NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

MEASURING AND MODELING POTABLE WATER DEMAND IN THE UNITED STATES VIRGIN ISLANDS

by

Andrew J. Borgdorff

June 2020

Thesis Advisor: David L. Alderson Jr.
Second Reader: Daniel Eisenberg

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2020	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE MEASURING AND MODELING POTABLE WATER DEMAND IN THE UNITED STATES VIRGIN ISLANDS	**5. FUNDING NUMBERS** RNP4H
6. AUTHOR(S) Andrew J. Borgdorff	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Department of Homeland Security (DHS) Federal Emergency Management Agency (FEMA)	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

This thesis uses a data-driven and model-based approach to measure and estimate potable water flows in the United States Virgin Islands (USVI). The USVI potable water system is considerably more complicated than typical systems in the mainland United States. For this reason, we conduct a first-principles analysis on water production, storage, delivery, and use patterns to see if standard statistical models of water demand apply in the USVI context. We curate and combine historical weather, demographic, water use, and water delivery data to understand past water use and flows across St. Thomas and St. John. We identify statistically significant differences in water demand patterns on daily and seasonal time horizons. We further quantify how statistical models for normal operations may make poor predictions of water demands after disasters. Based on these analyses, we provide recommendations for USVI stakeholders focused on making the territory resilient to future disasters, such as hurricanes.

14. SUBJECT TERMS disaster response, USVI, mitigation, water demand modeling, time series modeling, ARIMA, SARIMA, United States Virgin Islands, USVI	**15. NUMBER OF PAGES** 113
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

MEASURING AND MODELING POTABLE WATER DEMAND IN THE UNITED STATES VIRGIN ISLANDS

Andrew J. Borgdorff
Ensign, United States Navy
BS, United States Naval Academy, 2019

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN APPLIED SCIENCE
(OPERATIONS RESEARCH)**

from the

**NAVAL POSTGRADUATE SCHOOL
June 2020**

Approved by: David L. Alderson Jr.
Advisor

Daniel Eisenberg
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis uses a data-driven and model-based approach to measure and estimate potable water flows in the United States Virgin Islands (USVI). The USVI potable water system is considerably more complicated than typical systems in the mainland United States. For this reason, we conduct a first-principles analysis on water production, storage, delivery, and use patterns to see if standard statistical models of water demand apply in the USVI context. We curate and combine historical weather, demographic, water use, and water delivery data to understand past water use and flows across St. Thomas and St. John. We identify statistically significant differences in water demand patterns on daily and seasonal time horizons. We further quantify how statistical models for normal operations may make poor predictions of water demands after disasters. Based on these analyses, we provide recommendations for USVI stakeholders focused on making the territory resilient to future disasters, such as hurricanes.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1 Introduction **1**

1.1 Overview of the U.S. Virgin Islands 1

1.2 Critical Infrastructure Services in the USVI 5

1.3 The Potable Water System in the USVI. 7

1.4 Vulnerability of the WAPA Pipeline System to Service Disruptions 15

1.5 Thesis Goals 18

2 Literature Review **19**

2.1 Past Work on Water-Use Patterns 19

2.2 Past Work on Water Demand Projections 23

2.3 Water Resources Availability. 26

2.4 2017 Hurricane Impacts and Their Implications 27

3 Data Preparation and Model **29**

3.1 NOAA Weather Data 29

3.2 Demographic Data. 32

3.3 WAPA Water System Data. 35

3.4 Understanding Water Demand 44

4 Analysis **45**

4.1 Water Usage by Day of Week 45

4.2 Water Usage by Month of Year 51

4.3 Impact of Rainfall on Water Usage 58

4.4 Impact of Temperature on Water Usage 63

4.5 Time Series Characteristics of Water Demand Pre- and Post-Storm. 65

4.6 Implications WAPA and Territorial Water Planning 70

5 Summary and Conclusion **75**

5.1 Summary 75

5.2 Conclusions and Future Work 77

Appendix: Time Series Models **79**
A.1 Linear Regression 79
A.2 Autoregressive Integrated Moving Average (ARIMA) 80
A.3 Seasonal Autoregressive Integrated Moving Average (SARIMA) 83

List of References **87**

Initial Distribution List **91**

List of Figures

Figure 1.1 USVI Map. 2

Figure 1.2 St. Thomas (STT) Land use Map. 3

Figure 1.3 St. John (STJ) Land use Map. 4

Figure 1.4 St. Croix (STX) Land use Map. 5

Figure 1.5 Overview of Basic Water System. 6

Figure 1.6 Private Home Cistern. 7

Figure 1.7 WAPA Water Assets. 11

Figure 1.8 Overview of the interactions in the USVI Potable Water System. . 15

Figure 2.1 Example of Business, Industrial, and Residential Demand Patterns. 20

Figure 2.2 Indoor and Outdoor Water Usage in Fort Collins, CO. 21

Figure 3.1 STT Mean Maximum Temperature March, 2006 to December, 2019. 30

Figure 3.2 STT Mean Maximum Temperature March, 2006 to December, 2019. 31

Figure 3.3 STT Monthly Precipitation March, 2006 to December, 2019. . . . 32

Figure 3.4 USVI Population. 33

Figure 3.5 WAPA Daily Report Example 36

Figure 3.6 Simplified WAPA System for STT & STJ 37

Figure 3.7 STT Reverse Osmosis (RO) Production. 38

Figure 3.8 STT Combined Tank Levels. 39

Figure 3.9 STT Monthly Water Truck Volumes January, 2012 to March, 2020. 40

Figure 3.10 STT Monthly Water Volume Delivered to “Town” January, 2012 to March, 2020. 41

Figure 3.11 STT Monthly Donoe Tank Levels January, 2012 to March, 2020. 42

Figure 3.12 STT Monthly Water Volume Delivered to East End January, 2012 to March, 2020. 43

Figure 3.13 STT Monthly STJ Water Transfer January, 2012 to March, 2020. 44

Figure 4.1 WAPA Total STT Water Volume by Day of Week. 46

Figure 4.2 Total Water Volume Sent to Town by Day of Week. 47

Figure 4.3 Total Water Volume Sent to East End by Day of Week. 48

Figure 4.4 Water Volume Transferred from STT to STJ by Day of Week. . . 49

Figure 4.5 WAPA Water Sales to Trucks on STT by Day of Week. 50

Figure 4.6 WAPA Total STT Water Volume by Month of Year. 52

Figure 4.7 Total Water Volume Sent to Town by Month of Year. 53

Figure 4.8 Total Water Volume Sent to East End by Month of Year. 54

Figure 4.9 Water Volume Transferred from STT to STJ by Month of Year. . 55

Figure 4.10 WAPA Water Sales to Truck on STT by Month of Year. 56

Figure 4.11 STT Accumulated Precipitation. 59

Figure 4.12 Correlation of Rainfall Accumulation with WAPA Water Volumes. 60

Figure 4.13 WAPA Truck Sales Rainfall Partition Tree. 62

Figure 4.14 WAPA Delivery and Truck Sales Correlation with Daily Mean Maximum Temperature. 64

Figure 4.15 Single Differenced Correlation Plot for Autoregressive Integrated Moving Average (ARIMA) Parameter Selection. 66

Figure 4.16 Annually Differenced Correlation Plot for Seasonal Autoregressive Integrated Moving Average (SARIMA) Parameter Selection. . . . 67

Figure 4.17 SARIMA(1,0,0)(0,1,1,12) STJ Pre-Storm. 68

Figure 4.18 SARIMA(1,0,0)(0,1,1,12) STJ Post-Storm. 69

List of Tables

Table 2.1 USVI Cistern Size Requirements. 23

Table 3.1 USVI Water Purchases By Source from 2010 Census Data. 34

Table 3.2 STT Household Water Purchases by Source. 34

Table 3.3 STJ Household Water Purchases by Source. 35

Table 4.1 STT East End Weekend Categorical Regression. 51

Table 4.2 STT East End Month Categorical Regression. 57

Table 4.3 STT Water Sales to Water Trucks Rainfall Regression. 61

Table 4.4 Summary of Partition Tree Results for Truck Volumes. 63

Table 4.5 Mean Absolute Deviation of Pre- and Post-Storm STJ Predictions with SARIMA(1,0,0)(0,1,1,12). 70

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

ADF augmented Dickey-Fuller test

ANN Artificial Neural Network

ARMA Autoregressive Moving Average

ARIMA Autoregressive Integrated Moving Average

CID Center for Infrastructure Defense

DoD Department of Defense

DOE Department of Energy

DPNR Department of Planning and Natural Resources

DPW Department of Public Works

FEMA Federal Emergency Management Agency

GIS Geographic Information System

MG Mega Gallons

MGD Mega Gallons/Day

mm millimeter

MAD Mean Absolute Deviation

MLE maximum likelihood estimation

NASA National Aeronautics and Space Administration

NOAA National Oceanic and Atmospheric Administration

NPS Naval Postgraduate School

PRV pressure regulating valve

RO reverse osmosis

RWC rainwater catchment

RQ1 Research Question 1

RQ2 Research Question 2

RQ3 Research Question 3

SARIMA Seasonal Autoregressive Integrated Moving Average

STJ St. John

STT St. Thomas

STX St. Croix

US EPA United States Environmental Protection Agency

USG United States Government

USGS United States Geological Survey

USN US Navy

USVI U.S. Virgin Islands

UVI University of the Virgin Islands

VITEMA Virgin Islands Territorial Emergency Management Agency

WAPA Virgin Islands Water and Power Authority

WMA Virgin Islands Wastewater Management Authority

Executive Summary

The United States Virgin Islands (USVI) is a collection of Caribbean islands about 1,100 miles southeast of Florida that became a Territory of the United States (US) in 1917. The three main islands—St. Thomas (STT), St. John (STJ), and St. Croix (STX)—account for nearly the entire territorial landmass of 135 square miles and population of approximately 105,000 people. The modern history of the territory dates back to European settlement in the 1600s, and much of the territory's infrastructure is derived from its colonial roots. The distinct history, geography, and demographics of the territory have yielded a patchwork of infrastructure systems that has developed on the islands over time.

Due to the territory's small size, varied geography, and isolated location, the provision of critical infrastructure services like electric power, water, transportation and mobility, telecommunications, and healthcare in the USVI is different than in the rest of the US. In general, there is no single entity with complete situational awareness of this complex set of operational and economic interactions across critical infrastructure systems. Moreover, there is no central planner for managing critical infrastructure protection.

Since 1989, the USVI has been hit by four of the forty costliest hurricanes in US history: Hugo (1989), Marilyn (1995), Irma (2017), and Maria (2017). Hurricanes Irma and Maria were both category-5 hurricanes that struck the territory within a two-week period in September 2017. They caused widespread outages to the water system and other critical infrastructures. Recovery efforts are ongoing today, more than 2.5 years later.

This thesis addresses water security and water provision within the USVI. We focus on potable water provision and use prior to wastewater treatment, discharge, and drainage. In most mainland US systems, water is acquired from a source, treated so that it is usable for drinking and commercial purposes, stored for future use, and then transported to customers for end use. The USVI potable water system is comprised of the same essential elements, but is considerably more complicated due to the additional widespread use of residential rainwater catchment systems and the use of water truck delivery that complements a pipeline delivery system with only partial coverage of the population. The mixture of water production and multiple consumption choices for all users in the system creates a complicated set

of interactions that this thesis seeks to disentangle. Specifically, we use a data-driven and model-based approach to explore the relationships between water flows, weather, and time factors like day of week, month of year, and pre- and post-hurricane.

Our analysis begins with data curation for the combined system serving STT and STJ. We first compile and curate a data set comprised of meter flow data, water storage data, weather data including temperature and rainfall, and demographic data about population and water use. We carry out an exploratory data analysis on the water flow data on STT comprised of use in "Town" (the area around the capital of Charlotte Amalie), the East End, transfer to STJ, storage tank levels on STT, and weather data with rainfall and temperature information.

We identify several factors involving water flows and use in the USVI.

Water usage patterns by day of week suggest that there is higher likelihood of an unusually high demand occurring on Monday, Wednesday, or Friday. A system operator may consider maintenance occurring on Tuesday or Thursday to minimize risk of failing to meet demand or depleting storage reserves more than necessary. The consistent daily patterns for water demand suggests the potential to develop different demand distributions by day of week; these could be used in conjunction with Monte-Carlo simulation techniques to generate synthetic demand scenarios and assess the capability of the system to handle them.

The water usage pattern for month of year is not as clear as for the day of week. Long-term planning can consider August-February to be lower demand months. But planners have to keep in mind occasions where unexpected increases occur, particularly around September. In general, simple statistical models like categorical linear regression can provide estimates of Town, East End, and Total water flows.

In contrast, water truck sales appears highly correlated to weather. Average daily precipitation over the last 45-days provides a good estimate of the demand for delivery by water truck. A small decrease in average daily rainfall is correlated to a large increase in demand for truck water sales. Water system operators can potentially use this to anticipate demand volumes. Water truck sales also appear correlated to other weather features like temperature. However, temperature does not seem to play a currently quantifiable role in influencing water demand at the aggregate level. Correlation between temperature and water use seems to be an artifact of wet and dry seasons.

Time series analysis of water transfer from STT to STJ shows that standard forecasting models may predict STJ water demands. Specifically a standard parameterization of the Seasonal Autoregressive Integrated Moving Average (SARIMA) model provides a good fits and prediction of flow data to STJ.

We quantify how well statistical models predict demand after a disaster using the SARIMA model for STJ. Our analysis shows stark differences in water flow patterns before and after the September 2017 hurricanes. The mean absolute deviation from the predicted water flows nearly doubled to 1.25 million gallons per month. This change is caused by over predicting demands when pipeline systems are non-operational and under predicting demands after water service is restored. The under prediction can be caused by many unknown factors, including increased demand for water due to changes in behavior, leaking pipes due to hurricane damage, or some shifts in the seasons water is most required.

This thesis opens the doors for additional work on understanding and improving the water security for the USVI. Future work could consider additional exogenous variables in forecast models that improve time-series forecasts of water demands. For example, we have not considered the periodic presence of cruise ships in STT, which can deposit tens of thousands of patrons on STT and STJ on a single day. Cruise ships follow regular schedules, so it is reasonable to think that these patterns could be identified and incorporated into our analysis.

Finally, there is an opportunity to combine the work here with models that consider cistern storage and water use from first principles. That is, there is an opportunity to develop demand-driven models that incorporate household details — such as the size of each cistern, water usage by household members, and the way in which rainfall and temperature affect the demand for water. Understanding these interactions will be important for assessing how systematic disruptions (e.g., drought, earthquakes) affect both the demand for water and the ability of the system to provide it. Integrating such analyses into hazard mitigation planning for the territory is a topic of ongoing work.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

To my thesis advisors, Dr. David Alderson and Dr. Daniel Eisenberg, thank you for for the opportunity to work on a larger project doing meaningful work. I appreciate the time and space to learn through my own mistakes, with the appropriate guidance and mentorship. I hope this is not just the end of my time as an Naval Postgraduate School (NPS) student analyst, but a stepping stone to continuing meaningful work.

To Dr. Samuel Buttrey, thank you for your time and willingness to get on Zoom to discuss time series models. Go Tigers!

To our colleagues at the University of the Virgin Islands, thank you for taking me into the project on our visit to the US Virgin Islands (USVI). I appreciated the care taken to explain the way of life in the territory and how it was altered by the hurricanes in September, 2017. Moreover, thank you for sharing a taste of the culture in the USVI.

To our colleagues at the Water and Power Authority, thank you for the data you shared that made this work possible. Furthermore, to Brian Leonard, thank you for the detailed and extensive tour of the water facilities and stories of the US Navy's presence in the USVI.

Finally, to my family, thank you for the support as I worked through the end of my time at NPS. I could not have done this alone, without the support provided to me by all involved.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

The U.S. Virgin Islands (USVI) is a collection of Caribbean islands in northern portion the Lesser Antilles—located between the Caribbean Sea and the North Atlantic Ocean, about 1,100 miles southeast of Florida—that became a US territory in 1917. The modern history of the territory dates back to European settlement in the 1600s, and much of the territory's infrastructure is derived from its colonial roots. The distinct history, geography, demographics of the territory have yielded a patchwork of infrastructure systems that has developed on the islands over time.

1.1 Overview of the U.S. Virgin Islands

The USVI is made up of three main islands and two smaller islands. The three main islands—St. Thomas (STT), St. John (STJ), and St. Croix (STX)—account for nearly the entire Territorial landmass of 135 square miles and population of approximately 105,000 people. The USVI also include two smaller islands—Water Island and Hassel Island. These two islands are comprised mainly of beaches and national parks and do not have significant human activity.

The territory is located 40 miles East of Puerto Rico and to the immediate west of the British Virgin Islands. STT and STJ are three miles apart and STX sits 40 miles to the south of STT. The geography and activity is distinct on each island (US Virgin Islands Bureau of Economic Research 2014). A map of the three main islands is shown in Figure 1.1.

Figure 1.1. USVI Map. Map showing the geographic relationship between STX, STT, and STJ. STT and STJ are approximately 40 miles north of STX. Source: Parsons Brinckerhoff and Jaredian Design Group (2014).

St. Thomas (STT)

St. Thomas (STT) is a hilly island with 31 square miles of land that has a population of approximately 51,000. The capital of the USVI, Charlotte Amalie, is on STT. The island is the tourist hub, and it serves as the center for commerce, trade, finance, and government for

the USVI (US Virgin Islands Bureau of Economic Research 2014). The land use map for the island is shown in Figure 1.2.

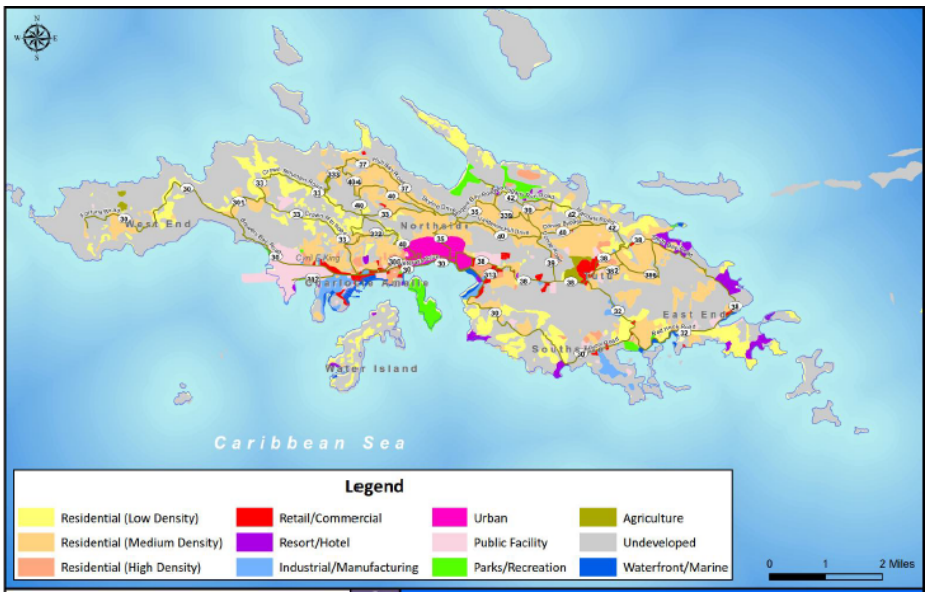

Figure 1.2. STT Land use Map. The island has a number of resort areas with several high density residential areas among undeveloped forested and hilly landscape. There is limited agricultural, industrial, and manufacturing activities. Source: Parsons Brinckerhoff and Jaredian Design Group (2014).

St. John (STJ)

Dominated by national parkland, St. John (STJ) is three miles east of STT. Two thirds of the island's 20 square miles belong to the national park as shown in the land use map in Figure 1.3. The small island's population is approximately 4,000. The pristine beauty and protected coastline make the island a popular destination for tourists and divers. The private land on the island has a number of luxurious homes and hotels making it a destination for rich and famous vacationers (US Virgin Islands Bureau of Economic Research 2014).

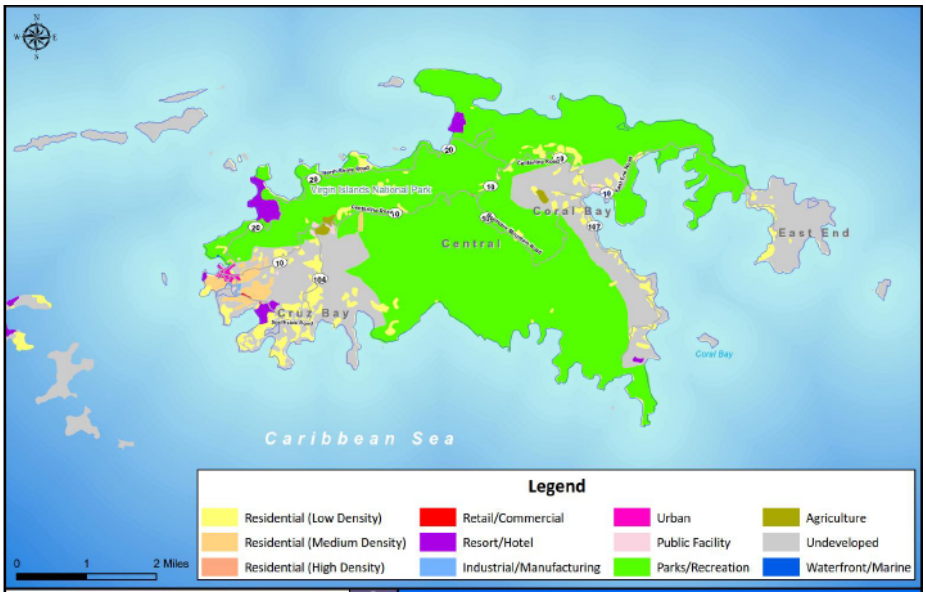

Figure 1.3. STJ Land use Map. STJ is dominated by national park land with two main population centers on either end of the island. There are also resorts scattered along the coastline. Source: Parsons Brinckerhoff and Jaredian Design Group (2014).

St. Croix (STX)

St. Croix (STX) is the largest island by land area at 84 square miles and has a population comparable to STT at approximately 50,000. The geography of STX has more variety than that of STT and STJ. STX has mountains in the northwest and cliffs on the east end of the island. The remainder of the island is rural terrain with little variation in elevation. Figure 1.4 shows land use on STX. Economically, the island has historically been a producer of oil and rum with some traditional agriculture and tourism. However, the island's oil refinery closed in 2012 causing economic damage to local residents and businesses (US Virgin Islands Bureau of Economic Research 2014).

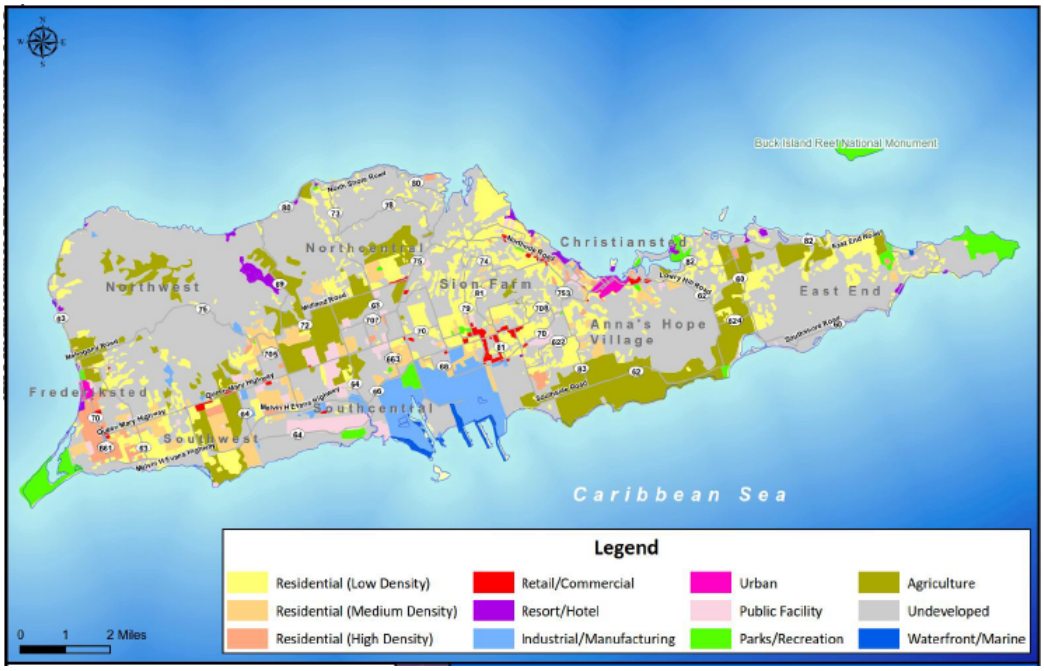

Figure 1.4. STX Land use Map. STX lacks the hilly geography of the other two islands. As a result, there is greater agricultural development and less dense residential areas. There is also less tourist traffic than there is to STT and STJ. Source: Parsons Brinckerhoff and Jaredian Design Group (2014).

1.2 Critical Infrastructure Services in the USVI

Due to the islands' small size, varied geography, and isolated location, the provision of critical infrastructure services like electric power, water, transportation and mobility, telecommunications, and healthcare in the USVI is different than in the rest of the United States. In general, there is no single entity with complete situational awareness of this complex set of operational and economic interactions across critical infrastructure systems. Moreover, there is no central planner for managing critical infrastructure protection in the territory.

This thesis focuses on water security and water provision within the USVI. For more information regarding other critical infrastructure systems and services in the USVI, we refer the reader to recent reports on electric power (Wille 2019), water delivery (Bunn 2018; Wille 2019), transportation and supply chain management (Good 2019), and telecommunications (Moeller 2020; Wine 2020).

The Virgin Islands Water and Power Authority (WAPA) is the drinking water utility for the territory. The potable water systems in the USVI were originally designed by the USVI Department of Public Works (DPW), but were taken over by WAPA who now operates and maintains drinking water pipelines for all islands. WAPA collaborates with the DPW to determine long-term water planning and provision. The USVI Department of Planning and Natural Resources (DPNR) manages natural resources in the territory and maintains list of approved producers and deliverers of water. Wastewater and wastewater infrastructure is managed by Virgin Islands Wastewater Management Authority (WMA) for treatment and discharge and surface drainage is maintained by the DPW. Together, these public agencies interact with a complex system of public and private water production, storage, delivery, treatment, and discharge.

We focus on potable water provision and use prior to wastewater treatment, discharge, and drainage. Figure 1.5 shows a general overview of a potable water system provided by the United States Environmental Protection Agency (US EPA) (2020). Water is acquired from a source. Next, the water is treated so that it is usable for drinking and commercial purposes. After treatment, water is stored while it waits to be used. From storage, water is transported to customers for end use. As we detail next, the key elements of the water system in the USVI are more complicated.

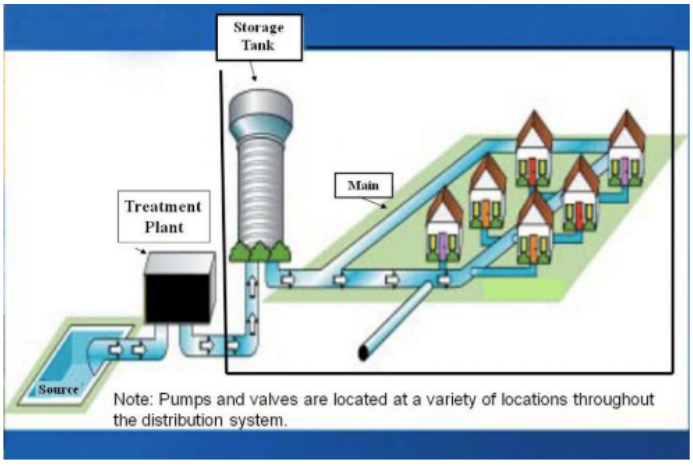

Figure 1.5. Overview of Basic Water System. Example of a general basic water system. Water is treated, then stored, and then delivered to the end user. Source: US EPA (2020).

1.3 The Potable Water System in the USVI

Figure 1.5 is often sufficient to describe many real water systems in the US. The USVI potable water system is comprised of the same essential elements. However, the number and interactions of key elements in the USVI potable water system is much larger and more complex than many water systems in the US. In fact, the USVI potable water system may be one of the most complicated water systems in the world.

1.3.1 Water Sources and Production

While the majority of communities in the US receive water from a single source location or type (e.g., a reservoir), the USVI has several distinct sources of potable water.

Rainwater

The islands rely heavily on rainwater catchment cisterns that date back to colonial times and still serve as a primary source of water for many homes and businesses. Historically, all buildings depended on cisterns for their water supply. Since 1964, all buildings are required to have a cistern by law, unless they have a water pipe connection (CDR Maguire 2019). Figure 1.6 presents an example of a private home with a cistern on STT.

Figure 1.6. Private Home Cistern. Private home cistern on STT in the USVI. Source: Solomon (2009).

Cisterns are not required to comply with Clean Water Act standards (CDR Maguire 2019) and are often used for non-drinking purposes (e.g., showers, toilets). Still, cistern water has to

be treated before use and many residents drink cistern water. Chemical treatment is the most common method for purification. DPNR provides recommendations for treating cisterns with a prescribed amount of bleach based upon amount of water stored (United States Virgin Islands Department of Planning and Natural Resources Division or Environmental Protection 2020).

Seawater

Most of the territory's drinking water is produced by Seven Seas Water Company at facilities located on STX and STT. These facilities use reverse osmosis (RO) to pressurize seawater over membranes, which separates the seawater into freshwater and brine. The brine is pumped back into the ocean, and the freshwater is kept for use. Minerals are then added to the freshwater for drinking. WAPA purchases all of the water it delivers from Seven Seas.

Seven Seas Water Company has two primary facilities that produce RO water for WAPA and other customers. Seven Seas Water Company started producing water at its facility at Sub Base on STT in 2013. The facility produces 1.9 Mega Gallons/Day (MGD) with the capability to produce up to 2.3 MGD. Also in 2013, Seven Seas started producing water for WAPA on STX at two plants at the Richmond Power Plant Facility. These facilities produce 2.9 MGD on average and have the capability to produce up to 3.6 MGD (CDR Maguire 2019).

In addition to Seven Seas Water Company, several large hotels and resorts in the territory have their own desalination facilities. Over time, hotels and resorts have sought to produce their own water due to unreliable service from WAPA. Hotels, as well as grocery stores, compete with WAPA for water sales.

Groundwater

The USVI has many individual wells for groundwater access, but relatively few groundwater resources. Groundwater sources in the USVI have been contaminated over time by salt-water intrusion and pollution from development on the islands. Some companies and individual households on the islands have private wells. When private companies sell from their well, they are limited by water safety regulations enforced by DPNR including the Clean Water Act and Territorial permitting regulations. Individuals with private well access will have a

reduced need for water from other sources.

1.3.2 Water Storage

Once water is produced it is stored for future use. Water storage can be conceptualized as a "water battery" that can be "charged" from a water source and "drained" when water is consumed. Many communities in the mainland US tend to have few, centralized storage locations for potable water such as tanks or water towers. The USVI has a more complicated system involving cisterns at homes and businesses, large storage tanks managed by WAPA, and bottled water and self-serve water vending machines at grocery stores.

Cisterns

Cisterns located at USVI homes and businesses are sized for storage for many days. Figure 1.6 shows an example of a home with a cistern that can support a family with most water needs for weeks to months depending on the season. Thus, the frequency of rainwater catchment and the volume of a cistern can offset a customer's need to purchase water from other sources for long periods of time. Many customers never connect to other USVI water delivery systems (e.g., via pipes) for this reason.

Large Storage Tanks

In addition to cisterns sized for individual buildings, there are several large storage tanks that store water for entire communities. All large storage tanks are owned and operated by WAPA. Across STT and STJ, WAPA has 7 storage tanks. Tank #2, Tank #3, and Tank #4 serve as the primary storage. These three tanks have a capacity of 10.5 Mega Gallons (MG) and are located adjacent to the WAPA offices and the Seven Seas RO facility. On STX, WAPA has 8 storage tanks. The primary tank is the Richmond tank with a capacity of 10 MG. On STX, the Kingshill and Mountain tanks each store up to 5 MG. WAPA does not currently have a way to keep the water fresh while it is stored long term, so the tanks must be emptied and filled periodically.

Grocery Outlets and Bottled Water

Another form of storage is water sold at grocery stores. This water is primarily stored in bottles and sold to customers. Some grocery stores in the territory that produce drinking

water for residents also have self service vending machines.

1.3.3 Water Delivery

When water is needed for consumption, water is moved from storage to water customers. Water delivery in the mainland US is almost entirely achieved by underground water pipelines that deliver pressurized water to taps in buildings. In the USVI, there are multiple, competing ways water can be delivered to customers.

Water Delivery by Pipeline

The potable water system—originally developed by the USVI DPW in the 1950s— is a water pipeline system similar to the pipelines found in the mainland US. WAPA maintains and operates the all drinking water pipeline systems in the USVI. WAPA first took responsibility of USVI water pipelines in 1988.

The current configuration of each pipeline system on STT, STJ, and STX is presented in Figure 1.7 (CDR Maguire 2019). WAPA water pipes do not reach all of the residents on the island. WAPA water mainly supplies low-lying population centers because there is limited pumping capability on the islands and there is limited incentive to extend systems to communities that historically relied on rainwater catchment. The largest customers for WAPA are schools, government buildings, and hospitals that require large quantities of potable water but lack (Alderson et al. 2018).

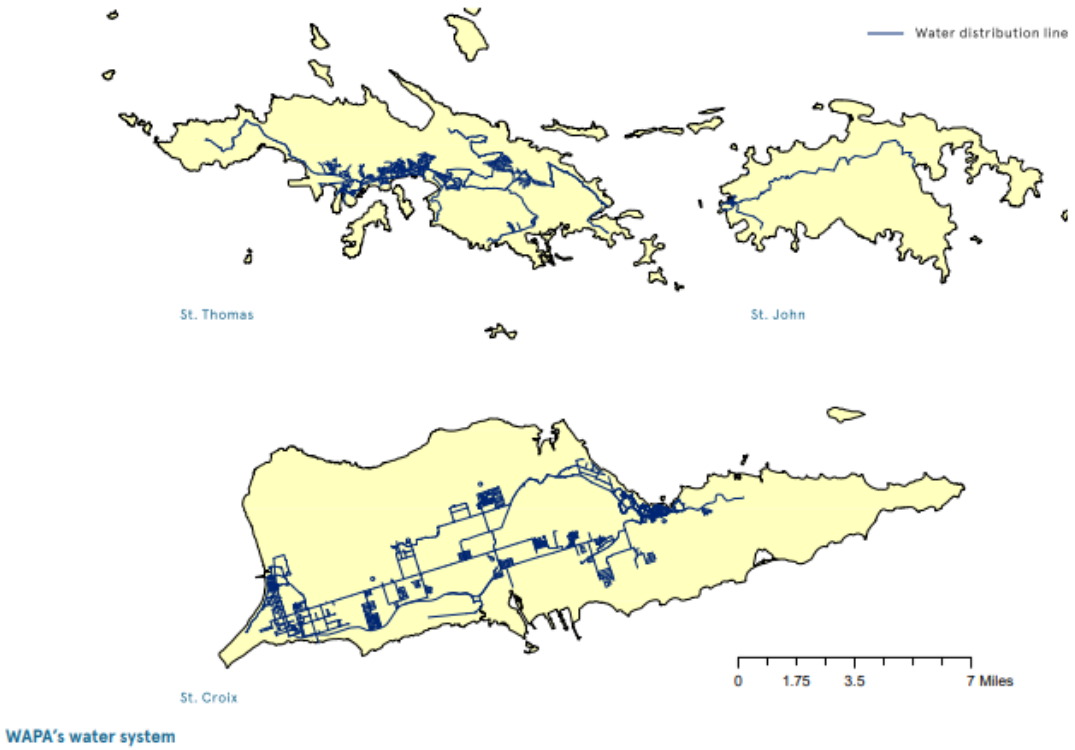

Figure 1.7. WAPA Water Assets. Map of WAPA water pipes on the three islands. It is apparent that the WAPA pipe network does not cover large regions of each island. Source: USVI Hurricane Recovery and Resilience Task Force (2018)

The WAPA water pipeline for STT and STJ is a single system serving both islands. Water produced on STT is pumped via underwater pipes to STJ. Collectively, the combined system on STT and STJ has approximately 82 miles of water mains, three pumping stations (two on STT and one on STJ), 10 booster stations, and six pressure regulating valve (PRV) stations. These two islands are divided into 30 metering districts. Under normal operations a 20 MG supply is sufficient for about ten days of operations (CDR Maguire 2019).

The WAPA water pipeline for STX is separate from STT and STJ. It has 110 miles of system water mains, three main pump stations, and one booster pump station. The STX system is at risk of losing pressure to Frederiksted, one of the largest population centers on the island (CDR Maguire 2019).

Water Delivery by Roadway

In the mainland US, water delivery by roadway is almost exclusively accomplished by individual residents purchasing water from grocery outlets. While USVI residents acquire water in the same way, a much larger volume of water is sold by water trucks, referred to as *water haulers*. A water hauler is a company that has one or more trucks that can hold roughly enough water to fill a cistern tank. Residential and business customers purchase from water haulers during seasons with low rainfall and/or when they do not want to use water delivered by WAPA pipeline.

DPNR maintains a list of licensed water haulers in the territory. There is a mix of large and small water trucks. The small water trucks are able to serve hard to reach homes accessible by roads that large water trucks cannot navigate due to their size.

Water Delivery by Standpipe

A unique feature of the USVI potable water system is the delivery of water by standpipe. A standpipe is a large pipe on the of a WAPA water main or private well that delivers large volumes of water. Standpipes are not located at final point of use, so customers who purchase by standpipe generally finish water delivery by road. Thus, standpipes are a combined delivery method, that combines pipeline and roadway.

Standpipes in the USVI are operated by WAPA, Seven Seas, and other commercial and private water producers. The primary customers for USVI standpipes are water truck haulers who want to purchase large volumes of water and have an efficient means of transport. For this reason, WAPA generally refers to standpipe sales as water truck sales. Still, some individual customers also purchase water from standpipes to save money by purchasing water in bulk. For example, WAPA has a water sales card that individual customers can use to access standpipe water and fill private tanks (similar to purchasing water in bulk at a grocery store).

1.3.4 Water Customers and Use

Potable water serves multiple health and commercial purposes, including drinking, sanitary use, and commercial processing. These different purposes relate to different demand patterns. For example, water use at a household tends to follow a diurnal pattern related

to increasing and decreasing water needs throughout a day. For this reason, potable water systems tend to separate customers into few characteristic types based on use.

In the USVI, customers are separated into three categories often used in the mainland US — residential, commercial, and industrial customers. However, each customer has unique features such that their associate demand patterns differ from those normally identified in US water systems.

Residential Users

Residential customers in the USVI typically follow a diurnal use pattern similar to most households in the US that involves high use in the morning and at night. However, USVI households have multiple ways to fill their water needs throughout the day, and often receive water from a combination of delivery and storage systems involving pipelines and cisterns. Residential customers can have any combination of a cistern, connection to the WAPA pipeline system, or have both. Moreover, residential cisterns can be filled by a combination of rainwater catchment or by water haulers. Many residential customers also purchase drinking water from grocery stores. Some residences may also use private wells for their water needs.

Commercial Users

Commercial businesses in the USVI also have similar use patterns to the mainland US, characterized by steady water use during the work day and low use during non-business hours. Again, businesses can access water in multiple ways, and are connected to water pipes, have a cistern, or make their own water. In addition, some businesses produce water in excess of their needs and sell to water haulers via a private well or standpipe.

Industrial Users

Industrial customers in the USVI are also similar to mainland industrial customers and are characterized by constant demand all day and night. Industrial customers are the largest customers for the WAPA pipeline system and put large demand on the water system. They also have been the most reliable for paying bills to WAPA, making them key source of revenue. One of the most important customers is the Diageo Rum distillery on STX. A

proposed distillery expansion would increase water demand by two to three times (CDR Maguire 2019).

1.3.5 Interactions Across the USVI Potable Water System

In total, the USVI potable water system is far more complicated than the one depicted in Figure 1.5, and possibly more complicated than other water system in the US. Figure 1.8 summarizes the main routes that water takes from production to consumption in the USVI.

USVI potable water has numerous sources, storage options, means of delivery, and potential use. Water is sourced by rainwater at residences, businesses, and used for all potable water needs and agriculture. Seawater is treated by RO systems, sold to WAPA to supply their pipe delivery system, bottled by grocery stores for residential and business customers, and sold at the WAPA standpipe. Water production via RO is also done by hotels and resorts, as well as industrial users, to maintain a consistent supply of water independent from WAPA. Finally, groundwater is trucked to residential and government users. Finally, customers receive water directly from their cistern, transported by water haulers, purchased and moved by the consumer from grocery stores, delivered by WAPA pipes, and produced by private businesses.

All consumers are supplied by at least two sources of water leading to complicated interactions affecting water demand. Therefore, consumers have choice, and their selection is influenced by factors normally not considered by normal water users like price, taste, and availability. For example, a high price for water from trucks could push residential users to draw more water from the WAPA pipe system during dry seasons. The taste of rainwater catchment (RWC) water may lead many households to use the cistern for all household uses other than drinking.

Overall, the USVI potable water system has a unique set of interactions affecting final water use and demand not seen elsewhere in the US. One of the main goals of this thesis is to quantify some of these interactions to guide operations and planning. In particular, we focus on understanding how external factors such as time and weather affect flows of water across different parts of the WAPA pipeline system to improve operations.

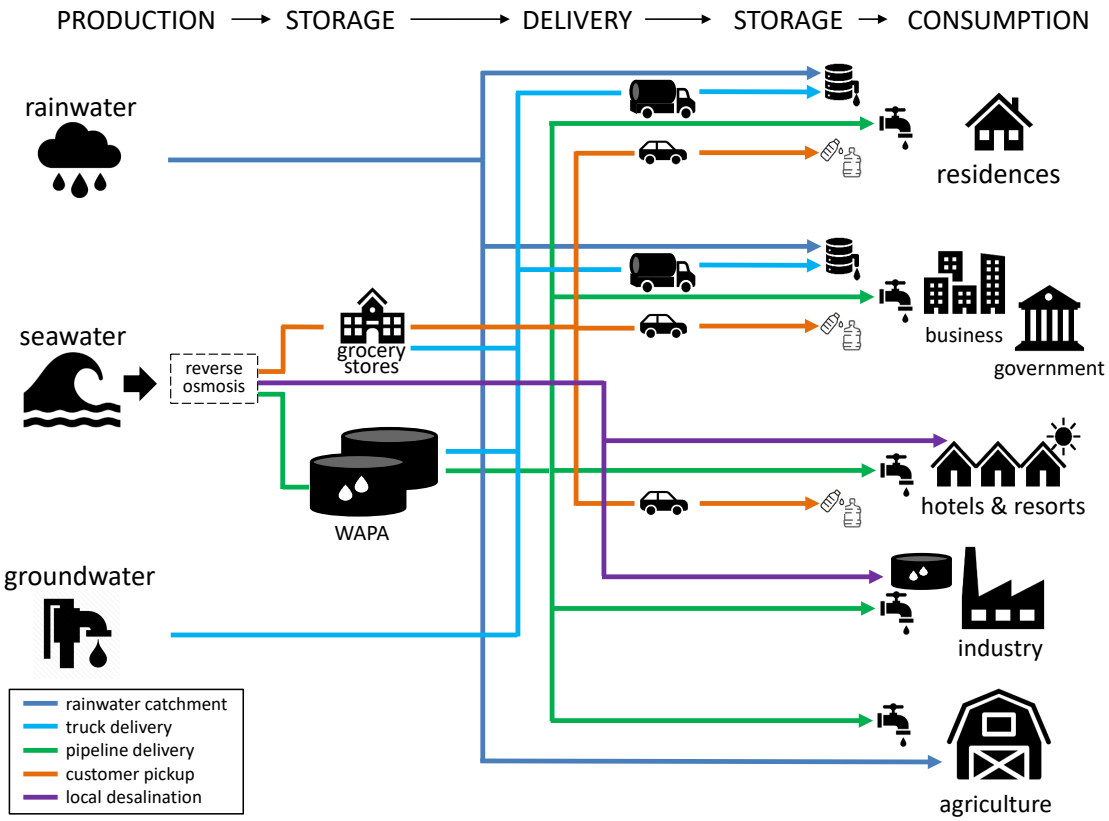

Figure 1.8. Overview of the interactions in the USVI Potable Water System. Diagram of how water moves through the water system in the USVI. The colors highlight different water distribution methods. Dark blue is rainwater catchment direct to the end user storage, light blue is transported by a water hauler, green through WAPA pipes, orange is water carried by individual private consumers, and purple is RO done by businesses.

1.4 Vulnerability of the WAPA Pipeline System to Service Disruptions

This work focuses on untangling interactions in the USVI potable water system to improve WAPA operations *and water security*. In general, potable water systems are vulnerable to a number of factors like natural disasters and terrorist attacks. It is important for water providers like WAPA to prepare for disasters and ensure communities do not lose water. However, the complex interactions in the USVI potable water system make it difficult to

predict when, where, how, or why water will be needed on a normal day. This makes it even more difficult for WAPA to plan for crisis situations.

Here, we present a description of several of interactions and associated vulnerabilities that challenge the long-term operations and protection of the WAPA pipeline system.

1.4.1 Water Quality and Ongoing Maintenance

The WAPA water system has known water quality issues, such as low chlorine residuals as well as turbidity resulting in places of low water pressure. In addition, there are widespread leaks in the territory's aging pipe system. Pipes installed in the 1950s and 1960s with 40-year service lives are still in use 50-60 years later. Approximately 85% of the non-revenue water on STX are attributed to leaks (CDR Maguire 2019).

In addition to the age of the pipe system, some tanks are more than 50 years old. There are concerns about the potential failure of a large storage tank in a storm or earthquake.

The aging system requires significant maintenance. This is expensive, but WAPA does not have steady cash flow to support a regular maintenance program. WAPA lacks a steady cash flow due a lack of large consistent customers and problems billing their smaller customers. The Government of the Virgin Islands is WAPA's largest customer and does not always pay its bills (Alderson et al. 2018).

1.4.2 Hurricanes

The Caribbean is at risk of getting hit by hurricanes and suffering severe damage to infrastructure. Since 1989 the USVI has been hit by four of the forty costliest hurricanes in US history: Hugo (1989), Marilyn (1995), Irma (2017), and Maria (2017).

Late on September 6, 2017 and into the morning of September 7, 2017 Hurricane Irma hit STT and STJ. The storm was a category-5 hurricane on the Saffir-Simpson scale (Taylor et al. 2010). Residents on STT and STJ lost power and water following the storm. The USVI responded by bringing emergency response assets from STX. This was helpful until two weeks later when on September 20, 2017 a second category-5 storm, Hurricane Maria, hit STX and Puerto Rico. And according to the National Hurricane Center it was the costliest storm to hit the USVI and Puerto Rico at an estimated combined $93.6 billion of damage

in 2019 dollars (National Centers for Environmental Information and National Hurricane Center 2020).

Hurricane Irma killed four USVI residents and caused damage to critical infrastructure to include the energy grid, telecommunication, wastewater, and water systems. This damage was in addition to flooding and damage to buildings. Hurricane Maria killed three USVI residents while causing structure and flood damage. The impact of Maria was exacerbated by the emergency supplies having already gone to the rest of the territory (Alderson et al. 2018).

The fragility of the water system was made apparent by Hurricanes Irma and Maria. Some residents in the territory did not regain services for over 100 days (USVI Hurricane Recovery and Resilience Task Force 2018). Service problems persist more than two years after the storms.

1.4.3 Droughts

Prior to the hurricanes of 2017, the most prevalent threat to the water system was drought. "Since the 1950s, the region has experienced at least seven major droughts. A severe drought that occurred from 2014–2016 resulted in water rationing for 1.2 million people and over $14 million in agricultural losses" (U.S. Geological Survey 2018).

Droughts are expected to both increase demand for water and limit the supply of water coming from rainwater catchment. This means that the demand for purified seawater and groundwater will increase in a drought scenario. Water demand projections are useful to understand if supply is capable of meeting demand.

1.4.4 Earthquakes

Earthquakes are a concern fresh on the mind of those in the USVI due to the magnitude 6.4 earthquake that struck Puerto Rico in January 2020 (van der Elst et al. 2020). Water storage tanks are vulnerable to damage from earthquakes. Catastrophic damage to one large tank on any island would reduce storage capacity by at least a third. This is of particular concern on STT where the ground is already crumbling in places under the tanks. The tanks are also positioned such that if one tank collapses it could take out more than one tank,

further damaging water storage capacity. This reduction in capacity is in addition to likely significant damage to the land surrounding the tank. Earthquakes also have the potential to cause widespread damage to the cisterns across the territory (thereby disrupting the ability of residents to store water) as well as the pipeline systems that deliver water.

1.5 Thesis Goals

As the territory looks to "build back better" and invest in a manner that addresses water security and resilience, there is a need to understand both the demand for water among the population and the capability of potable water systems to serve them. In addition to the threat of future hurricanes, there are a variety of other potential hazards that could cause problems, such as droughts and earthquakes. It is also possible that the population will grow or ongoing land development will lead to a rise in demand for potable water.

A model for production, storage, and delivery of potable water in the USVI does not exist. The lack of a demand model means there is no way to estimate who will have fresh water when parts of the system are in degraded service conditions. Moreover, without a demand model, there is no means to coordinate water delivery across the multiple Territorial stakeholders when future disaster strikes. The purpose of this thesis is to curate and integrate different sources of data, build a model, and perform analysis in support the following questions:

- **Research Question 1 (RQ1):** What factors are significant for understanding water demand in the USVI?
- **Research Question 2 (RQ2):** Can standard time series models capture these factors and appropriately predict water use?
- **Research Question 3 (RQ3):** Did water use in the USVI change following the September, 2017 hurricanes?

This thesis is in support of Federal Emergency Management Agency (FEMA) recovery activities and part of a broader FEMA-funded effort by the Naval Postgraduate School (NPS) and Department of Energy (DOE) to assess and improve the resilience of interdependent USVI lifeline infrastructure systems (Alderson et al. 2018). This thesis additionally supports several other complementary efforts with the University of the Virgin Islands (UVI) to develop a next-generation Hazard Mitigation and Resilience Plan for the territory.

CHAPTER 2:
Literature Review

Understanding water demand in the USVI requires pulling together standard urban water connections, non-traditional water connections, industrial water, and business water demand patterns. WAPA has lacked data and coordination to make these predictions in the past. Even with the data, the combination of customers that live in the territory has not been integrated in a model. Therefore, methods developed in prior literature require enhancement.

Here, we begin by discussing past work on water demand patterns. Then, we consider past work on water demand projections. In both of these sections we include a discussion of past work in these areas in the USVI. Finally, we discuss our approach and how they aid in understanding the implications of Hurricanes Irma and Maria in the USVI.

2.1 Past Work on Water-Use Patterns

The data available for water systems is typically lower resolution than the desired prediction interval. For example, a water utility may only have long term data available on a daily scale, when they want to make hourly demand predictions. In order to work with this data, a model's projections are fit to typical demand patterns that are lower resolution than the inputs to the model (Vassiljev et al. 2015).

Demand is typically broken into residential, industrial, and business patterns (Letting et al. 2017). Figure 2.1 is an example plot showing their different use patterns. Residential demand follows a diurnal patter. The diurnal pattern has a peak in the morning and in the evening. Residential demand is moderate through the day and low at night. Commercial, or business, demand is steady through the work day and low at night. Industrial demand is steady throughout a 24-hour period.

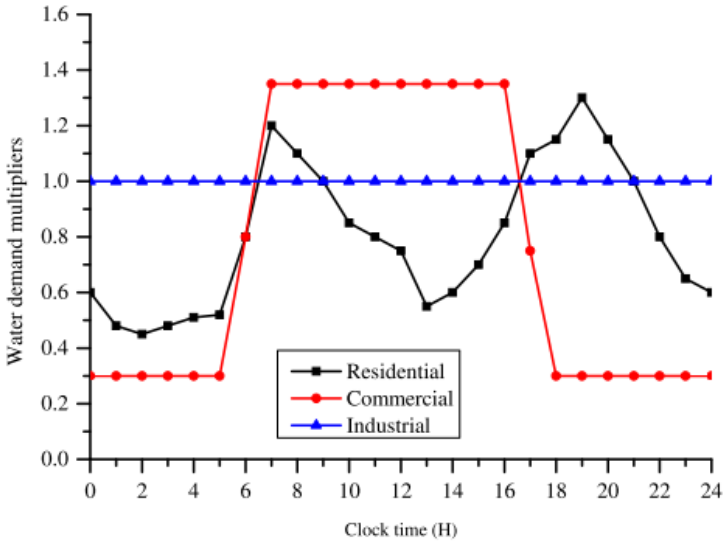

Figure 2.1. Example of Business, Industrial, and Residential Demand Patterns. Business demand is constant during the working day and zero outside it. Industrial demand is nearly constant over 24 hours. Residential demand pattern is diurnal with peaks in the morning and at night. Source: Letting et al. (2017)

Studying the city of Fort Collins, CO, Sharvelle et al. (2017) found that the indoor water demand was steady, but the outdoor water demand varied significantly by season (Figure 2.2 visualizes the difference). In Greece, Kossieris and Makropoulos (2018) also found a large variation between summer and winter seasonal water demand. In a conservation study, Mayer (2016) found that small changes in some customers outdoor water-usage can have a comparatively large impact on the overall demand.

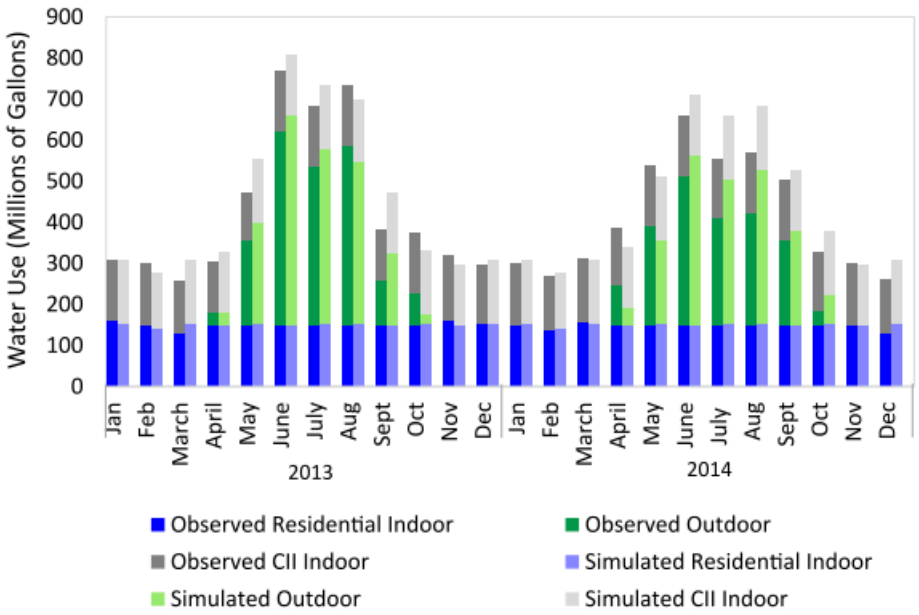

Figure 2.2. Indoor and outdoor Water Usage in Fort Collins, CO. Simulated and measured water usage shows a steady indoor demand along with significant seasonal variation in outdoor water-use. Outdoor water-use is high in the summer and low in the winter. Source: Sharvelle et al. (2017)

Arouna and Dabbert (2010) found that houses connected by pipe to the water system have inelastic demand with respect to the price of water. Jain and Ormsbee (2001) modeled water demand in Lexington, KY and found that drought characterized by period of time without two millimeter (mm) of rain was a significant indicator for water demand predictions up to five days out. Xue et al. (2017) found that residential water demand spikes upon return from vacation periods. Kossieris and Makropoulos (2018) found that Weibull and Gamma distributions were the best fits for predicting periods of nonzero demand on 15-minute and hour-long time scales, respectively.

Systems with non-homogeneous water sources often require customers to make a discrete choice to which supply they use to satisfy their water demand. Mu et al. (1990) found in Kenya that both price and time to collect the water also affect demand when there is a choice of several sources. Mu et al. (1990) used a stochastic term in a discrete choice model that was influenced by cost and time to collect the water. Overall, these choices do not affect water demand patterns. In a three-year South African case study, it was found that these

homes still follow the diurnal pattern for their overall water demand (Pretorius et al. 2019).

Water-Use Patterns in the USVI

In the USVI, we must consider more than the typical residential water system because of the many types of consumers. There are typical pipe connected users, users that only have cisterns with or without rainwater catchment, and users that have a combination of both. The residential demand on WAPA water pipes is complicated by the households that use cisterns. The cistern users likely follow a diurnal water-use pattern, but some of their water demand is filled by rainwater catchment, or by water haulers.

When houses have access to a cistern and WAPA water, there is potential for non-homogeneous water choices. An example of this, would be when a home with a cistern uses their cistern water for cooking and cleaning, but not drinking. Lucas et al. (2010) found that the presence of cisterns alters the diurnal demand pattern. This makes sense because the house is able to pull from the cistern at times which will reduce its pipe demand.

Time to collect water is not a significant consideration for water consumers in the USVI. If water is not delivered by pipe, it is either caught through rain water catchment, or delivered by water truck. Therefore, we can better understand the choice from the volume of recent rainfall influencing the choice between cistern water, trucking to fill the cistern, and pipe water.

Buscemi (2003) studied water usage on STT and predicted typical demand. In his study he considered that 50% of the population on STT does not have access to WAPA water pipes. This proportion of the population is reliant on their cisterns that are filled through rainwater catchment or by a water trucker. In the study the catchment volume for customers was estimated using building data.

Solomon (2009) claims that under regular conditions in the USVI (i.e. no drought), some cisterns never require filling from a water truck. These households are able to use rainwater catchment only, and not depend on WAPA for their water demand. These households may still purchase bottled water for drinking.

In 1964, the USVI government passed a law requiring all buildings not connected to WAPA water to have cisterns (CDR Maguire 2019). The size of the cisterns was dictated by the

square footage of the roof. The goal of the legislation was to increase the amount of water captured through rainwater catchment in the territory (Solomon 2009). The cistern size requirements are shown in Table 2.1.

Table 2.1. USVI Cistern Size Requirements. Rainwater catchment requirements in the USVI. Source: Adapted from Buscemi (2003).

Type of Structure	Required Cistern Size (gallons per square foot roof area)
Single Story Dwelling	10.00 gal./ft.^{2}
Multistory Dwelling	15.00 gal./ft.^{2}
Churches and Warehouses	4.50 gal./ft^2

In addition to rainwater catchment, some businesses in the territory also use wastewater capture. Hotels, condominiums, and industry in the territory use their wastewater for irrigation, habitat restoration, and industrial cooling (Cheslek 2003).

2.2 Past Work on Water Demand Projections

Water demand projections range from short-term to long-term, and can come from the supply or demand side. Short-term projections are used for tactical operator decisions in the coming days, weeks, or months. Medium- and long-term projection, three months out to years, are used for strategic decisions by utilities. Decisions include when to expand capabilities, or the best way to expand existing capability (Donkor et al. 2014). Producer-focused projection uses data from the water producers. Consumer-focused models use data based on the consumers of the water.

Water demand projections can consider a number of different outcome variables. Some outcomes include: peak day demand, daily total system demand, monthly total system demand, annual per capita demand, annual demand by customer class, and revenue (Donkor et al. 2014). We consider total demand as the outcome for our modeling effort.

2.2.1 Producer-Focused Demand Forecasting

There are a variety of techniques that have been employed by utilities and researchers to estimate future water demand. Often the water producer has the most complete and accessible data. Despite this, public utilities, to include WAPA, have often relied on qualitative prediction through operator intuition and experience (Donkor et al. 2014). More complex projection techniques use this data, while also considering time and environmental factors to create producer- or consumer-focused models.

In a univariate time series, or autoregressive approach, the modeler assumes that past patterns will repeat. These patterns are then used to predict future demand (Donkor et al. 2014). This type of analysis is good for steady-state use, but cannot capture changes in demands due to disruption.

There are some examples of univariate time series analysis that consider additional factors. Brentan et al. (2014) used a reconstruction method that incorporated rainfall data. However, this model left the modeler wanting for accuracy and a Fourier series approach was applied. The Fourier series produced more accurate results in aggregate. But, it failed to capture large variations (Brentan et al. 2014).

Time-series regression models are a way to move beyond the univariate time series without the end use usage data. Time-series regression considers trends, seasonality, and other factors such as weather (Donkor et al. 2014). Polebitski and Palmer (2010) considered this on a census tract level in Seattle, Washington with bimonthly usage, economic, and weather data.

Artificial Neural Network (ANN), or fuzzy techniques, can be more effective than traditional method for capturing non-linear changes in demand. They are considered to be more useful and effective for short term than long term projection (Donkor et al. 2014). Firat et al. (2010) used an ANN in a study in Turkey and found that it is important to consider weather (rainfall) as well as commercial and industrial conditions in the region. Herrera et al. (2010) also used ANN in Spain and found that weather factors are significant for explaining weather demand. Zhang and Qi (2005) used ANN, but found that in order to capture seasonal trend variation the data required nontrivial preprocessing.

It is also possible to combine some of these techniques in a composite model. One way

to approach a composite model is by having a weighted average of multiple prediction techniques. Another example would be to split the demand into the steady state and the cyclic component. A modeler might use regression for the deterministic component and ANN for the cyclic component (Donkor et al. 2014). Zhou et al. (2002) splits the base demand as 10% of the average demand for weekdays and weekends, respectively. The ANN was then used to model the remaining demand.

Similar to a composite model, Box-Jenkins models—also called Autoregressive Integrated Moving Average (ARIMA) models—combine autoregressive and moving-average components to make quantifiable estimates of future demand. ARIMA models are part of a class of stochastic process models that allow modelers to quantify uncertainty associated with forecasted values (Donkor et al. 2014). The Seasonal Autoregressive Integrated Moving Average (SARIMA) model extends the ARIMA model by adding a seasonal component. SARIMA models have be used to represent both daily cycles (24/48-hour) with fine time resolution (15 to 60-minute) data, or quarterly/annual cycles with coarser time resolution (weekly or monthly) data. Modeling annual cycles and daily cycles requires too many initial parameter estimates and is generally not possible in current implementations.

All-in-one projection software, or decision support systems, include multiple models and then select the model to use based on the data available and the projection timeline of interest. This software may also be employed when there is limited data available to compare several imperfect models. The US Army Corps of Engineers Institute for Water Resources and Maddaus Water Management developed two commonly used decision support systems (Donkor et al. 2014).

All of these techniques may suffer from a lack of data availability. Kang and Lansey (2009) handle this problem by bundling groups similar users to reduce the number of unknowns in the model. Do et al. (2017) employed an optimization algorithm that used averages and demand multipliers to assign demands to nodes that exceed the number of inputs to the model.

2.2.2 Consumer-Focused Demand Forecasting

A stochastic process model provides a consumer-focused, and finer-resolution alternative. These models expand on time series modeling by quantifying uncertainty. Donkor et al.

(2014) include a stochastic term in these models representing the fixture in the home that dispenses the water. In order to do this, the modeler requires extensive end use data for each residential fixture, such as that found in DeOreo et al. (2016). A more general consumer-focused approach is to take average water uses per person, or household, and multiply by the population to get a sense of the expected water demand. Time and environmental factors applied to supply-focused forecasting can also be implemented in a consumer-focused demand forecast.

Water Demand in the USVI

Water demand projection in the USVI is difficult due to a unique mix of water users. Water demand projection is further complicated in the USVI by variable dry seasons and storms that cause demand disruptions. Buscemi (2003) implemented a heuristic projection technique for the USVI using general average use per person and multiplying by the population.

The USVI has had a relatively stable population over the last several decades, so consumer base growth is not as large a concern for WAPA as it might be for another utility. However, the arrival of large cruise ships significantly alters the population of the territory and have the potential to drive up demand for water. The number of cruise ships in port could be useful to develop a total demand projection for the USVI.

2.3 Water Resources Availability

Water in the USVI is available through WAPA pipes, from cisterns filled by RWC or water haulers, by individuals purchasing their own water from grocery stores, or a consumer or business producing their own water. All of these sources are susceptible to disruption from either changes in weather or natural disaster.

Long term droughts can deplete cisterns that depend on RWC and force the consumers they support to increase their use of WAPA water, or fulfill their demand by purchasing from a water hauler. Drought may also be associated with higher overall water-use creating a strain on water availability throughout the territory.

Natural disasters such as tropical storms and earthquakes can make it difficult or impossible to deliver water to some or all consumers. After the tropical storms in 2017, all services

were not restored for more than 100 days (Alderson et al. 2018). Also after the storms, WAPA has had more frequent pipe breaks and leaks, which means that more water has to be produced to meet the same demand. Storms also have the potential to knock out private water production and storage, forcing them to turn to other water sources. Natural events do not have to destroy infrastructure to make water unusable, the 2017 storms saw widespread contamination of cistern water from debris. Taken to the limit, or coupled with an increase in usage, demand could outstrip supply.

Earthquakes also pose a threat to water resource availability. An earthquake could severely damage water delivery and storage systems. A failure of the tanks or pipe system would cause many consumers to lose access to water. To some extent, this could be mitigated by water haulers, but the capacity does not exist to support the entire island. WAPA tanks as well as private cisterns could be damaged or destroyed in an earthquake. Tank loss would create a drop in pressure on the island meaning that water could not be delivered to all consumers who depend on water pipes. Destruction of cisterns would mean loss of existing supply and complicate the use of water haulers.

2.4 2017 Hurricane Impacts and Their Implications

The Hurricanes of 2017 had major impacts. First, they damaged infrastructure systems in a major way. Many of these are still recovering more than two years later. A first question is whether or not there is a change in the water-use following these storms. A second issue relates to the implications for hazard mitigation and management in the territory. What does this mean for water security in the near- and long-term?

In the following chapters, we consider the available data for the USVI prior to the 2017 storms, and develop a statistical model to determine which variables are significant for predicting water demand in the territory. We then apply the model to data after the 2017 storms to understand if the factors that drive water demand have changed following this natural disaster.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 3:
Data Preparation and Model

A variety of water and environmental data sets are available for the USVI. The National Oceanic and Atmospheric Administration (NOAA) records atmospheric data available down to thirty-minute intervals. WAPA has daily and monthly reports for the water production, meter flow reading, and storage. Many of WAPA's assets are mapped as Geographic Information System (GIS) data. Calibrated elevation data is available for the entire territory. The Census Bureau has records of the water sources used by households down to the estate level in the territory. United States Geological Survey (USGS) surveyed the groundwater resources in the 1970s.

In this chapter we discuss the available data, how we process it, and how the data motivates the model that we employ for the USVI.

3.1 NOAA Weather Data

NOAA has weather stations in the USVI that track temperature and precipitation. The weather data includes daily precipitation in inches as well as the minimum and maximum temperature in Fahrenheit.

For temperature, we consider the maximum daily temperature from March, 2006 to December, 2019 and calculate a mean maximum temperature for each month. Figure 3.1 shows a box plot of the mean maximum monthly temperature for each month. The annual variation in mean maximum temperature is less than ten degrees Fahrenheit. The climate is consistently hot. High temperature can increase water use and cause an increase in evaporation.

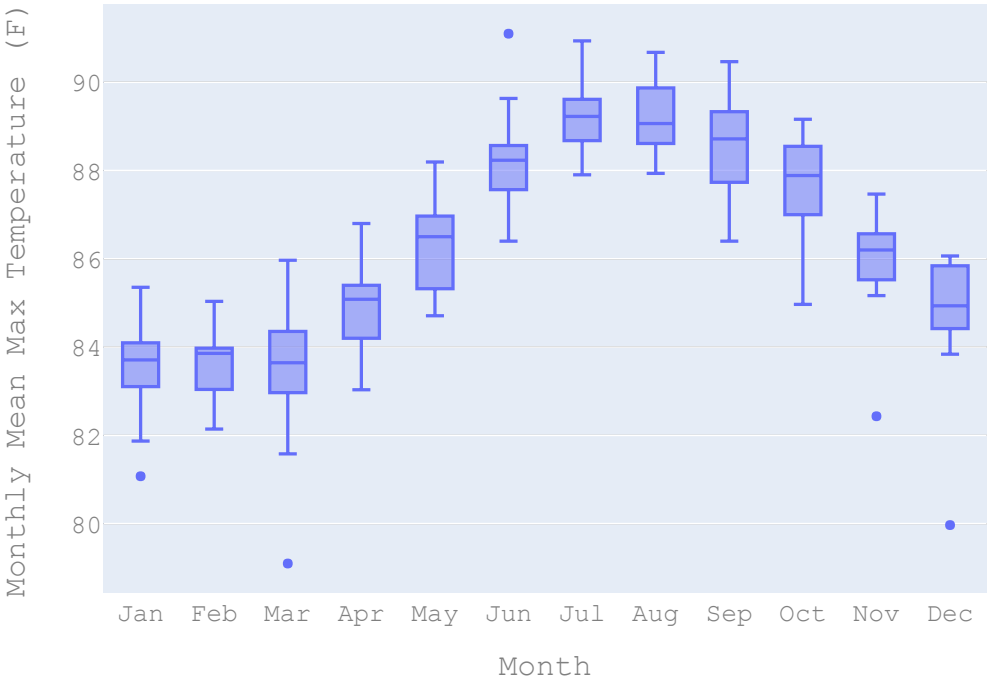

Figure 3.1. STT Mean Maximum Temperature March, 2006 to December, 2019. The mean maximum temperature varies by less than ten degrees Fahrenheit annually. The climate is consistently hot with mean highs rarely going under 80 degrees.

The changes in weather are historically consistent year to year, and cyclic. Figure 3.2 plots the monthly mean max temperature in Fahrenheit over 13 years showing that the same temperature pattern repeats with variation of less than 4 degrees Fahrenheit.

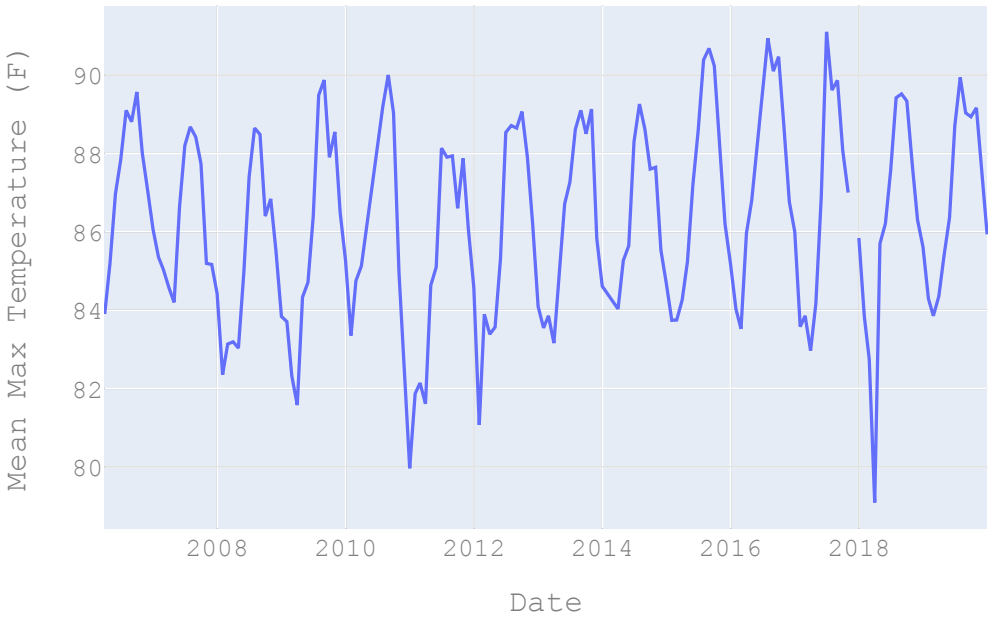

Figure 3.2. STT Mean Maximum Temperature March, 2006 to December, 2019. The mean maximum temperature in STT is cyclic and consistent year to year.

Rainfall is seasonal like temperature, but it does not follow as consistent a pattern. There is more rain in hurricane season (June-November), but there is more variation in when it comes. It may come all at once and may vary in total precipitation year to year. To understand annual rainfall, we sum the total rainfall in each month and then compare the total monthly rainfall for data going from March, 2006 to December, 2019. Figure 3.3 is a box plot of this rainfall data. We see there is usually less than three inches of rainfall per month outside of hurricane season (June-November). During hurricane season there is much greater monthly variation with as much as 16 inches of monthly rainfall and as little as less than an inch of rainfall in a month. The territory's population uses RWC for water, so the variation in rainfall could lead to shifts in the source used to meet water demands.

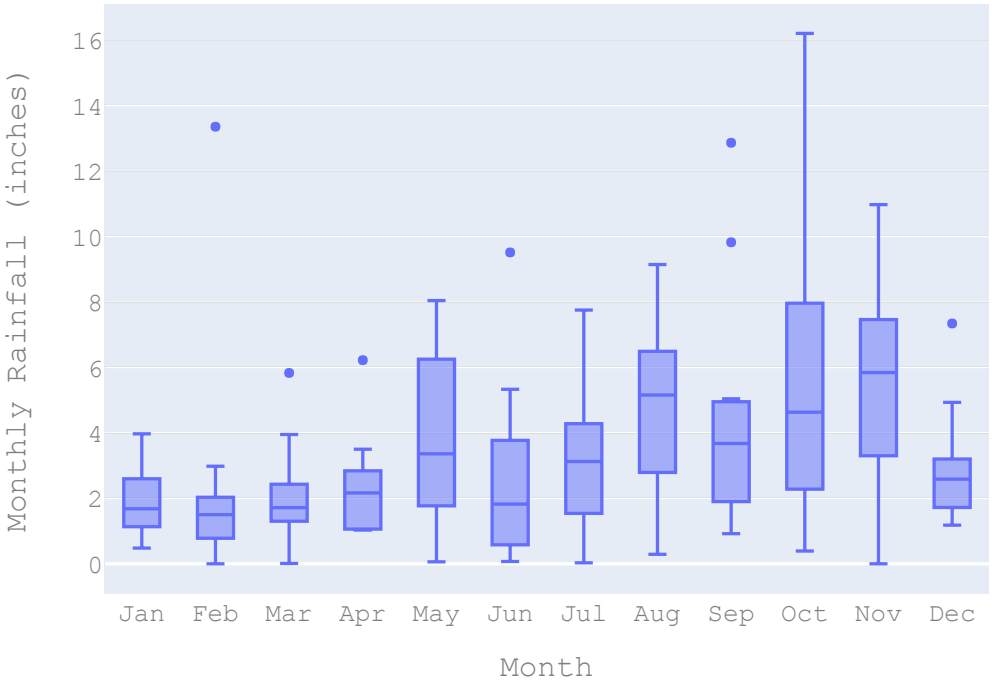

Figure 3.3. STT Monthly Precipitation March, 2006 to December, 2019. Total precipitation is typically less than three inches outside of hurricane season (June-November). Hurricane season can have months with greater than ten inches of accumulated rainfall mixed with months of little to no rainfall.

3.2 Demographic Data

The World Bank keeps annual population data for countries and territories around the world. The World bank data, plotted in Figure 3.4, shows that the population in the USVI is decreasing slightly. According to the World Bank, the population in 2002 was 108,510. This shrunk slightly in 2010 when the population as 108,358. The population shrunk again to 106,977 in 2018, the most recent year available (World Bank 2020).

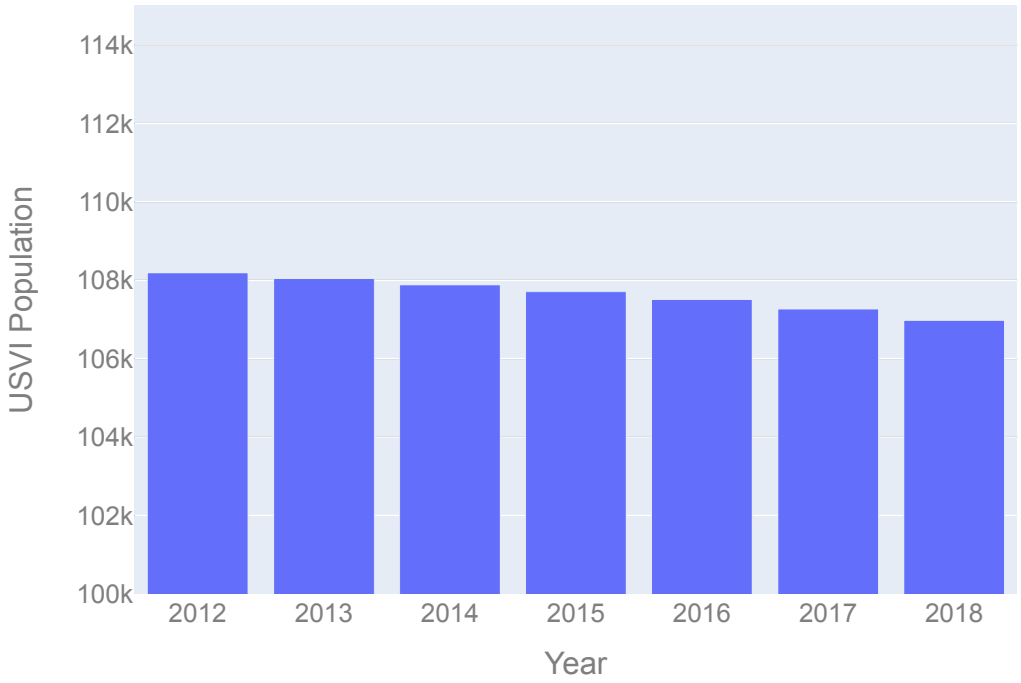

Figure 3.4. USVI Population, as reported annually from 2012 to 2018.

Although the resident population is relatively stable, the islands can see significant fluctuation due to tourist traffic. STT in particular can see its population of 51,000 increase by as many as 30,000 people when its cruise ports are full. This could potentially mean a significant increase in water use when visitors disembark and visit local restaurants and attractions.

The U.S. Census records, among other demographic data, what kind of vendor people use to acquire water in the USVI. Table 3.1 shows the percentages of water purchases on each island, and the territory overall, for each type of water vendor in both 2002 (data from Buscemi 2003) and in 2010. Since 2002, WAPA has increased pipe access across the territory. The data reflects the increasing reliance on the public system for water users on all three islands. Although the overall population is decreasing, the demands on WAPA's system could increase due to expanded pipe service.

Table 3.1. USVI Water Purchases By Source from 2010 Census Data. Numbers in parentheses are from Buscemi (2003) using US Census Bureau data in 2002. Across the territory use of the public system (WAPA pipes) has increased and cistern use has decreased.

	2002				2010			
Relevant Datum	**STT**	**STJ**	**STX**	**Total**	**STT**	**STJ**	**STX**	**Total**
Public System	24	3	25	24	33	11	28	29
Public System and Cistern	19	14	26	22	14	10	22	18
Cistern, Tanks, Drums Only	55	81	45	52	52	78	47	51
Public Standpipe	1	2	4	2	0	0	0	0
Other Source	0	0	0	0	1	1	3	2

The US Census also has records for which households in each subdistrict and estate have access to water from what sources. The subdistricts for STT are shown in table 3.2. There are fewer public system users away from the main storage tanks and Charlotte Amalie.

Table 3.2. STT Household Water Purchases by Source. The census data breaks down which households in each subdistrict uses water from what water source. There are fewer public system users away from Charlotte Amalie and the main storage tanks.

Relevant Datum by Subdistrict	**Charlotte Amalie**	**East End**	**North-side**	**South-side**	**Tutu**	**Water Island**	**West End**	**Total**
Public System Only	5427	1453	643	979	489	3	36	9030
Public System and Cistern	2370	224	262	414	477	11	36	3794
Cistern, Tanks, or Drums Only	1467	3370	4593	1533	1913	187	1050	14113
Public Standpipe	15	6	8	1	4	0	4	38
Other	64	75	30	17	8	2	2	198
Total	9343	5128	5536	2944	2891	203	1128	27173

Table 3.3 shows the same water source information by household for STJ. STJ has significantly fewer public system users. The island mostly uses RWC and private water production.

Table 3.3. STJ Household Water Purchases by Source. The island has relatively few public system users and large depends on RWC or privately produced water.

Relevant Datum by Subdistrict	Central	Coral Bay	Cruz Bay	East End	Total
Public System Only	39	45	308	0	392
Public System and Cistern	41	22	282	3	348
Cistern, Tanks, or Drums Only	528	471	1621	59	2679
Public Standpipe	0	1	8	0	9
Other	2	6	17	0	25
Total	610	545	2236	62	3453

3.3 WAPA Water System Data

WAPA is the only public water utility in the USVI. So, they have the most available data for the largest portion of water production and consumption in the territory. We use their reports to build a picture of water production and consumption in the USVI.

St. Thomas and St. John Daily Reports. We start with the data that WAPA maintains about their own water production and delivery. They have daily reports on an island-wide level for STT and STJ divided into two to three regions by island. These reports are nearly complete going back to 2010 and are sparse going all the way back to 2004. The most recent daily data available is for March, 2020.

Figure 3.5 shows an example of one of the daily reports. The reports are available as a spreadsheet, but are formatted for human readability and not easily parsed as a simple dataframe. Additionally, there is a separate data file for each month with a separate sheet for each day of the month. We used the `xlrd` package to read each data cell in the daily reports and create a single dataframe with all of the production, storage, and meter flow reading

information (Machin 2020). The resulting dataframe has over 4437 days of WAPA data.

VI WAPA					Activity From:		30-Dec-19	9:00 AM
Water Distribution Daily Summary Report:						To:	31-Dec-19	9:00 AM
ST. THOMAS				**Received**			**Distributed**	
Tank	Feet	Gallons		IDE:				
#2 10.5:	30.08	6,582,233		Gals	1,860,000		Town:	1,698,183
#3 10.5:	47.50	10,393,000		On	RO		East End:	315,000
#4 10.5:	21.92	4,795,367		Off	IDE		St. John	175,000
~~UG #1:~~	~~Off Line~~	~~0~~					Total:	2,188,183
UG #2:	~~Off Line~~	0						
UG #3:	~~Off Line~~	0						
UG #4:	~~Off Line~~	0						
~~Filt #1:~~	~~Off Line~~	~~0~~						
Filt #2:	5.42	270,833	**Location**		**Gallons**	**%**		
Sara Hill Storage:		21,931,433	**Sara Hill Net:**		**20,952,133**	**68.24%**		
Donoe:	33.92	3,900,417	**Donoe Net:**		**3,670,417**	**69.38%**		
System storage:		25,831,850	**System Net:**		**24,622,550**	**68.41%**	**Pumped:**	**24.00 hours**
Previous date storage:		26,160,033					**Plant S. S.**	0
Change in storage:		(328,183)					**Truck sales:**	1,000
ST. JOHN				**Received**			**Distributed**	
Tank	Feet	Gallons		IDE:				
Standpipe .75 Mil.:	34.25	713,530		Gals.	166,350		Cruz Bay:	160,193
Pastory .132M:	17.75	97,625		On:	STT		Truck sales:	24,130
Education .1M:	15.00	83,325		Off	IDE		Station Service:	60
Catchment .5M:	26.67	370,373					Total:	184,383
			Location		**Gallons**	**%**		
System storage:		1,264,854	**St. John Net:**		**1,173,302**	**84.38%**	Pumped:	24.00 hours
Previous date storage:		1,282,887						
Change in storage:		(18,033)						

Figure 3.5. WAPA Daily Report Example. WAPA has daily water reports for St. Thomas and St. John with storage tank levels. The reports also include daily water meter flow readings to two to three regions per island and the amount of water produced by RO for WAPA each day. We processed 4437 unique reports.

Figure 3.5 shows that STT is broken up into Town, East End, and water sent to STJ for storage and use. STJ is broken into Cruz Bay, Truck Sales, and Station service. The data has fine temporal resolution, but coarse spatial resolution.

The Daily Reports also provide information about the movement of water through a complicated network of pipes, junctions, and valves. Figure 3.6 depicts a simplified view of this system consistent with the WAPA Daily Water Reports. Water is produced by reverse osmosis and stored in Tanks #2, #3 and #4 at Sara Hill. Water is delivered from these tanks to "Town" (representing the greater area around Charlotte Amalie), and then passes to storage in the Donoe Tank, which in turn supplies the East End of St. Thomas as well as the undersea pipe transfer to STJ.

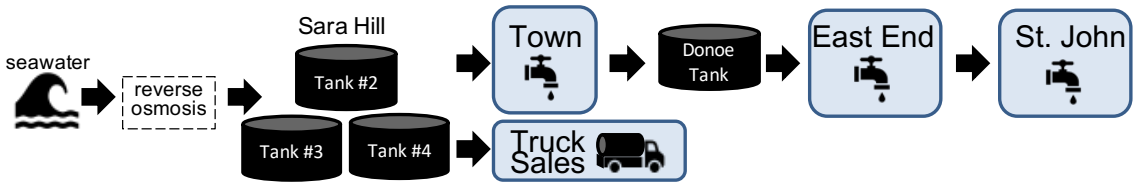

Figure 3.6. Simplified WAPA System for STT & STJ. Water is produced with RO on the left and moves through a network of pipes through the "Town" (representing the greater area around Charlotte Amalie) then the "East End" and finally to STJ. Truck sales are made from the standpipe fed directly by the tanks.

St. Thomas and St. John Monthly Water Meter Reports. WAPA has monthly data on a finer spatial resolution than the daily reports. On STT there is monthly data for 34 water meters, 27 of which have data sufficient for analysis. The monthly data goes back to March 2006 and up to December 2019. The monthly data comes from a single spreadsheet file spread across several sheets. Parsing the data requires knowledge about the idiosyncrasies of how the data was entered and maintained (including cases where values were mislabeled). The resulting monthly data frame contains 166 months of data.

We compare the monthly and daily data by aggregating the daily data into monthly data and plotting them together. By matching up the monthly and daily data we understand the fidelity of the daily data. In some cases a single meter does not match the water flow numbers reported in the daily WAPA data. We are able to calculate some of these values with our understanding of the WAPA network. This was not possible for Town, and we take the verification of the East End and STJ transfer data to mean that the Town data is also good.

RO Production. Figure 3.7 shows the RO production data from the daily and monthly WAPA report, from 2012 through the beginning of 2020. We observe that the daily and monthly reports generally match, but they diverge in October, 2017, immediately after Hurricanes Irma and Maria. We use the two data sets in tandem to verify the data points. For example, in October, 2017 the 85M gallon production reported by the monthly data set is more plausible than the zero gallon production reported by the daily data set.

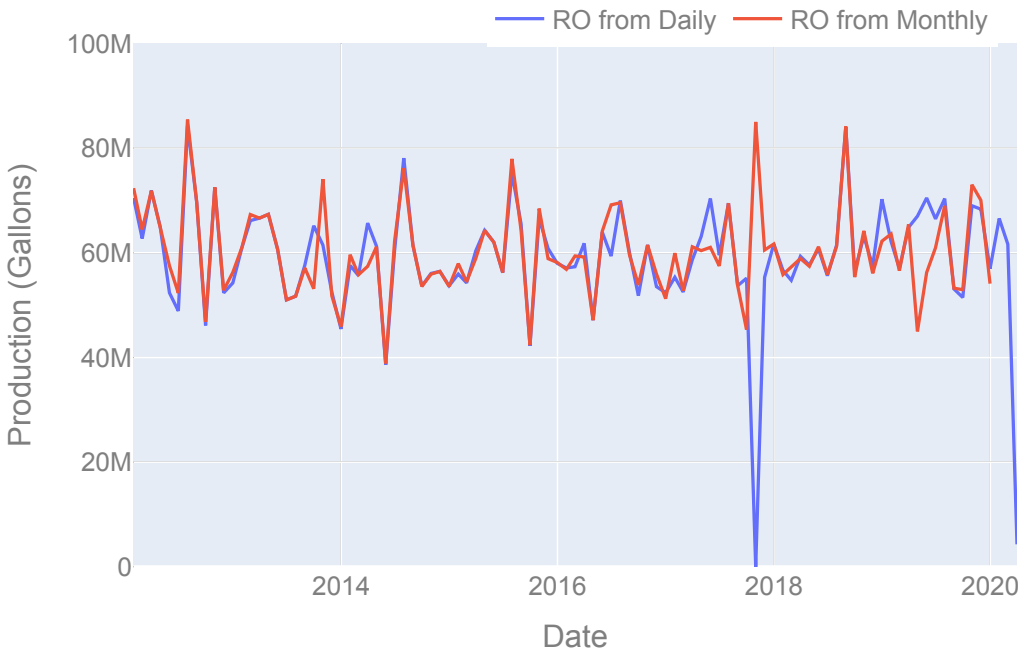

Figure 3.7. STT Reverse Osmosis (RO) Production. We see that the monthly and daily data start tracking each other closely in January, 2012 and continue without diverging until October, 2017. In this case, we can get rid of the zero value and consider the 85 million gallon production reported in the monthly data. The October, 2017 data problem is likely a result of the hurricanes in September, 2017.

Sara Hill Tank Storage. The WAPA daily water report also records the amount of water in each of its storage tanks. To the extent that these water tanks serve as a type of "water battery" for the delivery system, management of these tank levels is important for maintaining service. Figure 3.8 shows the fluctuation in the levels of Tanks 2, 3, and 4 over the period January, 2012 through March, 2020. We note that the tank levels are primarily a function of WAPA's operational policies and only an indirect measure of consumption. In general, WAPA has

the ability to increase the production rate from Seven Seas that directly fills the tanks at Sara Hill (including Tanks 2, 3, and 4).

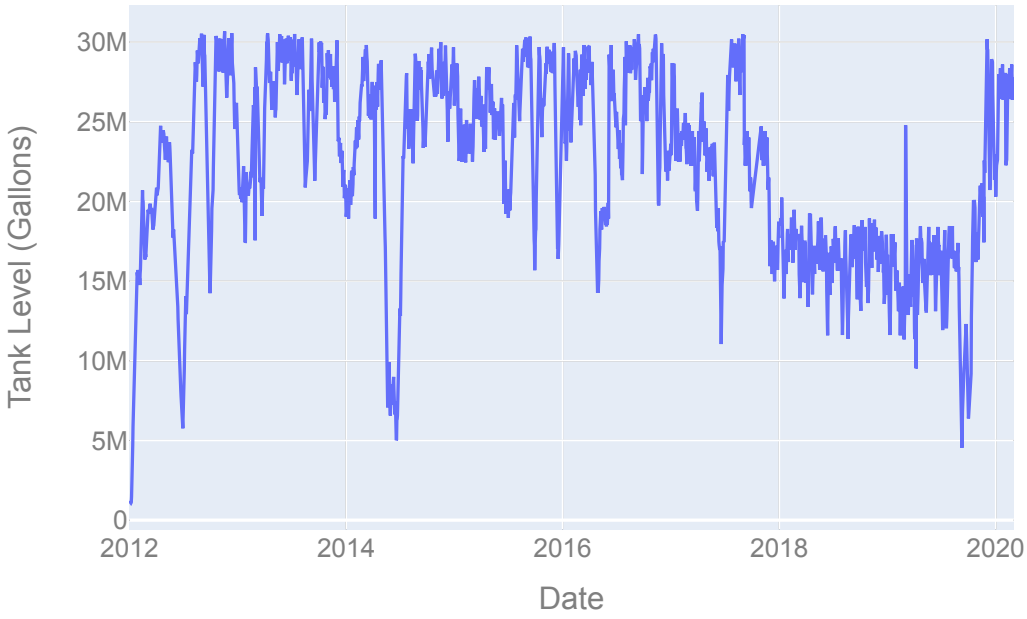

Figure 3.8. STT Combined Tank Levels. Daily water height reading for STT's three primary water storage tanks (#2, #3, #4) with an maximum capacity of 31.5 million gallons. Tank levels are only an indirect measurement of usage because it is influenced by operating policy in addition to production and consumption. The levels are WAPA's choice. After the hurricanes in 2017 WAPA was reduced to two thirds capacity for nearly two years.

The tanks have a maximum capacity of 31.5 million gallons and WAPA keeps them close to full. We also see that they keep storage from falling below 5 million gallons which can support 2-3 days of operation. Following the hurricanes in 2017, WAPA was reduced to two thirds storage capacity for nearly three years while they performed repairs to their tanks.

Water Truck Volume. Figure 3.9 shows the monthly and aggregated monthly from daily WAPA water truck sales. The data sets match up consistently. Truck volumes are consistently under 500,000 gallons. There are several years where there are monthly volumes in excess of a million gallons. Spikes in demand occurred in early 2013, mid 2015, mid 2016, mid 2018, and mid 2019.

Many water truck purchases are used to fill cisterns that otherwise depend on RWC. We likely see large spikes in demand due to long term drought, or reduced rainfall, that reduces RWC island wide and cisterns drain simultaneously. Similar to the storage tanks for the system, the cisterns are household or business "water batteries."

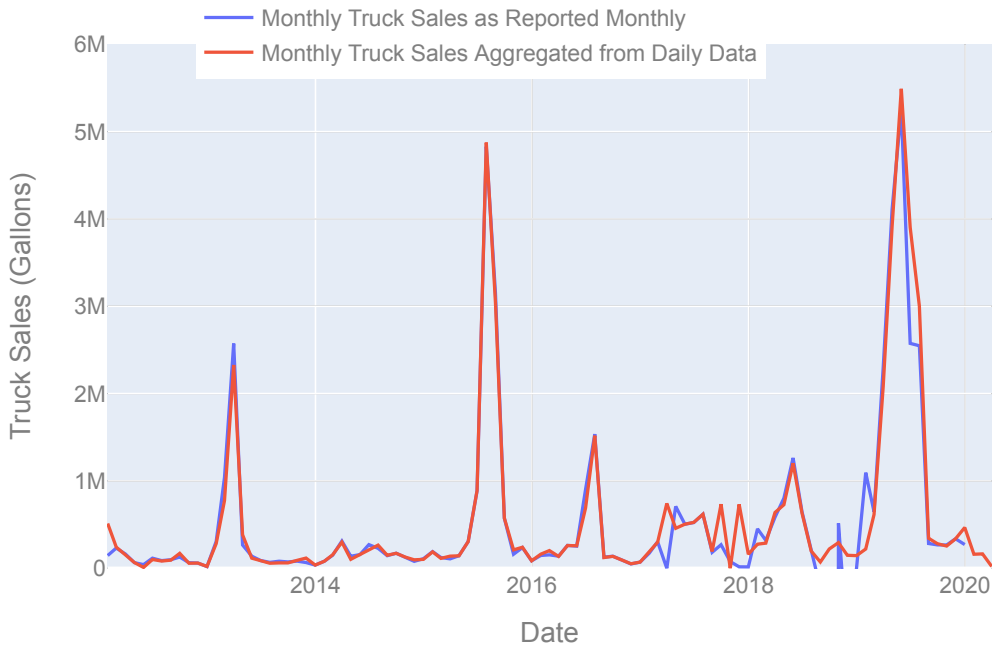

Figure 3.9. STT Monthly Water Truck Volumes January, 2012 to March, 2020. Monthly water truck volumes from the monthly and daily data sets are consistent. WAPA has a generally low level of truck volumes with multimillion gallon spikes in early 2013, mid 2015 and mid-2019.

Daily sales from WAPA to water trucks on STT range from 0 to 688,835 gallons, with a mean value of 26,916 gallons. Assuming a water truck with hauling capacity of 4,000 gallons, this means that WAPA fills between 0 and 172 trucks, with an average of 7 trucks daily.

Water Delivered To "Town." Figure 3.10 shows the monthly volumes of water delivered to Town (as aggregated from daily reports) and the same monthly volumes as calculated from our understanding of the monthly meter data. The monthly calculation was done by taking the inflow to Town and subtracting all of the readings that we know flow past Town. We see that there is still something missing in our calculation, but the Town data follows a similar trend to the overall volumes. Along with the alignment of our other segments of the system, we take this as evidence that the data for water delivered to Town is usable.

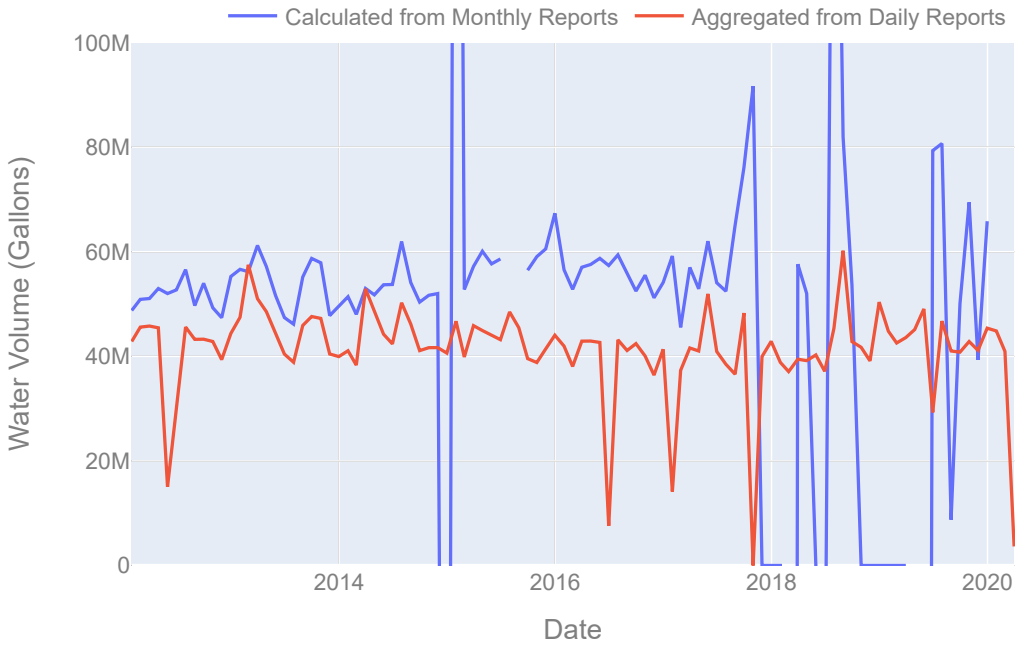

Figure 3.10. STT Monthly Water Volume Delivered to "Town" January, 2012 to March, 2020. When aggregated, the daily data follows a pattern to the volumes reported for monthly meters. The discrepancy between monthly and daily reports suggests there is something more in the monthly calculation that is not captured in the daily reports.

Donoe Tank. Figure 3.11 shows storage in the Donoe Tank. The Donoe tank supplies the East End of STT and the water transfer to STJ from STT. The tank has a maximum capacity of 5.4 million gallons and we can see that WAPA keeps it close to full as often as possible. From the monthly data we infer that WAPA's operating policy is to keep it over a minimum of two million gallons.

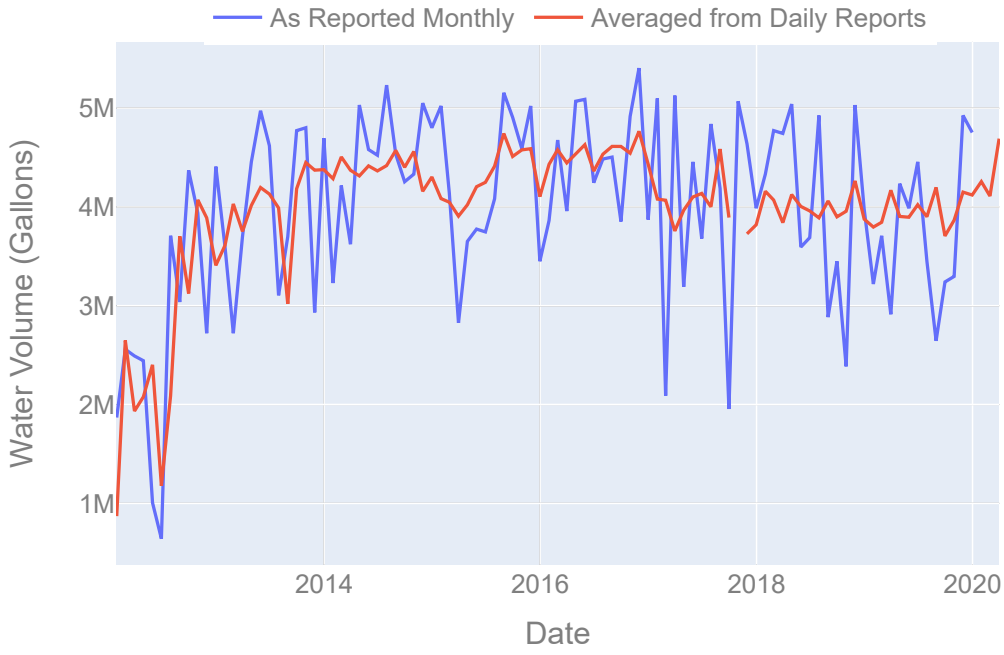

Figure 3.11. STT Monthly Donoe Tank Levels January, 2012 to March, 2020.

Water Delivered to East End. Figure 3.12 shows the volume of water delivered to East End on STT as reported by WAPA. East End is supplied by the Donoe Storage Tank. We calculate this monthly volume by taking the volume delivered for East End and subtracting the water transferred to STJ. There are some missing values, but the data is mostly complete prior to the 2017 hurricanes. The data matches closely for the period starting January, 2012 and continuing through July, 2015. The monthly data becomes unreliable after Hurricanes Irma and Maria. This shows that the daily volume numbers for East End are usable.

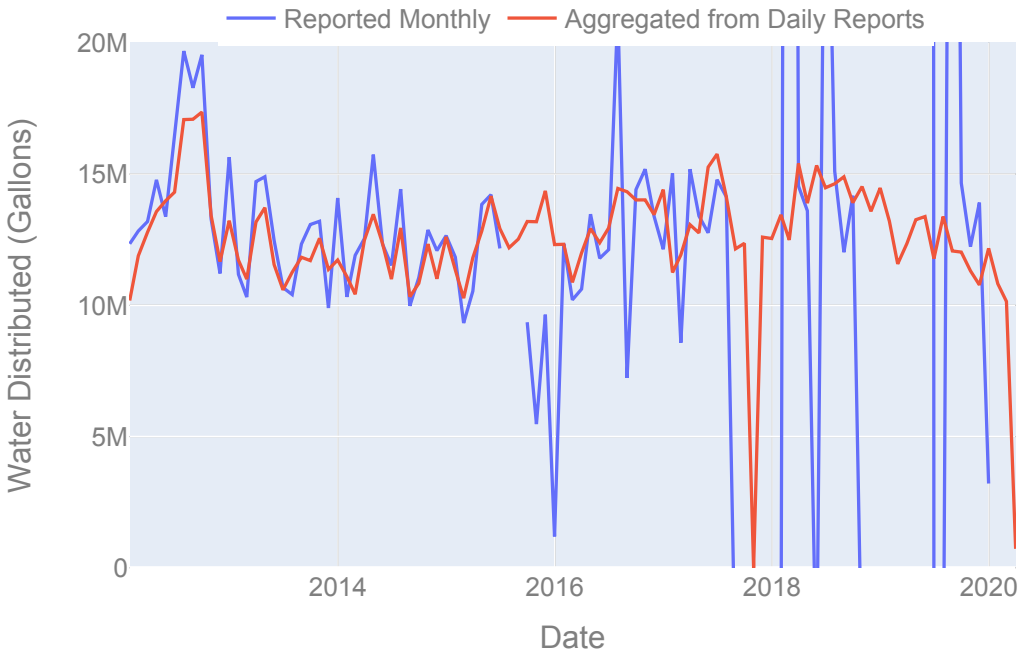

Figure 3.12. STT Monthly Water Volume Delivered to East End January, 2012 to March, 2020. The aggregated East End meter flow data from the monthly reports matches closely with the calculated monthly data prior to the 2017 hurricanes. The monthly post-storm data is unreliable.

Water Delivered to STJ. Figure 3.13 shows the monthly and daily aggregated readings for water transfer to STJ. No calculation was required because there is a meter measuring the submarine transfer of water to STJ. The transfer data matches for the entire period save for two deviations. The monthly data appears to have pulled data from July to August, 2016 and the daily data has a zero reading for October, 2017. This shows that we can use the daily data for our analysis.

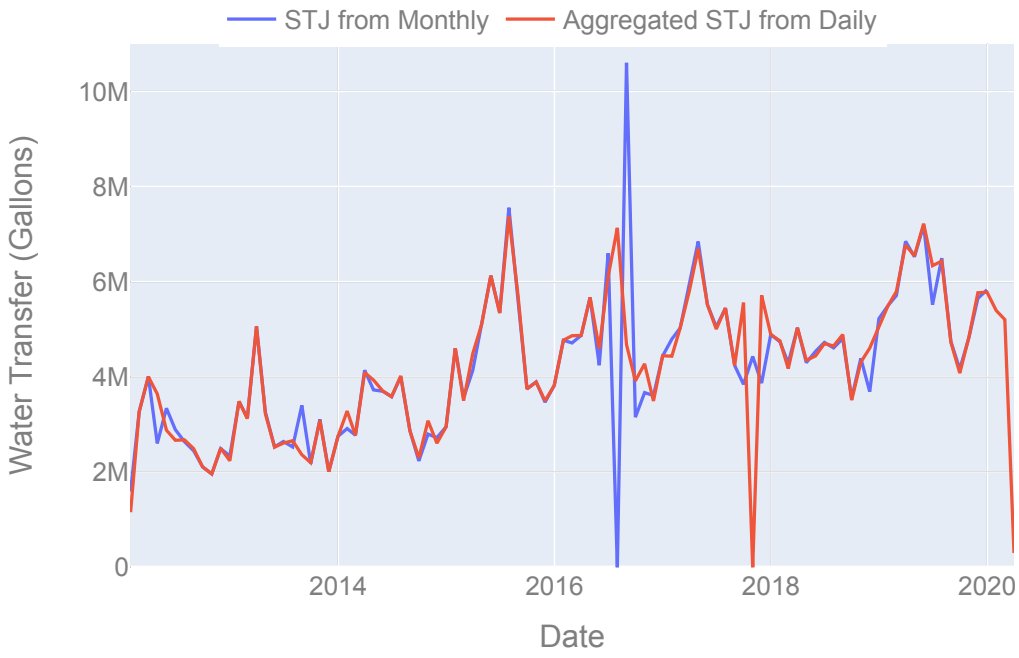

Figure 3.13. STT Monthly STJ Water Transfer January, 2012 to March, 2020. The transfer data matches for the entire period save for two deviations. The monthly data appears to have pulled data from July to August, 2016 and the daily data has a zero reading for October, 2017.

3.4 Understanding Water Demand

To the extent that water demand is driven primarily by per capita use, the relatively stable population of the USVI as reported in Figure 3.4 would suggest that water demand is regular and predictable. However, reporting by WAPA shows significant fluctuations in water use over days, months, and years. In particular, water delivered by trucks in Figure 3.9 shows extreme spikes that indicate the potential for surges in demand that might not be solely attributable to per capita consumption. Understanding these potential relationships in more detail requires additional investigation.

CHAPTER 4:
Analysis

This chapter investigates the statistical relationships between water demands in the USVI and the variables that potentially drive them. It also considers associated forecasting models that may predict flows of water across STT and STJ.

Data Selection. Our previous analysis of daily and monthly data for WAPA water volumes suggests that the most complete and reliable data begins in January 2012. In our analysis here, we focus on data from January 2012 through August 2017 (just prior to the hurricanes in September 2017), yielding more than 2,000 data points.

4.1 Water Usage by Day of Week

We begin by considering weekly patterns in water usage. Figure 4.1 shows the total usage of water in the WAPA system (aggregated across Town, East End, and STJ) by day of the week. From the box plot (Figure 4.1 top), we observe that the mean total usage is very similar on Mondays-Fridays around two million gallons per day, but is noticeably lower on the weekends, Saturday and Sunday. The violin plot of the same data (Figure 4.1 bottom) additionally shows similarities in the statistical distribution of water demands on Mondays, Wednesday, and Fridays. We also observe Tuesdays and Thursdays appear similar to one another, but distinct from the other days of the week.

To the extent that water usage reflects the weekly patterns of work and life, it is reasonable to suspect that the traditional five-day work week could cause differences in water use between the work week and the weekend. However, water usage across STT and STJ is not expected to be uniform. We therefore consider the patterns for water use across different parts of the system.

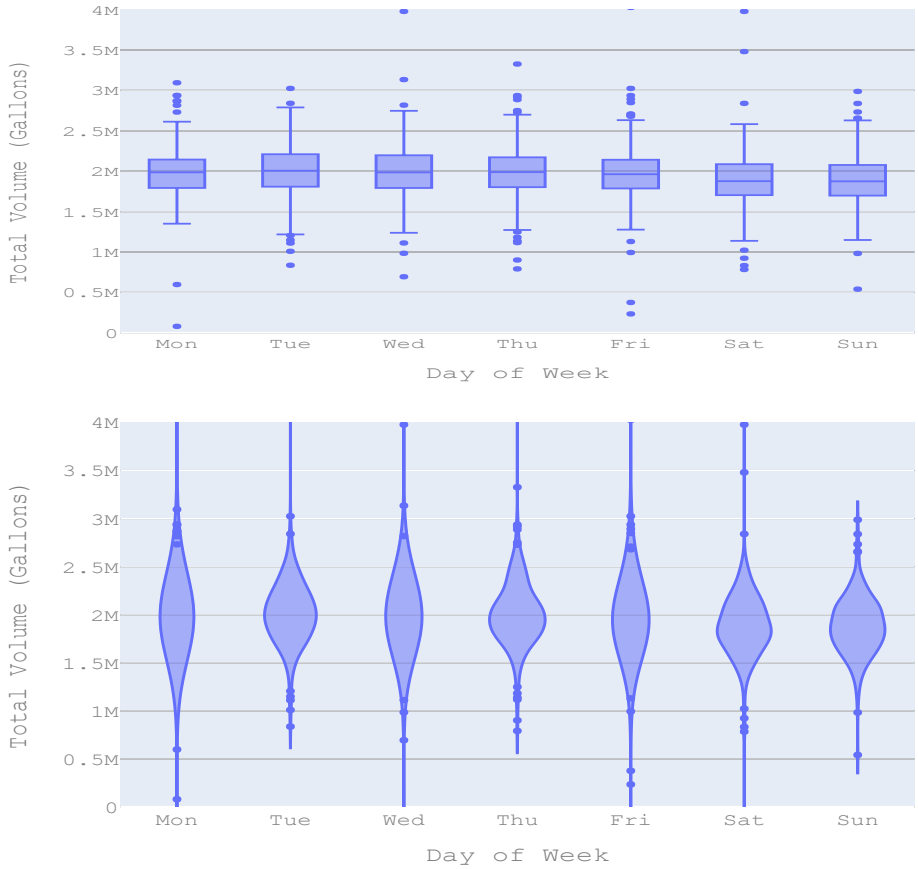

Figure 4.1. WAPA Total STT Water Volume by Day of Week. Top: The box plot show the mean with upper and lower quartiles. The minimum and maximum are plotted as whiskers with suspected outliers shown as points. Bottom: the violin plot also includes the kernel density estimate illustrating the shape of the statistical distribution. We observe the mean total volume each day of the week is close to two million gallons. But, the mean volume is noticeably lower on the weekend, Saturday and Sunday. Moreover, the weekly pattern on Mondays, Wednesdays, and Fridays, looks different from Tuesdays and Thursdays.

Figure 4.2 shows the total volume of water sent by WAPA to Town by day of the week. The mean is approximately 1.4 MGD and consistent Monday-Thursday, but slightly lower on Friday, and then lower even yet on Saturday and Sunday. The violin plots indicate similar distributions of water use as in Figure 4.1, but with lower overall volumes.

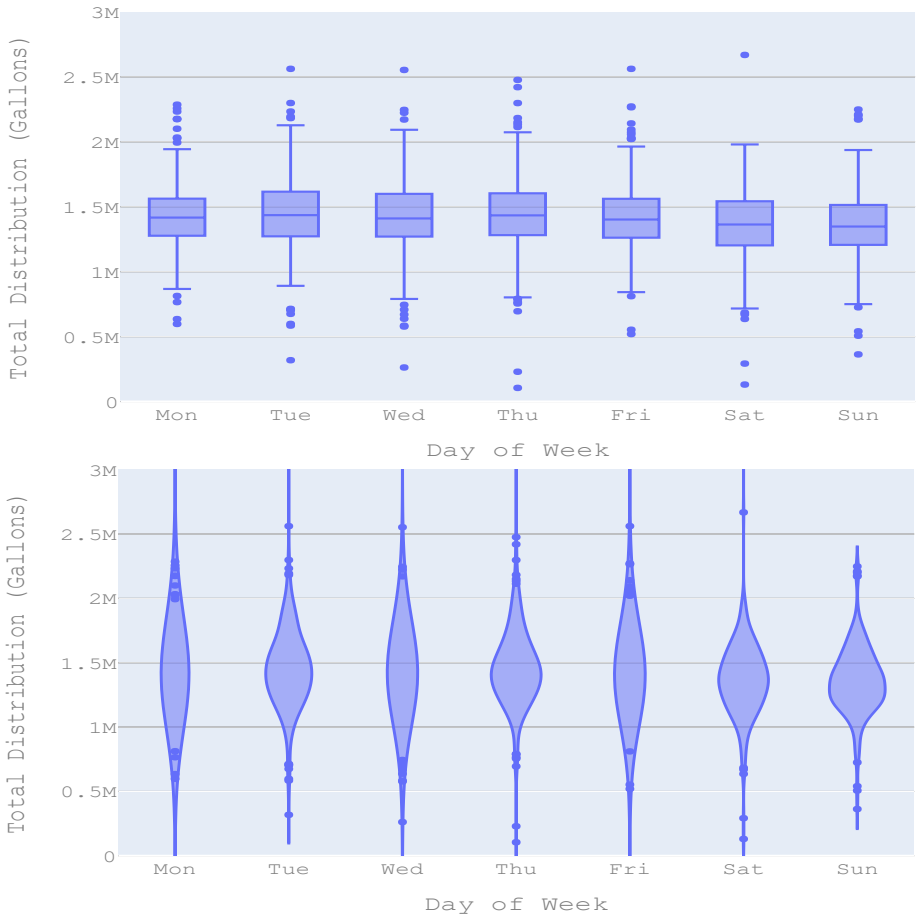

Figure 4.2. Total Water Volume Sent to Town by Day of Week. Top: The box plot shows the mean with upper and lower quartiles. The minimum and maximum are plotted as whiskers with suspected outliers shown as points. Bottom: the violin plot also includes the kernel density estimate illustrating the shape of the statistical distribution. The box plot shows that mean volume delivered by each day of the week is very close to 1.5 MGD gallons. But, this volume is noticeably lower on the weekend, Saturday and Sunday, and slightly lower on Friday. The violin plot shows similarity in water usage on Monday, Wednesday, and Friday. Tuesday and Thursday, and Saturday and Sunday also appear similar.

Figure 4.3 shows that the statistical distribution of water volume delivered daily to the East End of STT follows a similar pattern to that for all of STT and Town. The mean is just over 0.4 MGD and decreases on the weekend.

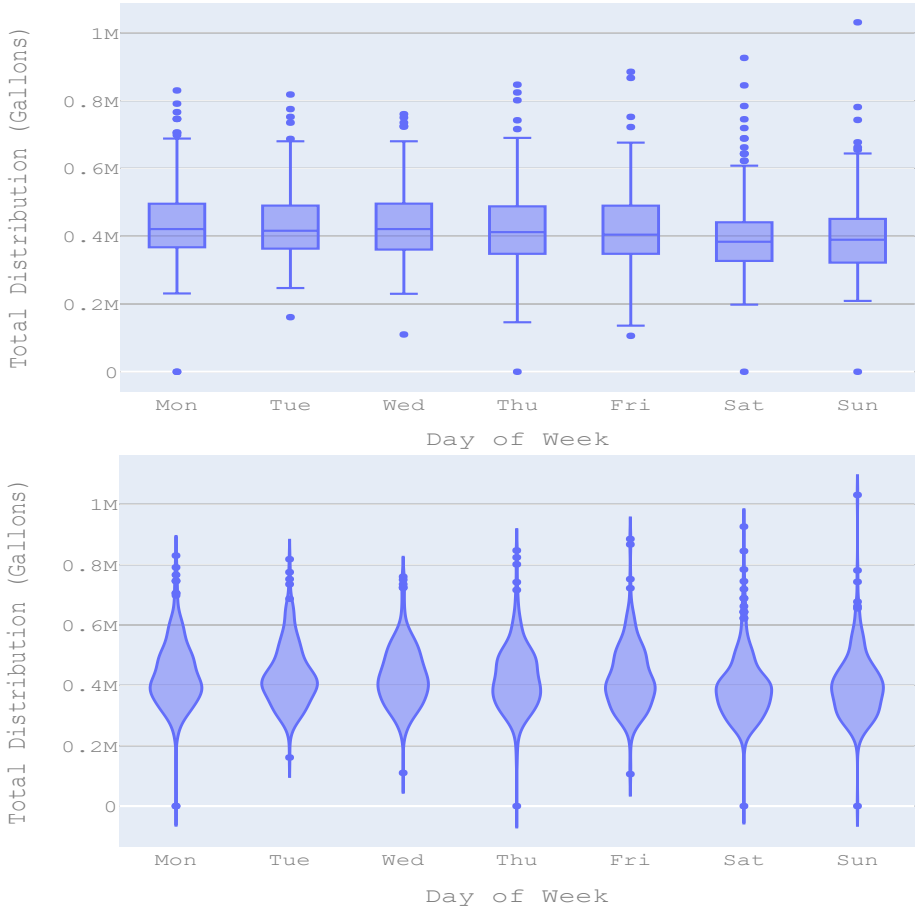

Figure 4.3. Total Water Volume Sent to East End by Day of Week. Top: The box plot show the mean with upper and lower quartiles. The minimum and maximum are plotted as whiskers with suspected outliers shown as points. Bottom: the violin plot also includes the kernel density estimate illustrating the shape of the statistical distribution. The box plot shows that mean water volume sent to East End each day of the week is very close to 1.5 MGD gallons. But, the mean water usage is noticeably lower on Saturday and Sunday. The violin plot shows similarity in the statistical distribution of water usage Monday, Wednesday, and Friday. Again Tuesday and Thursday are similar, but less distinct from Monday, Wednesday, and Friday. Saturday and Sunday also appear similar.

Figure 4.4 shows the statistical distribution of water transferred to STJ by day of week. The mean is approximately 120,00 gallons per day, but we do not see the drop on the weekend that we see with the STT delivery numbers.

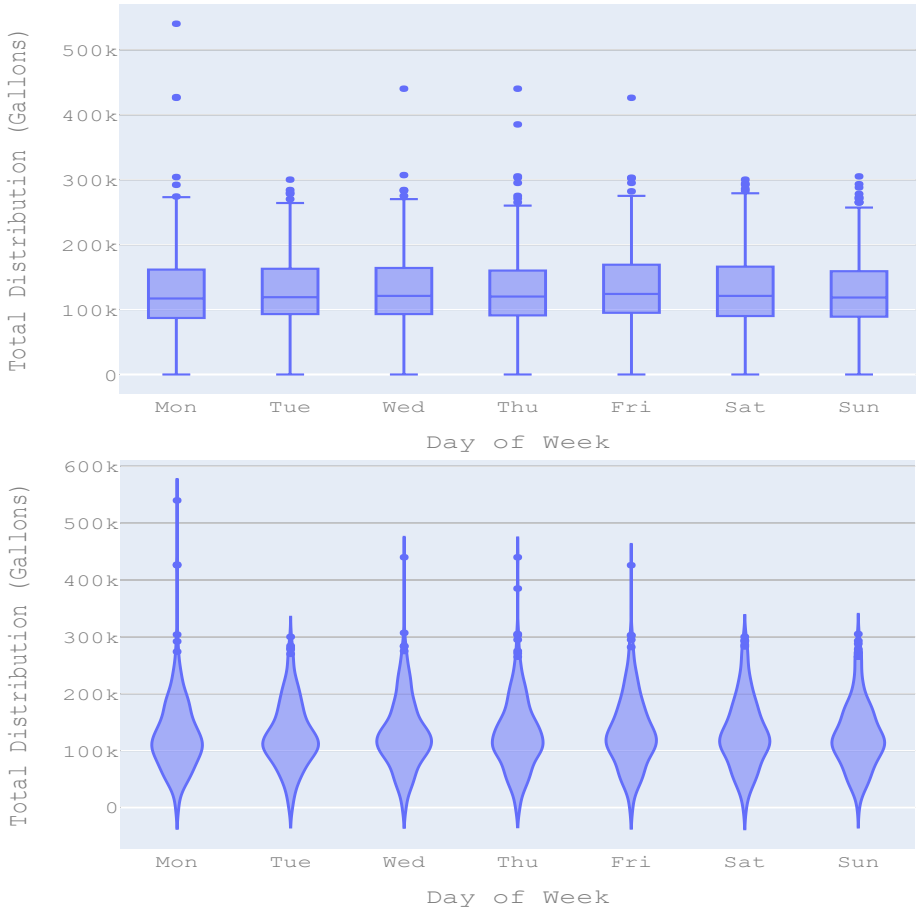

Figure 4.4. Water Volume Transferred from STT to STJ by Day of Week. Top: The box plot show the mean with upper and lower quartiles. The minimum and maximum are plotted as whiskers with suspected outliers shown as points. Bottom: the violin plot also includes the kernel density estimate illustrating the shape of the statistical distribution. The box plot shows a mean around 0.1 MGD of water transferred to STJ daily with a slight decrease on Saturday and Sunday. Unlike the distributions of total water for STT, the violin plot shows similarity Monday through Friday. Saturday and Sunday are also similar and less distinct from Monday through Friday.

Figure 4.5 shows the statistical distribution of daily water sales to trucks by day of week. The mean is low for all days of the week and larger purchases appear spread across the week. It appears that there may be fewer large purchases made on Mondays than any other day of the week.

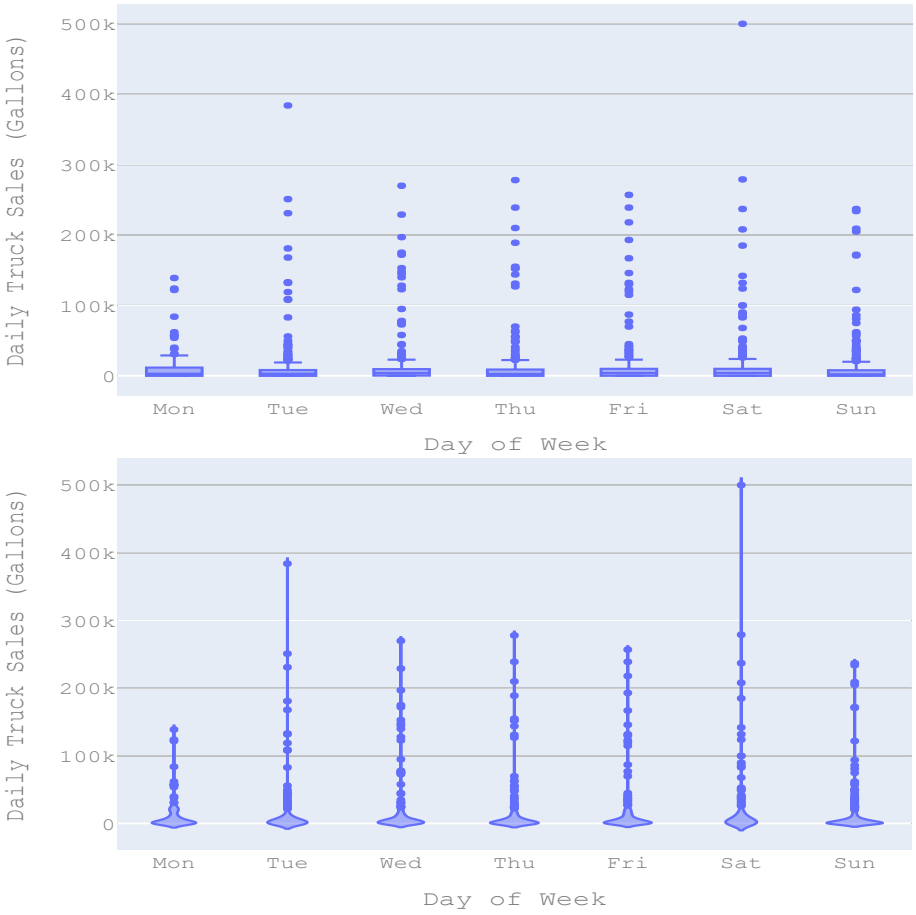

Figure 4.5. WAPA Water Sales to Trucks on STT by Day of Week. Top: The box plot show the mean with upper and lower quartiles. The minimum and maximum are plotted as whiskers with suspected outliers shown as points. Bottom: the violin plot also includes the kernel density estimate illustrating the shape of the statistical distribution. The box plot shows mean daily truck sales total is consistent and low across the week. Although high purchase days are spread through the week, Monday does not have as many large purchase days as the rest of the week. The violin plot shows that Saturday's truck sales looks distinct from the rest of the week.

The daily patterns for water demand in WAPA's pipeline system, reflected in Figures 4.1-4.4 suggest the potential for a regression model that explains water use by day of week. Focusing on water distributed to East End (where the data appears the most "clean"), we make two attempts. First, we use categorical variables to represent individual days of the week (Monday-Sunday), and we perform linear regression analysis on daily water usage values. In the resulting regression, the coefficients for Monday through Friday were not significant, so we made a second attempt in which we re-categorized water use in terms of either weekday (Monday-Friday) or weekend (Saturday-Sunday). Specifically, we use the regression equation

$$Y_{Usage} = \beta_0 + \beta_{Weekend} X_{Weekend} + \varepsilon \tag{4.1}$$

where $X_{Weekend}$ is a categorical variable taking on value 0 for a weekday and 1 for a weekend. Table 4.1 shows the results for this regression. We interpret the intercept as the mean usage on a weekday. We observe with 95% confidence that the weekend mean usage is 11,700 to 37,700 gallons per day lower than weekday usage. This regression has an R^2 value of only 0.007. This is small, but not unexpected for a single categorical explanation of daily water demand. While it does not capture the whole picture, it does help us understand the difference between weekdays and weekends.

Table 4.1. STT East End Weekend Categorical Regression. Categorical regression shows with a 95% confidence interval that the weekend usage mean is between 11,700 and 37,700 gallons lower than it is during the week.

	Coefficient	**Standard Error**	**[0.025**	**0.975]**
Intercept	4.237E5	2509	4.19E5	4.29E5
Weekend	-2.471E4	6640	-3.77E4	-1.17E4

4.2 Water Usage by Month of Year

We next consider the seasonal patterns of water use by month of year. Figure 4.6 shows the daily total water volume delivered to STT for each month. We observe the mean values in water use fluctuate over the year, with the mean lower in August-February than in March-July. This complements the view of monthly rainfall in Figure 3.3, showing that water use

tends to be lower in the months during and immediately after the rainy season. According to the violin plots, the statistical distribution of water usage for January, May, and June appear to be different than those for the rest of the months in the year.

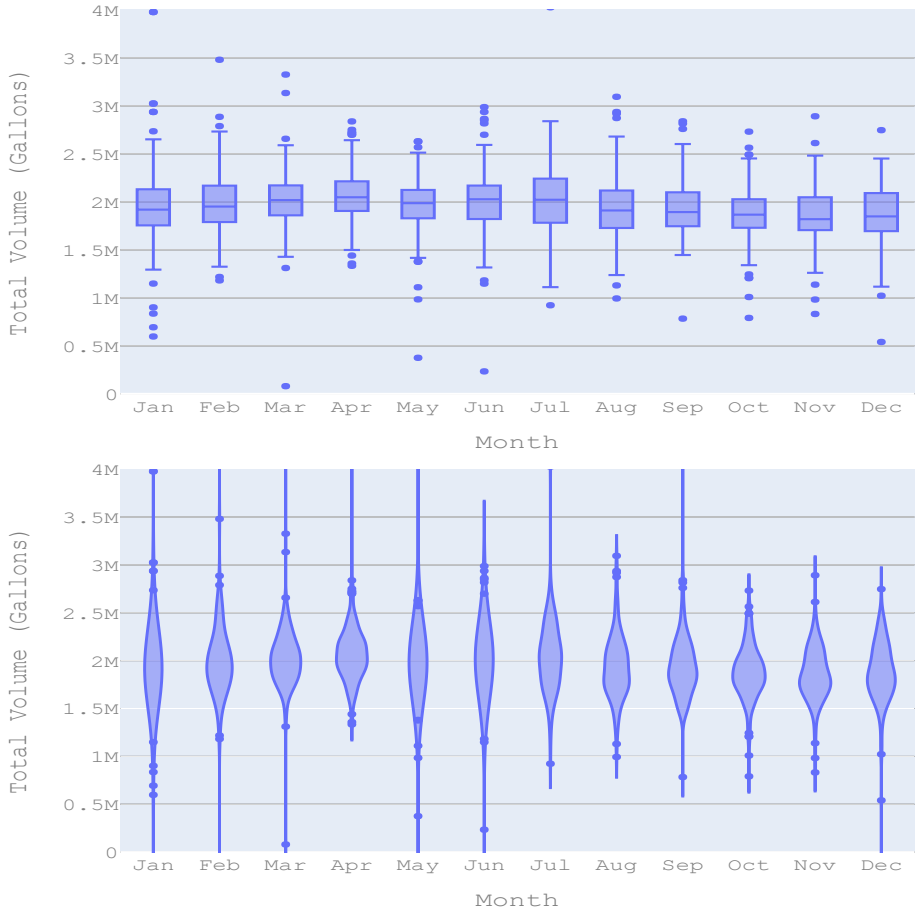

Figure 4.6. WAPA Total STT Water Volume by Month of Year. The box plot (top) shows that the mean total daily volume does not vary much from two MGD through the year. However, the mean appears to be lower August-February. The violin plot (bottom) shows the statistical distribution of daily water volumes in January, May, and June are broader than the statistical distributions for the other months in the year.

The total volume of delivered water for STT is made up of the water delivered to Town and to East End. The box plots for East End and Town show similar patterns. Figure 4.7 shows the statistical distribution of daily volumes to Town by month. The mean is around 1.4 MGD and similar to the total volume for STT the mean is steady throughout the year. The mean is

slightly lower August-December. Figure 4.7 also shows that, like the total volumes for STT, the volumes delivered to Town have a broader distribution in January, May, and June.

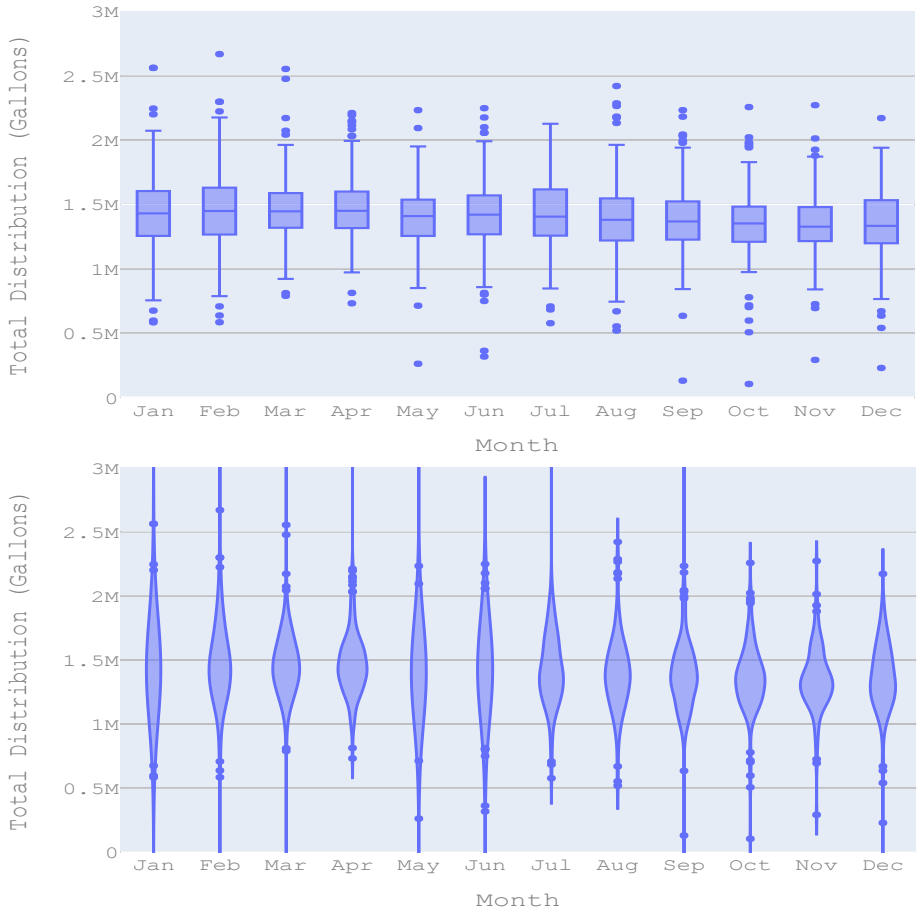

Figure 4.7. Total Water Volume Sent to Town by Month. The box plot (top) shows that mean daily total volume is very close to 1.4 MGD for each month. But, the mean water usage is slightly lower August-December. Similar to total volumes for STT, the violin plot (bottom) shows a broader statistical distribution in January, May and June.

Figure 4.8 shows the daily volume of water delivered to the East End of STT by month. Volumes of water delivered to the East End have greater monthly variation in the first and third quartile values than the volumes of water for all of STT. As with the monthly volumes for all of STT, it appears that the daily means are lower August-February. But, there is a rise in daily volumes in September for the East End. Moreover, the volumes of water delivered

to East End have broader statistical distribution in the warmer months, May-September, differing from those for both total STT and Town water usage.

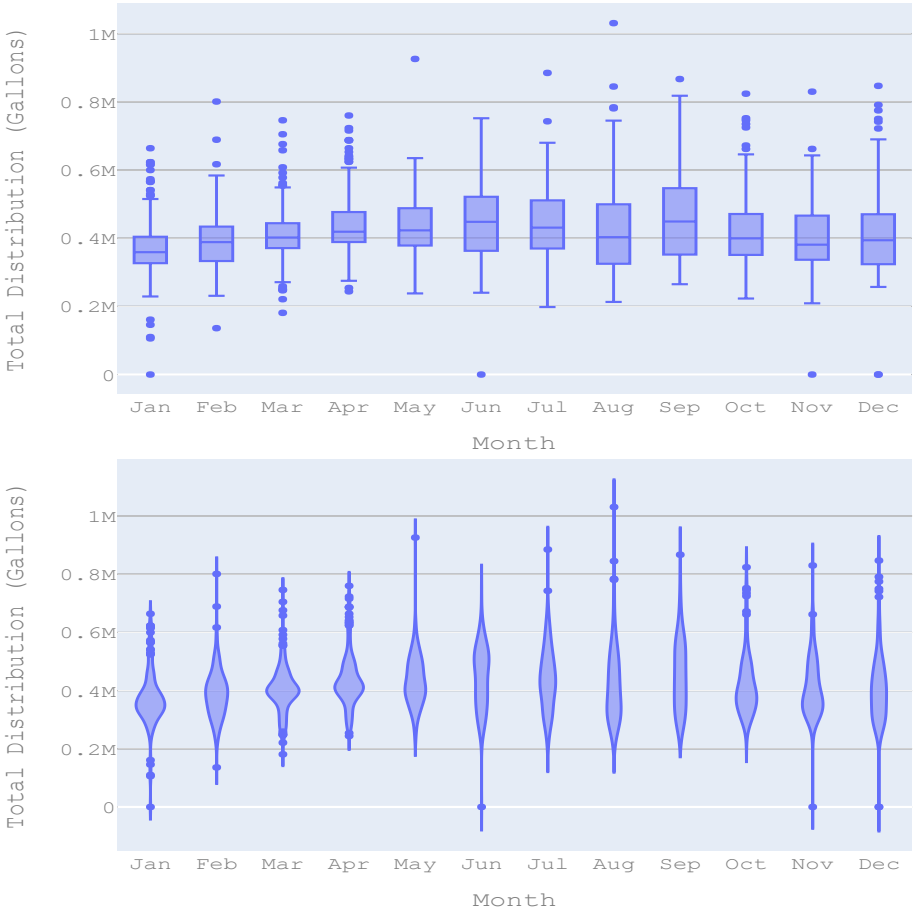

Figure 4.8. Total Water Volume Sent to East End by Month of Year. The box plot (top) shows that mean daily volume in each month is in the range of 0.4 MGD. Similar to total water for STT the mean appears slightly lower August-February. But, there is greater spread in the first and third quartile values than there is in the volumes for all of STT. The violin plot (bottom) shows broader statistical distribution in volumes during May-September.

The lower means are more distinct in daily water transfer to STJ than they are for all of STT. Figure 4.9 shows that there is a similar pattern of lower usage figures from August-February. The violin plot shows similar distributions throughout the year. But, there are fewer high usage days at the end of the year and statistical distributions of daily use skew higher mid-year and in January.

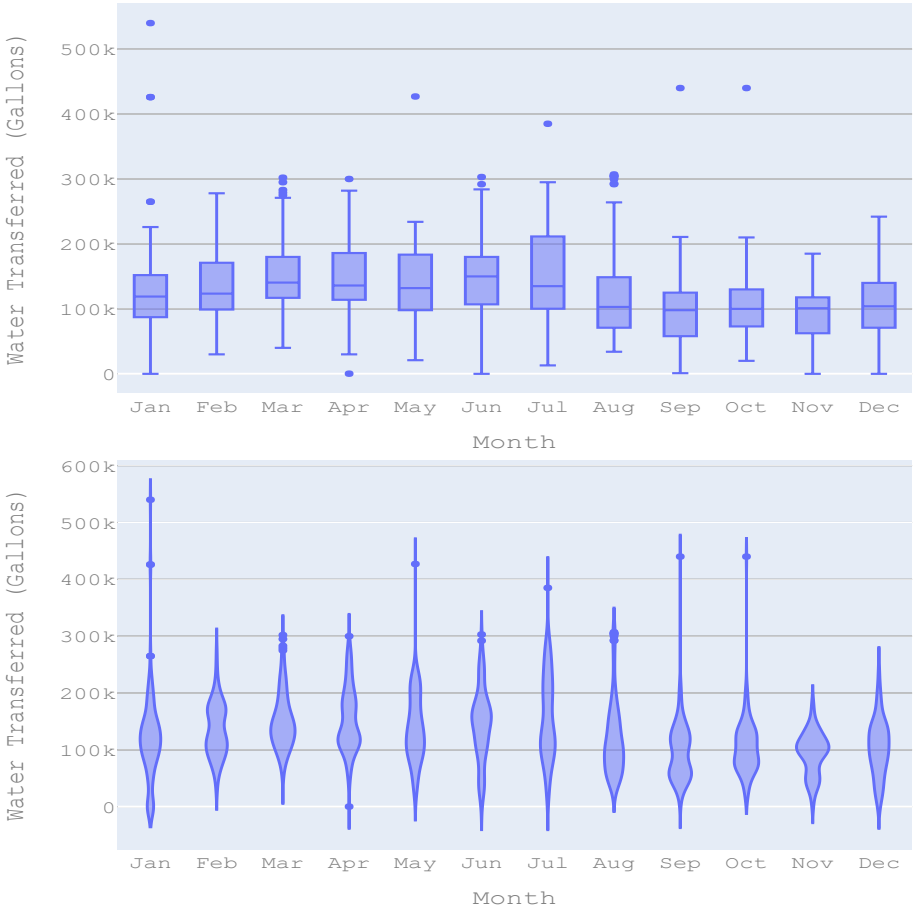

Figure 4.9. Water Volume Transferred from STT to STJ by Month of Year. The box plot (top) shows that mean daily transfer from STT to STJ each month is in the range of 0.4 MGD. Similar to pattern observed for total water use in STT the mean appears slightly lower August-February. But, there is greater spread in the first and third quartile values than there is in the total usage. The violin plot (bottom) shows that statistical distributions are similar throughout the year. But, there are fewer high usage days at the end of the year and statistical distributions skew higher mid-year and in January.

There is variation in the truck sales data that we do not see in the data for water volumes delivered by pipe. Figure 4.10 shows that there is a low mean for daily sales across all months. But, July holds the most sales significantly above the mean. The violin plot is dominated by infrequent large sales.

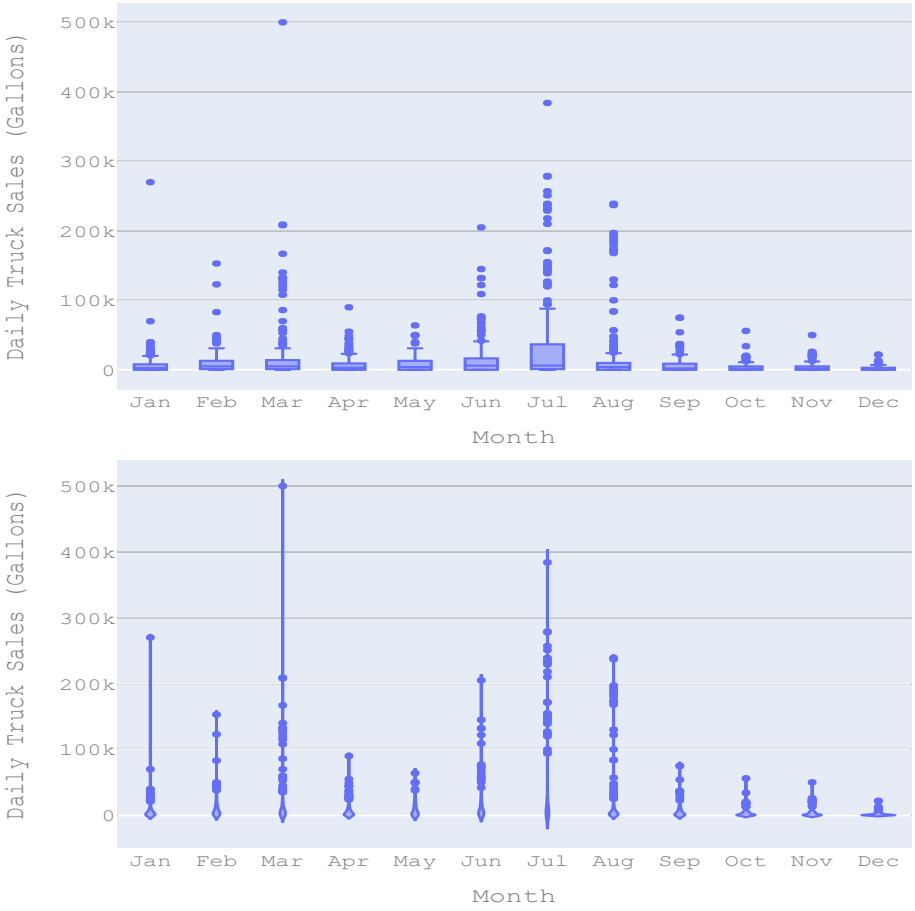

Figure 4.10. WAPA Water Sales to Truck on STT by Month of Year. The box plot (top) shows that mean daily truck sales are very low for all months. But, we do see many of the very high sales days occurred in July. The violin plot (bottom) does not show a clear distribution as it is dominated by infrequent large sales days.

The monthly patterns for water demand in WAPA's pipeline system, reflected in Figures 4.6-4.9 suggest the potential for a regression model that explains water use by month. Focusing on water delivered to East End (where the data appears the most "clean"). We use categorical variables to represent individual months (January-December), and we perform linear regression analysis on daily water usage values. Specifically, we use the regression equation

$$Y_{Usage} = \beta_0 + \Sigma_{i=2}^{12} \beta_i X_i + \varepsilon \tag{4.2}$$

where each X_i is a categorical variable taking on value 0 for and 1 for the month it represents (February (2)-December (12)). Table 4.2 shows the results for this regression. We interpret the intercept as the mean usage in January and each other entry is the difference between January and the respective month's mean. We observe that September has the highest mean daily water volume delivered for the East End of STT. The regression has an R^2 value of 0.045. This is low, but not unexpected for a categorical regression for daily usage because it has large variation.

Table 4.2. STT East End Month Categorical Regression. Categorical Regression by month with January represented by the intercept. Coefficient by month represents the difference in mean from January for each month. We observe that September has the largest mean daily water volume used.

	Coefficient	**Standard Error**	**[0.025**	**0.975]**
Intercept	3.742E5	7701	3.59E5	3.89E5
February	1.600E4	1.11E4	-5697	3.77E4
March	3.348E4	1.08E4	1.23E4	5.47E4
April	6.190E4	1.06E4	4.12E4	8.26E4
May	5.984E4	1.12E4	3.78E4	8.19E4
June	6.892E4	1.10E4	4.73E4	9.05E4
July	6.66E4	1.08E4	4.54E4	8.78E4
August	4.862E4	1.09E4	2.73E4	6.99E4
September	8.023E4	1.13E4	5.80E4	1.02E5
October	4.473E4	1.13E4	2.26E4	6.69E4
November	3.441E4	1.15E4	1.18E4	5.70E4
December	3.634E4	1.14E4	1.40E4	5.87E4

4.3 Impact of Rainfall on Water Usage

It is useful to think of rainfall in terms of accumulated precipitation over a previous period. This perspective is common across the US, as rain refills reservoirs and recharges underground aquifers, but is used by communities over longer time frames and the effects on water demand are delayed. Additional factors influence the timing of rainwater use in the USVI because rainfall also refills RWC cisterns on individual households. Here, water can be stored locally for later use, however too much rain over short periods can overflow cisterns and be lost, so excess rain does not always contribute to household use.

Figure 4.11 shows the moving average of precipitation (in inches) on STT over daily, 14-day, 45-day, 60-day, and 90-day periods. We can see rainfall is more variable than temperature. The 90-day rainfall totals vary from three inches to 30 inches. In addition to the overall variation, the year-to-year pattern can be very different. For example, 2009 had no 90-day periods with less than 15 inches of cumulative rainfall, while 2012 had no periods with at least 15 inches of rainfall. The period 2014-2016 was considered an intense drought in the U.S. Caribbean (Holupchinski et al. 2020).

We now consider the extent to which accumulated rainfall impacts water usage on STT. Figure 4.12 shows correlations of different periods of rainfall accumulation with volume of water delivered to all of STT, to Town around Charlotte Amalie, to the East End of the Island, as well as water transfer to STJ.

We observe several relationships in Figure 4.12. First, we observe that aggregated (Town + East End + STJ) volume does not appear correlated with any period of rainfall accumulation. Not surprisingly, this aggregated volume is positively correlated with its individual components, (i.e., volume to Town). However, the correlation is weaker as we get to East End which is beyond Donoe Tank and has more residents using alternative sources of water. The volume to East End shows correlation with the volume to STJ, which again makes sense given that water flows through East End before being transferred to STJ. However, the volume to Town does not appear correlated with the volume to STJ. It is not surprising that water consumption in STJ might be uncorrelated with water consumption in Town.

Figure 4.11. STT Accumulated Precipitation, 01 March 2013 through 30 June 2016. The daily average rainfall over single-day, 14-day, 45-day, 60-day, and 90-day periods shows that there is significant variation in total rainfall from year to year.

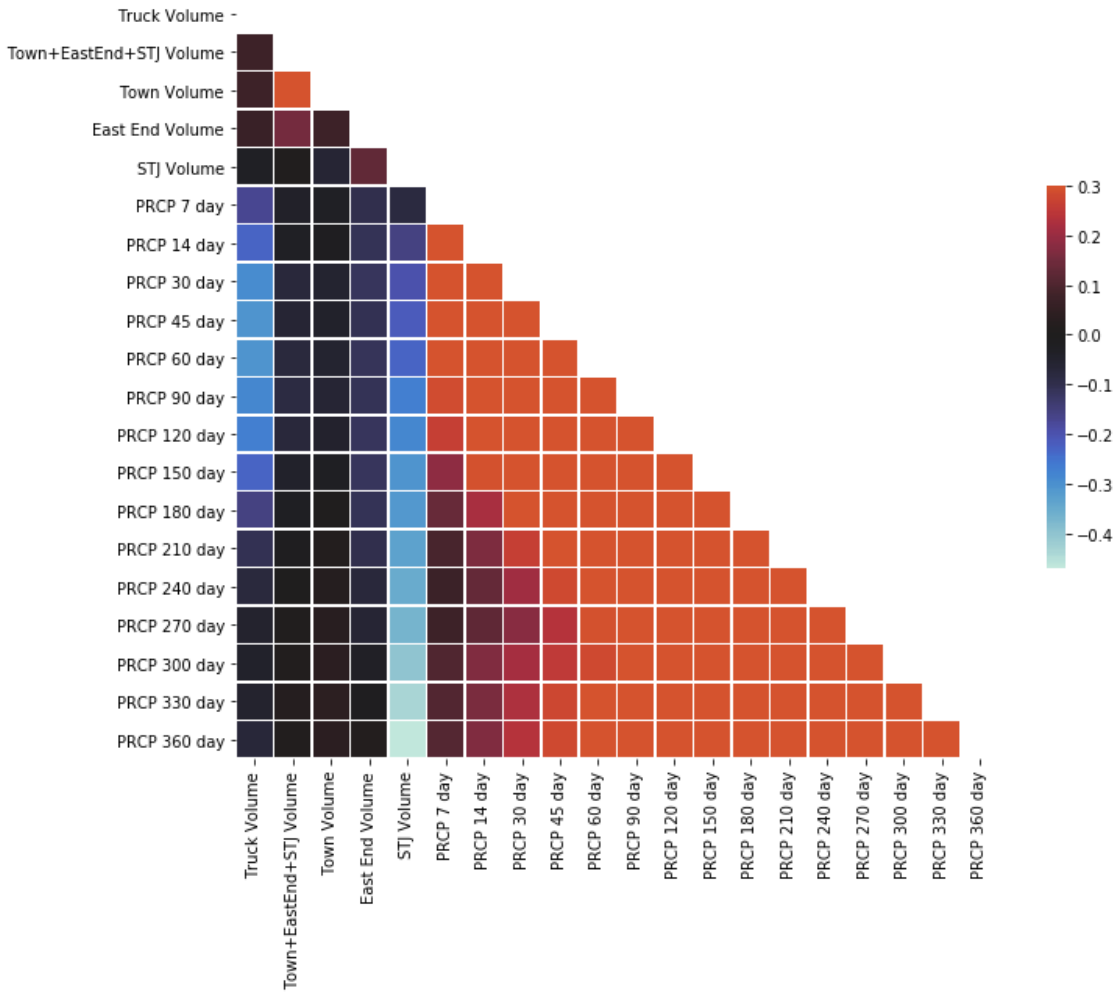

Figure 4.12. Correlation of Rainfall Accumulation with WAPA Water Volumes. Truck volumes are negatively correlated with accumulated volumes less than 180 days, and most strongly at 45-day and 60-day intervals. Volumes to STJ are negatively correlated over time as the length of accumulation period increases. There does not appear to be much correlation between accumulated rainfall and aggregated volumes.

There are a number of places where we observe correlation between water usage and accumulated rainfall. We observe that volume to East End shows some negative correlation with accumulated rainfall over periods ranging up to 150 days. More prominently, we observe a negative correlation between volume to STJ and accumulated rainfall that increases with duration. In both cases, negative correlations for water usage and precipitation accumulation

are reasonable because we expect that as more RWC occurs, less water is needed from the public system.

Perhaps most interesting is the negative correlation observed between accumulated rainfall and Truck volumes. We observe the strongest negative correlation with 45-day and 60-day accumulated rainfall. We interpret this as as evidence that when the accumulated rainfall is low, there are more water truck sales as households need to replenish the "water battery" in their cisterns.

The correlation between water volume flow and rainfall accumulation suggests the potential for a regression model that explains water use by the amount of rainfall accumulation. Truck sales have the strongest correlation, and 45-day accumulation has the strongest negative correlation, so we use that as our explanatory variable. First, we do a linear regression analysis with rainfall as the explanatory variable for daily water truck sales. Specifically we use the regression equation

$$Y_{TruckSales} = \beta_0 + \beta_{45-DayAccumulation} X_{45-DayAccumulation} + \varepsilon \tag{4.3}$$

where $X_{45-DayAccumulation}$ is the total rainfall in the last 45 days measured in inches. Table 4.3 shows the results of this regression. We interpret $X_{45-DayAccumulation}$ as the change in water sales to trucks per inch of rainfall accumulation in the last 45-days. We interpret the result with a 95% confidence interval to mean that truck sales decrease by between 1,577 and 2,375 gallons per inch of rainfall accumulation in the last 45 days. This regression has an R^2 value of 0.054 which is low, but not unexpected with such large variation in daily water demands. This still gives us an idea of how rainfall affects truck sales.

Table 4.3. STT Water Sales to Water Trucks Rainfall Regression.

	Coefficient	**Standard Error**	**[0.025**	**0.975]**
Intercept	2.16E4	1370	1.89E4	2.43E4
45-Day Accumulation	-1976	203	-2375	-1577

Next, we consider a partition tree to see if we can get more explanatory power out of setting rainfall thresholds for categorization. Figure 4.13 shows the resulting partition tree with

an R^2 value of 0.22. This is better than our previous value of 0.054 when we used linear regression analysis.

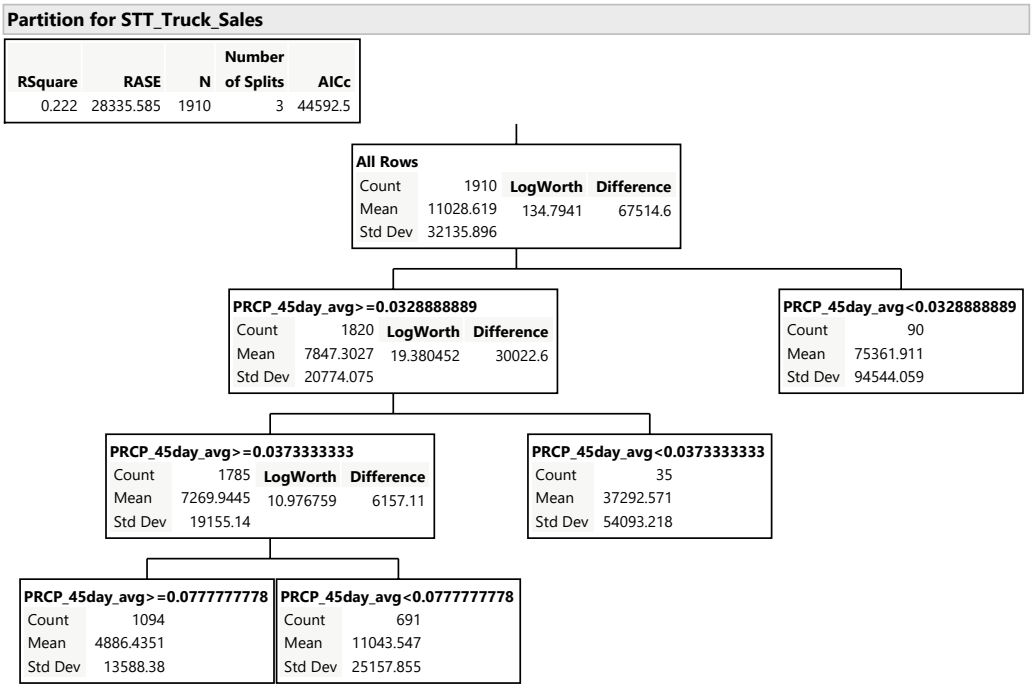

Figure 4.13. WAPA Truck Sales Rainfall Partition Tree. Partition Tree showing the role of 45-day average rainfall on WAPA Truck Sales. The tree shows that volumes to water trucks decrease as accumulated rainfall increases. This is likely a result of increase RWC reducing the need for water trucks.

Table 4.4 summarizes the results of the partition tree. Although there is high standard deviation, we observe that truck sales decrease as there is more rainfall. The large standard deviations are due to large daily fluctuation in truck sales. The results of the partitions concur with the linear regression, demand goes down as rainfall increases. Table 4.4 also shows the equivalent number of water truck deliveries per day based on a 4,000 gallon water truck.

Table 4.4. Summary of Partition Tree Results for Truck Volumes. The partition tree shows that water truck sales increase as average daily rainfall over the last 45 days decreases. The last column shows the equivalent number of 4000 gallon water trucks filled daily to deliver the mean volume.

Average Rainfall Over Previous 45-Days (inches)	Mean Daily Truck Sales (gallons)	Standard Deviation (gallons)	Equivalent Number of Water Trucks
0 to 0.0329	75,360	94,540	18.8
0.0329 to 0.0373	37,292	54,090	9.3
0.0373 to 0.0778	11,043	25,160	2.8
0.0778 and Greater	4,886	13,590	1.2

4.4 Impact of Temperature on Water Usage

Figure 4.14 shows correlation between water volumes and average daily mean maximum temperatures over different periods. Again, we consider separately water volumes to Town around Charlotte Amalie, to the East End of the STT, to STJ, and their aggregate, as well as water Truck volumes. We observe the strongest positive correlation between water volumes to STJ and high temperatures, with correlation increasingly positive with longer periods of high temperatures. Somewhat surprisingly, we observe negative correlation between high temperature and water volumes to East End (somewhat weaker) and also water Truck volumes (somewhat stronger), particularly for higher temperatures over 60-240 day periods. One possible explanation for this is that lower temperatures tend to occur during the winter months, which also happens to be the dry season with less rainfall, so perhaps we are observing the inherent relationship between temperature and rainfall.

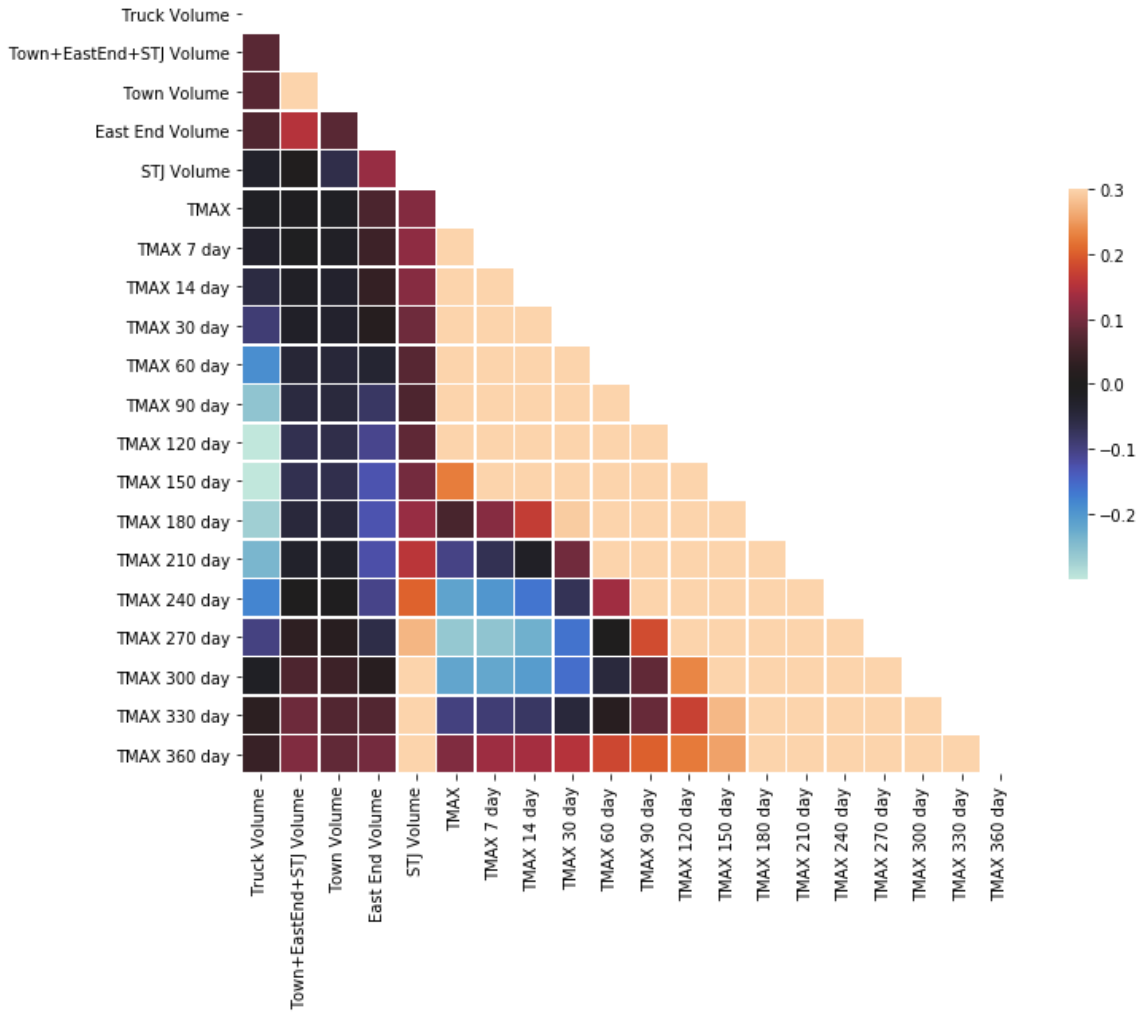

Figure 4.14. WAPA Delivery and Truck Sales Correlation with Daily Mean Maximum Temperature. Truck sales are negatively correlated with the mean maximum daily temperature with a 60-240 day window. Water transfer to STJ is positively correlated with increasing strength for increasing window size of mean maximum daily temperature.

The correlation plot for temperature does not suggest a physically meaningful interpretation to be used for regression analysis. The regression for STT daily truck sales based on 150-day moving average temperature has an R^2 of 0.004.

4.5 Time Series Characteristics of Water Demand Pre- and Post-Storm

The previous analysis suggests water demand data is correlated to accumulated weather, but it is unclear if these factors are sufficient to explain the patterns of water use. Here, we consider the potential for time series analysis to reveal, and perhaps even predict, patterns in water demand. Based on daily and monthly data and correlation with weather, we have two good candidates for time series modeling: water volume transfer to STJ and water volume sales to trucks. Here, we present a model of water transfer to STJ.

We investigate the use of the standard regression models ARIMA and SARIMA for time series analysis and forecasting. Specifically, we determine if there are significant autoregressive (lagging) factors due to seasonality and/or moving average trends that would not be captured by a linear regression model for STJ water demand. We fit the best model to STJ data and use the model to predict water demand pre- and post-hurricanes Irma and Maria. Together, this approach quantifies differences in water demand before and after the storm.

Appendix A.2 provides a review of the ARIMA model in relation to simpler linear regression models. Generally, the model is defined by parameters ARIMA(p, d, q), where d is the number of times the series is differenced to induce stationarity, and where p and q coefficients determine the number of lag terms used to make predictions. Parameter p is the number of previous terms used in an autoregressive manner, and q is the number of previous prediction errors used as part of the next terms fit.

Working with data for monthly water volumes to STJ, we difference the series once ($d = 1$) to make it stationary. Figure 4.16 shows the partial autocorrelation (left) and autocorrelation (right) plots for monthly transfer to STJ. We use partial autocorrelation to determine p, the autoregressive parameter. The partial autocorrelation plot suggests that we use a value $p > 0$ and the data has significant lags at 5, 8, 10, and 11. We use the autocorrelation plot to suggest q, the moving average parameter. The plot suggests that we use a value $q > 0$ with a significant lag at 12. The combination of the plots both suggest parameters greater than intended in ARIMA modeling (e.g., $p + q >> 10$). The strong lags near and around 12 further suggest that the series would be better modeled with an annual differencing.

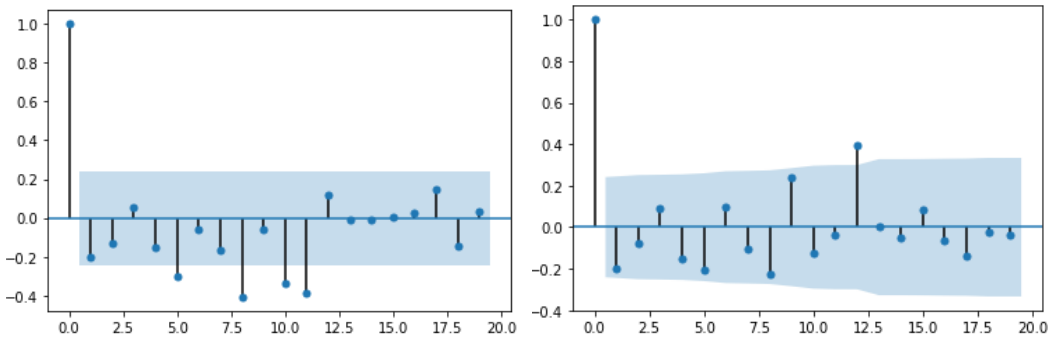

Figure 4.15. Single Differenced Correlation Plot for ARIMA Parameter Selection. The partial autocorrelation (left) and autocorrelation (right) plots are shown for water transfer volume to STJ by month with a single difference ($d = 1$) . While both plots appear stationary, the resulting plots indicate that an ARIMA model without seasonal variation may not be appropriate. The partial autocorrelation plot suggests the autoregressive term should be much greater than zero ($p >> 0$) due to positive initial correlation and high-order lags at 5, 8, 10, and 11. Moreover, the autocorrelation plot suggests a moving average term greater than zero ($q > 0$). The combination of long lag significance in both plots suggests that the data may have seasonal variation.

To capture annual differencing, we use SARIMA. Appendix A.3 provides a review of the SARIMA model in relation to simpler linear regression models and ARIMA models. Generally, the model is an extension of ARIMA and is defined by parameters SARIMA$(p, d, q)(P, D, Q, m)$, where (p, d, q) are the same parameters used in ARIMA and (P, D, Q, m) are hyperparameters for seasonal regression. Here, m is the seasonal period relating SARIMA hyperparameters to ARIMA parameters, D is the number of times the time series is differenced seasonally, P and Q are seasonal autoregressive and moving average parameters. For example, with our monthly data, we use $m = 12$ to capture autoregressive, differencing, and moving average effects on an annual cycle, such that $D > 0$ would apply differencing on y_i with observation y_{i-12}. P and Q would incorporate similar coefficients for the data, y_{i-12}, and errors, ε_{i-12}, respectively.

Working with data for monthly water volumes to STJ, we seasonally difference ($D = 1$) the series once to make it stationary. Figure 4.16 shows the partial autocorrelation (left) and autocorrelation (right) plots for monthly transfer to STJ with annual differencing ($m = 12$). Both plots appear stationary and indicate a standard SARIMA model should fit the data. In par-

ticular, positive initial autocorrelation and positive partial autocorrelation suggests that the autoregressive term should be greater than zero ($p > 0$), but no greater than 2. The additional peaks at low lag values in the partial autocorrelation imply the seasonal moving average term should also be greater than zero ($Q > 0$) with possible moving average terms greater than zero ($q \geq 0$). Again, these will be more reasonable values of 1 or 2. Based on these plots, possible SARIMA models that will fit the data include: SARIMA$(p,0,q)(0,1,Q,12)$. Best practices in time series forecasting using a standard SARIMA parameterization with the following parameters SARIMA$(p,0,0)(0,1,1,12)$ (Nau 2014). This parameterization avoids overdifferencing the data (inducing bad stationarity) and including additional lag factors that overfit the model and may embed unrealistic effects (e.g., too many moving average terms can product inappropriate long-term trends).

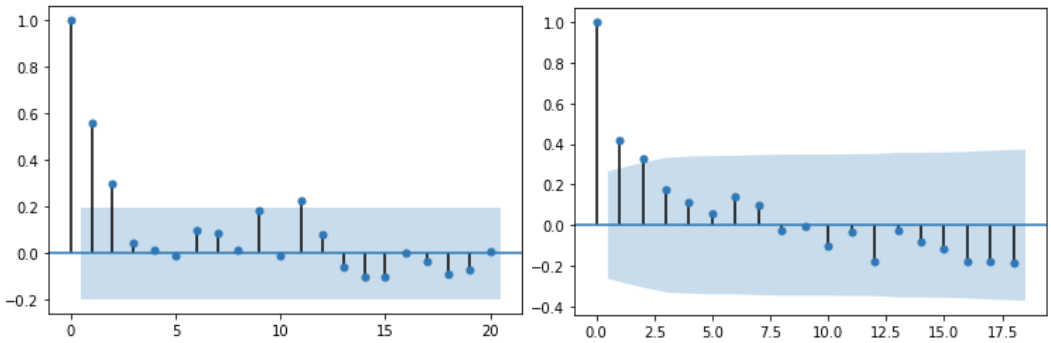

Figure 4.16. Annually Differenced Correlation Plot for SARIMA Parameter Selection. The partial autocorrelation (left) and autocorrelation (right) plots are shown for water transfer volume to STJ by month with a single annual difference for monthly data ($d = 0, D = 1, m = 12$). Here, both plots are appear stationary and indicate a standard SARIMA model should fit the data. In particular, the combination of positive initial autocorrelation with slow sign change over lags combined with a positive initial partial autocorrelation suggests that the autoregressive term should be greater than zero ($p > 0$). The additional peaks at low lag values in the partial autocorrelation imply the seasonal moving average term should also be greater than zero ($Q > 0$) with possible moving average terms greater than zero ($q \geq 0$). Best practices based on these plots suggest fitting the data with a SARIMA model with the following parameters SARIMA$(p,0,0)(0,1,1,12)$.

Using SARIMA$(1,0,0)(0,1,1,12)$, we fit the model on pre-storm data from January, 2012 to June, 2016. We then made predictions out to July, 2017. The result is shown in Figure

4.17. We observe a good fit that captures the seasonality of the data once there are greater than 12 observations (the annual seasonal period). With respect to the predicted data, the first 6 months of forecast data track very closely with the actual monthly flows. After the initial prediction period. the data continues with the seasonal cycle captured by the model.

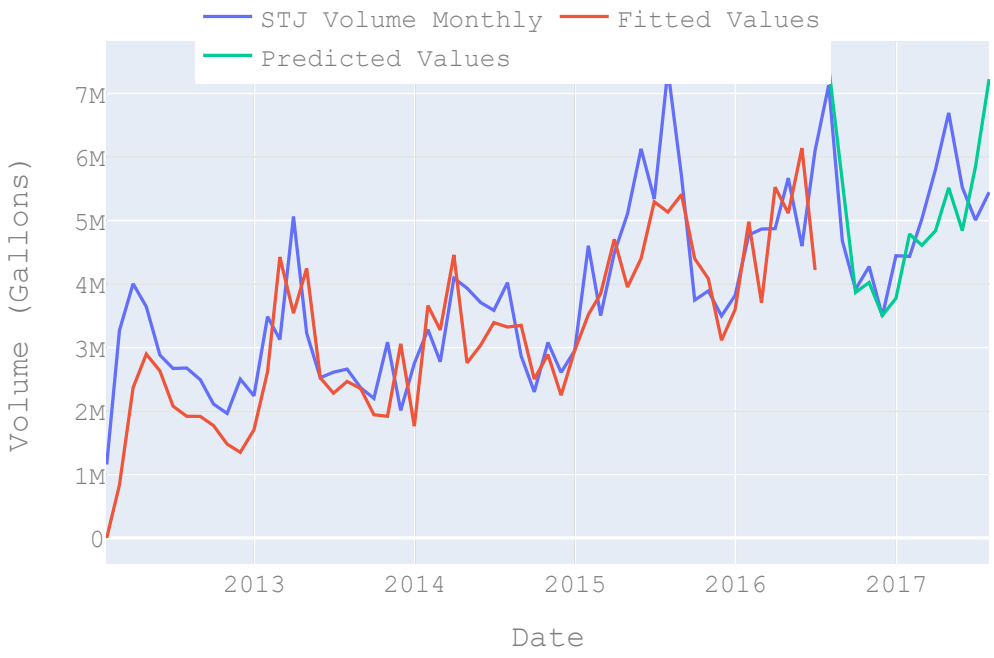

Figure 4.17. SARIMA(1,0,0)(0,1,1,12) STJ Pre-Storm. The fitted values fit better once there are more than 12 observations (the seasonal period) fit. The model captures the seasonality of the data and as a result, the predictions follow the true values.

We are satisfied with the fit of the model, so we refit with more of the pre-storm data and make predictions following the September, 2017 hurricanes to compare pre- and post-storm flows. The result is plotted in Figure 4.18. We again observe that the model provides a good fit for the training data once there are enough observations to cover the seasonal period of 12 months. We use the trained model to forecast all data starting from October 1, 2017 to capture post-storm system demand. We observe that the model initially under predicts water

flow to STJ immediately after the storm, followed by an over prediction of demand during the first half of 2018. Then, the model under predicts demand once the system returns to normal operational capacity (late-2018).

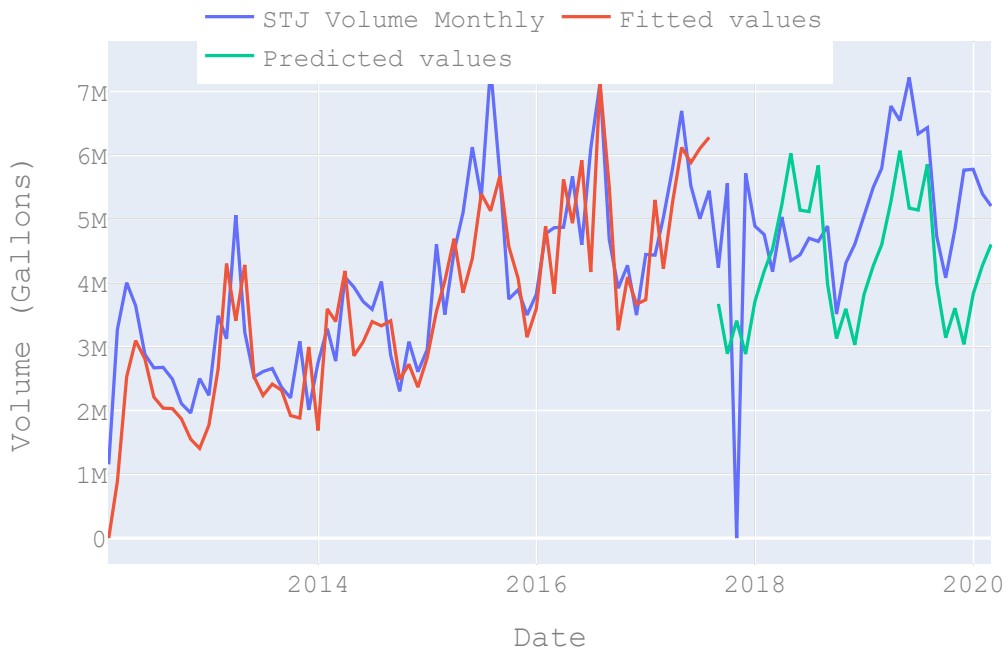

Figure 4.18. SARIMA(1,0,0)(0,1,1,12) STJ Post-Storm. The fitted values fit better once there are more than 12 observations (the seasonal period) fit. The model captures the seasonality of the pre-storm data. But, there are changes to the system do to the September, 2017 hurricanes. This causes the model to over predict immediately post-storm and under predict when the system returns to normal operations.

We compare pre- and post-storm fits to quantify possible changes in STJ flows. Table 4.5 presents the Mean Absolute Deviation (MAD) between the water flow readings to STJ and the fitted and predicted values for pre- and post-storm SARIMA models. We observe that the MAD of the pre-storm model decreases between training (January, 2012 to June, 2016) and testing (June, 2016 to July, 2017). This means the pre-storm forecast outperformed

the model fit on historical data. In contrast, we observe that the MAD of the post-storm model increases between training (January, 2012 to July, 2017) and testing (September, 2017 to February, 2020). This means the post-storm forecast under performed the model fit to historical data. Moreover, the MAD for the post-storm forecast for flows to STJ is nearly double the MAD pre-storm.

Taken together, this indicates that something may have fundamentally changed in the STJ portion of the WAPA system post-storm. The SARIMA$(1,0,0)(0,1,1,12)$ model pre-storm both fit and predicted real flows with high accuracy. However, the same model made poor predictions post-storm. This indicates the factors that dictate water demand and resulting flows to STJ may now be different than before the storm. This could be due to increased demand, demand shifting to a new season, pipe leaks leading to lost water flows, or other changes in behavior not quantifiable with historical data.

Table 4.5. Mean Absolute Deviation of Pre- and Post-Storm STJ Predictions with SARIMA(1,0,0)(0,1,1,12). The MAD with pre-storm data improves between training (fitting) and testing (predicting). In contrast, the MAD with post-storm data doubles between training and testing. This is potentially due to fundamental changes in demand following the storm, such as pipe leaks and changes in water use.

Time Frame	**Train Set Mean Absolute Deviation (Gallons)**	**Test Set Mean Absolute Deviation (Gallons)**
Pre-Storm	678,000	658,000
Post-Storm	651,000	1,252,000

4.6 Implications WAPA and Territorial Water Planning

The results identify the driving characteristics of water demand and provide valuable information to guide operations and disaster planning of the WAPA water delivery pipeline and related USVI water delivery systems. Each result relates to different time-horizons for which water demand must be predicted and planned for.

4.6.1 Weekly Water Use

Estimating water demand on a daily and weekly time-horizon is critical for the day-to-day operation of the WAPA system. In general, customer water demands fluctuate throughout a day and across a week with regularity — there are usually days of high water demand and days with reduced water demand. Understanding days of high and low water requirements help system operators optimize pump and valve schedules to reduce costs and maintain infrastructure.

Prior to this analysis, it was unclear if the USVI water demands would show regular weekly patterns. The widespread use of cisterns and WAPA water truck sales could displace or alter any regularity in daily use. Instead, we find regularity in daily water demand that can help guide day-to-day pipeline operations. In particular, there are at least three distinct groupings of daily demand patterns that WAPA should account for in operations:

- **Monday, Wednesday, and Friday** exhibit a broad distribution of daily usage with mean water demand. The longer tails shown in the data for these days implies that they are more likely to have higher than anticipated usage. WAPA should plan for filling tanks and moving water between tanks on these days, as total water requirements are reduced. Moreover, operators should be prepared for possible spikes in demand and high water use.
- **Tuesday and Thursday** exhibit a narrower distribution of daily usage with a higher mean demand than Monday, Wednesday, and Friday. Shorter tails mean these days are less likely to have higher than anticipated usage. However, the overall water usage on these days is expected to be higher on average. WAPA should plan for distributing water and prioritizing short maintenance jobs on these days. While the demand is higher, the shorter tails indicate a lower probability that demand spikes will happen. This means Tuesdays and Thursdays may be ideal days for having key infrastructure down for maintenance that could be completed within a 24-hr time period.
- **Saturday and Sunday (non-working days)** exhibit a narrower distribution of daily usage with a lower mean than working days during the week. This means weekends require the least water and are more predictable than other days of the week. In general, major maintenance that requires multiple days should be planned for the weekend rather than on weekdays. Moreover, if maintenance cannot be completed in a two-day time period, it is more effective to start later in the week (i.e., Thursday)

rather than on the weekend. That way the maintenance schedule will avoid the greatest number of days with long tails and possible spikes in water demand.

4.6.2 Seasonal Water Use

Estimating water use across months and seasons is also an important task for planning water delivery. Similar to weekly patterns, water demand tends to have seasonal patterns such as increased water use during warmer months for recreation. The widespread use of RWC in the USVI also means that water demand across STT and STJ may see significant variability across wet and dry seasons, as more customers may require water from the WAPA system or water haulers during droughts. Predicting monthly and seasonal patterns helps with water resource planning for a quarterly and annual cycle. This is necessary to determine future revenue and plan system upgrades that can improve water services and access. Moreover, this is necessary to ensure customers have access to water during unanticipated dry seasons.

Analysis of monthly water demand data identifies important features of seasonal water demands. The most important result for operations is that sub-regions within the WAPA system show different seasonal characteristics. Specifically, Town, East End, and Total water flows in the do not show significant month-to-month or seasonal variation, while Water Truck Sales and Transfer to STJ show strong seasonal variation. This means total water demands across the majority of the WAPA system remain the same over an annual cycle and can be estimated using categorical regression. However, sub-regions within the WAPA system require distinct methods for water planning.

Correlation and partition trees indicate how WAPA should plan for the seasonality of Water Truck Sales. Water Truck Sales seasonal variation are correlated with 45-60 day average rainfall and 120-150 day average maximum temperature. Average rainfall is a particularly strong indicator of water truck demands, and WAPA should prepare for sharp increases in water truck sales after 45 days without rain. Results show that periods with rain correspond to less than two water haulers purchasing water. In contrast, periods without rain show sharp increases demand to an average of 19 water trucks. For these reasons, WAPA should plan water production and delivery to the WAPA standpipe using historical rainfall and cumulative moving averages. 2-month rainfall forecasts from authoritative sources would also benefit operations that need to prepare for future water truck sales.

Time series decomposition and forecasting models are appropriate for estimating Transfer to STJ. Monthly water demand data for STJ shows variation over an annual cycle. Whereas Water Truck Sales shows strong correlation with weather, Transfer to STJ shows an increasing correlation over time. Since the temperature has clear seasonality throughout the year, this result is an indication of an underlying moving average, rather than a relationship to the weather. For this reason, WAPA should consider using time series forecasting with a SARIMA model to estimate STJ water demands. In particular, Transfer to STJ shows clear seasonal relationships between monthly data and annual cycles, with the standard model parameterization of SARIMA$(1,0,0)(0,1,1,12)$ providing a good fit. Fitting a SARIMA model with all available data provides a good model to forecast the next several months of water demands.

4.6.3 Post-Disaster Water Use

Finally, water providers must plan for disasters like Hurricane Irma and Maria. In general, it is difficult to predict the way water demand will change after a major disaster. Water demand may increase immediately after a disaster to support debris and waste management. However, water demand may also decrease due to community displacement when people evacuate or permanently leave a disaster-stricken region. Disasters can also impact the function and structure of water delivery systems by breaking or disabling components, leading to increased water flows to meet the same demands. Moreover, USVI water providers need to account for failures within their system (e.g., pipe breaks) *as well as in customer RWC systems*. Significant losses to cisterns across the USVI may also lead to increased water demands and water truck sales.

Time series analysis of pre- and post-storm water flows to STJ show that methods to estimate water demand during normal operations can lose half their predictive power after a disaster. Immediately after the storms, total water flows were reduced as WAPA water systems damaged in the disasters were offline. Here, statistical models will over predict water flows and may provide an estimate for how much water the community needed but could not receive due to incapacitated systems. In contrast, after the WAPA system recovered and began normal operations, post-storm flows to STJ showed statistically significant increases when compared to pre-storm flows. Here, statistical models fit to historical data will under predict total demand required by communities. Taken together, the combination of over and

under prediction led the MAD of the model to be twice as large after the disaster.

Operationally, this means WAPA should consider forecasting models for weekly and seasonal demands, but prepare for models to be characteristically wrong after a major disaster. We anticipate estimates to be at least 50% less accurate. We also anticipate estimation error will change from over prediction in the immediate aftermath of a disaster to under prediction once systems are functioning as normal.

It is important to note that this analysis is conducted with aggregate water flows through the WAPA system that hide the reasons for poor estimation. In particular, increases in demand after the storms can be caused by many interrelated factors, including unknown pipe breaks and leaks, infrastructure failures of customer RWCs, unaccounted for changes in customer use patterns, and large shifts in water demand seasonality. More analysis is needed to identify the factors that influence post-storm water demand. For this reason, we recommend WAPA to continue to analyze pre-storm data to determine the driving factors before the storm and to continue to collect compare these results to post-storm data.

CHAPTER 5:
Summary and Conclusion

This chapter summarizes the contributions of this thesis and identifies opportunities for future work.

5.1 Summary

The potable water system in the USVI is more complicated than water systems in the mainland US. Understanding "how the system works" requires a combination of data-driven and model-driven methods to determine the relationships between production, storage, delivery, and use. The characteristics that make the USVI potable water system so complicated include the widespread use of residential RWCs and cisterns, the customer choice to pay for water truck delivery instead rely on the WAPA pipeline system to receive water, and the existence of standpipes that move large volumes of water partially by pipe and partially by truck. This is complemented by the fact that the WAPA pipeline delivery system provides only partial coverage of the USVI population. Together, it is unclear of standard modeling approaches used for water delivery systems in the mainland US are appropriate to predict water demands the USVI.

We conduct statistical analyses to inform the efficacy of estimating USVI water demands. Our analysis begins with data curation. We obtain, parse, and validate data from a variety of sources: weather data on rainfall and temperatures, demographic data about population and water use, and data from WAPA about the production, storage, and delivery of water to the population. After processing and validating both daily and monthly data sets, we look to understand the water use patterns in the system.

We find statistical differences in WAPA water demand across days of the week. We observe that the mean water volume for the aggregate system is relatively consistent Monday through Friday, but is lower on the weekend. Although the mean is approximately the same through the five-day work week, the statistical distributions appear split into two groups. Monday, Wednesday, and Friday appear distinct when compared to Tuesday and Thursday. The statistical distributions for Saturday and Sunday appear similar to Monday, Wednesday, and

Friday, but with lower means. This result provides useful information for day-to-day WAPA operations and maintenance scheduling.

Using monthly water data, we also find significant differences in the flows of water within sub-regions of the WAPA system. The majority of flow regions within the WAPA system, including Town (Charlotte Amalie), East End, and Total water flows show minor variation over each year. We observe some differences in the patterns of water use between these regions, where the largest water use is in Town, which dominates the overall usage patterns for the system. Simple statistical models are appropriate to estimate and predict these water flows. We fit linear regression models to the day-of-week and month-of-year water demands. These indicate some quantitative differences, but they fail to capture the wide variability that can occur on any given day.

In contrast, assessing the correlation between monthly data and weather shows strong relationships between Water Truck sales and rainfall. We find that accumulated rainfall has a negative correlation with truck sales and water volume sent to STJ. The negative correlation with water truck sales is strongest around 45 days of rainfall accumulation, and the negative correlation between water volumes for STJ becomes stronger with longer durations of rainfall accumulation. Mean maximum temperature also appears correlated with water truck sales and water volume transferred to STJ. There is a negative correlation with 150-day mean daily maximum temperature. However, changing high temperatures could simply be following seasonal patterns, in which temperatures move with changes in precipitation accumulation.

We find that standard time series forecasting models for water demand are most appropriate for estimating water transfer to STJ. We decompose the time series characteristics of water flows to STJ and identify the standard SARIMA model parameterization of SARIMA$(1,0,0)(0,1,1,12)$ is appropriate for forecasting. This model fit and predicted pre-storm data with high accuracy.

Finally, we identify the potential changes in water demand after a major disaster. We compare the quality of fit for the SARIMA$(1,0,0)(0,1,1,12)$ model pre- and post-storm. We observe that the model fit is half as accurate when predicting water flows after Hurricanes Irma and Maria. We expect similar errors to occur when using statistical models to predict water demands after future disasters.

5.2 Conclusions and Future Work

The monthly and daily WAPA data numbers are largely consistent beginning in 2012. Missing or erroneous values can often be addressed by cross-referencing daily and monthly values, using interpolation, or review of the paper records maintained by WAPA. The use of computational preprocessing tools helps to automate what would otherwise be tedious data curation.

Water usage patterns by day of week suggest that there is higher likelihood of an unusually high demand occurring on Monday, Wednesday, or Friday. A system operator may consider maintenance occurring on Tuesday or Thursday to minimize risk of failing to meet demand or depleting storage reserves more than necessary. The consistent daily patterns for water demand suggests the potential to develop different demand distributions by day of week; these could be used in conjunction with Monte-Carlo simulation techniques to generate synthetic demand scenarios and assess the capability of the system to handle them.

The water usage pattern for month of year is not as clear as for the day of week. Long-term planning can consider August-February to be lower demand months. But planners have to keep in mind occasions where unexpected increases occur, particularly around September. Overall, there is not a consistent monthly seasonal component for long-term demand modeling.

Average daily precipitation over the last 45-days seems to provide a rough estimate of the number of water trucks expected at a stand pipe. A small decrease in average daily rainfall can lead to a large increase in demand for truck water sales. Stand pipe operators can potentially use this to anticipate demand volumes.

Temperature does not seem to play a quantifiable role in influencing water demand at the aggregate level. Future work could examine the effect on smaller scales. The household daily diurnal cycle could be influenced by the progression of temperature throughout a day.

SARIMA models can capture and carry out seasonal cycles in prediction. However, they are limited to only one seasonality. Our model showed that water flow patterns for water sent to STJ change following the September, 2017 hurricanes. The MAD from the modeled value increased by nearly 600,000 gallons per month. This change could be caused by increased demand for water due to changes in behavior, leaking pipes due to hurricane damage, or

some other unquantified change in the system.

Models such as BATS (De Livera et al. 2011) may be considered for capturing multiple seasonal cycles (weekly, monthly, annual). Future work could attempt to identify exogenous variables that improve time-series model of water demands. For example, we have not considered the periodic presence of cruise ships in STT, which can deposit tens of thousands of patrons on STT and STJ on a single day. Cruise ships follow regular schedules, so it is reasonable to think that these patterns could be identified and incorporated into our analysis.

Finally, there is an opportunity to combine the work here with models that consider cistern storage and water use from first principles. That is, there is an opportunity to develop a model that incorporates the size of household cisterns, water usage by household members, and the way in which rainfall and temperature affect the demand for water from WAPA. Understanding these interactions will be important for assessing how systematic disruptions (e.g., drought, earthquakes) affect both the demand for water and the ability of the system to provide it. Integrating such analyses into hazard mitigation planning for the territory is a topic of ongoing work.

APPENDIX: Time Series Models

We provide a brief review of the models we use to understand the relationships between these data sets. Model formulations are adapted from Hyndman and Athanasopoulos (2018).

A.1 Linear Regression

We began our analysis by considering linear regression. This gives us a point of comparison for the performance of our more complex models. Linear regression assumes the outcome (dependent) variable as if it is linearly dependent on one of more causal (independent) variables. A linear regression model with k explanatory variables is written as

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t \tag{A.1}$$

where

y_t = observation at time t
β_0 = a constant, sometimes referred to as the intercept term
β_i = coefficient for the i^{th} explanatory variable
$x_{i,t}$ = explanatory variable i at time t
ε_t = error in prediction for observation at time t

When using linear regression to determine linear models we assume that the errors have mean zero, are not autocorrelated, and are not correlated to the predictor variables. If the errors do not have mean zero the prediction will be biased. If the errors are correlated to themselves or the predictor variables, there could be more information to add to the model or a helpful transformation to apply to the data.

The linear model is fit such that squared error is minimized. This penalizes large prediciton

errors more heavily and is written as

$$\min \sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} \left(y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \cdots - \beta_k x_{k,t} \right)^2 \tag{A.2}$$

where all parameters are defined the same way as Equation A.1 and T is the total number of observations predicted at their respective time steps, t.

A.2 Autoregressive Integrated Moving Average (ARIMA)

Box-Jenkins, or ARIMA models are used to predict time series. They are a combination of three modeling techniques: autoregression, integration or differencing, and moving average. We address each component of the model separately. Then we discuss how they are combined.

The integration, or differencing, must be done first to make the time series stationary. A time series is stationary if its statistical properties do not depend on time. This means that the series will have constant mean and standard deviation. Time series are commonly not stationary. There are a several techniques to make a time series stationary. Differencing is one that is used in Box-Jenkins modeling. Differencing means to calculate the change between sequential observations. Written in equation form, a difference is

$$y_t' = y_t - y_{t-1} \tag{A.3}$$

where

$$\begin{aligned} y_t' &= \text{differenced observation} \\ y_t &= \text{original observation at time t} \\ y_{t-1} &= \text{original observation at time t-1.} \end{aligned}$$

More than one difference may be necessary to make a time series stationary. This is referred

to as the order of differencing in the model.

Next, we consider the autoregressive component of the model. An autoregressive model uses a linear combination of previous observations in a time series to predict future values. Autoregressive models are expressed as an order p representing the number of previous terms used to predict the next term. An autoregressive model of order p is written as

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{A.4}$$

where

$$\begin{aligned} y_t &= \text{observation at time t} \\ c &= \text{a constant} \\ \phi_i &= \text{coefficient for the } i^{th} \text{ previous observation} \\ y_{t-i} &= \text{the } i^{th} \text{ previous observation} \\ p &= \text{the order of the autoregressive model} \\ \varepsilon_t &= \text{error in prediction for observation at time t.} \end{aligned}$$

The constant, c, adds a drift to the model.

Finally, we consider a moving average model. Instead of using past observations or predicted values, a moving average model uses past forecast error to predict the next observation. The moving average model is parameterized by q, representing the number of previous error terms used to predict the next observation. Similar to the autoregressive model, we refer to q as the order of the model. A model of order q is written as

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{A.5}$$

where

y_t = observation at time t

c = a constant

ε_t = error in prediction for observation at time t

ε_{t-i} = error in prediction for the i^{th} observation preceding time t

θ_i = coefficient for the error in prediction for the i^{th} observation preceding time t

q = order of the moving average model.

Now that we have defined the components, we define the ARIMA model. An ARIMA model is parameterized by (p, d, q) where

p = order of autoregressive part

d = order of differencing of time series

q = order of moving average part

then an ARIMA model (p, d, q) is defined as

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{A.6}$$

where

y'_t = differenced observation at time t

c = a constant

ϕ_i = coefficient for the i^{th} previous observation

y'_{t-i} = the i^{th} previous differenced observation

θ_i = coefficient for the error in prediction for the i^{th} observation preceding time t

ε_{t-i} = error in prediction for the i^{th} observation preceding time t

ε_t = error in prediction for observation at time t.

This is truly an Autoregressive Moving Average (ARMA) model where the time series is differenced beforehand to produce a stationary series. If $d = 0$ then the model is literally just an ARMA model.

Similarly, if $q = 0$ and $p > 0$, it is only an autoregressive model. And if $p = 0$ and $q > 0$, then it is only a moving average model.

The coefficients in an ARIMA model are estimated similar to the linear models. But, instead of minimizing squared error, ARIMA modeling uses maximum likelihood estimation (MLE). MLE produces the same estimate as least squares for linear models.

In this thesis ARIMA models are referred to by their parameters in the form: ARIMA(p, d, q).

A.3 Seasonal Autoregressive Integrated Moving Average (SARIMA)

SARIMA models are an extension of ARIMA models that allow for the inclusion of a seasonality. SARIMA models are parameterized by $(p, d, q)(P, D, Q, m)$. The seasonality is defined by parameter m. For example, in this thesis we employ SARIMA with monthly data, so our use of $m = 12$ indicates seasonal cycles.

Similar to ARIMA, the time series is first differenced to induce stationarity. Unlike ARIMA

models, there are two choices for differencing. The first is to set parameter d the same way that it is done for ARIMA modeling. The additional option for SARIMA models is parameter D. Instead of differencing sequential observations, D differences observations separated by the seasonality parameter, m. Seasonal differencing with parameter m are represented as

$$y'_{tm} = y_t - y_{t-m} \tag{A.7}$$

where

$$\begin{aligned} y'_{tm} &= \text{seasonally differenced observation} \\ y_t &= \text{original observation at time t} \\ y_{t-m} &= \text{original observation at time t-m.} \end{aligned}$$

Parameters P and Q are defined in the same way as parameters p and q in the ARIMA model, but they operate on observation y'_{tm} instead of y'_t.

The (1,0,0)(0,1,1,12) SARIMA model used in this thesis is expressed as

$$y'_{tm} = c + \phi_1 y'_{tm-1} + \theta_1 \varepsilon_{t-12} + \varepsilon_t \tag{A.8}$$

where

y'_{tm} = seasonally differenced observation at time t

c = a constant

ϕ_1 = coefficient for the previous seasonally differenced observation

y'_{tm-1} = the previous seasonally differenced observation

θ_1 = coefficient for the error in prediction for the t-12^{th} seasonally differenced observation preceding time t

ε_{t-12} = error in prediction for the t-12^{th} seasonally differenced observation preceding time t

ε_t = error in prediction for observation at time t.

In this thesis SARIMA models are expressed as SARIMA$(p, d, q)(P, D, Q, m)$.

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

Alderson DL, Bunn BB, Eisenberg DA, Howard AR, Nussbaum DA, Templeton II J (2018) Interdependent Infrastructure Resilience in the U.S. Virgin Islands: Preliminary Assessment. Technical Report NPS-OR-18-005, Naval Postgraduate School.

Arouna A, Dabbert S (2010) Determinants of domestic water use by rural households without access to private improved water sources in benin: A seemingly unrelated tobit approach. *Water Resources Management* 24(7):1381–1398.

Brentan BM, Ribeiro LC, Luvizotto Jr E, Mendonça DC, Guidi JM, et al. (2014) Synthetic reconstruction of water demand time series for real time demand forecasting. *Journal of Water Resource and Protection* 6(15):1437.

Bunn B (2018) *An Operational Model of Interdependent Water and Power Distribution Infrastructure Systems*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Buscemi PE (2003) *Integrating Water Resources Management: Analysis of the St. Thomas, US Virgin Islands, Water Market*. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

CDR Maguire (2019) Virgin Islands Water and Power Authority Water Distribution System Master Plan. Technical report, Virgin Islands Water and Power Authority.

Cheslek HA (2003) *Water Reuse and Conservation in the United States Virgin Islands*. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

De Livera AM, Hyndman RJ, Snyder RD (2011) Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American statistical association* 106(496):1513–1527.

DeOreo W, Mayer P, Dziegielewski B, Kiefer J (2016) Residential end uses of water, version 2 (Executive report). Denver, CO: Water Research Foundation.

Do N, Simpson A, Deuerlein J, Piller O (2017) Demand estimation in water distribution systems: solving underdetermined problems using genetic algorithms. *Procedia Engineering* 186:193–201.

Donkor EA, Mazzuchi TA, Soyer R, Alan Roberson J (2014) Urban water demand forecasting: review of methods and models. *Journal of Water Resources Planning and Management* 140(2):146–159.

Firat M, Turan ME, Yurdusev MA (2010) Comparative analysis of neural network techniques for predicting water consumption time series. *Journal of Hydrology* 384(1-2):46–51.

Good J (2019) *An Operational Model of Food Supply Chain and Transportation Systems in the U.S. Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Herrera M, Torgo L, Izquierdo J, Pérez-García R (2010) Predictive models for forecasting hourly urban water demand. *Journal of Hydrology* 387(1-2):141–150.

Holupchinski E, Alvarez-Berrios N, Gould W, Fain J (2020) Drought impacts to crops in the US caribbean. Technical report, US Department of the Interior and US Geological Survey, https://www.usgs.gov/land-resources/climate-adaptation-science-centers/drought-impacts-crops-us-caribbean.

Hyndman RJ, Athanasopoulos G (2018) *Forecasting: principles and practice* (OTexts).

Jain A, Ormsbee LE (2001) A decision support system for drought characterization and management. *Civil Engineering Systems* 18(2):105–140.

Kang D, Lansey K (2009) Real-time demand estimation and confidence limit analysis for water distribution systems. *Journal of Hydraulic Engineering* 135(10):825–837.

Kossieris P, Makropoulos C (2018) Exploring the statistical and distributional properties of residential water demand at fine time scales. *Water* 10(10):1481.

Letting LK, Hamam Y, Abu-Mahfouz AM (2017) Estimation of water demand in water distribution systems using particle swarm optimization. *Water* 9(8):593.

Lucas S, Coombes P, Sharma A (2010) The impact of diurnal water use patterns, demand management and rainwater tanks on water supply network design. *Water Science and Technology: Water Supply* 10(1):69–80.

Machin SJ (2020) xlrd. Https://xlrd.readthedocs.io/en/latest/api.html.

Mayer PW (2016) Water research foundation study documents water conservation potential and more efficiency in households. *Journal-American Water Works Association* 108(10):31–40.

Moeller B (2020) *Synthetic Network Generation and Vulnerability Analysis of Internet Infrastructure Systems in the United States Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Mu X, Whittington D, Briscoe J (1990) Modeling village water demand behavior: a discrete choice approach. *Water Resources Research* 26(4):521–529.

National Centers for Environmental Information and National Hurricane Center (2020) Costliest u.s. tropical cyclonse. Technical report, National Oceanographic and Atmospheric Administration.

Nau R (2014) Lecture notes on forecasting introduction to arima models, nonseasonal, seasonal.

Parsons Brinckerhoff, Jaredian Design Group (2014) 2040 Comprehensive Transportation Master Plan Report. Technical report.

Polebitski AS, Palmer RN (2010) Seasonal residential water demand forecasting for census tracts. *Journal of Water Resources Planning and Management* 136(1):27–36.

Pretorius D, Crouch ML, Jacobs HE (2019) Diurnal water use patterns for low-cost houses with indigent water allocation: a south african case study. *Journal of Water, Sanitation and Hygiene for Development* 9(3):513–521.

Sharvelle S, Dozier A, Arabi M, Reichel B (2017) A geospatially-enabled web tool for urban water demand forecasting and assessment of alternative urban water management strategies. *Environmental Modelling & Software* 97:213–228.

Solomon H (2009) Effects of Rainwater Harvesting and Desalination on Supply and Demand of Public Water in US Virgin Islands. *Proceedings of the 4th IASME/WSEAS Int. Conference on Water Resources, Hydraulics & Hydrology (WHH'09)*.

Taylor HT, Ward B, Willis M, Zaleski W (2010) The Saffir-Simpson Hurricane Wind Scale. *Atmospheric Administration: Washington, DC, USA* .

United States Virgin Islands Department of Planning and Natural Resources Division or Environmental Protection (2020) Public Water System Supervision Program Safe Drinking Water Emergency Operations Plan. Technical report.

US Geological Survey (2018) Drought in the U.S. Caribbean. Climate Adaptation Science Centers Website, https://www.usgs.gov/land-resources/climate-adaptation-science-centers/drought-us-caribbean.

US Virgin Islands Bureau of Economic Research (2014) The United States Virgin Islands 2015 Comprehensive Economic Development Strategy. Technical report, U.S. Virgin Islands Bureau of Economic Research.

US EPA (2020) Drinking Water Distribution Systems. Six-Year Review of Drinking Water Standards Website, https://www.epa.gov/dwsixyearreview/drinking-water-distribution-systems.

USVI Hurricane Recovery and Resilience Task Force (2018) USVI Hurricane Recovery and Resilience Task Force Report 2018. Technical report, USVI Hurricane Recovery and Resilience Task Force.

van der Elst NJ, Hardenbeck JL, Michael AJ (2020) Potential Duration of the Aftershocks of the 2020 Southwestern Puerto Rico Earthquake. Technical report, US Department of the Interior and US Geological Survey.

Vassiljev A, Koor M, Koppel T (2015) Real-time demands and calibration of water distribution systems. *Advances in Engineering Software* 89:108–113.

Wille D (2019) *Simulation Optimization for Operational Resilience of Interdependent Water-Power Systems in the US Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Wine W (2020) *Analyzing Cell Phone Network Resilience in the US Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

World Bank (2020) Population, total. Http://datatopics.worldbank.org/world-development-indicators/themes/people.html.

Xue P, Hong T, Dong B, Mak C (2017) A preliminary investigation of water usage behavior in single-family homes. *Building Simulation*, volume 10, 949–962 (Springer).

Zhang GP, Qi M (2005) Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research* 160(2):501–514.

Zhou S, McMahon T, Walton A, Lewis J (2002) Forecasting operational demand for an urban water supply zone. *Journal of Hydrology* 259(1-4):189–202.

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California