# A picture containing text, clipart Description automatically generated

**CUSP 2022 Capstone**

**Urban Science Intensive (USII) GX 5005**

**HARDENING NEW YORK CITY WATER/WASTEWATER INFRASTRUCTURES AGAINST EXTREME WEATHER EVENTS AND CYBERATTACKS**

**Proposed by:**

Charalampos Avraam, PhD

Yury Dvorkin, PhD

Seung Hwa Lee

Vaishnavi Muthukrishnan

Wonchan Lee

**DISCLAIMER**

The data and analysis that were used in this report are gathered from accessible open sources. Some of the information portrayed are extrapolated and estimated to the best of our abilities. Information published on this report is for general knowledge purpose only, and reliability and accuracy aren’t assured, therefore, created databases and analysis performed on these databases should not be utilized for any official purposes.

You may visit our GitHub for more information on how we made our databases and codes.

**TABLE OF CONTENTS**

[Abstract 2](#_Toc111940151)

[Introduction 3](#_Toc111940152)

[Literature Review 4](#_Toc111940153)

[Problem Statement 4](#_Toc111940154)

[Data 5](#_Toc111940155)

[Data Sources 5](#_Toc111940156)

[Data Handling 9](#_Toc111940157)

[Limitations 10](#_Toc111940158)

[Scenario #1: Impact of Extreme Weather Event on Wastewater System 11](#_Toc111940159)

[Description 11](#_Toc111940160)

[Discussion of Model 12](#_Toc111940161)

[Analytic Approaches 14](#_Toc111940162)

[Limitations 17](#_Toc111940163)

[Findings and Interpretations 17](#_Toc111940164)

[Policy Implications 20](#_Toc111940165)

[Scenario #2: Water Distribution Network Analysis 20](#_Toc111940166)

[Description 20](#_Toc111940167)

[Discussion of Models 21](#_Toc111940168)

[Limitations 22](#_Toc111940169)

[Findings and Interpretations 22](#_Toc111940170)

[Policy Implementations 23](#_Toc111940171)

[Conclusions 23](#_Toc111940172)

[References 24](#_Toc111940173)

[Appendix 26](#_Toc111940174)

# Abstract

Extreme events stress New York City’s (NYC’s) interdependent energy and water infrastructures; impact human livelihood; and can disrupt local ecosystems. The dependence of water and wastewater operations on electricity implies that a blackout, coupled with backup system components’ failures, can force the discharge of untreated wastewater into NYC’s waterways, and result in a public health emergency. Data-driven and optimization techniques can leverage publicly available data to reveal vulnerabilities in electricity and water/wastewater infrastructures. Our analysis can aid policy design against natural hazards and cyberattacks, and thus inform the modernization of interdependent urban electricity and water/wastewater infrastructures.

# Introduction

Water requires a tremendous amount of energy to move from a reservoir or well, through the treatment process, and out into a distribution system. In addition, energy is required to process wastewater and recycle or discharge it [1]. Water’s dependence on energy opens it up to great risk of structural breakdown whenever there is a disruption in the power sector. Natural hazards like hurricanes, floods, and cyberattacks stress regional water and electricity infrastructures beyond their limits at an increasing rate and intensity [2] and threaten human livelihood and the environment.

Hurricane Ida hit New York City (NYC) in September 2021 and caused more than 57,000 power outages [3], at least 10 deaths, and total damage of ~$95 billion [4]. Even a recent snowstorm in January 2022 left tens of thousands of households without electricity from Georgia to Massachusetts [5]. As natural events are also exacerbated by climate change, these disruptions are only expected to increase in frequency and intensity. [6]

In [October](https://us-cert.cisa.gov/sites/default/files/publications/AA21-287A-Ongoing_Cyber_Threats_to_U.S._Water_and_Wastewater_Systems.pdf) 2021, the FBI, National Security Agency, Cybersecurity and Infrastructure Security Agency (CISA) and Environmental Protection Agency jointly issued a warning about ongoing malicious cyber activity targeting the information technology (IT) and operational technology (OT) networks, systems, and devices of the U.S. [Water and Wastewater Systems (WWS) Sector facilities](https://www.cisa.gov/water-and-wastewater-systems-sector). About attempts to compromise system integrity via unauthorized access—threatening the ability of WWS facilities to provide clean, potable water to, and effectively manage the wastewater of, their communities [7]. In February, a hacker or hackers breached the water-treatment system in [Oldsmar, Fla.,](https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2021/03/10/florida-hack-exposes-danger-to-water-systems) and attempted to raise the level of sodium hydroxide, or lye, in the water more than 100-fold — from 100 parts per million to [11,100 parts per million](https://www.techrepublic.com/article/fbi-secret-service-investigating-cyberattack-on-florida-water-treatment-plant/). Sodium hydroxide, used to control water acidity, is poisonous at high levels [8].

Thus, the automation of technology in water utilities implemented over the past two decades to save money and increase efficiency has also exposed them to malicious cyber activity that could disrupt or manipulate services. Combined with strong dependence of water supply chains on electricity, the current water infrastructure exposes itself to potentially calamitous breakdown. To be better prepared against such scenarios, federal agencies like NERC are organizing preparation exercises. On November 16-17, 2021, more than 700 planners participated in GridEx VI to exercise their response and recovery plans in the face of simulated, coordinated cyber and physical attacks on the North American bulk power system and other critical infrastructure [9].

The motivation of this project is to study the current infrastructure of the water system in New York City and identify its vulnerability points against electricity failure and cyber-attack and analyze the best prevention and/ or mitigation strategies against natural and man-made disruptions to the system.

The project examines two specific scenarios: impact on extreme events on wastewater systems and water distribution network analysis. Each scenario will observe the vulnerabilities using machine learning methods and visualize the results to open possible future implications and recommendations.

# Literature Review

Water distribution data is highly sensitive and often openly unavailable, so to build our water distribution network, we took inspiration from a work that developed a model that allows the topology of an urban water infrastructure to be mapped using freely available data [10]. The work leverages the parallelism of water and urban transportation infrastructures to identify the topology of a network by applying optimization methods.

Another study aimed to develop a method that locates critical nodes without hydraulic analysis of every failure scenario with the Flow Distribution method [11][12]. The Flow Distribution method is the application of the gravity model, typically used to predict traffic flows in transportation engineering, to a distribution system.

# Problem Statement

This project aims to identify supply chain vulnerabilities of New York City’s physical water and wastewater infrastructure; and inform resilience policies against natural disasters. The project will:

* Provide a water and wastewater management framework which integrates publicly available databases.
* Perform a scenario for each water and wastewater infrastructures using different machine learning methods.
* Visualize the results to clearly display the impact from the vulnerabilities.
* Pose possible future implications or recommendations.

Assess vulnerabilities and inform water and wastewater infrastructure policy design in New York City against extreme events and potential cyberattacks.

# Data

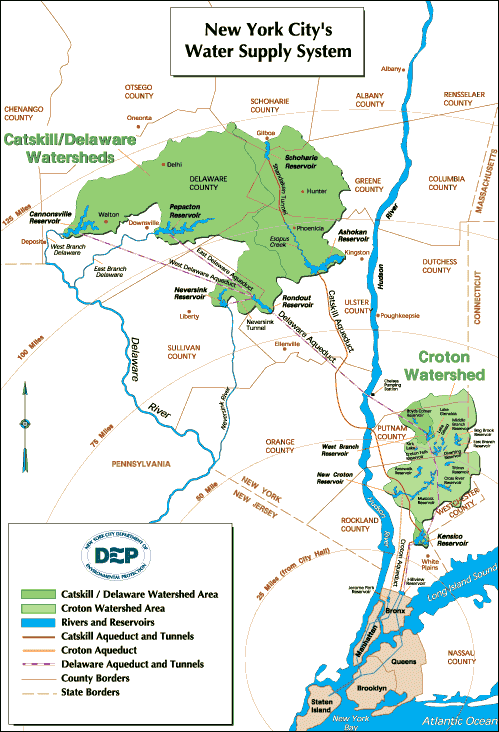
## Data Sources

**Reservoir Database**

The reservoir database contains 5 columns: Reservoirs, Watershed, Storage (BG), Capacity (BG), and Outflow (BGD). There are two main watersheds from Catskill/Delaware and Croton System, each having its designated reservoirs. Catskill/Delaware has 6 specific reservoirs: Ashokan, Schoharie, Rondout, Neversink, Pepacton, and Cannonsville. Data for Storage and Outflow were collected from NYC Open Data, where the current reservoir levels and details were available [13], and the Capacity was taken from the NYC government database [14]. For the Croton System, two reservoirs were collected: New Croton and Jerome Park. The New Croton Capacity data was accessible from the NYC government database [15]and Outflow data was taken from USGS [16]. New Croton Reservoir data was mainly selected since it is the collecting point from all reservoirs in the Croton Watershed. Jerome Park Reservoir was built to help distribute water and also as an additional source of water storage within the city. From the NYC government database [17], the capacity data was available, whereas the Outflow data had to be extrapolated.

**Water Treatment Plant Database**

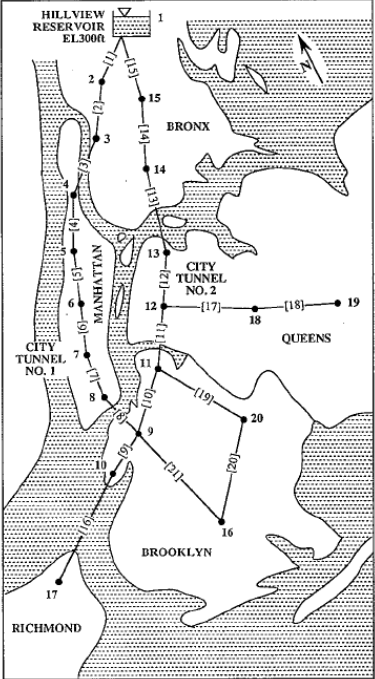
The water treatment plant database contains 4 columns: Water Treatment Plants, Capacity (BG), Inflow (BGD), and Outflow (BGD). There are two water treatment plants: Catskill-Delaware Water Ultraviolet Disinfection Facility and Croton Water Filtration Plant. For clarification, the Capacity portrays how much water a plant can filter in on a daily basis. Catskill-Delaware Water UV Disinfection Facility’s Capacity was available from an article from Water-Technology [18], and Capacity for Croton Water Filtration Plant was accessible from the NYC government database [19]. While Inflow data is essentially the sum of reservoirs’ Outflows, there was no attainable data with each of the water treatment plant’s Outflow; therefore, we had to make a deduction by formulating a calculation of water demand and each plant’s capacity.



*Figure 1: New York City Water Supply System*

**New York City water distribution tunnels**

The New York City primary water supply tunnel system is shown in Figure 1. [20] It is a gravity flow system that draws water (2017.5 feet3(s or 57,129.5 Lis) from the Hillview Reservoir at node 1. The network has two city tunnels. City tunnel no. 1 extends from Hillview Reservoir to node 16 in Brooklyn by way of Manhattan. City tunnel no. 2 extends between Hillview Reservoir and Queens.



*Figure 2: New York City Water Supply Tunnels*

**Precipitation data**

An average uniform rate of rainfall was considered for different scenarios of low, moderate, and heavy rainfall according to the United States Geological Survey [21]. The values considered for different scenarios are listed below.

|  |  |
| --- | --- |
| Rainfall Scenario | Precipitation rate (inches/ hour) |
| Low | 0.1 |
| Moderate | 0.3 |
| Heavy | 1 |
| Hurricane Ida | 3 |

**Combined Sewer Outfall points and shed**

The over 450 combined sewer outfalls in New York City, with information about their spatial point locations, the annual volume of CSO discharge, number of events per year, wastewater recovery facility they are connected to, were sourced from Open Sewer Atlas [22]

**Wastewater Recovery Plants**

New York City has 14 Wastewater Recovery Plants situated around the five boroughs. The spatial locations and daily inflow capacity of each plant was utilized for creating the wastewater treatment system database for the project and was sourced from New York City Department of Environmental Protection. [22]

**Water Demand**

New York City’s estimated daily water demand data by county, zip code, and type of industries (Commercial, Industrial, Residential). To create this estimated data set, input data were derived from US county-level water and energy use data from United States Geological Survey (USGS) and NYC Open Data.

**Sewer Complaint**

Historical data of New York City’s sewer complaints reported from 2010 to present days by zip code. The data set originated from NYC Open Data’s 311 Service Requests from 2010 to present days.

## Data Handling

**Reservoir Database**

For the Catskill/Delaware watershed, we filtered and sorted out the required columns and renamed the names based on the data dictionary provided along with the dataset. Storage and Capacity were already in the desired measurement of billion gallons (BG), and we converted the Outflow measurement to billion gallons per day (BGD) from million gallons per day (MGD) in order to keep the measurements consistent and easy to follow. From the dataset, it was difficult to view all the corresponding data for each of the reservoirs, so we transposed the data columns, portraying a vivid display for later calculations or necessary extractions for modeling.

**Water Treatment Plants Database**

There were two water treatment plants that were handled: Catskill-Delaware Water Ultraviolet Disinfection Facility and Croton Water Filtration Plant. The Capacity (BG) was in the desired measurement, The Inflow data measured in billion gallons per day (BGD) was calculated by adding up the Outflows from the reservoirs associated with their designated plants. The Outflow (BGD) data was not available, so we had to extrapolate this data from existing databases like the Water Demand Database.

Catskill-Delaware Water UV Disinfection Facility Outflow:

*(Demand \* (CapacityCatskill-Delaware / CapacityCatskill-Delaware + CapacityCroton))*

Croton Water Filtration Plant Outflow:

*(Demand \* (CapacityCroton / CapacityCroton + CapacityCatskill-Delaware))*

The Demand was acquired from the Water Demand Database’s total daily water use (kgal/d) and Capacity (BG) for both water treatment plants were already available. After taking the sum of the daily demand, it was converted to BGD and calculated with the Capacity values to derive to our Outflow data.

**Precipitation**

The precipitation data was considered as a uniform rainfall rate in inches for each rainfall scenario. For our analysis purposes, we needed to convert it into Million Gallons per Hour based on the area of the zip code. The area in square inches was sourced from New York City Survey data. The following equation was used to convert the data into Million Gallons per Hour.

**Water Demand**

There were two water demand data collected as input data. First one was 2015 US county level water use data from USGS. The dataset has water use by county in Mgal/day. The second data was 2020 annual NYC water consumption data for privately owned buildings over 25,000 ft2 and City-owned buildings over 10,000 ft2. The dataset was annual water use of 2020 in zip code level by types of industries (Industrial, commercial, residential). Although 2020 annual water consumption was at zip level, 40 % of the data was unavailable. So, we calculated percent water use by industry from the 2020 zip level data set and applied it to 2015 county level data which covers larger boundaries of industry and locations to create our final zip level data set. The final output data is also an estimate, and it can be way off from input data. Therefore, more effort to collect actual data to inform decision making is needed.

Also, there are large discrepancies between two input datasets we used. The two data sets deviate from each other by 466676.62060191977 at county level.

**Sewer Complaint**

Number of sewer complaints reported during hurricane Ida was used as a representative of affected citizens during hurricane Ida. Therefore, in order to use this data in modeling, prediction, and spatial analysis, we first filtered complaints reported during hurricane Ida and normalized the number of complaints in each zip code by their populations.

## **Limitations**

**Water Treatment Plant Database**

A discrepancy was found when there were values from the Inflow data that exceeds the Croton Water Filtration Plant’s Capacity. This did not make sense since the daily inflow of water cannot go over the plant’s capacity, so further research was needed to explain this possible inconsistency. One finding provided the information about the existence of Jerome Park Reservoir, which act as an additional source of water storage, receiving water from the Croton watershed [17]. Having a capacity of 770 million gallons, water traveling from New Croton to the Croton Water Filtration Plant can make an extra stop at the Jerome Park Reservoir in order to hold additional water to avoid overshooting of the plant’s capacity.

From the database, 438 out of 1094 Inflow values greater than the Capacity of 0.29 BG were converted to 0.29 BG. However, an addition was made in the Reservoir Database where Jerome Park Reservoir was added to the reservoirs in the Croton watershed. The only required column for the database was the Outflow, which was extrapolated by utilizing the IF statement in Excel performing a logical test where if an Output value from New Croton is greater than or equal to 0.29 BG, it would print out 0.29, but if it were, it would print out whatever cell value it has. For example, if the Outflow value from New Croton is 0.4115 BG, the cell value for Outflow from Jerome Park Reservoir would be 0.29 BG, whereas if it was 0.1502 BG, the cell value would just be 0.1502 BG.

Facing the limitation of insufficient data, further research and data extrapolation were necessary to derive on such an estimation, indicating a possibility of incredible information.

**Wastewater and Water Distribution**

Due to the sensitive nature of the information, most data about water distribution and wastewater treatment systems are not publicly available. Thus, this project relies on sparse publicly available data and makes assumptions to clean, modify and utilize it.

# Scenario #1: Impact of Extreme Weather Event on Wastewater System

## Description

Every year in New York City, around 20 billion gallons of untreated sewage and overflowed polluted water pass through the city’s sewage treatment plants and spread into the shoreline in all five boroughs. Additionally, according to the data provided by Notify NYC, there was nearly a 50 percent chance that water somewhere in New York City was not safe to drink or touch due to sewage overflow and pollution on any random day in 2019 [23].

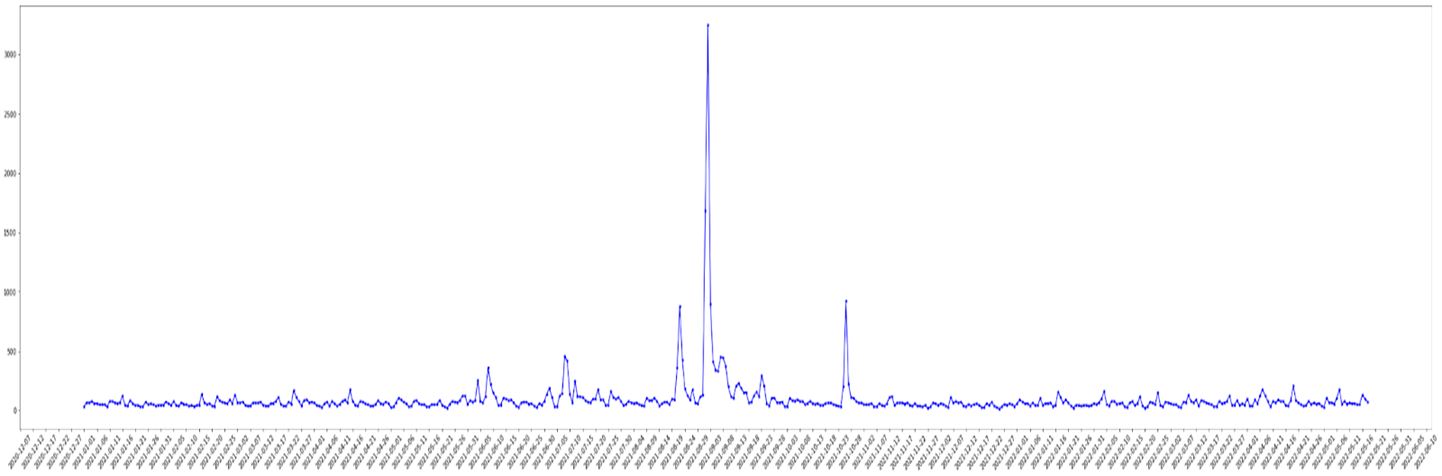
These overflows called combined sewer overflows can be caused by extreme weather conditions such as heavy rainfall or snowmelt. Increased volume of wastewater produced by significant rainfall or snow stresses capacity of the sewer system causing overflows, releasing untreated wastewater containing chemicals and other contaminants into waterways [24].

Bacteria, intestinal worms, protozoa, and viruses in combined sewer overflows (CSO) discharge can cause serious public health issues. Contact, inhalation, or ingestion of CSO discharge can cause diarrhea, nausea and other infections in the ear, respiratory system, and skins. In the worst case, exposure to these discharges can cause life–threatening diseases such as cholera, dysentery, and gastroenteritis. Moreover, if the overflowed water carries chemicals or toxins the consequences can be worse. A person who ingests chemicals or toxins in the water can be at risk of cancer, hormone disruption, altered brain function, damage to immune system, cardiovascular and kidney problems.

Additionally, since consequences of CSO can cause serious damage in public health, sewage treatment systems are at risk of cyber-attacks. Therefore, it is important to know where overflow is happening and is most vulnerable to heavy rainfall and cyber-attacks, when it is happening, and how severe are the impacts of sewage overflows to minimize the impact of CSO discharge and prevent future possible problems caused by CSO. The main goals in this scenario are finding areas that are most likely to be affected by CSO discharge, predicting possible impacted areas in the next extreme weather event due to certain factors such as location and discharge amount, and lastly further understanding demographics of impacted areas. We will approach statistically using Sewage overflow data during hurricane Ida, demographics data, and number of sewer complaints reported during hurricane Ida to answer the following questions.

## Discussion of Model

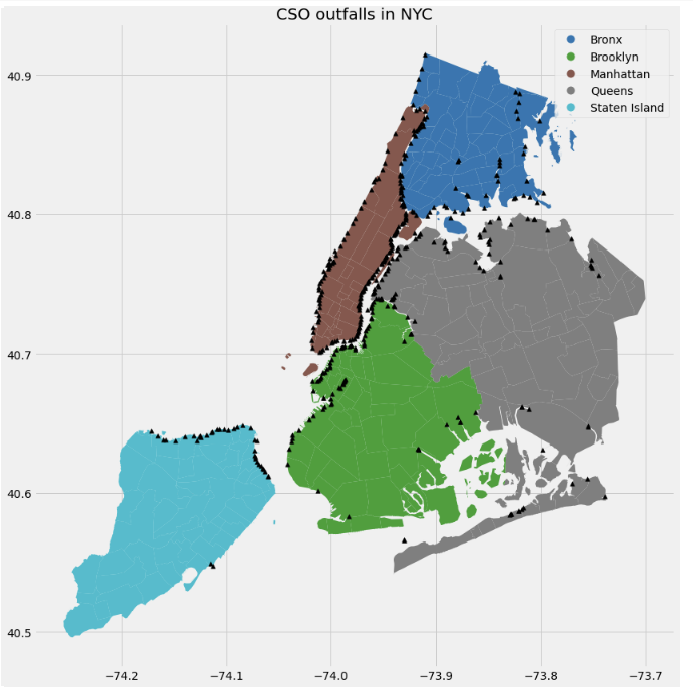
During hurricane Ida, the number of sewer complaints reported dramatically increased.



*Figure 3: Time plot of number of sewer complaints (2020/12-2022/6)*

In *Figure 3,* there has been a spike that started on 2021/8/26 which is the beginning of hurricane Ida and lasts until 2021/9/4. From this, high correlation between the number of sewer complaints and hurricane Ida could be found. Therefore, the number of sewer complaints were used as representative of the amount of impact from Ida throughout this scenario.

From visualization of CSO discharge points created from Combined Sewer Outfall points, sewer shed and Wastewater Recovery Plants data, an hypothesis was made that areas (zip codes) that are close to CSO outfall points will have relatively high number of sewer complaints reported during Ida.

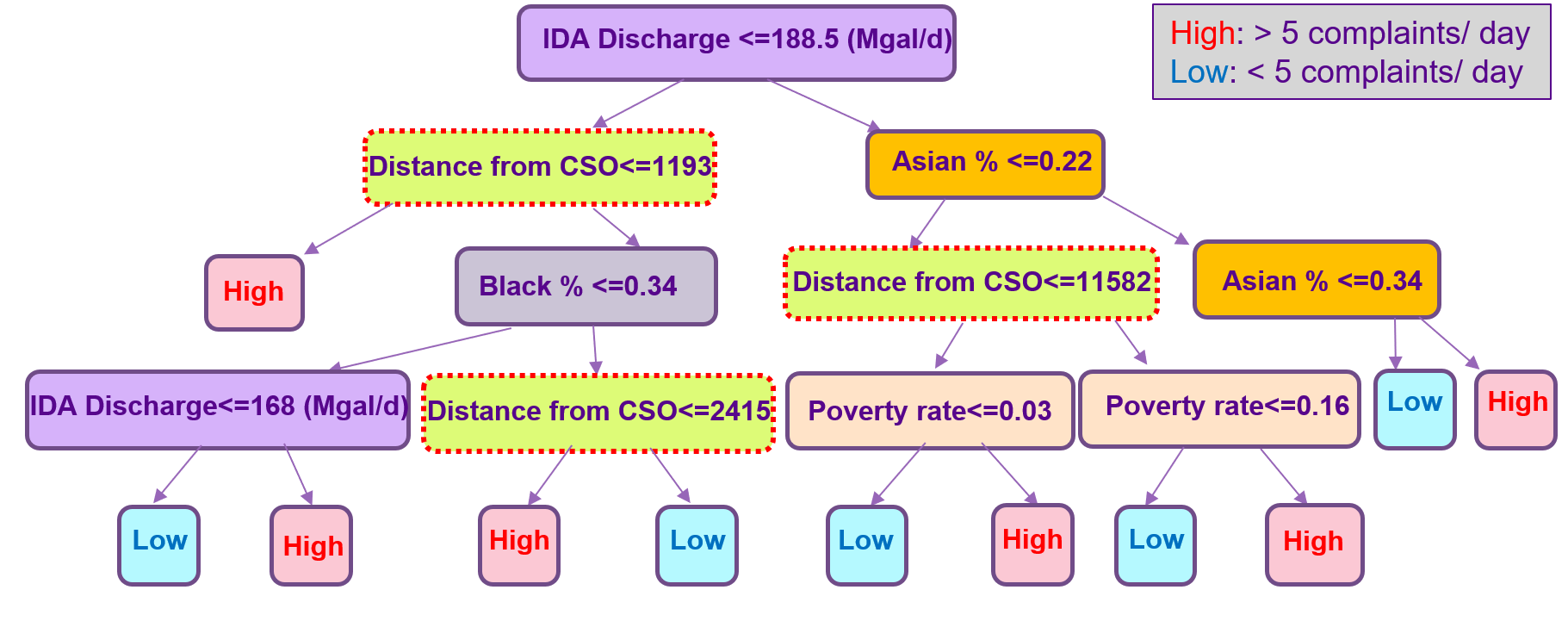


*Figure 4: CSO outfall points in NYC*

In order to prove the hypothesis, decision tree model and spatial analysis were used.

Decision tree model was trained to find the strongest factor to the number of complaints. Number of complaints during Ida was divided into two classes: low complaints and high complaints. The tree was trained on the amount of discharge during Ida, demographic features (race, poverty rate), and distance from nearest discharge point.

Each zip codes nearest distance from discharge point was calculated as Euclidean distance between center of each zip code and nearest discharge point.



*Figure 5: Decision tree model*

From the decision tree, it can be observed that distance from discharge point is a strong factor that leads to a high level of complaints. This observation supports the idea that zip codes closer to discharge points will be more affected during extreme weather events.

In the next step, further study on demographics of impacted areas were done.

## Analytic Approaches

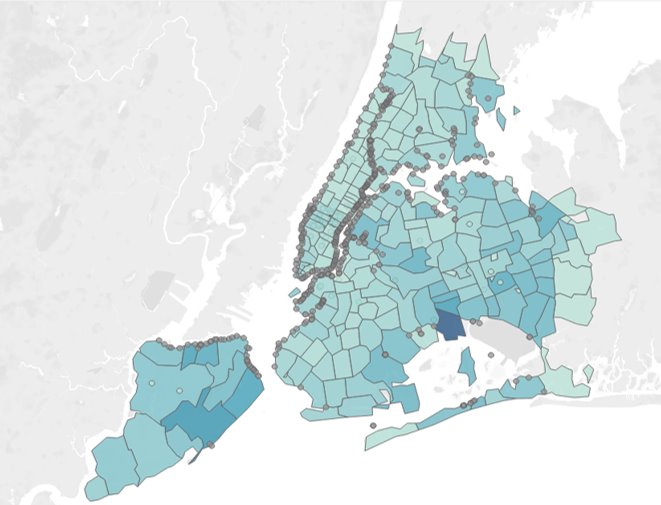
**Impact of rainfall on combined sewer outfall infrastructure**

To analyze the impact of different levels of rainfall on New York City’s sewer system, we used the average zip code wise precipitation data to understand how the system is vulnerable to pressure during different rainfall levels. This pressure level was understood by comparing the capacity of a treatment plant against the flow into the plant from all its drainage zips. We hypothesized that anytime the flow crosses the capacity, the overflow is directly released into the drainage waterway of each combined sewer outflow.

Four different rainfall scenarios were simulated in Tableau using the data collected for the end-to-end sewage infrastructure and precipitation to observe which wastewater treatment plants and corresponding drainage areas are most vulnerable in the city.

**Correlation between distance from discharge point and sewer complaints**

From *Figure 4,* the hypothesis that distance from CSO outfall points will affect the number of sewer complaints reported during hurricane Ida was made. In order to test this, distance from nearest CSO points and center of each zip codes were calculated in Euclidean distance.



*Figure 6. Distribution of sewer complaints during Ida*

Figure 6 is a map of NYC’s sewer complaints during Ida. Darker shade represents a higher number of sewer complaints reported. Generally, zip codes that have CSO points recorded a relatively higher number of sewer complaints, especially in Brooklyn and Staten Island.

**Analyzing demographic factors of zip codes adversely impacted by sewage overflow**

To test if the correlation between distance from discharge point and number of complaints found on decision tree model, spatial analysis method was implemented. After identifying zip codes with high and low levels of complaints, further analysis on the demographics of the impacted area was made in order to observe which demographics are mostly affected. Below plot is demographic information of zip codes with high level of complaints.



*Figure 7: Demographic of zip codes with high level of complaints*

Poverty rate was mostly concentrated at 12% and families' median income was concentrated at 90k per year. Nearly 40 % of the zip codes with high level of complaints had high percentage of black and Hispanic population. On the other hand, zip codes with low level of complaints did not have specific patterns.



*Figure 8: Demographic of zip codes with low level of complaints*

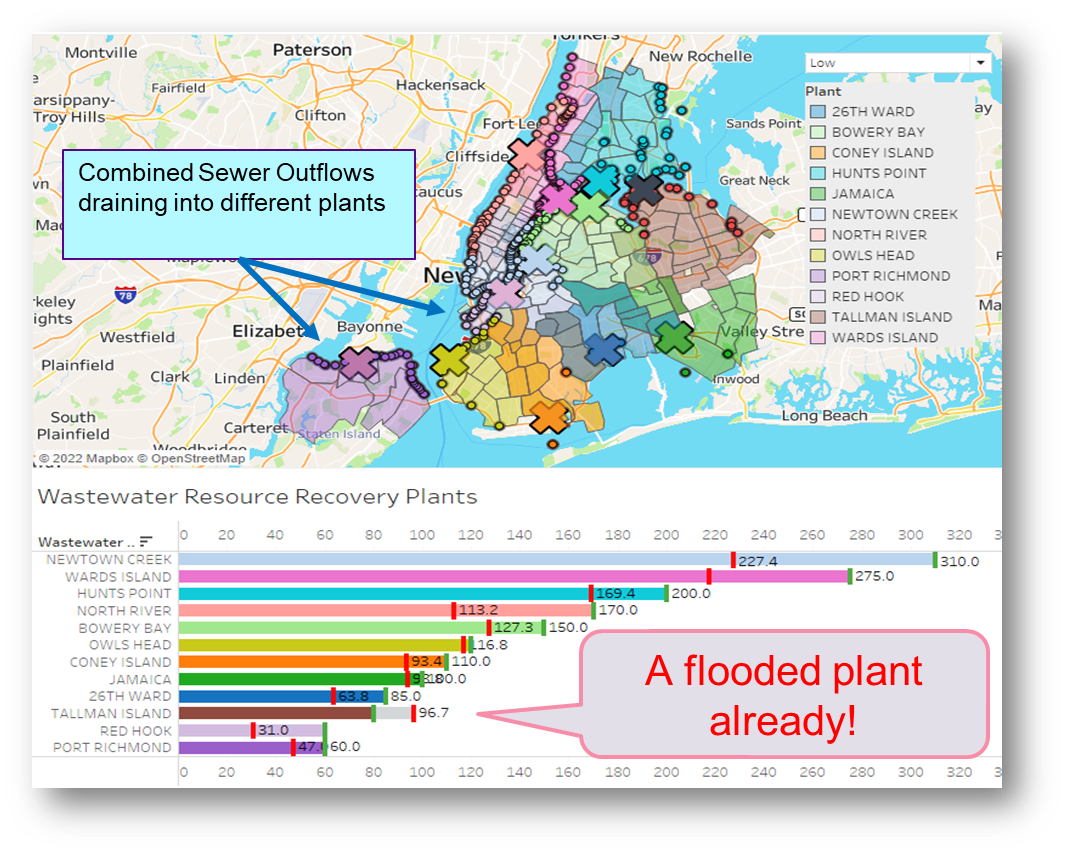
## Limitations

## Findings and Interpretations

**Impact of rainfall on combined sewer outfall infrastructure**

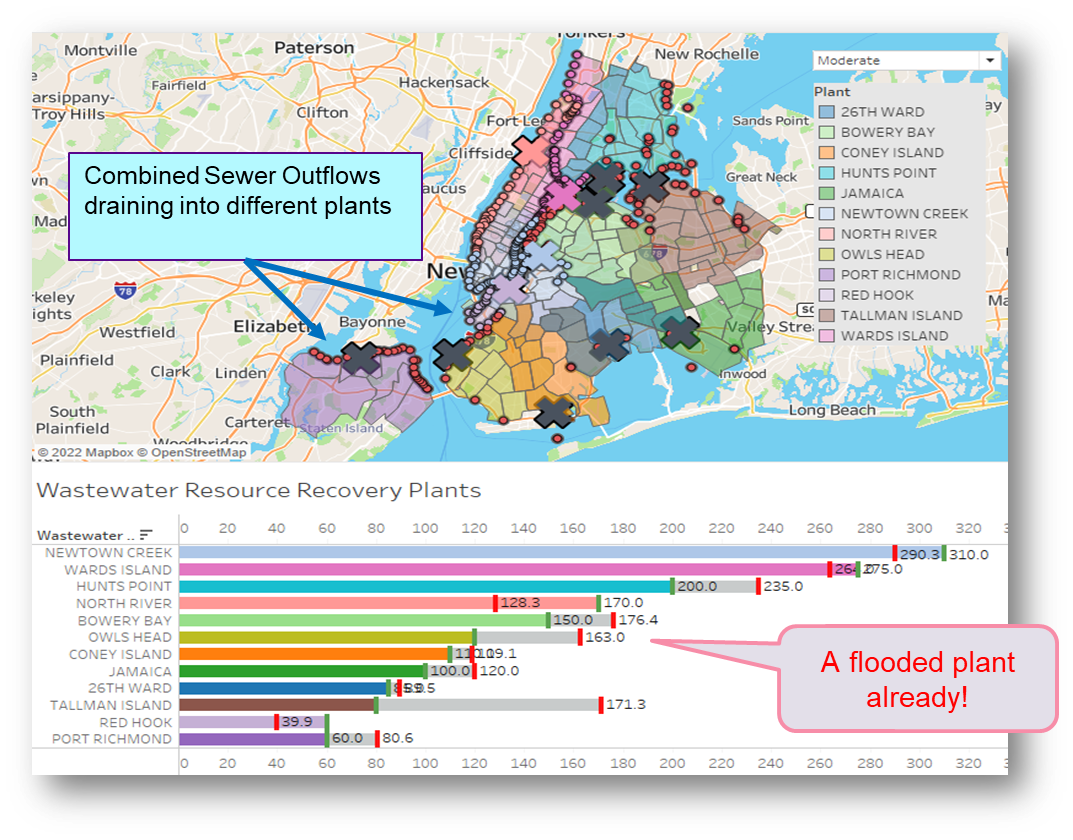
**Low Rainfall Scenario**

According to our simulation, in a low rainfall scenario i.e., 0.1 inches/hour rainfall, one of the treatment plants starts receiving more sewer flow than it can handle. Tallman Island plant’s capacity is only 80 MGD, while the overflow from its drainage zip codes in Queens is 96.7 MGD, so 16 MGD of water is released into the waterways near Queens.

****

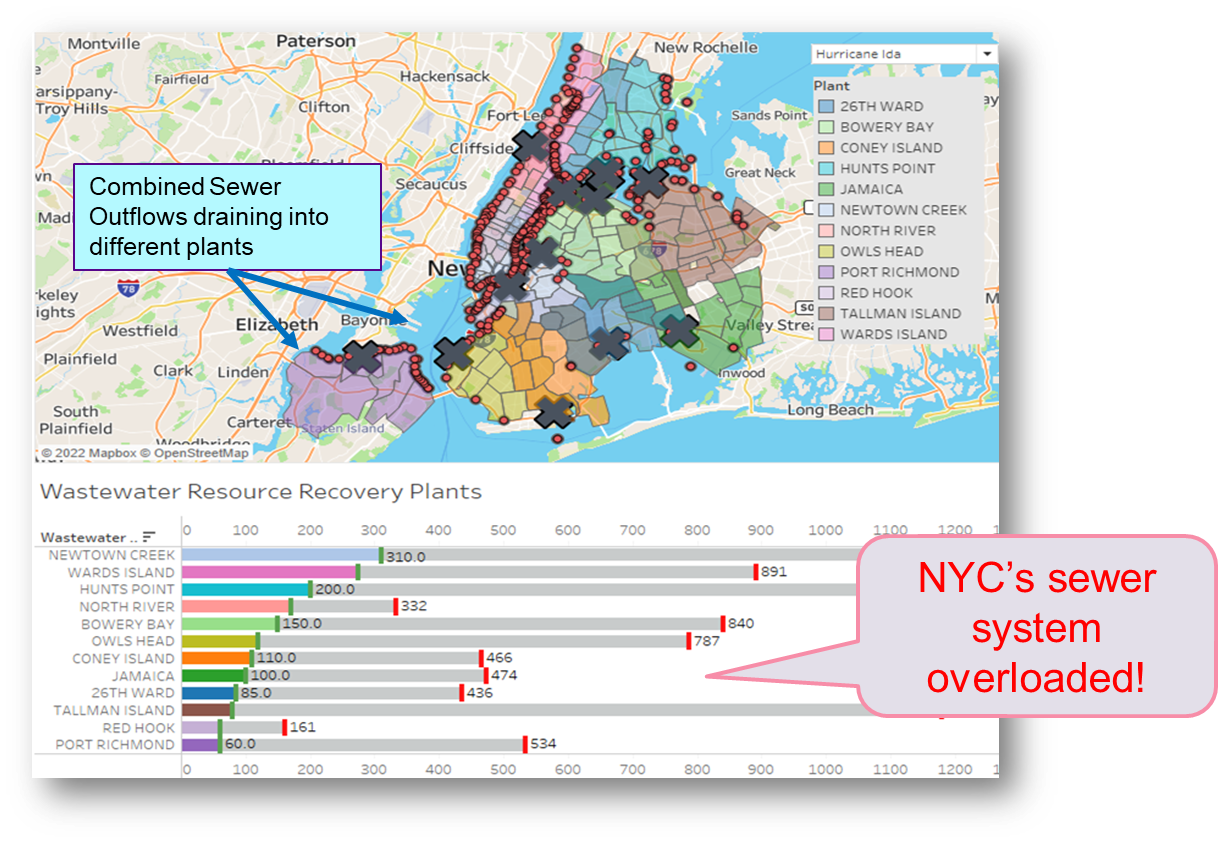
**Moderate Rainfall Scenario**

According to our simulation, for moderate rainfall (of 0.3 inches of rain/ hour), 50% of the treatment plants across Queens, Brooklyn and Staten Island are overwhelmed while the overflow from its drainage zip codes is released into the waterways.



**Ida level Rainfall Scenario**

During a heavy rainfall (1 inches/hour) or an Ida level rainfall i.e., 3 inches/hour, all treatment plants are shut down both due to flooding and loss of electricity, and all of the precipitation and consumption sewage is dumped into the open waterway, causing high levels of toxicity to be mixed into water around the city.



## Policy Implications

# Scenario #2: Water Distribution Network Analysis

## Description

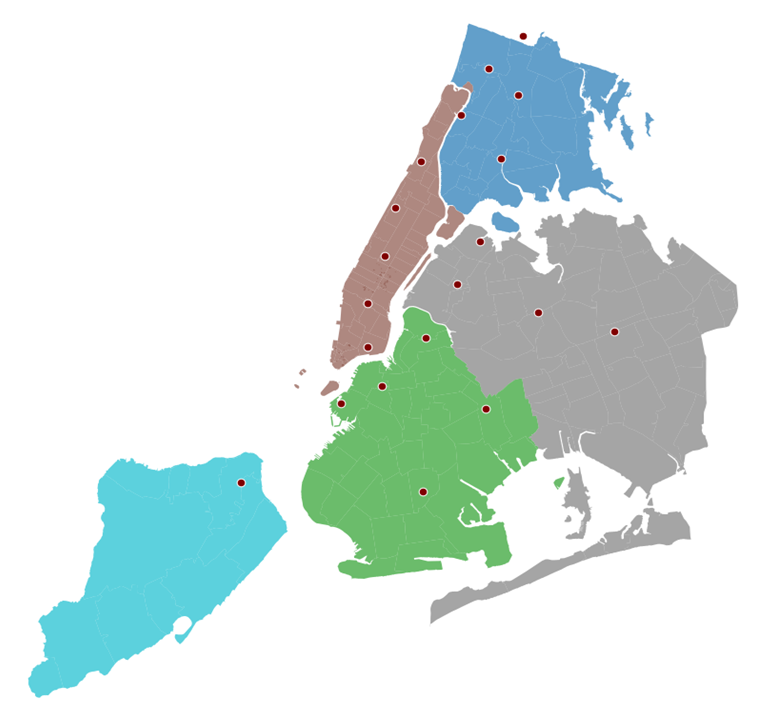
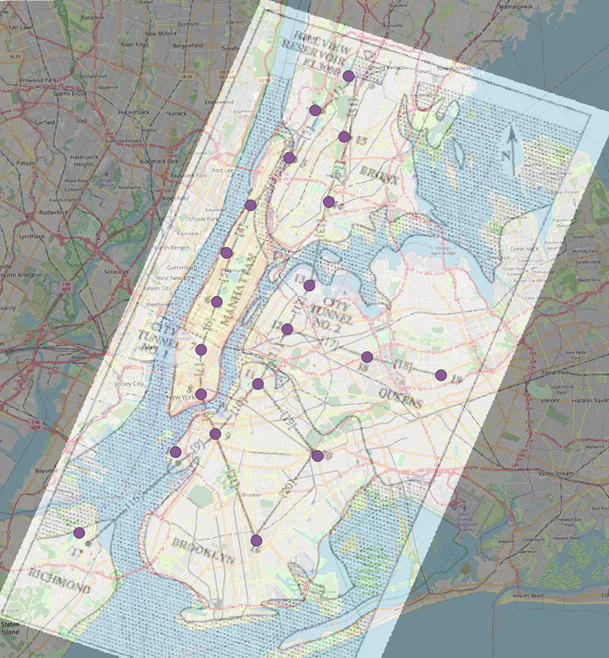
Water and Wastewater systems are considered critical infrastructures because of their essential role in modern urban livelihood [25]. High resilience of these infrastructures is of great importance and has brought these infrastructures into the focus of science and politics. New York City has one of the most complex water systems in the world with a network of 19 reservoirs and three controlled lakes that cover approximately 2,000 square miles of watershed land as far as 125 miles upstate. The City’s drinking water system is the largest unfiltered water supply in the world, delivering approximately one billion gallons of high-quality drinking water each day to nine million New Yorkers [26].

There is a need to gather end to end data about this vast expansive distribution system to study its vulnerabilities to both natural and manmade disruptions. However, water distribution data is highly sensitive and openly unavailable. So, this scenario focuses on performing network analysis with existing sparse data on the water distribution tunnels in New York City to extract information about their location and connected zip codes. The analysis will utilize the physical image of the city’s tunnels and perform an overlay analysis using spatial tools to get digital data about the location of tunnel nodes. Then gravity-based network analysis will be performed to get zip codes impacted by each tunnel and node.

## Discussion of Models

**Spatial Overlay Analysis**

Because the image was found to be scale accurate based on information provided at source, the image was geo-located *i.e.*, located over accurate geography based on common identifiable location markers on both the image and the geographic map using the “geo-location” tool on ArcGIS. The geo-located image was then used to identify point spatial geometries *i.e.*, latitude and longitude for each node on the tunnel network. This process was done using QGIS software’s “New Virtual Layer” tool



**Gravity based Network Analysis**

In transportation engineering, a gravity model is used to understand the flow of traffic. The gravity model assumes that the trips produced at an origin and attracted to a destination are directly proportional to the total trip productions at the origin and the total attractions at the destination [27]. The Flow Distribution method is the application of the gravity model to a distribution system. Flow distribution predicts the amount of demand and population that would be affected if any node in the system were disabled by solving for the distribution of each node’s outflow. To understand and build a water distribution network with zip codes as end points, we used the gravity network analysis method in the project.



## Limitations

Main limitation of the scenario revolves around scarcity of data for the analysis and dependence on data and images that might be outdated. Additionally, the network is currently built at a zip code level and further granularity of geography might benefit the purposes of the project.

## Findings and Interpretations

The main contribution of this scenario is the digitized water distribution network. This network can be utilized to further do analysis by simulating scenarios on disruption to water supply to quantify local level impact of natural and man-made disruptions to the water and electricity infrastructure. The zip code level data can be used for segmentation study of population demand and demographics impacted in different areas.

## Policy Implementations

This study can enable using detailed impact analysis in driving granular hardening policies against climate change and cyber-attacks, focusing on areas of highest impact and vulnerability. It can also be utilized to perform impact analysis of any maintenance and repair work being planned in a particular reservoir, water treatment plant or tunnel to quantify the impact to different regions in the city.

## Conclusions

This project has been an attempt to analyze the interdependent water and wastewater infrastructures in New York City and quantify their operations, measure their vulnerabilities and through them, build a wholistic vulnerability mapping of these critical infrastructures.

# References

[1] Young, R. (2015, June). *A survey of energy use in Water Companies - ACEEE*. Retrieved from https://www.aceee.org/sites/default/files/water-company-energy-use.pdf

[2] Acharya, S., Dvorkin, Y., & Karri, R. (2020). *Public plug-in electric vehicles + grid data: Is a new cyberattack vector viable?* IEEE Xplore. Retrieved from https://ieeexplore.ieee.org/document/9091609

[3] Jiménez, J., & Levenson, M. (2021, September 20). *Ida Drenches New York and New Jersey*. Retrieved from https://www.nytimes.com/live/2021/09/01/us/hurricane-ida-updates

[4] Puleo, M. (2021). *Death toll rises after 'most significant flash-flooding' in NYC history*. Yahoo! Retrieved from https://news.yahoo.com/deadly-disaster-nyc-idas-heavy-124648032.html

[5] Medina, E., Taylor, D. B., & Porter, C. (2022, January 17). *Winter Storm drops heavy snow in the northeast as the South digs out*. Retrieved from https://www.nytimes.com/2022/01/17/us/snowstorm-updates-northeast.html

[6] *How climate change makes hurricanes more destructive*. Environmental Defense Fund. (n.d.). Retrieved from https://www.edf.org/climate/how-climate-change-makes-hurricanes-more-destructive

[7] CISA. (n.d.). *Ongoing Cyber Threats to U.S. Water and Wastewater Systems*. Retrieved from https://www.cisa.gov/uscert/ncas/alerts/aa21-287a

[8] Bergal, J. (2021, March 10). *Florida hack exposes danger to water systems*. The Pew Charitable Trusts. Retrieved from https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2021/03/10/florida-hack-exposes-danger-to-water-systems

[9] *GridEx*. NERC. (n.d.). Retrieved from https://www.nerc.com/pa/CI/ESISAC/Pages/GridEx.aspx

[10] Hopkins, M. (2012, April). *Critical node analysis for water distribution system using Flow Distribution*. DigitalCommons@CalPoly. Retrieved from https://digitalcommons.calpoly.edu/theses/722/

[11] Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). *Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory*. CUNY Academic Works. Retrieved from https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1266&context=cc\_conf\_hic

[12] Rehm, I.-S., Friesen, J., Pouls, K., Busch, C., Taubenböck, H., & Pelz, P. F. (2021). A method for modeling urban water infrastructures combining geo-referenced data. *Water*, *13*(16), 2299. https://doi.org/10.3390/w13162299

[13] Dandy, G. C., Simpson, A. R., & Murphy, L. J. (1996). An improved genetic algorithm for pipe network optimization. *Water Resources Research*, *32*(2), 449–458. https://doi.org/10.1029/95wr02917

[14] (DEP), D. of E. P. (2022, August 8). *Current Reservoir Levels: NYC open data*. Current Reservoir Levels | NYC Open Data. Retrieved from https://data.cityofnewyork.us/Environment/Current-Reservoir-Levels/zkky-n5j3

[15] *Reservoir Levels*. Reservoir Levels - DEP. (n.d.). Retrieved from https://www1.nyc.gov/site/dep/water/reservoir-levels.page

[16] *New Croton Reservoir*. New Croton Reservoir - DEP. (n.d.). Retrieved from https://www1.nyc.gov/site/dep/water/new-croton-reservoir.page

[17] *USGS 01375000 CROTON R AT NEW CROTON DAM NR CROTON-ON-HUDSON NY*. USGS. (2022). Retrieved from https://waterdata.usgs.gov/nwis/uv?site\_no=01375000

[18] *$15.7 million upgrade completed at Jerome Park Reservoir in the Bronx*. NYC Environmental Protection. (2022, February 7). Retrieved from https://www1.nyc.gov/site/dep/news/22-004/-15-7-million-upgrade-completed-jerome-park-reservoir-the-bronx#/0

[19] *Catskill-Delaware Ultraviolet Water Treatment Facility, New York*. Water Technology. (n.d.). Retrieved from https://www.water-technology.net/projects/catskill-delaware-ultraviolet-water-treatment-facility/#:~:text=It%20is%20located%20downstream%20of,Environmental%20Protection%20owns%20the%20plant.

[20] *Croton Water Filtration Plant Activated*. NYC Environmental Protection. (2015, May 8). Retrieved from https://www1.nyc.gov/html/dep/html/press\_releases/15-034pr.shtml#.YwLoi3ZKjb0

[21] *Rainfall calculator, Metric Unitshow much water falls during a storm?* USGS. (n.d.). Retrieved from https://water.usgs.gov/edu/activity-howmuchrain-metric.html

[22] *Open Sewer Atlas NYC*. Tumblr. (n.d.). Retrieved from https://openseweratlas.tumblr.com/data

[23] Levine, L. (2020, February 24). *NYC's new plan would let massive sewage overflows continue*. NRDC. Retrieved from https://www.nrdc.org/experts/larry-levine/nycs-new-plan-would-let-massive-sewage-overflows-continue

[24] Malmassari, J. (2019, March 25). *The dangers of combined sewer overflows*. Municipal Sewer and Water. Retrieved from https://www.mswmag.com/online\_exclusives/2019/04/the-dangers-of-combined-sewer-overflows\_sc\_003d9

[25] *Water and wastewater systems sector*. Cybersecurity and Infrastructure Security Agency CISA. (n.d.). Retrieved from https://www.cisa.gov/water-and-wastewater-systems-sector

[26] *Water management*. Water Management - Mayor's Office of Sustainability. (n.d.). Retrieved from https://www1.nyc.gov/site/sustainability/initiatives/water-management.page

[27] University of Idaho. (2003). *The Gravity Model*. Travel Demand Forescasting. Retrieved from https://www.webpages.uidaho.edu/niatt\_labmanual/chapters/traveldemandforecasting/theoryandconcepts/GravityModel.htm

# 

# Appendix

Code available for all data processing and analysis performed in the project on [github](https://github.com/vaishu1396/Hardening_NYC_Infrastructure.git)