Proceedings of the IISE Annual Conference & Expo 2026
Y. Xiang, D., Yu, R. Thiesing, eds.

Optimizing Park Locations with Imperfect Information

Gabriel Griffin
University of Oklahoma
Norman, Oklahoma, USA

Anabel Fletcher
Smith College
Northampton, Massachusetts, USA

Dakota Goto
Fordham University
New York, New York, USA

Martha L. Sabogal De La Pava, Melika Jahan Beikloo, David White, Emily L. Tucker
Clemson University
Clemson, South Carolina, USA

Abstract

Multiple studies have focused on optimal locations for the preservation and development of greenspace. These studies are wide-ranging in their objectives and represent the breadth of benefits that greenspace provides for a community and its environment. Communities and subpopulations vary in their priorities, and these differences impact which locations are selected. In practice, publicly available databases provide different indices of benefit and access. Our work attempts to model the context in which park planners make decisions based on the data presented to them by these databases and to understand the regret of making decisions on curated data that may imperfectly reflect priorities. We create mutually exclusive service areas for all potential parks based on residents for whom it would become the primary park if created. We then define a utility function for each interest with common parameters whose values differ depending on the priorities of the community, as described by the indices. Using these, we present an optimization model that minimizes regret and evaluates against varying levels of planner knowledge. We conduct numerical experiments on a case study of Greenville County, South Carolina, to recommend selections of open lots for the development of new parks. Data on community priorities are gathered from the CDC and Trust for Public Land.

Keywords
Park Planning, Uncertainty, Regret Optimization, Facility Location, Public Policy Optimization

1. Introduction

The benefits offered by greenspace are well studied, but their breadth and distribution across communities are not currently adequately measured [1]. Consequently, park planners have historically relied on data that represents only a subset of these benefits to make choices to distribute greenspace across communities.

The limitations of available data bring challenges. Uncertainty in true priorities and needs can be understood as the gap between what is known and what should be known to make optimal decisions while minimizing risk exposure [2]. Risk can be quantified and managed with probabilistic modeling methods, though uncertainty may be more difficult to communicate to decision-makers when it arises due to ambiguities in the available data. Developing tools to improve how imperfect data are used to guide park decisions could improve park access and corresponding community benefit.

In recent years, work has been done to develop analytical practices and mathematical models that optimize park and greenspace placement [3]. However, these models rely on parameter choices aligned with specific objectives, such as preserving land for wildlife conservation or reducing walking distances from residents to selected parks [3–5]. This paper develops a modeling framework that treats an imperfect data source (hereafter, index choice) as a source of uncertainty about which data source best represents the needs of a community. Because parks ultimately serve all residents regardless of the intent behind their selection, we can quantify the regret of optimizing to an index that is not the best representation of need as the difference in utility between optimal choices of that index and any other given from multiple publicly available datasets.

2. Problem Description

Park planners must select a subset of open lots as candidates for development subject to budget constraints. Each potential park location serves a surrounding residential population; for the purpose of this research, we assume that residents are served only by their closest, or primary, park. The goal is to select parks that maximize overall community benefit as measured through data describing social, environmental, and greenspace needs.

In practice, benefit is represented through publicly available indices published by external organizations. They differ in their underlying assumptions, spatial resolution, and emphasis on specific population characteristics. Therefore, park planners are not only challenged by the absence of data, but also by competing data sources. Making choices based on a particular index could represent opportunity costs for residents not prioritized. A strategy for understanding these possible sub-optimalities and making park location choices to minimize them will aid park planners in making more holistic decisions.

3. Related Research

Urban greenspace placement has been widely studied using Geographic Information Systems (GIS) and optimization methods. Early work focuses on accessibility and spatial distribution of urban parks [4], while later studies incorporate multi-objective optimization to balance social, environmental, and health goals [3, 6]. Research also emphasizes the role of urban morphology and demographic patterns in park design [1, 7, 8]. These studies show the value of formal optimization over heuristics but typically assume that planner priorities can be captured through a single, well-defined objective. Models are simplifications, and their recommendations depend on the assumptions used [9, 10]. Solutions that optimize a single objective may under-perform if the assumptions do not reflect reality. Regret-based methods, which evaluate decisions against alternative outcomes, have been applied in facility location and planning [11], though they have not yet been applied in park siting. There is the potential to apply these ideas in park siting to allow planners to make decisions that remain robust across multiple indices of community need. Research has been conducted to develop composite indices to represent social, environmental, or health priorities when allocating park resources [12–14]. These indices differ in their assumptions, spatial resolution, and emphasis on population characteristics, yielding conflicting policy recommendations. Most models assume that the chosen index accurately reflects community need, without evaluating the consequences of selecting one index over another [1, 4]. This paper introduces the first approach to consider regret-optimization in park siting decisions.

4. Methodology

There are multiple data sources for community scoring systems, i.e., indices that represent how changes in organizational focus can change outcomes in park selection. For this analysis, we consider three: the Trust for Public Land's (TPL) Park Score index, the Centers for Disease Control and Prevention's (CDC) Social Vulnerability Index (SVI), and the CDC's Environmental Justice Index (EJI) [15–17].

Equation (1) defines the priority of a park (deviation from 0) within an index as follows. It is the priority score from an index of all residential locations it serves, q_{il}, multiplied by the number of people in those residential locations, n_l, the value of the land for creating a park, v_{kl}, and a binary variable to only include the information of the people for whom it would become the primary park, s_{kl}. This is an adaptation of the approach in [3]. Any consideration of park underutilization due to sociological factors was considered outside the scope of the project.

$$p_{ik} = \sum_{l \in L} [(q_{il} n_l (1 - v_{kl})) s_{kl}] \qquad \forall i \in I, k \in k \tag{1}$$

4.1 Indices

The parameter q_{il} $\forall i \in I, l \in L$ is determined as an aggregate of chosen parameters, which are different for each index. For the purposes of this study, we considered resident locations to be census blocks, which was the most granular level for which data was available. Where data at the block level did not exist, we assumed that each block shared the same characteristics as the block group or tract to which they belonged. This assumption risks falling victim to the scale effect of the modifiable area unit problem but was necessary given the lack of data at the block level. The methods for calculating the aggregate values for all three indices are based on the technical documentation for their respective datasets.

For the TPL index of each location $l \in L$, $q_{TPL,l}$ represents an aggregate demographic score composed of six rankings: population density, heat, air pollution, people of color per acre, percent of low-income households, and health score. The data for $q_{SVI,l}$ at the tract level are available directly through the Social Vulnerability Index dataset, available through the CDC's Agency for Toxic Substances and Disease Registry (ATSDR). The methods for calculating q_{il} were taken from the technical documentation for the SVI dataset. The SVI groups population characteristics around four themes: socioeconomic status, household characteristics, racial and ethnic minority status, and housing type and transportation. Each characteristic is a percentage of a census tract.

Data for $q_{EJI,l}$ at the tract level are available directly through the Environmental Justice Index dataset, through the CDC's ATSDR. The methods for calculating q_{il} were taken from the EJI dataset's technical documentation. The EJI groups population characteristics around three modules: social vulnerability, environmental burden, and health vulnerability. Most characteristics are expressed as a percentage of a census tract. Social vulnerability here is related to the SVI that we previously used. However, there are some variables included in the Social Vulnerability Index that are not included in the EJI's Social Vulnerability module.

4.2 Model

Table 1: Notation.

Sets and Indices	
I	set of indices
K	set of all park locations
L	set of resident locations
Parameters	
p_{ik}	the priority value of park location $k \in K$ according to index $i \in I$ for all resident locations for which it is the primary park
s_{kl}	$= \begin{cases} 1 & \text{if park } k \in K \text{ is the primary park of resident location } l \in L \\ 0 & \text{otherwise} \end{cases}$
q_{il}	aggregate demographic emphases of resident location $l \in L$ according to data set $i \in I$
n_l	number of people living in resident location $l \in L$
v_{kl}	value of park location $k \in K$ for resident locations $l \in L$ for which it is the closest park
c_k	cost of park location $k \in K$
b	budget allowed for final park selections
Decision Variables	
y_{ik}	$= \begin{cases} 1 & \text{if park location } k \in K \text{ is selected for construction by index } i \in I \\ 0 & \text{otherwise} \end{cases}$
R_i	the max regret of parks chosen by index $i \in I$

We consider a two-phase decision framework. In the first phase, a knapsack problem is solved independently for each index using their priority set and a fixed budget. That yields an optimal park decision set for each index and budget scenario. The second phase applies a minimax regret framework for the optimal decision sets of each index for all budget scenarios. Regret is defined as the difference in objective value between an index's optimal decision set and the value obtained when decisions are optimized using the priorities of a different index. This represents a decision-making context where park planners must choose a single index to make decisions beforehand and can only choose an index to recommend upon rather than intelligently choose a set of parks which lowers regret to a global minimum.

Phase 1 (Knapsack for each index i):

$$\max_{y} \quad \sum_{k\in K} p_{ik}\, y_{ik} \tag{2}$$

Subject to:

$$\sum_{k\in K} c_k\, y_{ik} \le b \tag{3}$$

$$y_{ik} \in \{0,1\} \qquad \forall k \in K \tag{4}$$

Phase 2 (Regret Minimization):

$$\min_{R} \quad \sum_{i\in I} R_i \tag{5}$$

Subject to:

$$R_i \;\ge\; \sum_{k\in K} p_{i'k} y_{i'k} - \sum_{k\in K} p_{i'k} y_{ik} \qquad \forall i, i' \in I \tag{6}$$

$$R_i \ge 0 \qquad \forall i \in I \tag{7}$$

The models were solved for 5,000 budget cases from \$0 to \$8,000,000 (step size of \$1600) in Gurobi version 12.0.1.

5. Results

Figure 1 shows the distribution of priority across Greenville County, South Carolina, by index. These indices differ in key areas. First, their granularity is different. Therefore, making recommendations from comparisons is difficult without reducing the ability of some indices to represent stakeholders as intended. However, recognized need often spans across index. From the parks considered, Figure 1 shows the parks chosen using the priorities of each index, at an example budget of \$333,568.

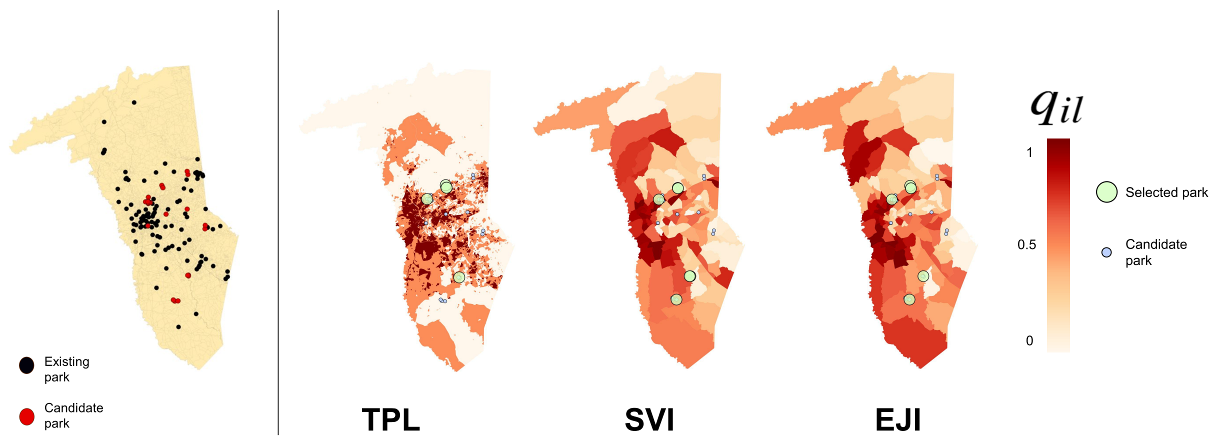

Figure 1: Priority Map and Chosen Park by Index

Figure 2 shows how the index chosen has little effect in most budgets considered. The percent of cases of budgets considered where index choice made no difference whatsoever was highest, followed by other ties and then TPL.

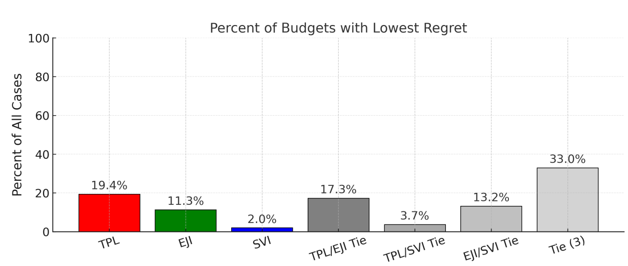

Figure 2: Percent Times Index had Lowest Regret Across Budgets Considered

Figure 3 shows the distribution of regret across the indices and budgets considered. The mean values have little difference across indices, but their maximum regret can be very large.

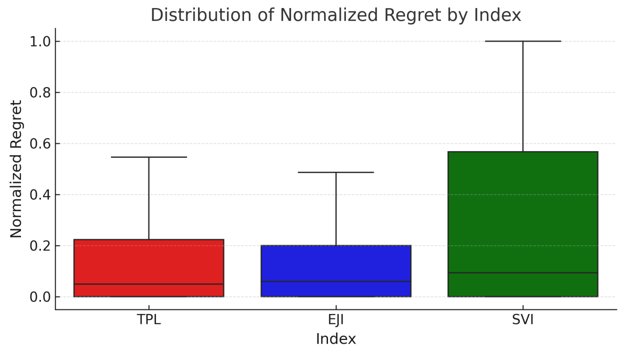

Figure 3: Regret Values for Each Index Across Budgets Considered

6. Conclusions

This work presents a regret-based optimization framework for the selection of park locations that treats index choice as a source of uncertainty rather than a fixed modeling assumption. By modeling how park planners make decisions using publicly available indices, we show that different, defensible representations of community need can lead to different park selections and, consequently, different distributions of benefit across residents.

Our results demonstrate that while index choice often does not change community access at many budget levels, there are scenarios in which regret can be substantial. These high-regret cases highlight the potential cost of committing to a single index that imperfectly reflects community priorities. The proposed minimax regret approach provides a principled way for planners to make decisions that remain consistent across competing representations of need, rather than optimizing narrowly to one perspective.

The case study of Greenville County, South Carolina illustrates how this framework can be applied in practice using widely available datasets from the CDC and the Trust for Public Land. By integrating service areas, demographic priorities, and budget constraints, the model produces park selections that balance tradeoffs between social, environmental, and health-oriented objectives.

Future work could extend this framework by incorporating partial or weighted combinations of indices, uncertainty in assigning populations to primary parks, or dynamic planning horizons where park development occurs over time. Furthermore, it could attempt to model a decision-making context in which park planners have knowledge of regret while choosing park locations. More broadly, this approach demonstrates how optimization models can support public planning decisions when uncertainty arises not from random variability but from competing priorities for what should be optimized.

Acknowledgements

Special thanks to Clemson IE Human-Centered Operations Research and Engineering (H-CORE) Research Experiences for Undergraduates (REU) program for support during the research project. We are grateful for the site leadership of Drs. Thomas Sharkey and Jessica Avilés as well as the members of the Tucker Research Group. This material is based upon work supported by the National Science Foundation under Grant Number 2348774. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] M. H.E.M. Browning et al. “Measuring the 3-30-300 rule to help cities meet nature access thresholds”. In: *Science of The Total Environment* 907 (Jan. 2024), p. 167739. ISSN: 0048-9697. DOI: 10.1016/J.SCITOTENV.2023.167739. URL: https://www.sciencedirect.com/science/article/pii/S0048969723063660.

[2] Efstratios. Nikolaidis, Dan M.. Ghiocel, and Suren. Singhal. "Engineering design reliability handbook". In: (2005). URL: https://books.google.com/books/about/Engineering_Design_Reliability_Handbook.html?id=gdHKBQAAQBAJ.

[3] Anisa Young et al. "An optimization approach to improve equitable access to local parks". In: *Socio-Economic Planning Sciences* 92 (Apr. 2024), p. 101826. ISSN: 0038-0121. DOI: 10.1016/J.SEPS.2024.101826. URL: https://experts.illinois.edu/en/publications/an-optimization-approach-to-improve-equitable-access-to-local-par/.

[4] Kyushik Oh and Seunghyun Jeong. "Assessing the spatial distribution of urban parks using GIS". In: *Landscape and Urban Planning* 82.1-2 (Aug. 2007), pp. 25–32. ISSN: 0169-2046. DOI: 10.1016/J.LANDURBPLAN.2007.01.014. URL: https://www.sciencedirect.com/science/article/pii/S0169204607000345.

[5] Alessandro Rigolon. "A complex landscape of inequity in access to urban parks: A literature review". In: *Landscape and Urban Planning* 153 (Sept. 2016), pp. 160–169. ISSN: 0169-2046. DOI: 10.1016/J.LANDURBPLAN.2016.05.017. URL: https://www.sciencedirect.com/science/article/pii/S0169204616300846.

[6] M. N. Neema and A. Ohgai. "Multi-objective location modeling of urban parks and open spaces: Continuous optimization". In: *Computers, Environment and Urban Systems* 34.5 (Aug. 2010), pp. 359–376. ISSN: 0198-9715. DOI: 10.1016/J.COMPENVURBSYS.2010.03.001. URL: https://www.sciencedirect.com/science/article/pii/S0198971510000207.

[7] Steffen Lehmann. "Research Methods in Urban Design: A Framework for Researching the Performance and Resilience of Places †". In: *Buildings* 13.6 (June 2023). ISSN: 20755309. DOI: 10.3390/BUILDINGS13061548.

[8] Huaqing Wang and Louis G. Tassinary. "Effects of greenspace morphology on mortality at the neighbourhood level: a cross-sectional ecological study". In: *The Lancet Planetary Health* 3.11 (Nov. 2019), e460–e468. ISSN: 2542-5196. DOI: 10.1016/S2542-5196(19)30217-7. URL: https://www.sciencedirect.com/science/article/pii/S2542519619302177.

[9] M Pidd. "Why modelling and model use matter". In: *Journal of the Operational Research Society* (2010), pp. 14–24. DOI: 10.1057/jors.2009.141. URL: www.palgrave-journals.com/jors/.

[10] Simon French. "From soft to hard elicitation". In: (). DOI: 10.1080/01605682.2021.1907244. URL: https://www.tandfonline.com/action/journalInformation?journalCode=tjor20.

[11] Jean Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. "Regret in online combinatorial optimization". In: *Mathematics of Operations Research* 39.1 (Feb. 2014), pp. 31–45. ISSN: 15265471. DOI: 10.1287/MOOR.2013.0598;CSUBTYPE:STRING:PERIODICAL;TAXONOMY:TAXONOMY:ACM-PUBTYPE;PAGEGROUP:STRING:PUBLICATION. URL: https://dl.acm.org/doi/10.1287/moor.2013.0598.

[12] Johanna Kujala et al. "Engaging with the Natural Environment". In: *Proceedings of the International Association for Business and Society* 30 (2019), pp. 73–81. DOI: 10.5840/IABSPROC20193010. URL: https://philpapers.org/rec/KUJEWT.

[13] D. Caparros-Midwood, S. Barr, and R. Dawson. "Optimised spatial planning to meet long term urban sustainability objectives". In: *Computers, Environment and Urban Systems* 54 (Nov. 2015), pp. 154–164. ISSN: 0198-9715. DOI: 10.1016/J.COMPENVURBSYS.2015.08.003. URL: https://www.sciencedirect.com/science/article/pii/S0198971515300090.

[14] Andrew T. Kaczynski et al. "ParkIndex: Validation and application of a pragmatic measure of park access and use". In: *Preventive Medicine Reports* 20 (Dec. 2020), p. 101218. ISSN: 2211-3355. DOI: 10.1016/J.PMEDR.2020.101218. URL: https://www.sciencedirect.com/science/article/pii/S2211335520301765.

[15] *ParkScore Home: Trust for Public Land*. URL: https://www.tpl.org/parkscore.

[16] *Social Vulnerability Index | Place and Health - Geospatial Research, Analysis, and Services Program (GRASP) | ATSDR*. URL: https://www.atsdr.cdc.gov/place-health/php/svi/index.html.

[17] *Environmental Justice Index | Place and Health - Geospatial Research, Analysis, and Services Program (GRASP) | ATSDR*. URL: https://www.atsdr.cdc.gov/place-health/php/eji/index.html.