	3601.48	196250.19	6006	2089	476.05	195618.5	601	11
17	25	47	40740	2615	214.05	196287	200	0	47.02	196287	11	0
18	25	47	43355	2729	2423.77	202027	3425	0	117.11	202042.25	63	4
19	25	47	40604	2601	571.5	200380	1929	0	94.45	200380	64	0
20	30	47	51472	3282	3601.58	247086.1	1948	743	629.95	245960.86	722	37
21	10	52	26838	2194	10.95	177957	3	0	11.8	177964.5	0	1
22	10	52	26353	2136	2.59	189079	0	0	8.5	189079	0	0
23	10	52	26830	2125	2.52	150819	0	0	6.59	150819	0	0
24	10	52	23639	1923	2.48	149488	0	0	7.36	149488	0	0
25	10	52	28766	2292	3.59	183871	0	0	9	183871	0	0
26	15	52	41709	3340	12.05	256705	0	0	20.08	256705	0	0
27	15	52	37656	3008	16.84	235739	9	0	25.95	235741.87	5	4
28	15	52	37540	3014	861.28	262359	895	0	355.3	262385.06	130	17
29	15	52	42335	3417	23.06	285499	5	0	36.24	285524.5	0	1
30	15	52	39800	3261	222.97	272767	84	0	167.72	272767	51	0
31	20	52	49678	3928	11.59	278094	0	0	21.84	278114	0	1
32	20	52	50353	3993	1385	306120	859	0	615.95	306146.26	209	16
					575.05				112.76			

Bibliography

[1] Commonwealth Scientific and Industrial Research Organisation. http://www.csiro.au/en/Research/OandA/Areas/Assessing-our-climate/CAWCR/ACCESS/. Accessed: 2016-09-10. 81

[2] Beijing Climate Center, China Meteorological Administration. http://forecast.bcccsm.cma.gov.cn/htm/. Accessed: 2016-09-10. 81

[3] Climate and Global Dynamics Laboratory (CGD) at the National Center for Atmospheric Research (NCAR). http://www.cesm.ucar.edu/models/ccsm4.0//. Accessed: 2016-09-10. 81

[4] Euro-Mediterranean Center on Climate Change. http://www.cmcc.it/models/cmcc-cm. Accessed: 2016-09-10. 81

[5] United States Environmental Protection Agency (EPA) green infrastructure website. http://water.epa.gov/infrastructure/greeninfrastructure/, . Last accessed February 11, 2016. 2

[6] United States Environmental Protection Agency (EPA) operation and maintenance considerations for green infrastructure. https://www.epa.gov/G3/operation-and-maintenance-considerations-green-infrastructure, . Last accessed July 31, 2018. 12

[7] Institute of Atmospheric Physics, Chinese Academy of Sciences, State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics. http://www.lasg.ac.cn/. Accessed: 2016-09-10. 81

[8] Geophysical Fluid Dynamics Laboratory (Princeton University). https://www.gfdl.noaa.gov/earth-system-model/. Accessed: 2016-09-10. 81

[9] Green infrastructure types. http://www.greenbelt.ca/gi_rain_garden_and_bioretention. Last accessed July 31, 2018. 12

[10] Institut Pierre Simon Laplace. http://icmc.ipsl.fr/. Accessed: 2016-09-10. 81

[11] Max Planck Institute for Meteorology. http://www.mpimet.mpg.de/en/science/models/mpi-esm.html. Accessed: 2016-09-10. 81

[12] Meteorological Research Institute (MRI) of the Japan Meteorological Agency. http://ds.data.jma.go.jp/tcc/tcc/products/gwp/gwp7/html_e/model.html. Accessed: 2016-09-10. 81

[13] National Association of City transportation Officials. https://nacto.org/, . Accessed: 2018-11-19. 12

[14] Bioretention cell sizing. https://nacto.org/publication/urban-street-stormwater-guide/stormwater-elements/bioretention-design-considerations/bioretention-cell-sizing/, . [Last accessed on November 19, 2018]. 12

[15] National centers for environmental information (ncei). https://www.ncdc.noaa.gov/. Last accessed July 31, 2018. 14

[16] Multi-institutional, Coordinated Climate Research in Norway. http://folk.uib.no/ngfhd/EarthClim/. Accessed: 2016-09-10. 81

[17] Tennessee Stormwater Training. http://tnstormwatertraining.org/index.asp. Accessed: 2018-10-16. 13, 42

[18] Park, trails and greenway framework. http://arnoldmo.org/download/arnold-comprehensive-plan-parks-trails-and-greenway-framework/?wpdmdl=4912. Accessed on 9 February 2016 at 08:00AM. 36

[19] New york water. http://wwf.panda.org/?204446/New-York-water/, 2012. Accessed on 9 February 2016 at 08:49AM. 36

[20] EPA issues guidance for cities on managing sewer overflows with green infrastructure. 2014. Bloomberg News Report issued on [10 March 2014]. 2

[21] Federal Emergency Management Agency. Flood Insurance Study, KNOX COUNTY, TENNESSEE, AND INCORPORATED AREAS . 24, 47, 61

[22] Egon Balas. An additive algorithm for solving linear programs with zero-one variables. *Operations Research*, 13(4):517–546, 1965. 53

[23] Masoud Barah, Anahita Khojandi, Xueping Li, Jon Hathaway, and OluFemi Omitaomu. Optimizing green infrastructure placement under precipitation uncertainty. 46

[24] Masoud Barah, Abbas Seifi, and James Ostrowski. Train scheduling problems. https://github.com/mbarah/Train_Scheduling_Problems, 2017. 60

[25] Masoud Barah, Abbas Seifi, and James Ostrowski. Decomposing train scheduling problem into integer-optimal polytopes. *Transportation Science*, 2018. viii, xi, 53, 55, 57, 58, 60, 63

[26] Mark A Benedict and Edward T McMahon. Green infrastructure. *Island, Washington, DC*, 2006. 36, 37

[27] Geoffrey M Bonnin, Deborah Martin, Bingzhang Lin, Tye Parzybok, Michael Yekta, and David Riley. Precipitation-frequency atlas of the united states. *NOAA atlas*, 14 (2), 2006. x, 14, 15, 17

[28] Edward H Bowman. The schedule-sequencing problem. *Operations Research*, 7(5): 621–624, 1959. 53

[29] Boyd Center Population Projections. Boyd Center Tennessee Population Projections: 2016-2070. http://tndata.utk.edu/sdcpopulationprojections.htm. Last accessed July 31, 2018. 44

[30] Ulf Brännlund, Per Olov Lindberg, Andreas Nõu, and J-E Nilsson. Railway timetabling using lagrangian relaxation. *Transportation science*, 32(4):358–369, 1998. 53

[31] MJ Cambez, J Pinho, and LM David. Using swmm 5 in the continuous modelling of stormwater hydraulics and quality. In *11th International Conference on Urban Drainage, Edinburgh, Scotland, UK*, 2008. 18

[32] Alberto Caprara, Matteo Fischetti, and Paolo Toth. Modeling and solving the train timetabling problem. *Operations research*, 50(5):851–861, 2002. 53

[33] Maria Raquel Catalano de Sousa, Franco Andre Montalto, and Patrick Gurian. Evaluating green infrastructure stormwater capture performance under extreme precipitation. *Journal of Extreme Events*, 3(02):1650006, 2016. 11

[34] Chein-Chi Chang, Kimberly Digiovanni, Gong Zhang, Xiahua Yang, and Shao-Hong You. sustainability. *Water Environment Research*, 87(10):1208–1255, 2015. 2

[35] Kyung-sook Choi and James E Ball. Parameter estimation for urban runoff modelling. *Urban Water*, 4(1):31–41, 2002. 6, 22, 82

[36] Tom Daniels. *Environmental Planning Handbook*. Routledge, 2017. 6

[37] Allen P Davis, William F Hunt, Robert G Traver, and Michael Clar. Bioretention technology: Overview of current practice and future needs. *Journal of Environmental Engineering*, 135(3):109–117, 2009. 12

[38] M De Luis, JC González-Hidalgo, J Raventós, JR Sánchez, and J Cortina. Distribución espacial de la concentración y agresividad de la lluvia en el territorio de la comunidad valenciana. *Cuaternario y Geomorfología*, 11(3-4):33–44, 1997. 14

[39] Anthony Dietrich, Rahul Yarlagadda, and Cyndee Gruden. Estimating the potential benefits of green stormwater infrastructure on developed sites using hydrologic model simulation. *Environmental Progress & Sustainable Energy*, 36(2):557–564, 2017. 18

[40] Loc Nguyen Doan and Allen P Davis. Bioretention–cistern–irrigation treatment train to minimize stormwater runoff. *Journal of Sustainable Water in the Built Environment*, 3(2):04017003, 2017. 6

[41] Stuart Echols and Eliza Pennypacker. *Artful rainwater design: creative ways to manage stormwater*. Island Press, 2015. 12

[42] Bradley Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992. 21

[43] Environmental Protection Agency (EPA). Opti-tool: Epa region 1's stormwater management optimization tool. https://www.epa.gov/tmdl/opti-tool-epa-region-1s-stormwater-management-optimization-tool, 2016. Accessed: 2018-06-22. 13, 43

[44] Thom Epps and Jon Mitchell Hathaway. sing spatially-identified effective impervious areato target green infrastructure retrofits: a modeling study in knoxville, tn. Submitted to Journal of Hydrology. 3

[45] Thom Epps and Jon Mitchell Hathaway. Establishing a framework for the spatial identification of effective impervious areas in gauged basins: Review and case study. *Journal of Sustainable Water in the Built Environment*, 4(2):05018001, 2018. 3

[46] Texas A& M AgriLife Extension. Rain Gardens for Stormwater Management . Technical report. 13

[47] Michael KH Fan, Andre L Tits, and John C Doyle. Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics. *IEEE Transactions on Automatic Control*, 36(1):25–38, 1991. 4

[48] C.B. Field. *Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change*. Cambridge University Press, 2012. 1

[49] George Elmer Forsythe, Cleve B Moler, and Michael A Malcolm. Computer methods for mathematical computations. 1977. 43

[50] Michael R Garey and David S Johnson. *Computers and intractability*, volume 29. wh freeman New York, 2002. 53

[51] Heather E Golden and Nahal Hoghooghi. Green infrastructure and its catchment-scale effects: an emerging science. *Wiley Interdisciplinary Reviews: Water*, 5(1):e1254, 2018. 6

[52] Paul J Goulart and Eric C Kerrigan. Relationships between affine feedback policies for robust control with constraints. *IFAC Proceedings Volumes*, 38(1):608–613, 2005. 4

[53] James CY Guo and Ben Urbonas. Volume-based runoff coefficients for urban catchments. *Journal of Irrigation and Drainage Engineering*, 140(2):04013013, 2013. 20

[54] Rui Guo and Yiping Guo. Discussion of "green infrastructure recovery: Analysis of the influence of back-to-back rainfall events" by bridget m. wadzuk, conor lewellyn, ryan lee, and robert g. traver. *Journal of Sustainable Water in the Built Environment*, 4(3):07018001, 2018. 61

[55] Kiichiro Hayashi. Economic incentives for green initiatives in nagoya city, japan. http://www.eea.europa.eu/atlas/teeb/economic-incentives-for-green-initiatives/, December 2010. 36

[56] Forrest M Hoffman, William W Hargrove Jr, David J Erickson III, and Robert J Oglesby. Using clustered climate regimes to analyze and compare predictions from fully coupled general circulation models. *Earth Interactions*, 9(10):1–27, 2005. 3

[57] Collin Homer, Jon Dewitz, Limin Yang, Suming Jin, Patrick Danielson, George Xian, John Coulston, Nathaniel Herold, James Wickham, and Kevin Megown. Completion of the 2011 national land cover database for the conterminous united states–representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5):345–354, 2015. x, xi, 24, 47

[58] Floyd A Huff. Time distribution of rainfall in heavy storms. *Water Resources Research*, 3(4):1007–1019, 1967. 14

[59] Anant Singh Jain and Sheik Meeran. A state-of-the-art review of job-shop scheduling techniques. Technical report, Technical report, Department of Applied Physics, Electronic and Mechanical . . . , 1998. 53

[60] Anant Singh Jain and Sheik Meeran. Deterministic job-shop scheduling: Past, present and future. *European journal of operational research*, 113(2):390–434, 1999. 53

[61] Suhyung Jang, Minock Cho, Jaeyoung Yoon, Yongnam Yoon, Sangdan Kim, Geonha Kim, Leehyung Kim, and Hafzullah Aksoy. Using swmm as a tool for hydrologic impact assessment. *Desalination*, 212(1-3):344–356, 2007. 18

[62] Kimberly M Jarden, Anne J Jefferson, and Jennifer M Grieser. Assessing the effects of catchment-scale urban green infrastructure retrofits on hydrograph characteristics. *Hydrological Processes*, 30(10):1536–1550, 2016. 6

[63] W. Junker. Flash flooding: A lesser known, but leading weather killer. The Washington Post, July 2013. Last accessed on February 14, 2016 https://www.washingtonpost.com/news/capital-weather-gang/wp/2013/07/11/flash-flooding-a-lesser-known-but-leading-weather-killer/. 2

[64] Aleksandra Kazmierczak and Jeremy Carter. Adaptation to climate change using green and blue infrastructure. a database of case studies. 2010. 36

[65] Hyomin Kim, Dong-Kun Lee, and Sunyong Sung. Effect of urban green spaces and flooded area type on flooding probability. *Sustainability*, 8(2):134, 2016. 2

[66] Pieter Klaassen. Financial asset-pricing theory and stochastic programming models for asset/liability management: A synthesis. *Management Science*, 44(1):31–48, 1998. 4

[67] Jia Liu, David J Sample, Cameron Bell, and Yuntao Guan. Review and research needs of bioretention used for the treatment of urban stormwater. *Water*, 6(4):1069–1099, 2014. 2

[68] Wen Liu, Weiping Chen, and Chi Peng. Assessing the effectiveness of green infrastructures on urban flooding reduction: A community scale study. *Ecological Modelling*, 291:6–14, 2014. 2

[69] Wen Liu, Weiping Chen, and Chi Peng. Influences of setting sizes and combination of green infrastructures on community's stormwater runoff reduction. *Ecological Modelling*, 318:236–244, 2015. 2

[70] Hugo A Loáiciga, K Majid Sadeghi, Sarah Shivers, and Shahram Kharaghani. Stormwater control measures: Optimization methods for sizing and selection. *Journal of Water Resources Planning and Management*, 141(9):04015006, 2015. 4

[71] St. Louis County Department of Planning. St. Louis County Meramec River Greenway Concept plan. https://www.stlouisco.com/Portals/8/docs/Document, 2003. Accessed on 9 February 2016 at 08:15AM. 36

[72] Matthew McCutcheon and Derek Wride. Shades of green: Using swmm lid controls to simulate green infrastructure. *Journal of Water Management Modeling, R246-15*, pages 289–301, 2013. 18

[73] R.I. McDonald, P. Green, D. Balk, B.M. Fekete, C. Revenga, M. Todd, and M. Montgomery. Urban growth, climate change, and freshwater availability. *Proceedings of the National Academy of Sciences*, 108(15):6312–6317, 2011. 2

[74] Jacob J McKee, Amy N Rose, Edward A Bright, Timmy Huynh, and Budhendra L Bhaduri. Locally adaptive, spatially explicit projection of us population for 2030 and 2050. *Proceedings of the National Academy of Sciences*, 112(5):1344–1349, 2015. xi, 43, 44

[75] Sara Meerow and Joshua P Newell. Spatial planning for multifunctional green infrastructure: Growing resilience in detroit. *Landscape and Urban Planning*, 159: 62–75, 2017. 6

[76] Trisha L Moore, John S Gulliver, Latham Stack, and Michael H Simpson. Stormwater management and climate change: vulnerability and capacity for adaptation in urban and suburban contexts. *Climatic change*, 138(3-4):491–504, 2016. 18

[77] "COINNEWS MEDIA GROUP LLC (COIN NEWS)". Current us inflation rates. https://www.usinflationcalculator.com/inflation/current-inflation-rates/. Last accessed July 31, 2018. 13, 43

[78] Sarah Nicholls and John L Crompton. The impact of greenways on property values: Evidence from austin, texas. *Journal of Leisure Research*, 37(3):321, 2005. 37

[79] CC Obropta and N Del Monaco. Reducing directly connected impervious areas with green stormwater infrastructure. *Journal of Sustainable Water in the Built Environment*, 4(1):05017004, 2017. 6

[80] City of Knoxville-Stormwater Engineering Division. http://www.knoxvilletn.gov/government/city_departments_offices/engineering/stormwater_engineering_division/. Accessed: 2018-06-22. 24

[81] John E Oliver. Monthly precipitation distribution: a comparative index. *The Professional Geographer*, 32(3):300–309, 1980. 14

[82] B.L. Preston. Local path dependence of U.S. socioeconomic exposure to climate extremes and the vulnerability commitment. *Global Environmental Change*, 23(4): 719–732, 2013. 1

[83] ProximityOne. Demographic Trends 2010-2060. http://proximityone.com/demographics2060.htm. Last accessed July 31, 2018. 44

[84] M. Ramshani, A. Khojandi, X. Li, and O. Omitaomu. Optimal planning of the joint placement of photovoltaic panels and green roofs under climate change uncertainty. *Omega*, 2018. URL https://doi.org/10.1016/j.omega.2018.10.016. 4

[85] David J Rosa, John C Clausen, and Michael E Dietz. Calibration and verification of swmm for low impact development. *JAWRA Journal of the American Water Resources Association*, 51(3):746–757, 2015. 18

[86] Lewis A Rossman. *Storm water management model user's manual, version 5.1*. National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency Cincinnati, 2015. 12, 18, 46

[87] David C Rouse and Ignacio F Bunster-Ossa. *Green infrastructure: a landscape approach*. Number 571. 2013. 6

[88] Maria Grazia Scutellà and Raffaella Recchia. Robust portfolio asset allocation and risk measures. *Annals of Operations Research*, 204(1):145–169, 2013. 4

[89] William Shuster and Stephen Dadio. Soils investigation for infiltration-based green infrastructure for sewershed management (omaha ne), 2014. 12

[90] Lawrence V Snyder. Facility location under uncertainty: a review. *IIE transactions*, 38(7):547–564, 2006. 4

[91] Allen L Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973. 11

[92] Eric Strassler, Jesse Pritts, and Kristen Strellec. Preliminary data summary of urban storm water best management practices. *United States Environmental Protection Agency, Office of Water.¡ http://www. epa. gov/waterscience/guide/stormwater/# nsbd*, 1999. 13, 43

[93] Robert G Traver and Ali Ebrahimian. Dynamic design of green stormwater infrastructure. *Frontiers of Environmental Science & Engineering*, 11(4):15, 2017. 12

[94] Vassilios A Tsihrintzis and Rizwan Hamid. Modeling and management of urban stormwater runoff quality: a review. *Water Resources Management*, 11(2):136–164, 1997. 18

[95] Vassilios A Tsihrintzis and Rizwan Hamid. Runoff quality prediction from small urban catchments using swmm. *Hydrological Processes*, 12(2):311–329, 1998. 18

[96] Bureau of Labor Statistics" "U.S. Department of Labor. Cpi inflation calculator. https://www.bls.gov/data/inflation_calculator.htm. Last accessed July 31, 2018. 13

[97] Ioan Voicu and Vicki Been. The effect of community gardens on neighboring property values. *Real Estate Economics*, 36(2):241–283, 2008. 37

[98] Susan Wachter. The determinants of neighborhood transformation in philadelphia: Identification and analysis: The new kensington pilot study. *University of Pennsylvania, Wharton School*, 2004. 37

[99] Susan M Wachter and Kevin C Gillen. Public investment strategies: How they matter for neighborhoods in philadelphia. *The Wharton School, University of Pennsylvania*, 2006. 37

[100] Bridget M Wadzuk, Conor Lewellyn, Ryan Lee, and Robert G Traver. Green infrastructure recovery: analysis of the influence of back-to-back rainfall events. *Journal of Sustainable Water in the Built Environment*, 3(1):04017001, 2017. 61

[101] Bridget M Wadzuk, Conor Lewellyn, Ryan Lee, and Robert G Traver. Closure to "green infrastructure recovery: Analysis of the influence of back-to-back rainfall events" by bridget m. wadzuk, conor lewellyn, ryan lee, and robert g. traver. *Journal of Sustainable Water in the Built Environment*, 4(3):07018002, 2018. 61

[102] Harvey M Wagner. An integer linear-programming model for machine scheduling. *Naval Research Logistics Quarterly*, 6(2):131–140, 1959. 53

[103] World Health Organization (WHO). Global health observatory: Urban population growth. http://www.who.int/gho/urban_health/situation_trends/, 2014. Situation in trends and key indicators made available by the WHO at [accessed January 28, 2016]. 1

[104] Jian Zhang, Guofu Ding, Yisheng Zou, Shengfeng Qin, and Jianlin Fu. Review of job shop scheduling research and its new perspectives under industry 4.0. *Journal of Intelligent Manufacturing*, pages 1–22, 2017. 53

[105] Qipeng P Zheng, Jianhui Wang, and Andrew L Liu. Stochastic optimization for unit commitment—a review. *IEEE Transactions on Power Systems*, 30(4):1913–1924, 2015. 4

Appendices

A Notation

Table A.1: The sets

Set	Description
V	Set of sub-catchments
A	Set of sub-catchment connectivity arcs
G^I	Set of large-scale GI practices
G^{II}	Set of small-scale GI practices
$G = G^I \cup G^{II}$	Set of GI practices
L	Set of available levels of installation of GI practices
Ψ	Set of projected precipitation time series for the watershed, referred to as scenarios
$\hat{\Psi}$	Set of projected daily precipitation time series for the watershed, produced by the CGCMs

Table A.2: The Parameters

Parameter	Description
T	Length of the planning horizon in years
$\overline{T}$	The year in which a precipitation scenario is realized, referred to as time to realize a scenario
π^{ψ}	Probability of scenario $\psi \in \Psi$
$a_{i'i} \in A$	An arc indicating that the upstream sub-catchment $i' \in V$ is connected to the downstream sub-catchment $i \in V$
$Q_i^{\psi,t}$	Total baseline surface runoff under scenario $\psi \in \Psi$ over sub-catchment $i \in V$ in year $t \leq T$ when no GI practice is placed
$2q_i^{\psi,t}(\alpha)$	The width of the 100(1-α)% CI for average baseline surface runoff volume within sub-catchment$i \in V$ under scenario $\psi \in \Psi$ in year $t \leq T$
$\hat{Q}_{i,j,l}^{\psi,t}$	Surface runoff captured by GI practice of type $j \in G$ installed in level $l \in L$ within sub-catchment $i \in V$ under scenario $\psi \in \Psi$ in year $t \leq T$. We also define $\hat{Q}_{i,0,0}^{\psi,t} = 0$.
$2\hat{q}_{i,j,l}^{\psi,t}(\alpha)$	The width of the 100(1-α)% CI for the average surface runoff captured by GI practice of type $j \in G$ installed in level $l \in L$ within sub-catchment $i \in V$ under scenario $\psi \in \Psi$ in year $t \leq T$
$\beta_{t',i',j',l'}^{t,i,j,l}$	Runoff 'adjustment factor' over the downstream sub-catchment $i \in V$, when a GI practice of type $j' \in G^I$ in level $l' \in L$ is placed within upstream sub-catchment $i' \in V$ and no GI practice or a GI practice of type $j \in G^I$ in level $l \in L$ is placed within the downstream sub-catchment $i \in V$
$C_{i,j}^t$	Per square feet present total cost of placing GI practice of type $j \in G$ within sub-catchment $i \in V$ in year $t \leq T$
$c_{i,j}^t$	Per square feet construction cost of a GI practice of type $j \in G$ in sub-catchment $i \in V$ in year $t \leq T$
$\mathfrak{c}_{i,j}^t$	Per square feet annual maintenance cost of a GI practice of type $j \in G$ in sub-catchment $i \in V$ in year $t \leq T$
r	Average annual inflation rate
$\delta_{i,j,l}$	Corresponding area (in square feet) of GI practice type $j \in G$ installed in level $l \in L$, within sub-catchment $i \in V$
η	Precipitation coefficient of variability

Table A.3: The Variables

Variable	Description
$x^t_{i,j,l}$	First stage binary decision variable indicating whether or not a GI practice of type $j \in G^I \subset G$ in level $l \in L$ is placed within sub-catchment $i \in V$ in year $t \leq \overline{T} - 1$
$z^{t',i',j',l'}_{t,i,j,l}$	First stage binary variable indicating whether or not GI practices of types $j', j \in G^I$ in levels $l', l \in L$ are placed within sub-catchment $i', i \in V$ at times $t', t \leq T$, respectively. We also define $z^{t',i',j',l'}_{t,i,0,l} = 0$.
$y^{\psi,t}_{i,j,l}$	Second stage binary decision variables indicating whether or not a GI practice of type $j \in G^{II} \subset G$ in level $l \in L$ is placed within sub-catchment $i \in V$ year t, $\overline{T} \leq t \leq T$

B Adjustment in Surface Runoff Reduction Due to GI Placement in Connected Sub-Catchments

Note that we assume 'adjustments' over downstream sub-catchments are additive. Hence, without loss of generality, here we simply present adjusting the runoff over the downstream sub-catchment i when a large-scale GI practice is placed within the single upstream sub-catchment $i' \in V$, $a_{i',i} \in A$.

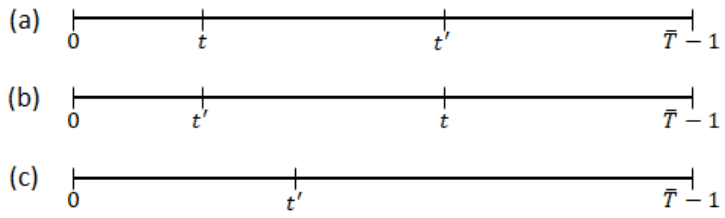

Figure B.1: The three cases to consider when accounting for surface run-off reduction over a downstream sub-catchment due to a GI practice placement upstream, where the downstream and upstream sub-catchments are placed in years t and t', respectively. Attention is restricted to large-scale practices only.

Figure B.1 presents the three cases to consider when accounting for surface run-off reduction over the downstream sub-catchment i due to placing a GI practice within the upstream sub-catchment i', $a_{i',i} \in A$, when accounting for large-scale practices only:

(a) GI practice of type j' in level l' is placed within upstream sub-catchment i' in year t' after GI practice of type j in level l is placed within downstream sub-catchment i in year t such that $0 \leq t \leq t' \leq \overline{T}-1$. In this case, run-off adjustment is needed only after the placement of a GI practice in the upstream sub-catchment i' in year t'. Hence the adjusted runoff reduction begins in year t';

(b) GI practice of type j' in level l' is placed within upstream sub-catchment i' in year t' before GI practice of type j in level l is placed within downstream sub-catchment i in year t such that $0 \leq t' \leq t \leq \overline{T}-1$. In this case, two levels of run-off adjustment are needed: The first adjustment is needed between years t' and $t-1$, and the second adjustment is needed on and after year t, i.e., after placing a GI practice in downstream sub-catchment i;

(c) GI practice of type j' in level l' is placed within upstream sub-catchment i' in year t' and no GI placed in downstream sub-catchment i by the beginning of year $\overline{T}$, i.e., $0 \leq t' \leq \overline{T}-1$: In this case, run-off adjustment over downstream sub-catchment i is needed on and after year t'.

First, consider the third term in the objective function (1.1), i.e.,

$$-\sum_{\{t''|\max\{t\cdot\mathbb{1}_{\{j\neq 0\}},t'\}\leq t''\leq T\}}\beta_{i,j,l}^{i',j',l'}\left(Q_i^{\psi,t''}\left(x_{i',j',l'}^{t'}-z_{t,i,j,l}^{t',i',j',l'}\right)+\hat{Q}_{i,j,l}^{\psi,t''}\cdot z_{t,i,j,l}^{t',i',j',l'}\right). \tag{B.1}$$

Equation (B.1) adjusts the run-off after placing a GI practice in the upstream sub-catchment i'. Also, consider the fourth term in the objective function (1.1), i.e.,

$$-\sum_{\{t''|t'\leq t''\leq t-1\}}\beta_{i,j,l}^{i',j',l'}\cdot Q_i^{\psi,t''}\cdot z_{t,i,j,l}^{t',i',j',l'}. \tag{B.2}$$

Equation (B.2) adjusts the run-off after placing a GI practice in the upstream sub-catchment i' if it occurs before placing a GI practice in the downstream sub-catchment i.

In case (a), the runoff adjustment over downstream sub-catchment i for the years in which GI practices are placed in both sub-catchments i' and i, i.e., in year t'' such that $t'\leq t''\leq T$ is given by equation (B.1), where the indicator function $\mathbb{1}_{\{j\neq 0\}}$ returns 1, and $\max\{t\cdot\mathbb{1}_{\{j\neq 0\}},t'\}$ returns t'. Note that $x_{i',j',l'}^{t'}-z_{t,i,j,l}^{t',i',j',l'}=0$ since both $x_{i',j',l'}^{t'}$ and $z_{t,i,j,l}^{t',i',j',l'}$ are equal to one. Therefore, in case (a) equation (B.1) simplifies as follows:

$$-\sum_{\{t''|t'\leq t''\leq T\}}\beta_{i,j,l}^{i',j',l'}\cdot\hat{Q}_{i,j,l}^{\psi,t''}.$$

Also, clearly, in case (a), equation (B.2) is not valid since $t\leq t'$.

Similarly, for case (b), for the years in which GI practices are placed in both sub-catchments i' and i, i.e., in year t'' such that $t\leq t''\leq T$, equation (B.1) simplifies as follows:

$$-\sum_{\{t''|t\leq t''\leq T\}}\beta_{i,j,l}^{i',j',l'}\cdot\hat{Q}_{i,j,l}^{\psi,t''}.$$

Also, for case (b), for the years in which the GI practice is placed in upstream sub-catchment i' and yet no GI is placed in downstream sub-catchment i, i.e., in year t'' such that $t'\leq t''\leq t-1$, Equation (B.2) is active and simplifies as follows:

$$-\sum_{\{t''|t'\leq t''\leq t-1\}}\beta_{i,j,l}^{i',j',l'}\cdot Q_i^{\psi,t''},$$

because $z_{t,i,j,l}^{t',i',j',l'}$ equals one.

Lastly, for case (c), for the years in which the GI practice is placed in upstream sub-catchment i', i.e., in year t'' such that $t'\leq t''\leq T$, the equation (B.1) simplifies as follows:

$$- \sum_{\{t'' | t' \leq t'' \leq T\}} \beta_{i,j,l}^{i',j',l'} \cdot Q_i^{\psi,t''}.$$

The indicator function $\mathbb{1}_{\{j \neq 0\}}$ returns 0 and hence, $\max\{t \cdot \mathbb{1}_{\{j \neq 0\}}, t'\}$ returns t'. Note that variable $x_{i',j',l'}^{t'}$ is equal to one and variable $z_{t,i,j,l}^{t',i',j',l'}$ is equal to zero as $j = 0$.

Also, note that equation (B.2) is equal to zero since variable $z_{t,i,j,l}^{t',i',j',l'}$ equals to zero in this case.

C Coupled Global Circulation Models (CGCMs)

Table C.1: Ten coupled global circulation models used for projecting future precipitation

Model Name	Institution
ACCESS: The Australian Community Climate and Earth-System Simulator [1]	Commonwealth Scientific and Industrial Research Organisation
BCC-CSM: Beijing Climate Center Climate System Model [2], referred to as 'BCC' in the text	Beijing Climate Center, China Meteorological Administration
CCSM4: The NCAR's Community Climate System Model [3]	Climate and Global Dynamics Laboratory (CGD) at the National Center for Atmospheric Research (NCAR)
CMCC-CM: The Centro Euro-Mediterraneo sui Cambiamenti Climatici Climate Model [4], referred to as 'CMCC' in the text	Euro-Mediterranean Center on Climate Change
FGOALS: Flexible Global Ocean–Atmosphere–Land System [7]	Institute of Atmospheric Physics, Chinese Academy of Sciences, State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics
GFDL-ESM2M: Geophysical Fluid Dynamics Laboratory Earth System Model [8], referred to as 'GFDL' in the text	Geophysical Fluid Dynamics Laboratory (Princeton University)
IPSL-CM5A: The Institut Pierre Simon Laplace Climate Model [10], referred to as 'IPSL' in the text	Institut Pierre Simon Laplace
MPI-ESM-MR: Max-Planck-Institute Earth System Model Mixed Resolution [11], referred to as 'MPI' in the text	Max Planck Institute for Meteorology
MRI-CGCM3: Japanese Meteorological Research Institute Coupled Global Climate Model [12], referred to as 'MRI' in the text	Meteorological Research Institute (MRI) of the Japan Meteorological Agency
NorESM1-M: Norwegian Earth System Model [16], referred to as 'NorESM' in the text	Multi-institutional, coordinated climate research in Norway

D Calculating Runoff Adjustment Factor Over a Downstream Sub-Catchment

We designed a set of experiments to calculate runoff adjustment factor over a downstream sub-catchment for any given pair of hydrologically connected sub-catchments. Consistent with the literature [35], we only use the most significant sub-catchment characteristics related to surface runoff in our experiments, namely, sub-catchments' percent of imperviousness, percent of slope, and Manning's n for overland flow over the pervious portion of the sub-catchment. Next, we use the values of these characteristics for the sub-catchments in the watershed of interest and calculate their corresponding quartiles. Accordingly, we stratify each characteristic into three categories of low, medium, and high, if the corresponding value is at or below the first quartile, between first and third quartiles, and above the third quartile. Table D.1 summarizes the combination of categories along with the number of observed sub-catchments within each one for our watershed of interest.

Consequently, we execute the SWMM model for all pairs of sub-catchments, given the average values for the categories in our watershed of interest. We run these simulations under the randomly selected precipitation events in the 'SWMM Simulation' step of the procedure described in Section 1.2.3. The runoff adjustment factor over a downstream sub-catchment is then estimated as the average difference in runoff coefficient in the sub-catchment over all precipitation events when a certain GI practice is placed within the upstream sub-catchment and no GI is placed there (i.e., no treatment). Figure D.1 presents the heat map of runoff adjustment factors over the downstream sub-catchment, given all observed combinations of sub-catchment characteristics' categories in the watershed of interest.

Table D.1: Summary characteristics of the 140 sub-catchments within our watershed of interest, categorized by percent of imperviousness, percent of slope, and Manning's n for overland flow over the pervious portion of the sub-catchment.

% of Imp.	% of Slope	Manning's n	Number of Sub-catchments	Average Area (acres)	Average Imp. (%)	Average Slope (%)	Average Manning's n
low	low	med	2	117.95	7.85	1.56	0.248
low	low	high	9	153.24	4.64	1.68	0.281
low	med	high	8	139.16	3.94	2.86	0.290
low	high	high	16	98.50	4.78	6.59	0.282
med	low	med	2	92.14	11.60	1.46	0.250
med	low	high	2	80.65	8.80	1.86	0.262
med	med	med	3	154.36	10.43	2.74	0.255
med	med	high	9	76.10	9.69	2.75	0.272
med	high	med	5	85.88	10.62	7.29	0.248
med	high	high	13	90.97	9.47	8.27	0.266
high	low	low	8	66.28	20.53	1.58	0.211
high	low	med	7	139.53	16.57	1.58	0.247
high	low	high	5	147.74	16.62	1.78	0.283
high	med	low	10	88.49	23.39	3.02	0.219
high	med	med	3	43.49	18.93	2.46	0.247
high	med	high	2	121.45	23.15	3.15	0.283
high	high	low	18	67.87	19.82	5.11	0.211
high	high	med	12	90.44	14.36	6.70	0.247
high	high	high	6	121.26	17.87	6.09	0.282

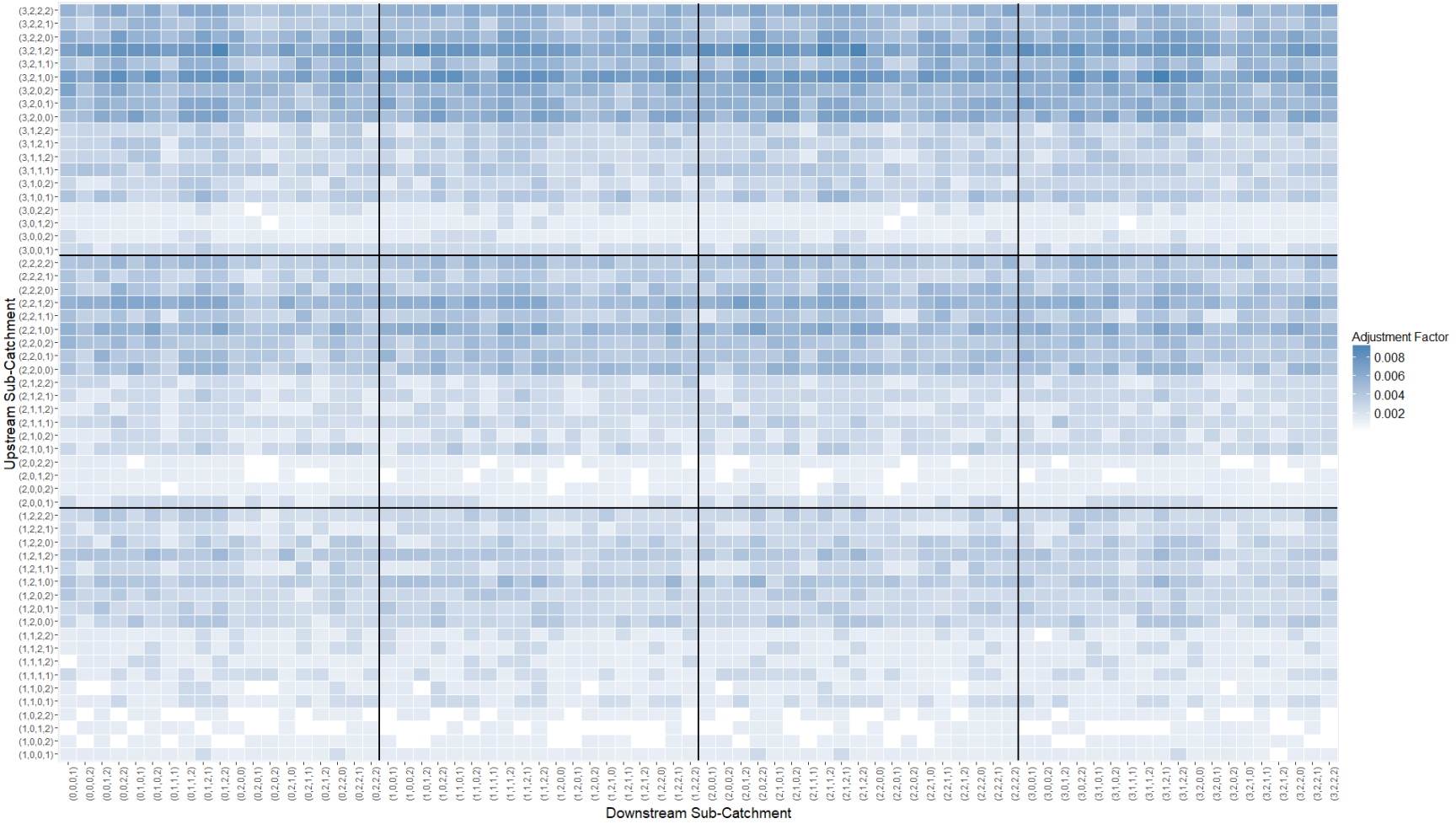

Figure D.1: Heat map of the runoff adjustment factors over the downstream sub-catchment, given all observed combinations of sub-catchment characteristics' categories in the watershed of interest. Sub-catchment characteristics are shown as tuples, where the first element corresponds to the GI level – 0 encodes no treatment and 1-3 refer to the levels low, medium, and large bioretentions, respectively. Elements 2-4 of the tuple correspond to percent of imperviousness, percent of slope, and Manning's n for overland flow over the pervious portion of the sub-catchment, each of which are categorized into three levels of 0-2, encoding low, medium, and high, respectively.

E Summary of the Characteristics of the Sub-catchments in First Creek, Knoxville, Tennessee

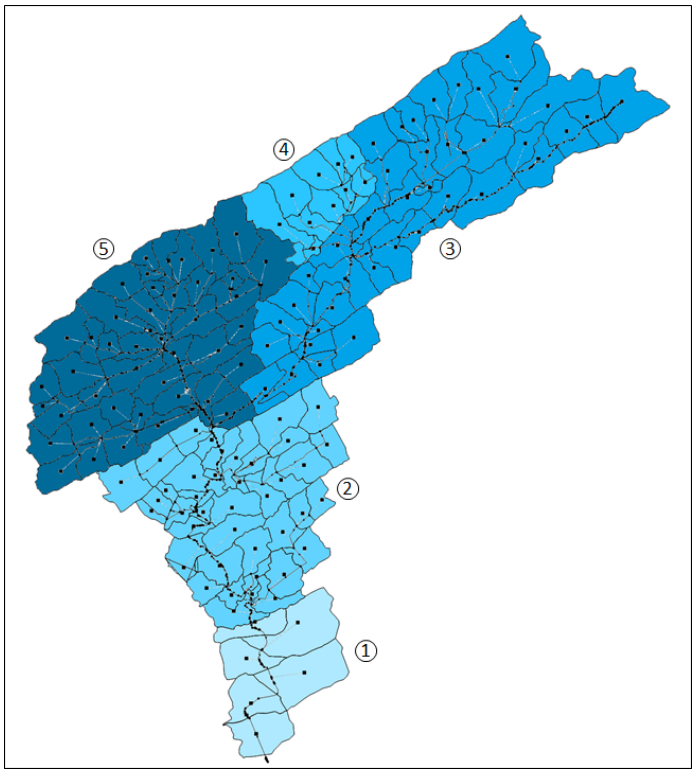

Figure E.1: Watershed of First Creek, grouped based on similarities in sub-catchment characteristics.

Table E.1: Summary of the characteristics of the sub-catchments in First Creek as labeled in Figure E.1.

Region	Total Area (Acres)	Average Impervious Area (%)	Average slope (%)
1	1292.05	23.36	3.65
2	3187.31	18.01	4.74
3	4915.84	8.07	3.83
4	807.43	8.31	6.12
5	3745.82	12.61	4.31

Vita

Masoud Barah was born on July 26, 1986 in Tehran, Iran. He graduated from the Khaje Nasir University of Technology in 2009 with a Bachelor of Science degree in Industrial Engineering. In Fall 2009, he attended Amirkabir University of Technology as a Master of Science student in Industrial and Systems Engineering and started working on various problems in operations research. Specifically, he concentrated on application of graph theory and integer programming in railway industry. His research work in the Amirkabir University of Technology led to one conference proceeding, and two book chapters on the application of graph theory in Industrial Engineering. In March 2012, he received a Master of Science degree in Industrial Engineering and started working as Operations Research Analyst in industry. In August 2015 he moved to the U.S. and started his Ph.D. program in Industrial Engineering in the University of Tennessee, Knoxville. At the same time, he started working as a Graduate Research Assistant in the University of Tennessee, Department of Industrial Engineering where he spent most of his time on research.

Masoud Barah's dissertation is on the application of operations research in environmental engineering, with an emphasis on improving urban infrastructure resilience under uncertainty. His research interests also lie in applied operations research and decision making under uncertainty with primary focus on data-driven healthcare problems and healthcare informatics. His methodological interests include data analysis, graph theory, integer programming, and Markov decision process, and stochastic programming.