Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 4, 2023
A Regional Conference of the Society for Industrial and Systems Engineering

Smart Water Meter Usage Data Visualization and Prediction

Matthew D'Orlando, Halana Desiervo, Jessica Jurcsak, Jared Pruskowski, Cole Zanni, and Yong Wang

Department of Systems Science and Industrial Engineering
Binghamton University
Binghamton, New York 13902

Corresponding author's Email: jprusko1@binghamton.edu

Author's Note: The project team consists of five student contributing members, listed above. The team would like to acknowledge our great appreciation for other contributors to this project, which include Dr. Yong Wang, and the Jamestown Board of Public Utilities; specifically, Mr. Frank Galeazzo who has helped us with getting the data necessary to conduct work on the project.

Abstract: The mission of this project is to aid the Jamestown Board of Public Utilities (BPU) with their business operations management regarding public water consumption. The goal of this project is to help the BPU with understanding the data they collected from smart meters, and the extent to which they can use the data to improve water quality and management, resulting in an improved standard of living for customers in Jamestown, NY. To accomplish this, the team presents a Power BI dashboard, a user-interactive ArcGIS map, and a predictive model using Python coding software. Suggestions are to be provided to Jamestown BPU based upon analytical solutions obtained using graphical visualization of data, system modeling, and software integration for improved efficiency. Projects like this represent an overall desire for environmental protection, public systems' utilities management, and a greater understanding of the potential of integration of machine learning (ML) and artificial intelligence.

Keywords: Data Analytics, Demand Forecasting, Service Quality, Smart Water Meter

1. Smart Water Meter Data Analytics

1.1 Background

The city of Jamestown, NY has installed smart water meters as part of an environmental protection initiative by Governor Kathy Hochul across New York State (The Governor's Press Office, n.d.). The city needed an updated meter reading system (Board of Public Utilities City of Jamestown, n.d.). The existing process of manually reading each meter in a city not only contributes to more labor costs and CO_2 emissions but has much more room for error (Post Journal, 2020). Misreadings result in inaccurate billing statements, thus monetary losses. By removing the need for a vehicle to drive around each meter, the carbon footprint of the city of Jamestown is decreased with the installation of these smart water meters. Steps like this are necessary on a greater scale, when considering the United States' commitment to the goal of net-zero government emissions (The White House, 2022). These new smart meters have been in use for a significant amount of time now, presenting a large data set containing information regarding public water usage (Board of Public Utilities, 2023). They can read the usage of five distinct types of water: residential (WATER-1), commercial (WATER-2), industrial (WATER-3), fire service (WATER-4), municipal metered (WATER-5), and present a wide range of metrics. Data flows into a central location and is logged into an Excel spreadsheet. This data set is difficult to understand efficiently, however it holds essential information about water usage and supplies insights on the water quality and where errors (leaks) may be occurring.

1.2 Problem Statement

The current problem is that the incoming meter data is difficult to understand and analyze. This project will transform the smart water meter data as presented to us into a dashboard design and interactive map. As part of the analysis, a predictive model will be developed to forecast water usage and make recommendations to the Jamestown BPU regarding billing residents and overall effectiveness of the newly installed smart water meters.

2. Interactive Dashboard/Map

The first step of the project was to visualize the data collected from the smart water meters. The data was received as an Excel spreadsheet but will be presented using a dashboard in Microsoft Power BI. This will also aid in the creation of the predictive model, which also serves as part of the analysis of the original data. Data visualization provides a deeper understanding of the water usage system in Jamestown, as well as an easier interface for the user to obtain information. An ArcGIS map will be presented containing the locations of each smart water meter and the usage at each. This will help the Jamestown BPU to respond to any leaks or consumer demands that require on-site remediation.

2.1 Power BI Visualization

To meet the listed requirements and customer expectations it was decided that Microsoft Power BI was the best software platform to effectively create a dashboard and visualize the provided data from Jamestown BPU. Power BI allows for its internal dashboards to be linked to ArcGIS to present findings in map form. Power BI also supports the creation of maps just using internal tools. It also allows for highly advanced graphics and metrics to be created with numerous dynamic filtering options which can be graphic-wide, page-wide, or file-wide. This level of filtering allows for the quick analysis of trends and scenarios by the end user of the dashboard.

The first deliverable the team developed was an effective water usage map. This map displays water usage by meter across Jamestown with larger circles representing higher water usage over the period filtered by the user:

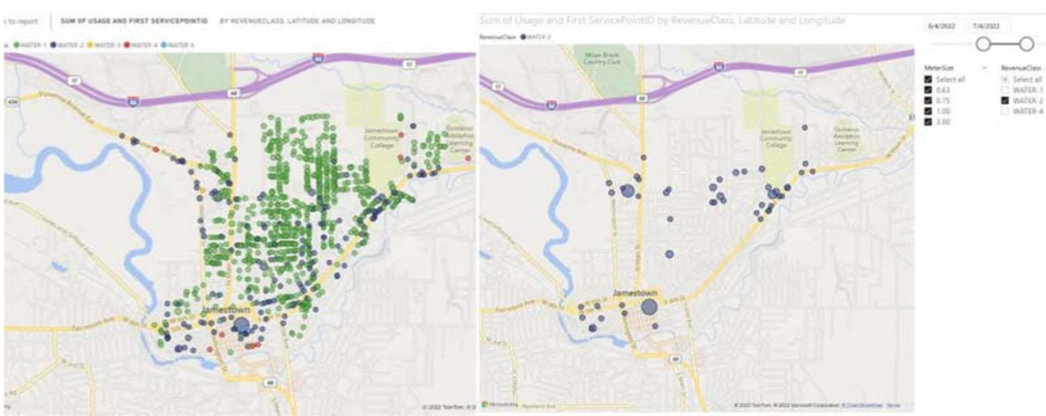

Figure 1: Unfiltered and Filtered Power BI Maps

Each of the meters are located at their exact latitude and longitude. The larger circled locations of the meters represent larger amounts of water usage. Most of the larger circles are shown as WATER-2, which is commercial usage of water. All visualizations can be easily filtered by meter size, revenue class, and date range. These filters are dynamic and globally affect all metrics. This allows for easy, real-time analysis and data exploration.

Next, the team wanted to create a rough dashboard displaying different attributes of water usage. This was then broken down by water type and meter size. This is the current iteration of the main water usage dashboard:

Figure 2: Unfiltered and Filtered Power BI Usage Dashboard

There are multiple useful graphics displayed on the dashboard. In order from top left to bottom right: An interactable filtering section, a tree map of water usage by customer type and meter size, gauge of the total water usage shown compared to the overall total usage, a line graph showing water usage per hour by customer type, and a bar chart showing water usage per month/day by customer type. All the previously mentioned filters are easily applicable to this dashboard as well.

Next the team went about showing the dashboard's capability to compare usage meter to meter. Multiple meters can be selected and displayed in a line graph of usage per hour to get a better understanding of the variation between them. Here is an example of 4 randomly selected meters on this line graph:

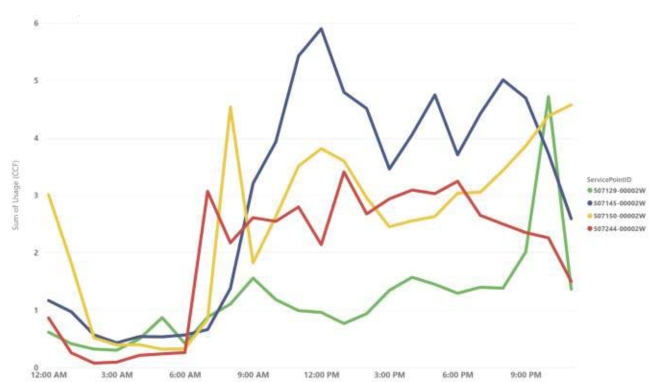

Figure 3: Power BI Hourly Sum Water Usage by Five Random Meters Comparison

The team also had the opportunity to begin working on device error analysis with the meters. Below is a bar chart of the frequency of different device errors. Each bar is then subdivided by the customer type whose meter experienced the error:

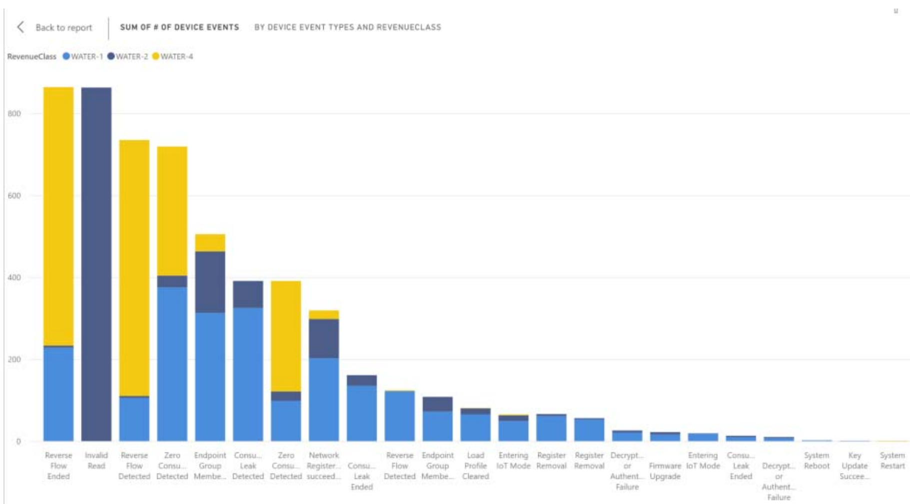

Figure 4: Power BI Device Error Frequency by Customer Type

3. Predictive Use Models

Power BI was used again, this time to create forecasting predictions to allow Jamestown to foresee future demand and better estimate what their system will experience in the coming days, weeks, and months. Specifically, Power BI uses an Exponential Smoothing (ETS) model to accurately forecast (Team, 2018). ETS modeling uses an error term, a trend component, and a seasonal component. For the seasonality of the forecast Power BI automatically selects what it believes to be the best season interval, but this may be manually adjusted (Team, 2018). The two metrics which the team decided to predict were daily average water usage - currently only for residential customers but the team will explore others in the future - and the total sum of daily usage in the entire system. Both these graphs can be seen below with three months forecast. The dark line is the specific predicted amount and the gray area surrounding it is a 95% confidence interval of the ETS model.

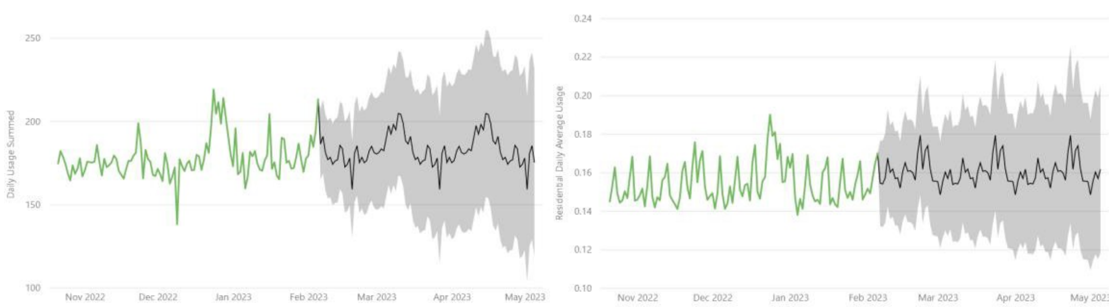

Figure 5: Forecasted Daily Usage Summed and Residential Daily Average Usage

Another point of analysis was to be able to detect leaks within the usage data. Leaks can be either one in which substantial amounts of water are used in short amounts of time, like a pipe that bursts, or one in which lesser amounts of water are used over longer amounts of time, like a drip. To detect leaks of substantial amounts of water, a python code was constructed to detect anomalies within the usage of one water meter. The hourly usage for one meter is and the average water usage was calculated. A threshold of five times the average usage was used to ensure they are truly outliers, and not a day where more water than usual was used. So, if the usage on that day was above the threshold, it was detected as an anomaly. The dates of the anomalies were then displayed, as seen in Figure 6. Figure 6 also displays a graph of the

ISBN: 97819384964-4-8

Usage, showing the average and thresholds values as horizontal lines. It was easy to visualize where the anomalies occur within the data.

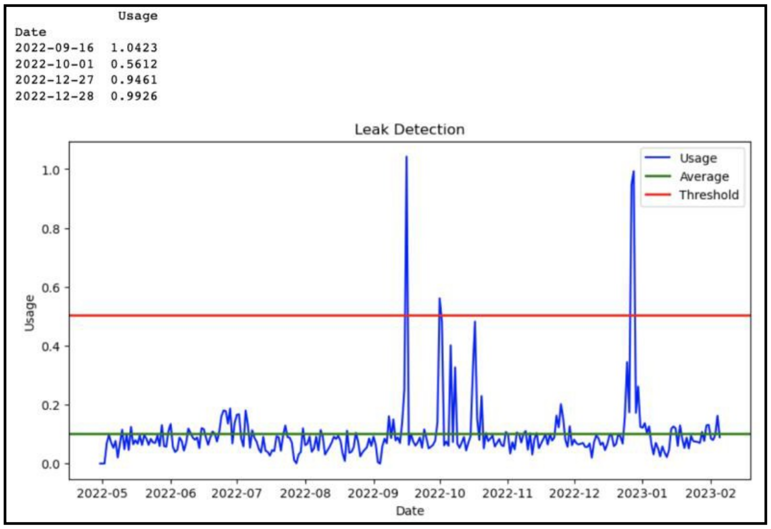

Figure 6: Large Leak Identification

For the smaller leaks, python code was written to detect when water was used for longer than 24 hours. A threshold of 0 CCF of water is used to identify for how long it was over 0 CCF. A period of 24 was defined so it can detect when usage was over 0 that period. If Usage was greater than 0 for the previous 24 hours, the column 0'leak_detected' in the data frame will have a value of 1. When no leak is detected, it has a value of 0. Figure 7 shows when the leaks occur using a red line, and when there is no leak detected using a green line.

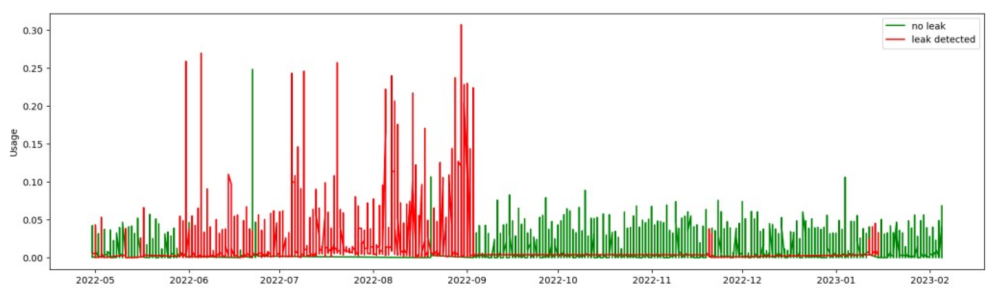

Figure 7: Small Leak Identification

4. Conclusion

Through data analysis including visualization, interactive mapping, predictive modeling, and detection of potential leaks, the team was able to study water consumption and determine forecasts of future consumption. The purpose of this project is to aid the Jamestown Board of Public Utilities in deciding how much water will need to be provided to consumers and aid in providing the basis for accurate water billing. The impacts of this project include environmental benefits from the decrease of carbon dioxide emissions from manual water readings, enhanced data and reporting, more accurate water billing, and elimination of waste from the system (excess supply of water). The final deliverables consist of an interactive dashboard including visualizations and device error analysis in Power BI and interactive maps in ArcGIS, predictive models on water consumption in Power BI, and potential leak detection of small and large leaks using Python.

In the future, the predictive model could potentially be integrated with weather data to provide more accurate readings, as well as with water conserving systems such as sprinkler systems and water efficient appliances to further conserve water waste.

5. References

Board of Public Utilities City of Jamestown, New York. (n.d.). Retrieved February 9, 2023, from https://www.jamestownbpu.com/DocumentCenter/View/116/Rates--Provisions-Applicable-To- All-Water-Consumers-PDF

Board of Public Utilities Proceedings by authority state of New York ... Jamestown BPU. (n.d.). Retrieved February 9, 2023, from https://www.jamestownbpu.com/AgendaCenter/ViewFile/Minutes/_07262021-128

Davidiseminger. (n.d.). *Tutorial: Build a machine learning model in power bi - power bi.* Power BI | Microsoft Learn. Retrieved March 28, 2023, from https://learn.microsoft.com/en-us/power- bi/connect-data/service-tutorial-build-machine-learning-model

Post Journal. (2020, July 18). *City BPU receives State grant. post*. Retrieved November 28, 2022, from https://www.post-journal.com/news/community/2020/07/city-bpu%E2%80%88receives-state-grant/

Team, T. P. B. I. (2018, December 10). *Describing the forecasting models in Power View*. Blog do Microsoft Power BI | Microsoft Power BI. Retrieved March 28, 2023, from https://powerbi.microsoft.com/pt- br/blog/describing-the-forecasting-models-in-power-view/

The Governor's Press Office. (n.d.). *Governor Hochul announces FY 2023 investments in clean energy infrastructure, climate resiliency and preservation*. Governor Kathy Hochul. Retrieved February 9, 2023, from https://www.governor.ny.gov/news/governor-hochul-announces-fy-2023-investments- clean-energy-infrastructure-climate-resiliency

The White House. (2022, November 17). *CEQ launches Global Net-Zero Government Initiative, announces 18 countries joining U.S. to slash emissions from Government Operations*. The White House. Retrieved April 14, 2023, from https://www.whitehouse.gov/ceq/news-updates/2022/11/17/ceq- launches-global-net-zero-government-initiative-announces-18-countries-joining-u-s-to-slash- emissions-from-government-operations/#:~:text=%E2%80%9CThe%20U.S.%20Government%20is%20now,urgency%20of%20the%20climate%20crisis.%E2%80%9D