RAND CORPORATION

Research Report

MICHELLE E. MIRO, SUSAN A. RESETAR, KELLY HYDE, JOSHUA STEIER, MICHAEL T. WILSON, ZARA FATIMA ABDURAHAMAN, VANESSA WOLF

Demographic and Geographic Characteristics of Green Stormwater Infrastructure Investments in Five U.S. Cities

A Machine Learning–Based Analysis

KEY FINDINGS

- Although no cities we examined were consistently and deliberately planning based on a co-benefits framework, areas that were more strongly associated with green stormwater infrastructure had higher percentages of Hispanic or Latino residents, lower percentages of White residents, and higher percentages of residents who self-reported coronary health challenges.
- City stormwater planning has predominately been consent decree–driven and publicly funded with a focus on reducing stormwater volumes. It is now, however, shifting toward an effort that is driven by compliance with municipal codes and that includes more private citizens. Ensuring that these more-distributed approaches carefully consider co-benefits could enhance the impact of stormwater management moving forward.
- Evaluating multiple machine learning approaches was advantageous in producing a better fit model for a given application and enhancing analytical flexibility. Modelers should also take care to understand the limitations, resolution, and relationships in their data and start simply before adding analytical complexity.
- There is a growing appetite for advanced analytics in municipal planning, but machine learning analysts must have expertise in the meaning and application of the datasets they work with, must be able to logically explain the results of their analyses, and should follow best practices for communicating about artificial intelligence (of which machine learning is a subset) to ensure the utility of the analysis and tools they produce.

Over the past few decades, stormwater managers across the United States have increasingly turned to green stormwater infrastructure to mitigate flooding and improve the quality of stormwater runoff. Green stormwater infrastructure offers an alternative to traditional gray infrastructure approaches, which typically involve the use of pipes, pumps, and other built or mechanical structures to capture, convey, and treat stormwater runoff. Gray infrastructure systems often include components such as underground storage tanks, detention basins, concrete channels, or sewer systems. Green stormwater infrastructure instead involves using natural or nature-inspired elements to manage stormwater, such as green roofs, rain gardens, or vegetated swales. This approach aims to mimic natural hydrological processes by replacing impervious (nonporous) surfaces with pervious soils or vegetation designed to slow and absorb more water. These changes to an urban landscape can meaningfully reduce the amount of stormwater runoff and improve the water quality of nearby waterways. For example, a city block largely composed of impervious surfaces can yield more than five times as much runoff as a comparably sized highly pervious area, and this runoff can carry pollutants, such as oil, pesticides, bacteria, and sediment, and nutrients into streams, rivers, and lakes.[1]

There are many types of green stormwater infrastructure. Although several categorizations of green stormwater infrastructure exist, the following represents a generalized variety of green stormwater infrastructure types. Figure 1 shows two examples of these types: a rain garden (left) and bioswale (right).

- *Rain gardens* are shallow depressions planted with native vegetation to capture and filter stormwater runoff, allowing it to infiltrate into the ground.
- *Bioswales* are vegetated channels or ditches designed to slow and filter stormwater runoff as it travels through the landscape.
- *Permeable pavement* is porous surfaces that allow water to pass through and infiltrate the ground, reducing runoff and preventing flooding.
- *Green roofs* are vegetated surfaces on rooftops that are designed to capture rainwater, reduce heat absorption, and improve air quality.
- *Urban trees and forests* involve the planting and maintaining of trees and forests in urban areas that can absorb rainwater, improve air quality, and provide shade.
- *Engineered wetlands* are engineered systems designed to replicate the functions of natural wetlands, effectively filtering pollutants and improving water quality.
- *Pervious open spaces* involve developing parks, plazas, and open areas designed to allow water infiltration and reduce runoff.

Green stormwater infrastructure can be implemented at various scales, such as individual properties or entire neighborhoods or cities, and by a variety of public and private actors. In addition to supporting stormwater management, green stormwater infrastructure can provide health, ecological,

Abbreviations

ACS	American Community Survey
AI	artificial intelligence
BWSC	Boston Water and Sewer Commission
CEJST	Climate and Economic Justice Screening Tool
CSO	combined sewer overflow
DEP	Department of Environmental Protection (New York City)
DOEE	Department of Energy and Environment
DWSD	Detroit Water and Sewerage Department
EPA	U.S. Environmental Protection Agency
GBDT	Gradient Boosted Decision Trees
LIME	Lasso Interpretable Model-agnostic Explanations
LOESS	locally estimated scatterplot smoothing
MS4	municipal separate storm sewer system
NPDES	National Pollutant Discharge Elimination System
PWSA	Pittsburgh Water and Sewer Authority
RMSE	Root Mean Square Error
SHAP	SHapley Additive exPlanations
SVM	Support Vector Machine
SVR	Support Vector Regressor
TMDL	Total Maximum Daily Load
WGS84	World Geodetic System 1984

FIGURE 1

Examples of Green Stormwater Infrastructure

SOURCES: Sonnenfeld, “Green Infrastructure (Water Retention), SUNY-ESF, Syracuse, NY”; and Psa1966, “Bioretention Area in Emeryville, California.”
NOTES: The left image shows a rain garden the State University of New York College of Environmental Science and Forestry in Syracuse, New York. The right image shows a bioswale in Emeryville, California.

and economic co-benefits to its surrounding areas. These include such co-benefits as those affecting culture and aesthetics, biodiversity and habitats, air purification, public education, and heat island reduction.[2] Figure 2 provides a more complete mapping of green stormwater infrastructure types with a broad set of co-benefits. Green stormwater infrastructure can also enhance equitable outcomes through the provision of co-benefits. For example, infrastructure installed in environmental justice communities is more likely to see enhanced co-benefits, because these areas often have fewer green spaces and suffer greater environmental burdens. Processes to plan and install green stormwater infrastructure can also be shaped to enhance structural and procedural equity.

Although green stormwater infrastructure is increasingly used to manage stormwater in the urban landscape, its co-benefits are often not a core consideration in decisionmaking processes for infrastructure siting.[3] Research on green stormwater infrastructure deployment across the United States has overwhelmingly shown that the driving force behind investments have been regulatory,[4] either through consent decrees on combined stormwater and wastewater systems or through municipal separate storm sewer system (MS4) National Pollutant Discharge Elimination System (NPDES) permits.[5] This finding means that, despite the co-benefits offered by green stormwater infrastructure, cities might be optimizing siting locations for green stormwater infrastructure investments based on regulatory requirements alone and therefore not realizing the full benefits these investments could provide. Additionally, by focusing only on stormwater requirements and cost-benefit calculations, investments might be worsening the existing inequities among communities.

Researchers have begun examining how green stormwater infrastructure planning and siting can be more equitable and more likely to holistically benefit communities beyond just stormwater improvements.[6] However, more work is needed to understand the existing landscape of green stormwater infrastruc-

FIGURE 2

Green Stormwater Infrastructure Types and Their Co-Benefits

Benefit	Reduces stormwater runoff											Improves community livability						
	Reduces water treatment needs	Improves water quality	Reduces Gray infrastructure needs	Reduces flooding	Increases available water supply	Increases groundwater recharge	Reduces salt use	Reduces energy use	Improves air quality	Reduces atmospheric CO_2	Reduces urban heat island	Improves aesthetics	Increases recreational opportunity	Reduces noise pollution	Improves community cohesion	Urban agriculture	Improves habitat	Cultivates public education opportunities
Practice										CO_2								
Green roofs	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Maybe	Yes	Maybe	Maybe	Yes	Yes
Tree planting	Yes	Yes	Yes	Yes	No	Maybe	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Maybe	Yes	Yes
Bioretention and infiltration	Yes	Yes	Yes	Yes	Maybe	Maybe	No	No	Yes	Yes	Yes	Yes	Yes	Maybe	Maybe	No	Yes	Yes
Permeable pavement	Yes	Yes	Yes	Yes	No	Maybe	Yes	Maybe	Yes	Yes	Yes	No	No	Yes	No	No	No	Yes
Water harvesting	Yes	Yes	Yes	Yes	Yes	Maybe	No	Maybe	Maybe	Maybe	No	No	No	No	No	No	No	Yes

Yes Maybe No

SOURCE: Reproduced from Center for Neighborhood Technology and American Rivers, *The Value of Green Infrastructure: A Guide to Recognizing Its Economic, Environmental, and Social Benefits*, 2010.

ture across the United States; specifically, if that infrastructure is located in areas that have the potential to achieve a broad variety of co-benefits, such as those locations that experience greater urban heat island effects, suffer from poorer air quality, or lack green spaces for recreation. Such an evaluation can inform future policy and planning aimed at enhancing these co-benefits.

Machine Learning for Policy Analysis

Artificial intelligence (AI), which includes machine learning, is emerging as a major disruptor to both public and private life. For policymakers and decisionmakers, AI has the potential to enhance the role of data and models in decisionmaking and offers the possibility for more-rapid and -novel insights into complex policy spaces. However, these methods can come at the cost of reduced interpretability or misleading conclusions drawn from noisy datasets. Still, decisionmakers are beginning to look to these methods to answer thorny policy questions, and a new generation of quantitative researchers have the training and technical capacity to support them.

Specifically, machine learning holds several unique advantages for policy analysis. Many machine learning approaches are adept at handling and characterizing a variety of functional relationships between input and output variables (e.g., linear, nonlinear), as well as addressing high collinearity between input variables. They can also be deployed when the number of input variables is large and a modeler wishes to examine the association among a larger space of predictors. Such a case might also be common in an exploratory setting when analysts might be working to understand the impacts of existing or planned policies.

In the urban landscape, AI—and more specifically machine learning—has been applied to several different policy spaces (e.g., urban planning; air quality monitoring; and water management, including stormwater and green stormwater infrastructure).[7] Much of the existing scientific literature has used machine learning for enhancing prediction or modeling for future planning with the aim of adding value to both science and practice. There has been relatively less work using machine learning in a retrospective capacity to evaluate existing plans, infrastructure, or policies. In both contexts, the literature does point out that perspectives on the adoption of AI or machine learning into urban policy and decisionmaking are still mixed and highlights that more work is needed to bridge the gap between AI and policy.[8]

Research Objectives

Our research aimed to evaluate the location of installed green stormwater infrastructure in the context of the demographic and land-use characteristics of the surrounding area. Our primary goal was to characterize which demographic and land-use characteristics are the most strongly associated with green stormwater infrastructure investments across a set of five case study cities. Such an understanding can be useful for evaluating if infrastructure has been located in areas that, in addition to stormwater reduction, stand to gain from the co-benefits these investments provide. It can also provide a basis for reprioritizing how and where green stormwater infrastructure investments are planned for and sited in the future.

At the same time, this work is supported by a RAND Corporation internal grant focused on enhancing the use of AI in applied policy analysis. Therefore, a second goal was to examine the utility of a broad set of AI approaches for addressing our research questions and to understand how AI tools can be used to enhance decisionmaking at the local level. We therefore also carried out an exploratory machine learning–based analysis of green stormwater infrastructure and city-level data across five cities in the United States. The goal of this portion of the work was to (1) understand which approaches offer the best fit for a given application and (2) examine any trade-offs in interpretability or communication.

This report presents the results of our analysis. In addition to this report, we produced an online visualization tool that demonstrates the types of information that (1) machine learning tools can provide to stakeholders and (2) leverage best practices in AI communication.

Data and Methods

We aimed to analyze if investments in green stormwater infrastructure were made in locations that stood to realize their co-benefits. To do so, we collected city-level data on the types, sizes, and locations of existing investments to understand how much and where cities had installed green stormwater infrastructure. We also collected information to understand local characteristics—the socioeconomic, geographic, and physical landscape—surrounding these investments in each city. This information included data, for example, on income, race, climate, and infrastructure condition. These datasets were cleaned, combined, and used as inputs to a variety of popular machine learning methods. Using each of these methods, a model was trained to quantify the relationship between green stormwater infrastructure and local characteristics. With this approach, we were able to examine the strength of the associa-

We aimed to analyze if investments in green stormwater infrastructure were made in locations that stood to realize their co-benefits.

tion between green stormwater infrastructure and each local characteristic that aligned with a type of co-benefit. By looking across a variety of machine learning approaches, we could understand the trade-offs in model fit and interpretability and examine which local characteristics were found to have strong associations across multiple approaches. This section describes the data we used and our approach to exploring a variety of machine learning methods.

City Selection

The first focus of our work was identifying cities with a significant number of green stormwater infrastructure assets and sufficient datasets detailing their location, type, and size for inclusion in our analysis. We started with a search of medium to large municipalities that had open-source information on their green stormwater infrastructure. Those we found to have accessible data were Boston, Chicago, Detroit, Los Angeles, Milwaukee, New York City, Norfolk, Philadelphia, Pittsburgh, Portland, San Francisco, Seattle, and Washington, D.C.[9] We assessed the coverage of information on green stormwater infrastructure projects and the relative completeness and recency of the data.[10] We eliminated cities with an insufficient number of data points or those that were not well documented or missing data.

From this work, we identified NYC, Boston, D.C., Pittsburgh, and Detroit for analysis. Although these cities have relatively similar climates compared with other parts of the country, they have substantial differences in their socioeconomic characteristics. Table 1 shows a sample of the cities' demographic and geographic characteristics with showing a population range from 382,047 for Pittsburgh to 8,443,689 for NYC, a land area range of 61 square miles for D.C. to 266 square miles for NYC, and a population density from 4,500 people per square mile for Pittsburgh to 31,795 people per square mile for NYC. From this, we can see that Pittsburgh has the smallest population and lowest population density, D.C. has the smallest land area, and NYC is the largest across the board. In terms of socioeconomic factors, these cities also have a wide variety of racial and economic characteristics. For example, the percentage of a city's popu-

TABLE 1

Demographic and Geographic Characteristics of Analysis Areas

Characteristic	NYC	Boston Area	D.C.	Pittsburgh	Detroit
Population	8,443,689	1,076,284	684,498	382,047	806,958
Land area (mi2)	266	103	61	85	172
Population density (persons/mi2)	31,795	10,482	11,223	4,500	4,697
Percentage of population other than non-Hispanic White	68%	47%	64%	33%	83%
Percentage of population with income below federal poverty line	37%	30%	30%	37%	60%
Green stormwater infrastructure types					
Bioretention and infiltration	96%	99%	25%	42%	80%
Permeable pavement	3%	1%	9%	11%	3%
Water harvesting	<1%	0%	31%	14%	5%
Green roof	<1%	0%	13%	23%	2%
Total number of green stormwater infrastructure assets	13,384	5,563	23,458	209	262

NOTE: These values correspond to the analytic data described further in the "Findings" section, generated from 2014–2018 American Community Survey (ACS) five-year estimates. The population of Boston also includes that of the nearby incorporated places of Somerville, Cambridge, Newton, Brookline, Dedham, and Milton. mi^2 = square miles.

lation of other than non-Hispanic White residents ranges from 33 percent in Pittsburgh to 83 percent in Detroit, and the percentage of the population with income below the federal poverty line ranges from 30 percent for Boston and D.C. to 60 percent for Detroit. Finally, Table 1 also shows the variety in the type and amount of green infrastructure in each city. D.C. has the most assets and Pittsburgh the least. Boston has nearly all of its investments in bioretention and infiltration, while D.C. and Pittsburgh have a wider distribution across asset types. For additional context, the appendix provides an overview of the stormwater policy environment in each city.

Data

We geospatially joined the locations of green stormwater infrastructure from the five cities considered in this study to data on local characteristics—socioeconomic conditions, demographics, population health, and features of the physical environment—that are related to the potential co-benefits of green stormwater infrastructure (e.g., air quality). These data were aggregated and merged at the census tract level, which is the unit of analysis for all methods presented in this report. We describe the data sources and the specific features we incorporated into our analyses from each respective data source below. Table 2 provides a list of all variables included in this analysis, along with a short description and their source.

Green Stormwater Infrastructure

For each of the five cities included in our analysis, we obtained publicly available shapefiles with the coordinates of green stormwater infrastructure installations. Because the typology of green stormwater infrastructure is not standardized across cities, to ensure comparability of data, we created a mapping of each city's green stormwater infrastructure types to a minimal set of fundamental types (bioretention and infiltration, permeable pavement, water harvesting, green roofs, and tree planting) which encompassed every type of green stormwater infrastructure we observed in city datasets (Table 3). We excluded tree planting from our analyses for comparability of the definition of green infrastructure across cities because it is only included in the D.C. dataset. We found that, across the five cities in our analysis, bioretention and infiltration is by far the most common type of green stormwater infrastructure; as a result, we do not present separate analyses by type because types other than bioretention and infiltration lacked sufficient sample size to draw meaningful conclusions. We used the World Geodetic System 1984 (WGS84) projection[11] for all shapefiles and transformed any shapefile that initially had a different projection.

TABLE 2
Overview of Data and Sources

Variable	Description (Census Tract Level)	Source
Green stormwater infrastructure	Location, type, and size of existing assets	City stormwater departments (see Table 3 for details)
Average age of residents	Approximate mean age in years of city residents	2014–2018 ACS five-year estimates
Average housing costs	Approximate mean monthly housing costs (including rent and mortgage payments) of city residents	2014–2018 ACS five-year estimates
Average number of housing units per building	Approximate mean number of housing units in city residential structures	2014–2018 ACS five-year estimates
Average year moved in (owners)	Approximate average year city resident homeowners moved into their current residence	2014–2018 ACS five-year estimates
Average year moved in (renters)	Approximate average year city resident renters moved into their current residence	2014–2018 ACS five-year estimates

Table 2—Continued

Variable	Description (Census Tract Level)	Source
Average year of residence built	Approximate average year built of city residential structures	2014–2018 ACS 5-year estimates
Economic inequality	Gini Index	2014–2018 ACS 5-year estimates
High redlining score	Binary variable which equals 1 if the census tract had a score of 3.25 or greater out of 4 (“Hazardous”) in Home Owners’ Loan Corporation maps from 1935 to 1940	Climate and Economic Justice Screening Tool
High summer mean temperature	Average intraday high temperature in the census tract during the summer	Climate and Economic Justice Screening Tool
Level of air quality contamination	Particulate matter (PM) 2.5 air pollution as a percentile relative to other census tracts	Climate and Economic Justice Screening Tool
Median household income	Median annual income in dollars for a city household	Tree Equity Score
Number of residents with coronary health challenges	Proportion of population with self-reported coronary heart disease, normalized in standard deviations from the national average	Tree Equity Score
Number of residents with mental health challenges	Proportion of population with self-reported subjectively poor mental health	Tree Equity Score
Percentage impervious land	Proportion of impervious land coverage in the census tract as a percentile relative to other census tracts	Climate and Economic Justice Screening Tool
Percentage lacking tree canopy	Proportion of land area in census tract that does not have tree canopy coverage as a percentile relative to other census tracts	Climate and Economic Justice Screening Tool
Percentage of American Indian and Alaska Native residents	Proportion of city population that is non-Hispanic American Indian or Alaska Native	2014–2018 ACS 5-year estimates
Percentage of area at risk of future flooding	Proportion of land area in census tract that is at risk of future flooding according to climate models, as a percentile relative to other census tracts	Climate and Economic Justice Screening Tool
Percentage of Asian residents	Proportion of city population that is non-Hispanic Asian	2014–2018 ACS 5-year estimates
Percentage of Black or African American residents	Proportion of city population that is non-Hispanic Black or African American	2014–2018 ACS 5-year estimates
Percentage of Hispanic or Latino residents	Proportion of city population that is Hispanic or Latino	2014–2018 ACS 5-year estimates
Percentage of Native Hawaiian and other Pacific Islander residents	Proportion of city population that is non-Hispanic Native Hawaiian or other Pacific Islander	2014–2018 ACS 5-year estimates
Percentage of residents in poverty	Proportion of city population with household income below 150 percent of the federal poverty line	Tree Equity Score
Percentage of residents reporting as other race	Proportion of city population that is non-Hispanic, other race	2014–2018 ACS 5-year estimates
Percent of residents reporting as two or more races	Proportion of city population that is non-Hispanic, multiracial	2014–2018 ACS 5-year estimates
Percentage of residents with asthma	Proportion of city population with self-reported asthma	Tree Equity Score
Percentage of White residents	Proportion of city population that is non-Hispanic White	2014–2018 ACS 5-year estimates
Population density	Millions of people living in city per square kilometer of area	Tree Equity Score

TABLE 3

Details on City Green Stormwater Infrastructure Data

City	Green Stormwater Infrastructure Types Covered	Source
NYC	Green/blue roofs, detention systems, drywells, engineered soil tree pits, permeable pavers, porous asphalt/concrete, cisterns, rainwater harvesting	Department of Environmental Protection (DEP) Green Infrastructure Program Map
Boston	Drywells, perforated pipes, leaching basins, Stormtech/Cultec chambers, tree pits/planter beds, rain gardens, permeable pavement, detention storage, bioswales, tank/injection wells, detention ponds, filtration basins	Boston Water and Sewer Commission (BWSC) Green Infrastructure Map
D.C.	Green roofs, bioretention, infiltration, permeable pavement, rainwater harvesting	Open Data D.C. Green Infrastructure Practices in the District map
Pittsburgh	Green roofs/walls, rain gardens, bioretention, porous pavement, stormwater wetlands, infiltration/storage trenches, cisterns/rain barrels, stormwater tree pits, stormwater planters, naturalized meadows	3 Rivers Wet Weather Green Infrastructure Atlas
Detroit	Green roofs, bioretention, bioswale, subsurface infiltration, stormwater/constructed wetlands, rain gardens, (sub)surface detention, downspout disconnections, permeable pavement, rain barrels, water harvesting	City of Detroit Open Data Portal Green Stormwater Infrastructure Locations

Tree Equity Score

The Tree Equity Score tool developed by American Forests, a national nonprofit organization focused on conservation and environmental equity, assesses the equity of tree canopy coverage across the United States. The score combines a *tree canopy gap*, calculated by comparing the current tree canopy coverage within a neighborhood with a biome- and density-adjusted goal for that neighborhood, and a *priority index*, which is based on income, employment, age, race, climate, and health within that neighborhood. Because our analysis also focused on equity and tree canopy coverage offers some similar benefits and co-benefits to green stormwater infrastructure, there is significant overlap between the features needed for our analysis and the components of the Tree Equity Score's priority index. Specifically, we integrated the Tree Equity Score's measures of poverty, median household income, and self-reported health challenges related to asthma, coronary heart conditions, overall physical health, and overall mental health, respectively. We aggregated the Tree Equity Score data from the census block group level to the census tract level by calculating a population-weighted average of each measure within each census tract. We also used the Tree Equity Score data on population and area in square kilometers to directly calculate population density at the census tract level.

2014–2018 American Community Survey Five-Year Estimates

We obtained additional information on census tract demographics, income inequality, and housing attributes from the 2014–2018 ACS five-year estimates. Specifically, our analyses included ACS data on the racial-ethnic composition of the census tract (i.e., the respective percentages of White, Black, Asian, American Indian and Alaska Native, and multiracial individuals, as well as the percentage of residents who identify as Hispanic or Latino); the Gini Index of income inequality; average monthly housing costs; average tenure of residency for homeowners and renters, respectively; average age of housing structures; and the average number of units in a housing structure.[12] Where the ACS reports results in bins (e.g., for year structure built, "between 2000 and 2009"), we used the midpoint of each bin to calculate an approximate average. We used the 2014–2018 ACS five-year estimates for consistency with the Tree Equity Score data, which derive poverty and median household income from the same source. We obtained ACS data from the U.S. Census Bureau Application Programming Interface using the *tidycensus* R package.[13]

Additional Climate Justice Variables

We integrated additional climate justice variables from the Climate and Economic Justice Screening Tool (CEJST), which identifies disadvantaged communities simultaneously facing climate, economic, and other burdens (e.g., high expected losses from natural hazards resulting from climate change) and the related socioeconomic burdens (e.g., high proportion of low-income residents). Specifically, we used CEJST measures at the census tract level on the percentage of properties at risk of future flooding, historical redlining, average intraday high summer temperatures, percent impervious land, percentage of land lacking tree canopy coverage, and average PM 2.5 air pollution. The time span of CEJST v1.0 data varies by specific measure, but is generally concentrated between 2014 and 2020 (with the exception of redlining, which links historical Home Owners' Loan Corporation maps between 1935 and 1940 with 2010 Census Bureau boundaries), thus overlapping with the Tree Equity Score and ACS data we incorporated.

Construction of Analytic Dataset

Our analytic dataset comprises all the preceding data sources (city green stormwater infrastructure asset locations, Tree Equity Score data, ACS data, and CEJST data) merged and aggregated at the census tract level. We aggregated the city green stormwater infrastructure asset locations, which were originally in point shapefile format, to counts of installations by census tract using the shapefile geographic identifiers. This count variable served as our outcome variable for all analyses presented in the report. Additionally, the Tree Equity Score data was originally exported at the census block group level. We aggregated these data to the census tract level by taking weighted averages of values for census block groups within each census tract. To construct these weighted averages for demographic variables, such as race-ethnicity, we used census block group populations as the weight. For physical environment variables, such as air pollution and impervious land surface, we used census block group geographic areas as the weight.

> To quantify the relationship between the location of green stormwater infrastructure and the variety of location characteristics, we developed an analytic approach.

Analytical Approach and Machine Learning Methods

To quantify the relationship between the location of green stormwater infrastructure and the variety of location characteristics, we developed an analytic approach that relates the number of green stormwater infrastructure assets within a census tract to the data we collected on its local characteristics (e.g., air quality, median household income). The following equation simplifies our modeling approach for each city:

$$\textit{Number of Green Infrastructure Assets} = f(\textit{Location Characteristics}, \beta) + e,$$

where β represents the relationship between a specific location characteristic (e.g., median household income) and the number of green stormwater infrastructure assets, and e represents error.

Using the β values or an approximation of these values, we can understand which location characteristics are prominent in locations with and without green stormwater infrastructure. For each city we then examined (1) the strength of the association between the given characteristic and the number of green stormwater infrastructure assets and (2) its direction, or whether or not a higher value for a local characteristic predicted more (a positive relationship) or fewer (a negative relationship) green stormwater infrastructure assets.

For this analysis, we considered several machine learning methods that are commonly used in water resources analyses, environmental health, or environmental modeling more broadly. These include both nonlinear approaches and those based on linear regression to be able to examine any added benefits in goodness of fit or trade-offs in interpretability. The selection of these methods ultimately hinged on our data characteristics and the specific objectives of the study; each method was evaluated based on its ability to predict outcomes accurately and interpret its results. These methods and their benefits and tradeoffs for this study are introduced below.

Lasso Regression

Lasso regression is a linear regression analysis approach that performs both variable selection and regularization. This approach enhances the prediction accuracy and interpretability of the produced statistical model, which proves useful for isolating the most relevant subset of variables connected to the outcome. This is useful in our context because our data, which comprise a single cross-section at the census tract level for each city we analyzed, contain a relatively small number of observations (ranging from 171 in Pittsburgh to 2,164 in NYC) while our analysis included a relatively large number of predictors (29). As a result, the variable selection in lasso regression prevented us from fitting a model with more predictors than can realistically be estimated given the amount of data, as would occur with traditional linear regression.[14]

However, Lasso is a linear regression technique and therefore does not address the possibility of nonlinear relationships between local characteristics and green stormwater infrastructure assets. Also, the Lasso method is limited when variables are colinear, as it tends to arbitrarily select one variable from among the highly correlated ones, thus introducing potential bias. This often results in unstable models where slight changes in the data could lead to very different selected variables.

The selection of these methods ultimately hinged on our data characteristics and the specific objectives of the study; each method was evaluated based on its ability to predict outcomes accurately and interpret its results.

ElasticNet

ElasticNet is a linear regression approach that is a compromise between Lasso and ridge regression methods that combines their respective properties of performing variable selection and avoiding multicollinearity.[15] Just like Lasso, ElasticNet enhances the prediction accuracy and interpretability of the produced statistical model, which can be useful for characterizing the most relevant variables connected to the outcome. However, ElasticNet has an advantage over Lasso in situations where variables are colinear. It is designed to handle multicollinearity, therefore offering a solution to one of the drawbacks of Lasso. This approach handles the bias introduced by Lasso when it arbitrarily selects one variable from among highly correlated ones.[16] This results in more-stable models, even with slight changes in the data.

Random Forest and Gradient Boosted Decision Trees

Random Forest and Gradient Boosted Decision Trees (GBDT) are related classes of algorithms known as ensemble methods that create collections of *decision* *trees*—sets of rules or guidelines based on the data—and then combine their outputs. A decision tree can be thought of as having multiple advisers, each giving their own opinion, and then making a final decision based on the majority opinion, an intuitive analogy that makes these methods relatively easy to explain to nontechnical audiences. Relative to using a single decision tree, which may be prone to overfitting and affected by outliers, the ensemble approach allows for better accuracy and less overreliance on a single set of rules or a single adviser. The primary difference between Random Forest and GBDT is the approach to building decision trees; Random Forests build decision trees in parallel on different random samples of the data, akin to a bootstrap,[17] while GBDT are built sequentially on the full dataset, with each subsequent tree attempting to reduce the prediction errors of the previous tree using gradient descent. Because of the bootstrap approach, Random Forests may be more robust to overfitting the training data than GBDT but may also be more computationally expensive.

In our context, decision tree–based algorithms are useful because they are effective at dealing with different types of data and identifying which pieces of information have the strongest relationship with the outcome variable.

In our context, decision tree–based algorithms are useful because they are effective at dealing with different types of data and identifying which pieces of information have the strongest relationship with the outcome variable. Our analytic dataset includes a mixture of continuous variables (e.g., median household income), percentiles, and binary categorical variables (e.g., high redlining score). Although parametric regression methods might struggle with these heterogeneous types of predictors with different underlying distributions, decision tree–based methods are less affected because they do not rely on any assumptions about the distribution of any parameter. The effectiveness of decision tree–based methods in determining the importance of each predictor in the model is also useful in our context because the relative strength of association between each of our demographic and climatic predictors and green stormwater infrastructure investment outcomes is the focus of our study. Because the branches of decision trees are based on splits in the data at arbitrary cut points in a predictor, the relative importance of a predictor can intuitively be estimated by looking at what happens to predictive accuracy each time that predictor is used at a branch (a measure known as *gain*).

Support Vector Regressor

The Support Vector Regressor (SVR) method, derived from the Support Vector Machine (SVM), is specifically designed for regression tasks. It is better suited for problems with complex domains where clear margins of separation exist within the data. It tends to underperform when the dataset is noisy but is a useful method for predicting continuous outcomes and can be especially useful when handling nonlinear relationships. In our context, many of the potential advantages of SVR overlap with the potential advantages of decision tree–based methods, including flexibility to deal with heterogeneous predictor types and nonlinear relationships between predictors and the outcome. However, the intuition behind the approach differs, and SVR may be more intuitively explained to some audiences. In principle, an SVR/SVM can be pictured as drawing lines on a scatter plot of the data, classifying all data points on one side of the line in one group and all data points on the other side in another group, and selecting the line that maximizes accuracy on the training set. This is why SVR performs best in cases where clear separations exist that are pertinent to the outcome.

Artificial Neural Networks

Neural networks essentially look for patterns in the data and can interpret highly complex nonlinear relationships, making them most useful when dealing with high-dimensional data, such as images or complex signals. However, these networks come with their share of downsides, such as their black-box nature, the need for copious amounts of data, and lengthy training times.

In our context, neural networks might be beneficial in a case in which extremely complex relationships that other methods cannot effectively capture exist between our demographic and climatic predictors and green stormwater infrastructure investment. Of the methods we explored, neural networks are the most model- and data-agnostic, because their approach of feeding training data through arbitrarily connected nodes (*neurons*) that can be assigned different weights and transformation functions to minimize prediction error on the training set can theoretically deal with any possible numerical relationship between predictor and outcome. However, neural networks are significantly less intuitive to communicate to nontechnical audiences than the other methods we explored, and the relatively small number of observations in our data (especially in cities other than NYC) might limit accuracy.

Model Training and Implementation

We used each of these methods to train a model that relates the number of green stormwater assets in a city to the variety of local characteristics we considered. To fit each machine learning model, we used various machine learning packages from the scikit-learn library in Python, including Lasso Regression, ElasticNet, GBDT, SVR, and Random Forest.[18] *NumPy* and *pandas* libraries were also used to handle numerical operations and data manipulation respectively.[19] The scikit-learn library was further used to import several functions for preprocessing and metric evaluation.

To set up the data for model training, data normalization was carried out using the StandardScaler from the scikit-learn library. This preprocessing step transformed each feature in our dataset by subtracting the mean (centralizing around zero) and then scaling to unit variance (standard deviation of one), ensuring all features have the same scale and avoiding any single feature dominating the model because of its scale.[20] A standard train-test split was performed using a 80:20 ratio for training and testing datasets, respectively. After fitting the model to the training data, the model was run on the test data. The model's performance was evaluated using R2 and a normalized root mean square error (RMSE) to enable comparison of model fit within and across cities.

Hyperparameter Optimization

Hyperparameter optimization is a crucial part of machine learning model construction. Hyper-

Neural networks essentially look for patterns in the data and can interpret highly complex nonlinear relationships, making them most useful when dealing with high-dimensional data, such as images or complex signals.

Best Practices on Communicating the Results of Machine Learning Analyses

As AI becomes more integrated into decisionmaking processes, it is important to ensure that people understand how AI works and its limitations and to have their concerns addressed so they can properly interpret the results and make well-founded decisions. Clear and accurate education is required to establish trust in any results from AI analyses. This section provides best practices to effectively educate about AI and inspire confidence in the AI results.

Effective education tailors communications about AI to different audiences by considering the varying levels of technical knowledge, priorities, or cultural perceptions of AI.[57] For example, it is helpful to convert AI terminology and jargon into plain language when talking with a nontechnical audience. Also, using compelling, relevant, and relatable titles, examples, and images spark interest and encourage discussion.[58] Relatable examples help make connections to new concepts from a familiar reference point.

When communicating the results of AI analyses, focus on the narrative around the results more than the AI methods themselves and proactively address any concerns the audience may have.[59] This is because how evidence is presented can influence how it is received and understood by the audience.[60] Although the rational narrative supported by quality data, synthesized evidence, and case studies has been the most popular approach for communicating AI results, it is also important to directly address possible ethical and moral concerns from the audience.[61] Clarify that human judgment, behavior, and decisionmaking are still highly valued and be clear and transparent about the scientific uncertainty and limitations of the AI.[62] One critique about AI is "tech-for-tech sake," which refers to "using a technology because you can, rather than because you should."[63] Discussing how the AI results might be an enhancement over previously used approaches and identifying where there are differences from the status quo can address this concern. Finally, in AI engagement or education, be patient and recognize that achieving full understanding about AI takes time and regular communication.[64]

The analysis itself should follow standard practices around technical robustness and data quality, transparency, security and privacy protection, fairness, and if applicable legal and regulatory requirements. For example, ensure the data used to train the AI model are representative, unbiased, and relevant to the task. Additionally, when communicating results, explain how the principles of Trustworthy AI (established by the European Commission's High-Level Expert Group on Artificial Intelligence that include explainability, safety and security, and privacy) were applied and how the analysis complies with professional and federal guidelines, recommended protocols, or regulatory and legal requirements.[65] This effort might involve highlighting the methods used to ensure interpretability and explaining how the model makes decisions.[66] For technical audiences, communications could cover the rigorous testing and validation methods used. Be transparent about procedures used to evaluate errors identified in outputs. Lastly, use model confidence whenever it is possible to describe the level of certainty of the AI results in comparison with any alternatives considered.[67]

For the results to be believed and used, it is critical to establish a level of the trust and understanding of the results, which can be accomplished by explaining the results, including the probabilistic aspects and describing any limitations so that decisionmakers and other consumers of the outputs can reasonably use the analysis results to enhance human judgment.

A good example of communicating AI results is a RAND Europe commentary that synthesizes machine learning research, findings, and key recommendations into a consistent narrative for government actors.[68] Additionally, the full report explains technical concepts and methods clearly and accessibly and highlights the level of accuracy of the model and interpretability to establish confidence in the results.[69] Another relevant example is past RAND research on using machine learning to address infant mortality.[70]

parameters are not learned from the data but are instead set before the learning process; they influence the speed and quality of learning. For example, a Random Forest model has hyperparameters such as *n_estimators* and *max_depth* that define the size and configuration of the model. To determine the values for these and comparable parameters, hyperparameter optimization was carried out using the GridSearchCV function from the *scikit-learn* library. GridSearchCV performs an exhaustive search over a specified parameter grid,[21] evaluating the model performance for each combination of hyperparameters and returning the parameters that give the best performance based on a scoring method (in this case, R2). For each combination of hyperparameters, a model was trained using the K-Fold cross-validation method with five splits.[22] The K-Fold cross-validator provides train-test indices to split data in train and test sets, and this is done in a way that each sample is used for validation exactly once. This minimizes the risk of overfitting and ensures that all data is used for both training and validation, thus optimizing the use of available data.

A technique called Hyperband was used for tuning the hyperparameters of the neural networks. *Hyperband* is an advanced method for hyperparameter tuning that is designed to explore efficiently in the hyperparameter space and to converge rapidly to a high-performing model. It is particularly suited for deep-learning contexts, where the computational cost of hyperparameter optimization can be high. This makes it a better choice than GridSearchCV for the neural network models because Hyperband does not search the entire hyperparameter space exhaustively. Instead, it employs a more efficient approach to sample a small fraction of the hyperparameter space and progressively narrows down the search to the best-performing configurations.

In each iteration of the Hyperband algorithm, a set of hyperparameters was randomly selected within the predefined ranges.[23] The models were then trained for a maximum of 100 epochs, and the performance of the models was evaluated using the validation loss, which is the root mean squared error calculated on a validation set excluded from the training data. This process was performed for a certain number of iterations, which is automatically determined by the Hyperband algorithm based on the total computational resources allocated for the hyperparameter tuning process. The best set of hyperparameters was identified as those that led to the model with the lowest validation loss. These optimal hyperparameters were then used to build the final models, which were trained on the full training data and evaluated on the test data.

Interpreting Modeled Relationships

As machine learning models become increasingly complex, the need for tools and techniques to interpret these models has grown. *Interpretability* in machine learning refers to the ability to understand and communicate the reasoning behind the decisions made by a machine learning model. It is about being able to explain why a model made a particular prediction or decision. Interpretability is crucial for several reasons. First, it fosters trust in machine learning models. If users such as data scientists, stakeholders, or the general public can understand why a model is making the decisions it is, they are more likely to understand and believe the results. Without this understanding, users might be hesitant to adopt or rely on AI systems, especially in applications where the consequences of a decision are great. Second, interpretability is essential for debugging and improving models. If a model makes an unexpected prediction, knowing the model's decisionmaking process can help identify where things might have gone wrong and provide insight into how to improve the model. Third, interpretability can help ensure fairness and prevent bias in machine learning models. If we can understand what factors a model is using to make its decisions, we can check whether it is relying on inappropriate

TABLE 4

Normalized Root Mean Square Error Values for Machine Learning Methods

Model	Boston	D.C.	NYC	Pittsburgh	Detroit
Lasso Regression	0.74	0.63	2.19	2.50	1.83
ElasticNet	0.74	0.65	2.19	2.50	1.86
GBDT	0.72	1.11	2.03	2.31	1.87
Random Forest	0.69	0.97	1.98	2.30	1.92
SVR	0.71	0.76	2.16	2.54	1.94
Shallow Neural Network	0.83	0.59	1.98	1.92	3.34
Deep Neural Network	0.99	0.81	1.96	1.51	3.00

NOTE: We normalized the RMSE values in this table by dividing the calculated RMSE (i.e., square root of the sum of squared differences between the method's predicted number of green stormwater infrastructure assets for each census tract and the actual number of green stormwater infrastructure assets in that census tract) by the average actual number of green stormwater infrastructure assets across each city's census tracts. The normalized RMSE is more comparable across cities because the nonnormalized RMSE tends to be larger in cities with more green stormwater infrastructure. This difference in size is because, all else equal, the differences between predicted and actual values tend to be larger when the actual value is larger (e.g., the difference between an actual value of 10 and a predicted value of 9, versus the difference between an actual value of 1 and a predicted value of 0.9).

or discriminatory factors and take steps to address this if necessary. Box 1 provides an additional set of best practices in communicating machine learning findings to policymakers and the public.

State-of-the-art methods in interpretability seek to strike a balance between model performance and ease of interpretability. One approach to interpretability is to select machine learning models in which the results are relatively straightforward, such as linear regression or decision trees. However, one important trade-off is that these simpler models can sometimes be outperformed by more-complex, less-interpretable models. This has led to the development of post-hoc interpretability methods to explain the decisions of complex models after they have been trained. Examples of these methods include Local Interpretable Model-agnostic Explanations (LIME), which provides explanations for individual predictions, and SHapley Additive exPlanations (SHAP), which provides a measure of the contribution of each feature to the prediction for each instance in the dataset. There are also several frameworks and libraries available designed to aid in model interpretation, such as Microsoft's Interpret ML, and IBM's AI Explainability 360 and OmniXAI.[24] These tools offer a variety of techniques for both training interpretable models and explaining black-box models.

FIGURE 3

Estimated Strength of Association with the Amount of Green Stormwater Infrastructure

Feature	Boston	Washington, D.C.	New York City	Pittsburgh	Detroit
Newer residences	0.018	1.000	-0.006	1.000	-0.204
Higher percentage of Hispanic or Latino residents	0.345	0.022	0.384	-0.033	0.306
Larger number of housing units per building	-0.063	-0.174	-0.277	0.321	1.000
Warmer summer mean temperature	-0.051	0.068	1.000	-0.038	-0.188
Higher percentage of area at risk of future flooding	0.027	0.006	0.014	0.151	0.451
Higher percentage of American Indian and Alaska Native residents	0.068	-0.004	0.081	0.172	0.288
Poorer air quality	1.000	0.019	-0.438	0.001	0.009
Higher percentage of residents reporting as two or more races	0.027	0.043	0.195	0.109	0.087
Higher percentage of Asian residents	-0.072	-0.005	0.047	0.425	0.032
Higher median household income	0.154	0.064	0.057	-0.001	0.097
More residents with coronary health challenges	0.699	-0.061	-0.362	-0.051	0.107
Less tree canopy coverage	0.349	-0.063	0.001	0.156	-0.112
Higher percentage of residents reporting as other race	0.088	0.091	0.036	0.002	0.087
Higher redlining score	0.162		0.018	0.073	0.006
Average age of older residents	0.147	0.052	0.029	0.035	-0.005
Higher percentage of residents with asthma	0.457	0.011	-0.284	0.008	-0.008
Owners who moved in more recently	0.251	0.054	-0.053	0.148	-0.223
Higher percentage of Black or African American residents	-0.126	0.050	0.153	0.012	0.057
Higher percentage of Native Hawaiian/other Pacific Islander residents	0.010	0.006	0.064	0.003	0.049
More residents living in poverty	0.097	-0.019	-0.056	-0.004	-0.013
Higher percentage of White residents	0.068	-0.043	-0.046	0.003	-0.034
More residents with mental health challenges	0.066	-0.035	0.041	0.009	-0.164
Renters who moved in more recently	-0.075	-0.116	0.020	-0.063	-0.193
Greater economic inequality	0.003	-0.257	0.028	0.003	-0.240
Higher percentage impervious land	-0.017	0.011	0.035	0.011	-0.514
Higher average housing costs	-0.602	0.011	0.010	0.021	-0.187
Greater population density	-0.180	-0.790	-0.086	0.016	-0.501

Strength of association

-1.000 1.000

SHAP Regressions

In our work, we adopted an interpretability approach that leverages SHAP values.[25] SHAP values are calculated as the average marginal contribution of a feature value over all possible combinations of feature values. To enhance their interpretability, we fit a linear regression using the SHAP values as inputs and the original values of each feature as outputs. In this way, we were able to quantify the relationship between a feature and the prediction. The regression coefficients that come from the trained linear regression model are interpreted as the average change in a SHAP value for each unit increase in the feature value, assuming all other features remain constant. Therefore, we could quantify how the contribution of a factor to the prediction changes as the value of the factor itself changes. The limitation with this approach was that we assumes the relationship between SHAP values and feature values are linear.

This methodology is conceptually similar to alternative methods of deriving summary statistics or functional forms from SHAP values commonly used in scientific literature, such as the mean absolute SHAP value as a measure of feature importance or using a locally estimated scatterplot smoothing (LOESS) curve to approximate the predictor-response function.[26] However, our regression-based coefficient approach preserves directionality of the relationship, unlike the mean absolute SHAP value, which is always positive. It also quantifies the relationship with a single easy-to-interpret number, unlike LOESS curves, which would require the reader to infer the relationship between the predictor and the outcome from a graph which may be complex or suggest an unclear, noisy relationship. Although it is possible for the relationship between predictor variables and outcomes to be complex, the use of more-complicated local estimation techniques has the risk of over-inferring from small slices of variation in the data, especially when there is not an underlying hypothesis which would suggest a complex functional form. Our simplified approach instead generates a global estimate of both feature importance and relationship directionality, reducing this risk at the expense of some flexibility to represent nonlinear relationships.[27]

Findings

We examined both the performance of the machine learning methods we used and what we could interpret about the strength of the relationship between census tract characteristics and our output variable, the number of green stormwater assets. Table 4 summarizes the performance of the six machine learning methods for each city. When considering all cities, Random Forest slightly outperforms all other approaches, but there is not an obvious method that performs best for all locations. For individual cities, Random Forest only produced the lowest normalized RMSE for Boston. In contrast, neural networks produced the least error for NYC, Pittsburgh, and D.C. but are the worst performers in Detroit. In Detroit, the Lasso Regression provides the lowest error. These

We examined both the performance of the machine learning methods we used and what we could interpret about the strength of the relationship between census tract characteristics and our output variable, the number of green stormwater assets.

Across all the cities, there are several similarities related to building age and duration of tenancy, population density, and racial and ethnic demographics.

inconsistencies in model performance suggest that some models handle the specific nonlinearities and distributions of input and output data for a given location better than others. It also suggests that evaluating multiple machine learning approaches, assuming they are appropriate to the analysis and data at hand, will produce more analytical flexibility and ultimately a better performing model. Because these findings show that simpler methods, such as Lasso Regression, can still outperform more advanced approaches, analysts would be wise to start simple before adding analytical complexity.

Using the normalized SHAP regression coefficients shown in Figure 3,[28] we could assess both the strength and direction of association between the amount of green stormwater infrastructure and each feature (e.g., median household income) of the surrounding census tract. In Figure 3, green boxes indicate positive associations between a given feature and the amount of green stormwater infrastructure and purple boxes indicate a negative, or inverse, relationship. A positive relationship means a higher value for a given feature would suggest more green stormwater infrastructure, and a negative relationship means a higher value for a given feature would suggest less green stormwater infrastructure. A value of 1.000 indicates the strongest positive association between a feature and the amount of green stormwater infrastructure for that city, and the largest negative value indicates the strongest negative association. For example, for D.C., the average year of residence built has the strongest positive relationship with the amount of green stormwater infrastructure, suggesting that areas with newer buildings have more green stormwater infrastructure.[29] In D.C., population density has the strongest negative relationship, meaning that the lower the population density is in a census tract, the more likely it is to have more green stormwater infrastructure.[30] We describe the specific findings relevant to each city below.

Across all the cities, we can see from Figure 3 that there are several similarities related to building age and duration of tenancy, population density, and racial and ethnic demographics. Specifically, for three of the five cities, more green stormwater infrastructure was associated with locations having newer residences. In two of these cities, D.C. and Boston, this association was also positive in places with fewer housing units per building and a lower population density. And overall, more green stormwater infrastructure was generally associated with less population density, particularly in Detroit and D.C.

Across all cities except for Boston, green stormwater infrastructure was more strongly associated with locations with a lower percentage of White residents. For all locations but Pittsburgh, green stormwater infrastructure was more strongly associated with locations with a higher percentage of Hispanic or Latino residents. Most cities had a positive association between higher percentages of American Indian and Alaska Native, Asian, and Black residents, and all cities had a positive association with higher percentages of Native Hawaiian or other Pacific Islander, as well as in locations where residents reported as two or more races or other races. In most cases, however, the strength of these associations was relatively weak.

At the same time, because the selected cities varied on the intent and maturity of their green stormwater infrastructure policy (and how and why they made investments to date) there were several observed differences. For example, Boston's initial goal with green stormwater infrastructure investments was to address climate change, whereas Detroit's was to reduce combined sewer overflows

(CSOs).[31] However, based on the findings in Figure 3, we can see that Detroit has a positive and strong association with green stormwater infrastructure in locations at risk of future flooding, Boston does not, and both have a negative association with percent impervious land. The sections below detail additional differences and high-level findings for the five cities.

Boston

In the Boston metropolitan area, poor air quality and more residents with coronary health challenges and asthma are the factors most strongly, positively associated with green stormwater infrastructure. In contrast, average housing costs has the strongest negative association with green stormwater infrastructure. Together, this suggests that green stormwater infrastructure is more associated with areas with poorer air quality, lower housing costs, and with areas where many residents face respiratory and cardiac health issues. Furthermore, Figure 3 also illustrates that green stormwater infrastructure is associated with locations with a higher percentage of Hispanic or Latino residents and a lower amount of tree canopy coverage.

Although we did not review these results with BWSC officials, from our own professional knowledge, there are likely to be several reasons why these results are likely accurately reflecting conditions on the ground. For example, East Boston, which is adjacent to Logan International Airport and several highways, has been the site of rapid recent redevelopment, has a high percentage of Hispanic or Latino residents, and has the least amount of tree canopy coverage of any city neighborhood.[32] As a whole, it appears that Boston's policies focused on the CSO consent decree and, coupled with redevelopment projects, have incidentally sited green stormwater infrastructure in locations that have a higher potential for community co-benefits.

Washington, D.C.

For D.C., only the average year of residence built has a strong, positive association with the amount of green stormwater infrastructure. Population density and, to a much lesser extent, economic inequality have a negative association with the amount of green stormwater infrastructure. For the district, this would suggest that green stormwater infrastructure is more associated with locations with newer buildings, less population density, and lower economic inequality.

Our interviews with Department of Energy and Environment (DOEE) officials aligned with these findings. They indicated that green stormwater infrastructure was disproportionately being built in areas of rapid recent redevelopment (e.g., along the Anacostia) and in the less-densely populated municipal separate storm sewer system (MS4) areas because of their stormwater credit program. Both neighborhood types are relatively economically homogenous.

Across all cities except for Boston, green stormwater infrastructure was more-strongly associated with locations with a lower percentage of White residents, and for all locations but Pittsburgh, it was more-strongly associated with locations with a higher percentage of Hispanic or Latino residents.

New York City

In NYC, high summer mean temperatures and a high percentage of Hispanic or Latino residents have a strong, positive association with the amount of green stormwater infrastructure, and poor air quality and number of residents with coronary health challenges have a negative association. To a lesser extent, the percentage of residents with asthma and the average number of housing units per building have a negative association with green stormwater infrastructure. Together, these factors indicate that more green stormwater infrastructure in NYC is associated with places with warmer summer temperatures, more Hispanic and Latino residents, better air quality, less-dense housing, and fewer residents with respiratory and cardiac health issues.

NYC DEP explained that their green stormwater infrastructure strategy was deployed on a neighborhood basis prioritized by the consent decree for individual CSO-impacted waterbodies, such as other cities' dynamics, but different in implementation given its level of saturation and extensive use of installations in public rights of way.[33]

> Following our machine learning analysis, we carried out a series of outreach engagements with city officials to validate our findings, understand the utility of our approach in a real-world context, and gain more insight into specific policies and contexts informing green stormwater infrastructure investment at the local level.

Pittsburgh

In Pittsburgh, only two features stood out of the analysis—average year of residence built and percent of Asian residents—both of which have a strong, positive association. No other variables appeared to have a strong association with green stormwater infrastructure. For Pittsburgh, this suggests that green stormwater infrastructure is more strongly associated with locations with newer housing stock and a higher percentage of Asian residents.

Rapid recent redevelopment in Oakland, Shadyside, Squirrel Hill, and East Liberty (among other neighborhoods) aligns with the findings of this analysis.[34] In addition, green stormwater infrastructure at the University of Pittsburgh, Phipps Conservatory and Botanical Gardens, and Carnegie Mellon University constitute a substantial portion of the rest of the city's green stormwater infrastructure. The demographics of these neighborhoods might also be driving the results.[35]

Detroit

For Detroit, the model suggests that the average number of housing units per building followed by the percentage of area at risk of future flooding are the factors most strongly, positively associated with green stormwater infrastructure and the percent impervious land and population density the most strongly negatively associated. This finding, on the surface, suggests contradictory results; it could mean that there is more green stormwater infrastructure in places with a larger number of housing units per building but a lower population density or a lower-percent impervious surface but a higher risk of future flooding.

Like Pittsburgh, local concentrations of green stormwater infrastructure can produce potentially incidental associations. In Detroit's case, a majority of installed green stormwater infrastructure (55 percent of projects) have disconnected paved surfaces and downspouts and particularly focused on a single sewershed (Upper Rouge Tributary) and the area immediately adjacent to the Detroit Riverfront. Both areas have a lower-percent impervious surface than nearby areas, but a higher future flood risk that is associated with regional surface conveyance and lake levels rather than the piped infrastructure of other cities in our sample.

City Outreach

Following our machine learning analyses, we carried out a series of outreach engagements with city officials to validate our findings, understand the utility of our approach in a real-world context, and gain more insight into specific policies and contexts informing green stormwater infrastructure investment at the local level. The goals of these engagements were to

- understand the policy and decisionmaking context for green stormwater infrastructure from the relevant department or authority's perspective, such as how green stormwater infrastructure policies came to be implemented, if their goals and priorities have changed over time, and if planners rely on any tools or datasets (beyond hydraulic and hydrologic modeling) when making siting decisions
- validate the findings of our machine learning analysis of green stormwater infrastructure data and inquire about what sort of data-driven or AI-supported tool might be useful to support green stormwater infrastructure policy moving forward.

We interviewed officials at DOEE, NYC DEP, and Pittsburgh Water and Sewer Authority (PWSA) using a semi-structured interview protocol. Information on Boston came from previous communications and professional engagements with the BWSC. Our understanding of Detroit was limited to two recent conversations with a local hydrologic consultant and nonprofit organization, rather than any Detroit Water and Sewerage Department (DWSD) officials.

Findings

City officials in NYC, D.C., and Pittsburgh described how city stormwater planning has been predominantly consent decree–driven and publicly funded with a focus on reducing stormwater volumes. As of 2023, however, efforts are beginning to refocus on compliance with municipal codes, leading to more active involvement with private developers and landowners. Cities, such as D.C., are additionally working to understand how site selection processes can enhance equitable outcomes.[36]

Like this trend, early success with green stormwater infrastructure implementation in the most challenging sewersheds and the renewed focus on

Early success with green stormwater infrastructure implementation in the most challenging sewersheds and the renewed focus on equity has encouraged cities to expand their citywide focus and to engage in larger, neighborhood-scale planning.

equity has encouraged cities to expand their citywide focus and to engage in larger, neighborhood-scale planning.[37] For example, Pittsburgh is initiating an effort to determine neighborhood levels of service of current drainage infrastructure under different design storm events and will assess the relative additional performance and benefits of proposed investments given current climate conditions.[38] These regionally significant infrastructures go beyond what has historically been considered green stormwater infrastructure installations, and, going forward, cities might need different tracking mechanisms to account for the much larger scale of investment associated with neighborhood-based approaches. For example, DOEE has developed a blue/green flood resilience strategy for Buzzard Point in D.C.,[39] and NYC DEP has constructed bluebelts on Staten Island. *Blue infrastructure*, such as NYC's bluebelts, is a concept akin to green stormwater infrastructure that focuses more on features such as streams, ponds, and wetlands that are integrated with permanent surface waterbodies. As of 2023, there are too few of these projects to analyze statistically, but we anticipate that future green stormwater infrastructure socioeconomic analyses should see if there are differences in location characteristics and their potential co-benefits.

Our interviews with city officials in D.C. and NYC also characterized the appetite for the types of tools, data, and insights possible with machine learning techniques. They were interested in understanding socioeconomic and other patterns in the placement of green stormwater infrastructure and using that understanding to progress. In general, the officials also understood how machine learning tools hold the potential to offer greater flexibility than other methods for characterizing nonlinear relationships and handling large amounts of variables. On the other hand, they felt that counterintuitive findings might be challenging to communicate with residents as compared with traditional planning processes, and that investment decisions would still need some element of ground truth to resonate with the community.[40]

Conclusion

Green stormwater infrastructure investments have gained momentum across the United States as a distributed means to manage stormwater in an urban setting. They also offer a variety of co-benefits to the surrounding community compared with traditional gray infrastructure, including reduced urban heat island, improved water quality, and enhanced aesthetics. In this study, we employed machine learning to evaluate the location of installed green stormwater infrastructure and the demographic and land-use characteristics of the surrounding areas in Boston, Detroit, NYC, Pittsburgh, and D.C. Our goal was to understand whether green stormwater infrastructure implemented by these cities has been located in areas that, in addition to stormwater reduction, stand to gain from the co-benefits these investments provide. We also examined a broad set of machine learning approaches and evaluated their utility in characterizing these dynamics.

From a methods standpoint, no single methodology performed the best for all cities and no single methodology performed the worst. Instead, we found inconsistencies in model performance across cities, according to metrics for goodness of fit. Evaluating multiple machine learning approaches could therefore be advantageous in producing the best model for a given application and enhancing analytical flexibility. At the same time, machine learning–based analyses do not always require the most advanced approach. Our findings show that, in some locations, approaches such as the Lasso regression can still produce the best performing model. Analytical approaches that include multiple machine learning methods should therefore start simply and add model complexity only as needed. Finally, data acquisition, cleaning, and transformation were substantial parts of the analysis; the form of the independent variable had a strong influence on model performance. It is critical that machine learning, or any data-driven empirical analysis, have modelers that understand the nuances of datasets to ensure that results are meaningful.

From our city-level machine learning analysis, we found several correlations between local characteristics and the prevalence of green stormwater infrastructure. In Boston, green stormwater infrastructure is more associated with areas with poorer air quality, lower housing costs, and many residents facing respiratory and cardiac health issues. To lesser extent, this is also true of locations with a higher percent of Hispanic or Latino residents and a lower amount of tree canopy. In D.C., green stormwater infrastructure is more associated with locations with newer buildings, less population density, and lower economic inequality. In NYC, green stormwater infrastructure has a stronger association with places with warmer summer temperatures, more Hispanic and Latino residents, better air quality, less dense housing, and fewer residents with respiratory and cardiac health issues. In Pittsburgh, green stormwater infrastructure is associated with locations with newer housing stock and a higher percentage of Asian residents. And, finally, in Detroit, areas with a higher number of housing units per building and a higher risk of future flooding have a stronger association with green stormwater infrastructure, with lower-percent impervious surface and lower population density affiliated with more green stormwater infrastructure. Although conversations with stormwater managers in these locations often aligned with these findings, it was also clear that no cities we examined were consistently and deliberately planning based on a co-benefits framework. This finding suggests that some of these findings might by correlations at best and cities might not have been intentionally planning to achieve co-benefits with their investments. Looking forward, cities could look to each of the city-level findings to determine how planning priorities could shift with respect to green infrastructure if they intend to realize more co-benefits in the future.

From our interviews with NYC, D.C., and Pittsburgh's stormwater managers, we learned city stormwater planning has predominately been consent decree–driven and publicly-funded with a focus on reducing stormwater volumes. It is now, however, shifting toward an effort that is driven by compliance with municipal codes, requiring more active involvement from private developers and other landowners. Although this work cannot decide if and how co-benefits were considered, we can see from the models that some cities, such as Boston, might be well positioned to derive more co-benefits from their green stormwater infrastructure investments because of the location characteristics of existing assets. For other cities, more work is needed to understand how to influence site selection by considering co-benefits, such as in D.C.[41]

Finally, to support cities with these efforts, analyses such as this one can generate tools, data or insights to help cities understand existing socioeconomic and other patterns in the placement of green stormwater infrastructure. Machine learning tools hold the potential to enable city planners to evaluate investments and adjust policies and practices accordingly. They offer greater flexibility than other methods for dealing with nonlinear relationships and offer the ability to consider a wide variety of factors to see which are associated with the amount of green stormwater infrastructure in that area. With a growing appetite for the role advanced analytics can play in municipal planning, machine learning analysts must have expertise in the meaning and application of the datasets they work with, must be able to logically explain the results of their analyses, and should follow best practices for communicating AI to ensure the utility of the analysis and tools they produce.

Appendix
City Green Stormwater Infrastructure Policy

For each of the selected cities, we characterized the green stormwater infrastructure policy context, including the priorities and intent of city green stormwater infrastructure–related policies. Among the five cities, city green stormwater infrastructure investments have primarily centered on reducing the number and severity of CSO events, with a secondary level of investment in managing TMDLs to maintain MS4 permit compliance. In addition, based on our analysis of city green stormwater infrastructure plans and the findings generated by this study, redevelopment processes appear to be the single largest driver of where green stormwater infrastructure has been installed. Some of our case study cities have specific policies for green stormwater infrastructure placement on publicly-owned land, typically in street right of ways. In this section, we describe these policies and identify the main drivers behind green stormwater infrastructure investment and siting decisions in each city.

New York City

Compared with other cities in our analysis, NYC has the largest sewershed in terms of drainage area, the number of CSOs, and the amount and size of treatment plants. Several parts of the city operate under MS4 permits from the state, and, in these cases, green stormwater infrastructure focuses more on improving water quality than reducing runoff quantity. Figure 4 shows green stormwater infrastructure installations relative to the combined versus separated sewersheds.

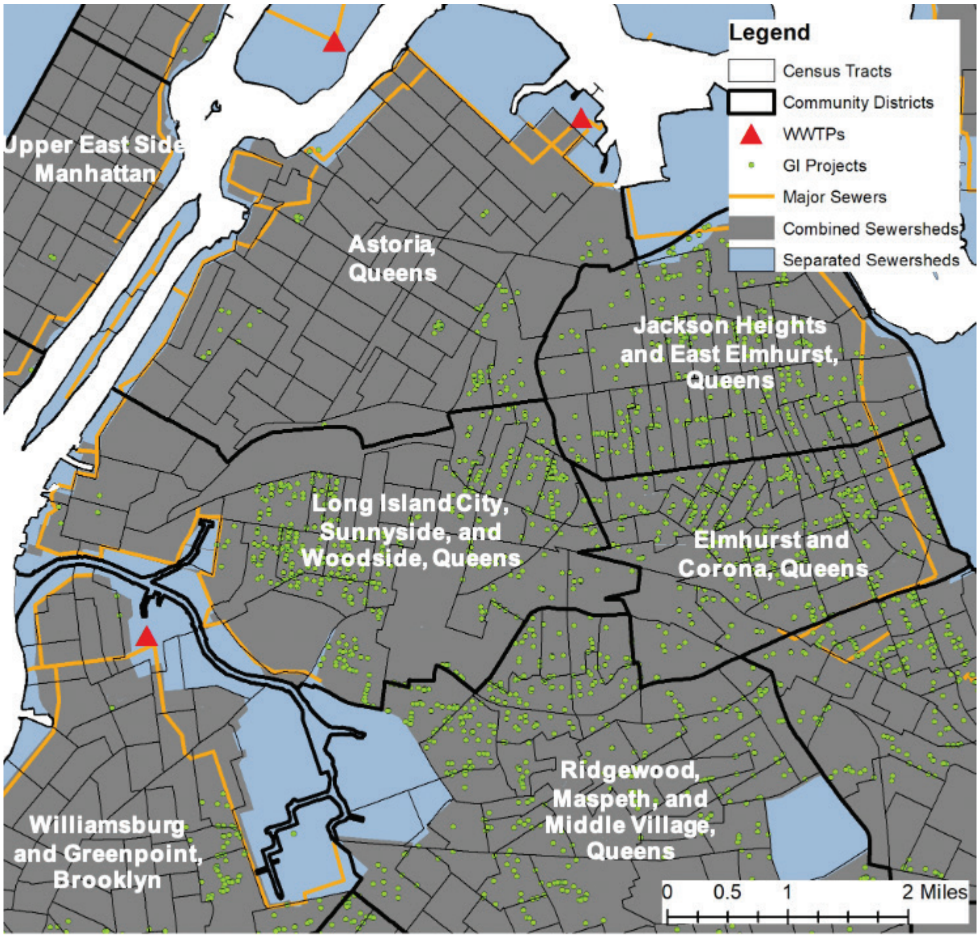
NYC operates its green stormwater infrastructure strategy through the DEP Green Infrastructure Plan, first unveiled in 2010 and incorporated into a 2012 consent decree.[42] There are several components to the plan, including standards for design and construction; a grant program to support installations; a review process to ensure standards are met and maintained; and ongoing education, outreach, monitoring, and evaluation. The 2010 plan was driven by regulatory compliance with the 2012 CSO Long Term Control Plan (LTCP).[43] The 2023 Citywide Green Infrastructure Modification,[44] however, acknowledged the deadline in the LTCP would not be met and delayed the annual CSO reduction goal of 1.67 billion gallons from 2030 to 2040, requiring DEP to expend an additional $2 billion on green stormwater infrastructure.[45] Future efforts will increasingly focus on other areas of the city that are covered by its MS4 permits and unique TMDL challenges,[46] such as addressing cloudburst events (a short, heavy rain event).

Boston

Beginning in the 1970s, the BWSC has made extensive capital investments toward cleaning up the harbor, especially by closing CSO outfalls and introducing new treatment facilities. More recently, a 2012 consent decree mandated several green stormwater infrastructure projects throughout the city, including a redesigned City Hall Plaza.[47] Since 2014, additional partnerships (such as one producing educational opportunities at public schools and one facilitating construction associated with redevelopment in MS4 areas)[48] have increased the amount of green stormwater infrastructure. In 2016, Boston officials codified the city's green stormwater infrastructure strategy as a part of the Environment Department's Climate Ready Boston planning process. In 2022, Mayor Michelle Wu appointed an inaugural director of green infrastructure to oversee coordination between different city agencies.[49] More broadly, all new development or substantial redevelopment must include infiltration capacity equivalent to the first 1.25 inches of rainfall on site, reducing the overall quantity and peak amount of runoff. Finally, the city has established a new utility fee that will take effect in 2024 that includes a stormwater charge and a stormwater credit and grant program. Not only does the fee apply to existing rate payers, but it also extends to any owner of impervious surfaces, including parking lots.[50] Together, this has amounted to 5,563 green stormwater infrastructure installations across the greater Boston area.

FIGURE 4

Green Infrastructure Installations Within Combined and Separated Sewersheds in New York City



SOURCE: Produced using data from "Open Sewer Atlas NYC"; and New York Department of City Planning, "Bytes of the Big AppleTM."
NOTE: WWTPs = wastewater treatment plants.

Washington, D.C.

D.C. operates its green stormwater infrastructure strategy through its Clean Rivers Project, launched in 2002. The strategy includes two components. The first involves a new capital-intensive interceptor system focused on CSO flow reduction led by DC Water.[51] The second is a distributed stormwater management system led by the DOEE. The latter was our primary focus. The city has issued extensive guidance for the private sector and a substantial portion of its areas in separated, MS4 sewersheds (many of which are undergoing rapid redevelopment).[52] In addition, redevelopment projects in combined sewersheds have the option of purchasing Stormwater Retention Credits rather than installing and maintaining green stormwater infrastructure on-site. Because this trading program can only place best management practices in the separated sewersheds, it also shifts the potential location-based co-benefits from one part of the city to another, which might have different socioeconomic characteristics.[53]

Pittsburgh

Pittsburgh's green stormwater infrastructure strategy arose out of the 2016 Citywide Green First Plan, which started as an effort to accomplish partial CSO reductions. Through the advocacy of such groups as 3 Rivers Wet Weather, which cataloged many of the 209 green stormwater infrastructure projects considered in this study, by 2020, the PWSA had launched a stormwater strategic plan and introduced a stormwater utility to collect fees. Akin to the district, where DC Water is leading with a gray-infrastructure CSO approach and DOEE is leading with distributed green stormwater infrastructure, the regional Allegheny County Sanitation Authority is building interceptors and a treatment plant expansion, whereas PWSA is focusing on neighborhood-level impacts and associated equity-based improvements.[54] One major challenge, however, is the hilly topography and clay-based soils that can require substantially engineered solutions. For example, planning to mitigate deadly flash flood events in Negley Run has required complex coordination between agencies ranging from the U.S. Army Corps of Engineers to the Pennsylvania Department of Transportation.[55]

Detroit

Detroit's Open Data Portal Green Stormwater Infrastructure Locations was started in 2013 and is administered by the DWSD. Many of the city's green stormwater infrastructure investments, typically in parks and city-owned rights of way, were developed with stakeholder input and have focused in the Upper Rouge Tributary area (approximately 27 percent of the city land area) where CSO discharges enter surface water. Other activities occurring across the city include tree planting and downspout disconnection. In addition, some neighborhoods have seen a reduction in impervious cover through demolition and the greening of vacant properties in accordance with the city's master plan. To incentivize these opportunities, the DWSD offers a drainage charge credit system. To drive additional improvements going forward the city is developing a new ordinance to require stormwater management and is restructuring its building code.[56] Though there are no stormwater fees to finance projects, there has been long-standing foundation-led funding of local nonprofits to ensure equity-based delivery of co-benefits.

Notes

[1] U.S. Environmental Protection Agency (EPA), "Protecting Water Quality from Urban Runoff."

[2] Elliot et al., "Identifying Linkages Between Urban Green Infrastructure and Ecosystem Services Using an Expert Opinion Methodology"; Earth Economics, *User Guide: Green Infrastructure Benefits Valuation Tool*.

[3] Diringer et al., *Incorporating Multiple Benefits into Water Projects: A Guide for Water Managers*.

[4] The Green Infrastructure Leadership Exchange, *The State of Public Sector Green Stormwater Infrastructure*.

[5] Consent decrees are issued by the EPA for violations of the Clean Water Act. Consent decrees require cities to act to reduce the impacts of stormwater on local waterways, often resulting in large infrastructure investments from cities under a consent decree (or multiple). NPDES permits require a Stormwater Management Program that has control measures in place to meet the Total Maximum Daily Loads (TMDLs), which are pollutant limits for impaired waterbodies covered by the NPDES federal program (that regulates point sources and nonpoint load allocations).

[6] Heckert and Rosan, "Developing a Green Infrastructure Equity Index to Promote Equity Planning, Urban Forestry & Urban Greening"; Grabowski, McPhearson, and Pickett, "Transforming US Urban Green Infrastructure Planning to Address Equity."

[7] For more on urban planning, see Jha et al., "A Review of AI for Urban Planning: Towards Building Sustainable Smart Cities." For more on air quality monitoring, see Castelli et al., "A Machine Learning Approach to Predict Air Quality in California"; Rybarczyk and Zalakeviciute, "Machine Learning Approaches for Outdoor Air Quality Modelling." For more on water management, see Doorn, "Artificial Intelligence in the Water Domain"; Xiang et al., "Urban Water Resource Management for Sustainable Environment Planning Using Artificial Intelligence Techniques." For more on stormwater, specifically, see Yang, Chang, and Chang, "AI-Based Design of Urban Stormwater Detention Facilities Accounting for Carryover Storage." For more on green stormwater infrastructure, see Kranjčić et al., "Machine Learning Methods for Classification of the Green Infrastructure in City Areas."

[8] Sanchez et al., "The Prospects of Artificial Intelligence in Urban Planning."

[9] Washington, D.C. will hereafter be referred to as *Washington* or *D.C.*

[10] Atlanta and Baltimore were excluded because they required fees to access their data.

[11] WGS84 is a highly accurate standard coordinate system used by the global positioning system that provides coordinates for a location using the Earth's center of mass as the origin.

[12] The *Gini Index* is a measure of income dispersion that is used to represent economic inequality.

[13] Walker, Herman, and Eberwein, "tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames."

[14] Lasso was not used specifically for the purpose of model selection, but instead for its regularization benefit.

[15] Ridge regression is like Lasso regression, except the method of regularization employed does not perform feature selection. All predictors included in a ridge regression have a nonzero coefficient, while Lasso frequently assigns a coefficient of zero to one or more predictors. In practice, this means that ridge regression might be more appropriate in situations where automatic feature selection is not desired (e.g., if the analyst has a strong prior or theoretical basis that all features have some effect on the outcome).

[16] Zou and Hastie, "Regularization and Variable Selection via the Elastic Net."

[17] A *bootstrap* is a statistical approach that repeatedly performs an estimation procedure on random subsamples of a dataset to approximate the distribution of estimates. This can be used to estimate standard errors and assess the stability of estimates.

[18] scikit-learn, homepage.

[19] Van der Walt, Colbert, and Varoquaux, "The NumPy Array."

[20] All features except high redlining score were continuous variables and were normalized. High redlining score was a binary variable.

[21] For example, for the Lasso Regression model, the "alpha" hyperparameter was tuned over the values [0.1, 0.5, 1.0, 5.0, 10.0]. For the ElasticNet model, the "alpha" and "l1_ratio" hyperparameters were tuned over the values [0.1, 0.5, 1.0, 5.0, 10.0] and [0.2, 0.5, 0.8], respectively.

[22] In the context of this work, K-Fold cross-validation was used to minimize the risk of overfitting, ensure that all data were used for both training and validation, thus optimizing the use of available data. Estimation procedures using cross-validation frequently use ten splits; however, because our data were relatively small, this would result in exceedingly small numbers of observations in each split. To prevent this, we used five splits.

[23] For the shallow neural network, the hyperparameters under consideration were the number of units in the single hidden layer, denoted as units, and the dropout rate, denoted as dropout. The number of units was varied in steps of 32 within a range of 32 to 1,024, and the dropout rate was varied in steps of 0.05 within a range of 0 to 0.5. For the deep neural network, similar hyperparameters were tuned but for two hidden layers instead of one. This resulted in four hyperparameters in total: units_1 and dropout_1 for the first hidden layer and units_2 and dropout_2 for the second hidden layer.

[24] Nori et al., "InterpretML"; Arya, et al. "One Explanation Does Not Fit All"; Yang et al. "OmniXAI."

[25] Because the focus of the study was on global feature contributions rather than trying to explain particular local predictions, SHAP was more suitable for our context than LIME and simpler and more accessible to implement than other libraries. Nori et al., "InterpretML"; Arya et al., "One Explanation Does Not Fit All"; Yang et al., "OmniXAI."

[26] For an example, see Wang et al., "Towards Better Process Management in Wastewater Treatment Plants."

[27] A limitation of our linear regression SHAP coefficient approach arose in the case of U-shaped relationships. For example, if green stormwater infrastructure investment in a city is a mix of public investment in lower-income areas and private investment by higher-income property owners, the relationship between green stormwater infrastructure and community income may be U-shaped, with more assets at the lower and higher end of the income distribution, respectively. In this case, the linear regression SHAP coefficient would be near zero and, thus, would not represent the meaningful nonlinear relationship between income and green stormwater infrastructure investment. In cases where U-shaped and other nonlinear relationships are of interest or explicitly hypothesized, an alternative approach including a square term in the regression and reporting this coefficient separately or instead using a LOESS curve might be more appropriate.

[28] SHAP regression coefficients for each city were normalized by the largest absolute value of SHAP regression coefficients within each city. This enabled us to understand which features might have the strongest association with the output variable, as well as compare across cities. SHAP regression coefficients were based on the Random Forest model for all cities.

[29] The strong positive association aligns with a redevelopment trend discussed during interviews with four environmental protection specialists at DOEE (DOEE environmental protection specialists, interview with the authors, June 21, 2023).

[30] The strong negative association could be because of an MS4-based stormwater credit described by environmental protection specialists at DOEE (DOEE environmental protection specialists, interview with the authors, June 21, 2023).

[31] *CSOs* are the discharge of combined storm and sanitary sewer systems into local waterways. They generally occur when stormwater runoff entering a combined system exceeds its capacity.

[32] For more information on East Boston, see Boston Planning and Development Agency, "PLAN: East Boston."

[33] For more information on how NYC saturated priority CSO tributary areas, see NYC DEP, "Green Infrastructure Contingency Plan."

[34] Other neighborhoods experiencing rapid recent redevelopment might not be suitable for green stormwater infrastructure installations because of clay soils, steep topography, or, in the case of the riverfronts, location in an MS4 area.

[35] Carnegie Mellon Heinz School Center for Economic Development, "Pittsburgh's Gateway Communities."

[36] DOEE environmental protection specialists, interview with the authors, June 21, 2023.

[37] A *sewershed* refers to the total land area that drains through a network of storm-sewer conveyances to a single wastewater treatment facility and discharge point.

[38] PWSA manager, interview with authors, June 21, 2023.

[39] DOEE environmental protection specialists, interview with the authors, June 21, 2023.

[40] NYC DEP director and engineer, interview with authors, May 16, 2023; DOEE environmental protection specialists, interview with the authors, June 21, 2023.

[41] DOEE environmental protection specialists, interview with the authors, June 21, 2023.

[42] State of New York Department of Environmental Conservation, "CSO Modification Order on Consent," 2012.

[43] NYC DEP, *NYC Green Infrastructure: 2022 Annual Report*.

[44] State of New York Department of Environmental Conservation, CSO Modification Order on Consent," 2023.

[45] Dulong, "State Proposes Changes to New York City's Green Infrastructure Program."

[46] These areas are referred to as separate storm sewers.

[47] EPA, "Settlement Requires Boston Water and Sewer Commission to Remedy Sewer and Stormwater Discharges."

[48] BWSC, *Compliance Report for the Period: July 1, 2022 Through December 31, 2022*; Metz and Kitchell, *Stormwater and Green Infrastructure Curriculum for Boston Public Schools*; Offshoots Productive Landscapes, "5 Boston Public Schoolyards."

[49] City of Boston "Kate England Appointed Boston's Inaugural Director of Green Design."

[50] BWSC, "A New Approach Funding Stormwater Management."

[51] An interceptor sewer receives wastewater from collectors and carries it to another interceptor, or the treatment facility.

[52] DOEE, *Stormwater Management Guidebook*.

[53] DOEE environmental protection specialists, interview with the authors, June 21, 2023.

[54] PWSA manager, interview with authors, June 21, 2023.

[55] PennPraxis and The Water Center, University of Pennsylvania, *Pittsburgh Water and Sewer Authority (PWSA) Strategic Plan for Stormwater*

[56] DWSD, *Green Infrastructure Progress Report Upper Rouge Tributary Area*.

[57] The Royal Society, *Communicating AI*.

[58] Centre for Strategic Communication Excellence, *Communicating AI*.

[59] Centre for Strategic Communication Excellence, *Communicating AI*.

[60] Bennett Institute for Public Policy, *How to Communicate Effectively to Policy Makers*.

[61] Centre for Strategic Communication Excellence, "Communicating Artificial Intelligence."

[62] Bennett Institute for Public Policy, *How to Communicate Effectively to Policy Makers*.

[63] Parker Software, "The Problem(s) with Tech for Tech's Sake."

[64] Centre for Strategic Communication Excellence, "Communicating Artificial Intelligence."

[65] U.S. Department of Health and Human Services, *Trustworthy AI (TAI) Playbook*. The National Institute of Standards has also identified the building blocks of responsible and trustworthy AI implementation that includes accuracy, explainability and interpretability, privacy, reliability, robustness, safety, security, and, importantly, the mitigation of harmful bias (see Information Technology Modernization Centers of Excellence, "AI Guide for Government"; High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*; Oxborough et al., *Accelerating Innovation: How to Build Trust and Confidence in AI*).

[66] Peet et al., *Machine Learning in Public Policy: The Perils and the Promise of Interpretability*.

[67] Deepchecks, "Model Confidence and How it Helps Model Validation."

[68] For the RAND Europe commentary, see Harris, Marcellino, and Slapakova, "Using Machine Learning to Detect Malign Information Efforts Online."

[69] Marcellino et al., *Human-Machine Detection of Online-Based Malign Information*.

[70] Schultz, Lovejoy, and Peet, *Using an Innovative Database and Machine Learning to Predict and Reduce Infant Mortality*.

References

Arya, Vijay, Rachel K. E. Bellamy, Pin-Yu Chen, Amit Dhurandhar, Michael Hind, Samuel C. Hoffman, Stephanie Houde, Q. Vera Liao, Ronny Luss, Aleksandra Mojsilović, Sami Mourad, Pablo Pedemonte, Ramya Raghavendra, John Richards, Prasanna Sattigeri, Karthikeyan Shanmugam, Moninder Singh, Kush R. Varshney, Dennis Wei, and Yunfeng Zhang, "One Explanation Does Not Fit All: A Toolkit and Taxonomy of AI Explainability Techniques," *arXiv*, September 14, 2019.

Bennett Institute for Public Policy, *How to Communicate Effectively to Policy Makers: A Guide for Academics*, University of Cambridge, August 2020.

Boston Planning and Development Agency, "PLAN: East Boston," webpage, undated. As of September 21, 2023: https://www.bostonplans.org/planning/planning-initiatives/plan-east-boston

Boston Water and Sewer Commission, "A New Approach Funding Stormwater Management" webpage, undated. As of August 2, 2023: https://www.bwsc.org/stormwater

Boston Water and Sewer Commission, *Compliance Report for the Period: July 1, 2022 Through December 31, 2022*, February 1, 2023.

BWSC—*See* Boston Water and Sewer Commission.

Carnegie Mellon Heinz School Center for Economic Development, "Pittsburgh's Gateway Communities," November 2003.

Castelli, Mauro, Fabiana Martins Clemente, Aleš Popovič, Sara Silva, and Leonardo Vanneschi, "A Machine Learning Approach to Predict Air Quality in California," *Hindawi*, August 4, 2020.

Center for Neighborhood Technology and American Rivers, *The Value of Green Infrastructure: A Guide to Recognizing Its Economic, Environmental, and Social Benefits*, 2010.

Centre for Strategic Communication Excellence, *Communicating AI: Building the Playbook 2019*, March 2019.

Centre for Strategic Communication Excellence, "Communicating Artificial Intelligence," webpage, undated. As of October 2, 2023: https://www.thecsce.com/resources/communicating-ai

City of Boston, "Kate England Appointed Boston's Inaugural Director of Green Design" press release, July 28, 2022.

Deepchecks, "Model Confidence and How it Helps Model Validation" blog, November 1, 2022.

Department of Energy and Environment, *Stormwater Management Guidebook*, January 2020.

Detroit Water and Sewerage Department, *Green Infrastructure Progress Report Upper Rouge Tributary Area*, April 1, 2022.

Dezeure, Ruben, Peter Bühlmann, Lukas Meier, and Nicolai Meinshausen "High-Dimensional Inference: Confidence Intervals, *p*-values and R-Software hdi," *Statistical Science*, Vol. 30, No. 4, November 2015.

Diringer, Sarah, Heather Cooley, Morgan Shimabuku, Sonali Abraham, Cora Kammeyer, Robert Wilkinson, and Madeline Gorchels, *Incorporating Multiple Benefits into Water Projects: A Guide for Water Managers*, Pacific Institute, June 2020.

DOEE—*See* Department of Energy and Environment.

Doorn, Neelke, "Artificial Intelligence in the Water Domain: Opportunities for Responsible Use," *Science of the Total Environment*, Vol. 755, No. 1, February 10, 2021.

Dulong, Mike, "State Proposes Changes to New York City's Green Infrastructure Program," Riverkeeper, blog, May 25, 2023.

DWSD—*See* Detroit Water and Sewerage Department.

Earth Economics, *User Guide: Green Infrastructure Benefits Valuation Tool*, last updated November 20, 2018.

Elliott, Robert M., Amy E. Motzny, Sudy Majd, Filiberto J. Viteri Chavez, Daniel Laimer, Benjamin S. Orlove, and Patricia J. Culligan, "Identifying Linkages Between Urban Green Infrastructure and Ecosystem Services Using an Expert Opinion Methodology," *Ambio*, Vol. 49, 2020.

EPA—*See* U.S. Environmental Protection Agency.

Grabowski, Zbigniew J., Timon McPhearson, and Steward T.A. Pickett, "Transforming US Urban Green Infrastructure Planning to Address Equity," *Landscape and Urban Planning*, Vol. 229, January 2023.

Green Infrastructure Leadership Exchange, *The State of Public Sector Green Stormwater Infrastructure*, 2022.

Harris, Ruth, William Marcellino, Linda Slapakova, "Using Machine Learning to Detect Malign Information Efforts Online," RAND Europe, undated.

Heckert, Megan, and Christina D. Rosan, "Developing a Green Infrastructure Equity Index to Promote Equity Planning, Urban Forestry & Urban Greening," Vol. 19, No. 1, September 2016.

High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, European Commission, April 8, 2019.

Information Technology Modernization Centers of Excellence, "AI Guide for Government," webpage, undated. As of September 21, 2023: https://coe.gsa.gov/coe/ai-guide-for-government

Jha, Avinash Kumar, Awishkar Ghimire, Surendrabikram Thapa, Aryan Mani Jha, and Ritu Raj, "A Review of AI for Urban Planning: Towards Building Sustainable Smart Cities," Delhi Technology University, January 5, 2021.

Kranjčić, Nikola, Damir Medak, Robert Župan, and Milan Rezo, "Machine Learning Methods for Classification of the Green Infrastructure in City Areas," *International Society for Photogrammetry and Remote Sensing International Journal of Geo-Information*, Vol. 8, No. 10, October 22, 2019.

Marcellino William, Kate Cox, Katerina Galai, Linda Slapakova, Amber Jaycocks, and Ruth Harris, *Human-Machine Detection of Online-Based Malign Information*, RAND Corporation, RR-A519-1, 2020. As of September 21, 2023: https://www.rand.org/pubs/research_reports/RRA519-1.html

Metz, Kristin, and Anne Kitchell, *Stormwater and Green Infrastructure Curriculum for Boston Public Schools*, April 2018.

Nori, Harsha, Samuel Jenkins, Paul Koch, and Rich Caruana, "InterpretML: A Unified Framework for Machine Learning Interpretability," *arXiv*, September 19, 2019.

New York City Department of City Planning, "Bytes of the Big AppleTM," webpage. As of August 1, 2023: https://www.nyc.gov/site/planning/data-maps/open-data.page

New York City Department of Environmental Protection, “Green Infrastructure Program Map” webpage, undated. As of August 2, 2023: https://nycdep.maps.arcgis.com/apps/webappviewer/index.html

New York City Department of Environmental Protection, “Green Infrastructure Contingency Plan,” June 30, 2021.

New York City Department of Environmental Protection, *NYC Green Infrastructure: 2022 Annual Report*, 2022.

NYC DEP—*See* New York City Department of Environmental Protection.

Offshoots Productive Landscapes, “5 Boston Public Schoolyards,” webpage, undated. As of August 2, 2023: https://www.offshootsinc.com/project/bwsc/

Open Sewer Atlas NYC, “Data,” webpage, undated. As of accessed August 1, 2023: https://openseweratlas.tumblr.com/data

Oxborough, Chris, Euan Cameron, Laurence Egglesfield, and Aldous Birchall, *Accelerating Innovation: How to Build Trust and Confidence in AI*, PricewaterhouseCoopers, 2017.

Parker Software, “The Problem(s) with Tech for Tech’s Sake,” blog, undated.

Peet, Evan D., Brian G. Vegetabile, Matthew Cefalu, Joseph D. Paine, Cheryl L. Damberg, *Machine Learning in Public Policy: The Perils and the Promise of Interpretability*, RAND Corporation, PE-A828-1, November 2022. As of September 21, 2023: https://www.rand.org/pubs/perspectives/PEA828-1.html

PennPraxis and The Water Center, University of Pennsylvania, *Pittsburgh Water and Sewer Authority (PWSA) Strategic Plan for Stormwater*, March 28, 2022.

Psa1966, “Bioretention Area in Emeryville, California,” 2013 (CC-BY-SA-4.0).

Public Law 115-436, Water Infrastructure Improvement Act, January 14, 2019.

Royal Society, *Communicating AI: The Role of Researchers in Enabling Public Conversations About AI*, May 2020.

Rybarczyk, Yves, and Rasa Zalakeviciute, “Machine Learning Approaches for Outdoor Air Quality Modelling: A Systematic Review,” *Applied Sciences*, Vol. 8, No. 12, December 11, 2018.

Sanchez, Thomas W., Hannah Shumway, Trey Gordner, and Theo Lim, “The Prospects of Artificial Intelligence in Urban Planning,” *International Journal of Urban Sciences*, Vol. 27, No. 2, April 2023.

Schultz, Dana, Susan L. Lovejoy, and Evan D. Peet, *Examining Interventions to Address Infant Mortality in Allegheny County, Pennsylvania*, RR-A858-1, 2020. As of October 27, 2023: https://www.rand.org/pubs/research_reports/RRA858-1.html

scikit-learn, homepage, undated. As of September 21, 2023: https://scikit-learn.org

Sonnenfeld, David A., “Green Infrastructure (Water Retention), SUNY-ESF, Syracuse, NY,” 2023 (CC-BY-SA-4.0)

State of New York Department of Environmental Conservation, “CSO Modification Order on Consent,” 2012.

State of New York Department of Environmental Conservation, CSO Modification Order on Consent,” 2023.

Tebyanian, Nastaran, Jordan Fischbach, Robert Lempert, Debra Knopman, Hong Wu, Lisa Iulo, and Klaus Keller, “Rhodium-SWMM: An Open-Source Tool for Green Infrastructure Placement Under Deep Uncertainty, *Environmental Modelling & Software*, Vol. 163, May 2023.

U.S. Department of Health and Human Services, *Trustworthy AI (TAI) Playbook*, September 2021.

U.S. Environmental Protection Agency, “Protecting Water Quality from Urban Runoff,” EPA 841-F-03-003, February 2003.

U.S. Environmental Protection Agency, “Settlement Requires Boston Water and Sewer Commission to Remedy Sewer and Stormwater Discharges,” press release, August 23, 2012.

van der Walt, Stéfan Johann, S. Chris Colbert, and Gael Varoquaux, “The NumPy Array: A Structure for Efficient Numerical Computation,” *Computer in Science & Engineering*, Vol. 13, No. 2, May 2011.

Walker, Kyle, Matt Herman, and Kris Eberwein, “tidycensus: Load US Census Boundary and Attribute Data as ‘tidyverse’ and ‘sf’-Ready Data Frames,” Version 1.4.4., webpage interface. As of September 18, 2023: https://CRAN.R-project.org/package=tidycensus

Wang, Dong, Sven Thunéll, Ulrika Lindberg, Lili Jiang, Johan Trygg, and Mats Tysklind, “Towards Better Process Management in Wastewater Treatment Plants: Process Analytics Based on SHAP Values for Tree-Based Machine Learning Methods,” *Journal of Environmental Management*, Vol. 201, January 1, 2022.

Wang, Jia, Jiahong Liu, Hao Wang, and Chao Mei, “Approaches to Multi-Objective Optimization and Assessment of Green Infrastructure and Their Multi-Functional Effectiveness: A Review, *Water*, Vol. 12, No. 10, September 28, 2020.

Xiang, Ziaojun, Qiong Li, Shahnawaz Khan, and Osamah Ibrahim Khalaf, “Urban Water Resource Management for Sustainable Environment Planning Using Artificial Intelligence Techniques,” *Environmental Impact Assessment Review*, Vol. 86, January 2021.

Yang, Shun-Nien, Li-Chiu Chang, and Fi-John Chang, “AI-Based Design of Urban Stormwater Detention Facilities Accounting for Carryover Storage,” *Journal of Hydrology*, Vol. 575, August 2019.

Yang, Wenzhuo, Hung Le, Silvio Savarese, and Steven C. H. Hoi, “OmniXAI: A Library for Explainable AI,” *arXiv*, June 1, 2022.

Zou, Hui, and Trevor Hastie, “Regularization and Variable Selection via the Elastic Net,” *Journal of the Royal Statistical Society, Series B*, Vol. 67, No. 2, 2005.

About This Report

Green stormwater infrastructure has been increasingly utilized across the United States over the past few decades. In some locations, cities have invested heavily in this type of infrastructure to reduce urban flooding and manage water quality. Green stormwater infrastructure also offers a variety of co-benefits to the surrounding community compared with traditional gray infrastructure. These include reduced urban heat island effect, improved water quality, and enhanced aesthetics. This report examines the results of an exploratory machine learning–based analysis of green stormwater infrastructure asset data across five cities in the United States. Within each city, we evaluated the location of installed green stormwater infrastructure based on the demographic and land use characteristics of the surrounding area. The goal of our analysis was to understand the local context surrounding green stormwater infrastructure investments. This evaluation can help cities understand the current potential for co-benefits of these investments and how future planning can enhance the co-benefits of green stormwater infrastructure.

Community Health and Environmental Policy Program

RAND Social and Economic Well-Being is a division of the RAND Corporation that seeks to actively improve the health and social and economic well-being of populations and communities throughout the world. This research was conducted in the Community Health and Environmental Policy Program within RAND Social and Economic Well-Being. The program focuses on such topics as infrastructure, science and technology, industrial policy, community design, community health promotion, migration and population dynamics, transportation, energy, and climate and the environment, as well as other policy concerns that are influenced by the natural and built environment, technology, and community organizations and institutions that affect well-being. For more information, email chep@rand.org.

Funding

Funding for this research was provided by gifts from RAND supporters and income from operations.

Acknowledgments

We would like to thank the various representatives of the cities of New York, Boston, Washington, D.C., Pittsburgh, and Detroit with whom we interacted over the course of the project. We would also like to thank Lisa Jaycox and Nancy Staudt from the RAND Corporation for supporting this work through the RAND Initiated Research program, as well as our peer reviewers Flannery Dolan and Sarah Diringer for their constructive and thorough review.

For more information on this publication, visit www.rand.org/t/RRA2846-1.

www.rand.org

RR-A2846-1