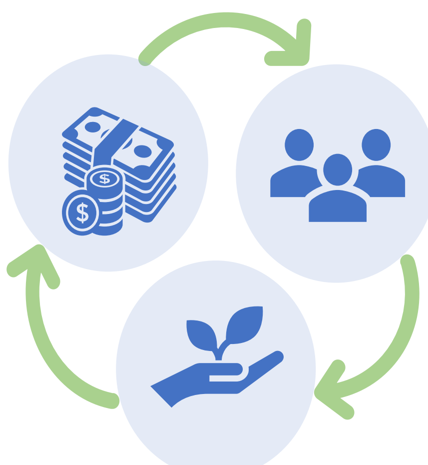
Collecting, Managing and Analyzing Water Usage Data



DEPARTMENT OF FISH AND GAME
Division of Ecological Restoration

dcr Massachusetts

Amy Vickers & Associates, Inc.
Water Planning, Policy, and Management

Tighe&Bond

TABLE OF CONTENTS

I. INTRODUCTION

Water is critical to both the natural and built environment. Drinking water supplies are finite resources, even in Massachusetts where there is a relative abundance of lakes, streams, and groundwater that supply our reservoirs and wells. For that reason, the state has established withdrawal limits while also promoting water use efficiency to ensure that water supplies in Massachusetts will be sufficient to support communities now and in the future.

To understand how to make the most of available water supplies, water managers first need to understand where and how water is used currently. Analyzing customer consumption data provides valuable insight as to how customers are using water, as well as identifying opportunities for demand reduction that may be needed for withdrawal permit compliance, supporting community growth or drought response. Although meter data is critical to virtually every operational element of a public water system, obtaining and analyzing it is challenging for many water systems.

This guidance document is designed to address common obstacles and provide insight into obtaining, managing, and evaluating water use data for the benefit of water suppliers, their customers, and the community. The first two sections provide general guidance on tracking and managing consumption and billing data. The third section demonstrates a series of analyses using a real sample dataset as a way to introduce the value of analysis to deepen understanding of your system, inform decision-making, and help communicate key system metrics to other stakeholders in visually compelling ways. The analyses we demonstrate are only a limited selection of what is possible with good underlying data, but we chose them with the belief that they are likely to be useful in addressing a broad range of system management challenges.

A. WATER USE DATA – THE KEY TO EFFECTIVE UTILITY MANAGEMENT

1. Why we collect water use data

Accounting for all water that enters and leaves your distribution system is critical for:

- Billing customers for water usage
- Determining unaccounted for / non-revenue water
- Managing and planning for peak demands
- Determining how much growth your system can support

2. Why we analyze water use data

Analyzing customer usage data can determine:

- What portion of overall demand appears to be discretionary
- How efficiently various customer classes use water
- How revenue is distributed and if a disproportionate share of revenue is associated with a particular user class or tier
- How weather or use restrictions might affect your cash flow
- How the continued installation of efficient water fixtures and appliances will influence future demands and planning forecasts
- How usage patterns are changing over time
- What impacts these changes are having on your system
- What effect these changes are having on your revenue over time

B. ANALYZING WATER USE DATA

Analyzing customer water usage data consists of three phases:

Phase 1: Data Collection and Management
Phase 2: Data Assembly and Preparation
Phase 3: Data Analysis and Interpretation

In this document, we provide specific guidance and tips for Phase 1 and 2 that should be applicable to most systems to help set them up for Phase 3. Since useful analyses will vary widely by system based on unique challenges, goals, and system profiles, we do not provide prescriptive guidance for Phase 3. Rather, we demonstrate through a particular case study *an example sequence of analyses* to help demonstrate what becomes possible with good data management and hopefully spur ideas for the analyses likely to be of greatest value to individual systems.

II. PHASE 1: DATA COLLECTION AND MANAGEMENT

Each system is unique in terms of the level of technology used for metering and billing, source water type, treatment requirements, and customer base composition. The guidelines discussed herein will apply to all water systems; however specifics may apply more to some systems than others. In general, the guidance will be most useful for systems with the following characteristics:

- Water meters are installed at every water account and meter readings are collected electronically
- A software package designed specifically for the purposes of creating water bills imports the water meter readings and uses them to generate customer bills
- The customer base is predominantly residential, in particular single family residential (SFR)

A. BEST PRACTICES FOR DATA COLLECTION AND MANAGEMENT

With any analysis, the results are only as good as the data. The following recommendations are intended to support the types of analyses described in this guidance document.

1. Meter all connections

All permanent connections to the distribution system should be metered and read at the same frequency. This includes non-revenue (municipal) accounts and bleeders if present.

2. Move towards monthly billing or meter reading

The majority of water customers are only made aware of their water use when they receive a bill. Monthly billing provides many benefits including:

- More consistent cash flow
- Reduced occurrence of high water bills due to leaking fixtures
- Increased customer awareness of water usage

Of course, increasing the billing frequency has a direct cost impact for water suppliers. A transitional option can be to read meters monthly, while maintaining a quarterly billing frequency.

Each water system should evaluate the cost and labor increases associated with monthly reading and billing and generate plans to achieve both within an appropriate timeframe.

3. Eliminate or reduce estimated reads

Estimated reads are less accurate than actual reads. Meter read data should be reviewed on a quarterly basis to monitor the number of estimated reads and work to eliminate them.

4. Correct incomplete or inaccurate meter reads

Erroneous meter reads for customers can distort the analysis of individual and total customer consumption data and excessive use of estimated usage erodes accuracy. To ensure accurate customer water use analysis, it may be necessary to adjust or omit such problem data to avoid distorting overall consumption patterns.

5. Investigate meter codes

Theft, leak, or tamper codes generated by water meters should be investigated on a regular basis. Leaks indicated on meters or generated by Advanced Metering Infrastructure (AMI) or customized meter software should be evaluated for frequency and estimated water loss.

6. Preserve meter data

All meter readings should be preserved as record data and maintained in perpetuity. Billing software packages vary in terms of data storage capabilities (see Appendix A for more information about billing software). Data should be backed up and preserved.

7. Data Maintenance

Erroneous reads discovered either as part of the billing review process or from a customer should be corrected in the source data and noted as a correction.

III. PHASE 2: DATA ASSEMBLY AND PREPARATION

This section focuses on extracting, organizing, formatting, and cleaning your data so it can be readily used to address a variety of management challenges you may face.

A. REQUIRED DATA

The analyses described in this document are based primarily upon consumption and billing data. However, additional customer data are generally required to draw meaningful conclusions. Here we group types of customer data into three categories that reflect the level of sophistication of analyses they can support. For almost all analyses, the data listed under Basic Customer Data are required. More complex analyses become available with intermediate data. Advanced data can be substantially harder to obtain but could be useful for specific purposes.

1. Basic Customer Data

Basic customer data consists primarily of the information required to issue water bills, which is required for almost any meaningful analyses. This includes:

- **Customer account number**
- **Meter size**

- **Customer class (e.g., residential, commercial, municipal)**
- **Meter read date**
- **Water usage**
- **Fixed billed amount**
- **Usage billed amount**

2. Intermediate Customer Data

A few additional pieces of readily available information can allow for substantially more depth of understanding of customer water use and revenue patterns. These include:

- **Customer subclass**. The residential class can consist of more than 100 usage types from single family homes to apartments and condominiums or dormitories. At a minimum, the residential class should be separated into single family and multi-family accounts.

- **Multi-Unit residential classification.** If known, the number of units per meter for multi-unit residential accounts can be identified to help estimate service population and per capita use.

- **Irrigation accounts**. If your system allows irrigation meters, then customers with irrigation meters should be designated as having them, and all irrigation meters should be identified and linked by address or other unique account identifier to those customers. To accurately report and evaluate each customer's total water demand, both the main (e.g., house) and irrigation meter data should be incorporated into that customer's consumption data.

- **Address.** If addresses are housed in a separate database for billing purposes, adding customer addresses to the consumption data set can assist with geographic analyses and with obtaining much of the information under Advanced Customer Data, below.

3. Advanced Customer Data

Advanced customer data consists of any other information that helps you add context to usage patterns. This can be thought of as a way to 'drill down' into more detailed information. The type of information you will need depends on what you intend to analyze. Examples include:

- **Household size.** Knowing household size can allow for calculations on a per capita basis. Some sophisticated "budget based" water rates also use individualized pricing tiers that are determined on the basis of household size. Note that on average, Single Family household sizes tend to be larger than those in Multi-family dwellings.

- **Year of construction or most recent kitchen or bathroom renovation**. Older plumbing fixtures tend to use more water. Identifying these properties may help explain patterns of what appears to be high indoor use. These properties may be candidates for fixture replacement rebate programs. Additionally, estimating the extent of older plumbing fixtures within your system can help anticipate future decreases in indoor demand, as fixtures are updated to more water-efficient models.

- **Lot size.** Outdoor usage can sometimes be correlated to lot size. Identifying larger properties may help explain patterns of what appears to be high outdoor use. These properties may be good targets for campaigns to improve outdoor watering efficiencies to reduce peak demand. High ratios of outdoor water use to lot size may also indicate irrigation leaks or particularly inefficient irrigation settings.

- **Water features**. Knowing which customers have irrigation systems, pools, and outdoor water features (e.g., fountains and artificial ponds and streams) can also help explain high outdoor water use patterns. These properties may be good targets for education campaigns on outdoor water-saving practices and information on any applicable water use restrictions.

B. DATA SOURCES

The primary source for the basic data listed above is the billing software system. Most systems have the ability to export data. (See Appendix A for more information.) If meter size is not tracked by the billing software, this can be obtained from a meter inventory by customer, which should only need updating as applicable after the initial effort. The intermediate data, if not already tracked, should also be available with an upfront investigation that should only require infrequent updating. Some of the advanced data can be obtained through the Standardized Assessors' Parcel database for all Massachusetts communities, available through MassGIS (see Appendix B for more information). Other advanced data may be available through the town clerk or building department.

C. DATA FILE FORMAT

In order to analyze data, the dataset must be in an acceptable file type and must also be properly formatted. The analyses demonstrated in this document were all performed in Microsoft Excel; however the guidelines here should apply to any software that can analyze water data.

1. Acceptable file type

Formats that are readily usable include: .xlsx, .csv, .txt, .dif, .prn, and .accdb (Microsoft Access). Adobe Portable Document Format (.pdf) files that were created electronically (versus scanned) can also be converted to Excel but may require additional work to format the data.

2. Data format

In order to analyze data, it generally must be in tabular format. Tabular format consists of data that is arranged in rows and columns, where the type of information in each column is consistent for all rows.

D. DATA CLEANING AND PREPARATION

Once the dataset is in a proper format, it must be cleaned or groomed. Before starting, make sure to work with a copy of the data so that the original data are preserved as a record. The actual steps and procedures that will need to be completed vary depending upon the source data and system characteristics; however, in general they consist of applying the following steps for each customer class or subclass[1] involved in the analysis.

[1] For brevity, user class is taken to mean the subclass (i.e., "Single Family Residential" vs. "Residential").

1. Review zero and negative values

Consumption datasets can contain a surprising number of zero and negative values. First, check that all such accounts are active. If inactive accounts are included in the dataset, make sure to incorporate a data field (column) that can be used to identify and filter them out. If the zero values are associated with a seasonal account, see guidance below. Negative values usually represent estimated meter read corrections and sometimes may reflect billing adjustments for leaks. Negative values should be excluded from analysis since they distort actual customer water use.

2. Review minimum and maximum use values for each user class

Filter or separate data by customer class, then using the data sort feature, arrange your consumption data in order of smallest to largest and then vice versa. This will help identify both outliers and miscoded data. For accounts that have consumption values that are an order of magnitude higher or lower than expected, look to see if there are other data fields that can help determine if this is a miscoded account. Address and lot size can be useful for this. Extreme values may also represent incorrectly classified metering units or decimal place coding, or a malfunctioning meter. Of course, the data could be legitimate and represent an actual outlier.

3. Identify seasonal accounts

For systems that have a large number of seasonal users, it is helpful to create a data field that identifies these accounts. This can be done by looking for consistent zero or near-zero consumption values for particular seasons (often the winter months), or if shutoff records are tracked, these may be reviewed for indications of seasonal occupancy.

4. Identify Irrigation meters

Communities that use water consumption records for sewer service billing sometimes allow separate irrigation meters or 'deduct' meters. Such customers should be identified with a data field and the consumption data from their separate meters should be able to be easily linked to these customers by address or unique customer account number.

5. Verify usage date

Water is billed in arrears, so a given read date represents the end of a usage period. The bill date can be a month or more later. Accurate analysis requires that the consumption values are accurately matched to the usage period. Some billing software systems do this correctly, but others do not, resulting in some months' or quarters' reported usage being shifted from the actual month or quarter when it was used to when it was billed. Such shifts distort the evaluation of seasonal and annual consumption patterns and should be corrected prior to analysis.

6. Assign season

The ability to subtotal customer usage by month and season is a key requirement for many of the analyses presented in this document. Classifying usage periods by season, as defined based on your billing cycles and particular analysis goals, helps set you up for such analyses.

7. Finalize data structure and calculate key analytic metrics

For most analyses, your dataset should be set up such that each row represents one customer account and all information associated with that account is in columns. For much of the cleaning and preparation from the previous steps, your data may have had multiple rows for each customer, each representing a different billing period. For many analyses, you may want to create a new set of calculated columns that summarize customer use and charges in various ways, depending on your analysis goals (e.g., annual totals, seasonal averages, monthly averages, annual minimums and maximums). These columns may be derived from a single year of data. However, for some analyses, you may want to generate them using multiple years of data, or perhaps separately for a dry year and a wet year. We do not get into prescriptive detail on data structure here, but recognize that this step of data preparation is complex.

IV. PHASE 3: DATA ANALYSIS AND INTERPRETATION

Choosing appropriate analyses depends on the composition of your customer base, your particular goals, and the challenges you are trying to address. The remainder of this document walks through a demonstration series of analyses, using a real dataset. It is less prescriptive than the previous sections and intended to illustrate how analyzing your data can deepen your understanding of your system, inform decision-making, and communicate key metrics to your local officials, customers, and other stakeholders.

A. DEFINE SYSTEM CHALLENGES AND PRIORITIES

Each system has different challenges, and reviewing some of the following common challenges will help guide your analyses.

1. Anticipate trends

Over the past 10 years, 60% of Massachusetts water systems have experienced a decrease in total water use, with an average decrease of 1.8% per year. The other 40% of systems experienced an increase in total water use, with an average increase of 1.7%[2]. As water system revenue is strongly associated with customer usage, while costs tend to be more fixed, a continuous decline in usage without appropriate rate restructuring can result in long-term revenue shortfalls. A steady increase in use may pose capacity challenges.

2. Assess available annual capacity

The difference between a system's annual available water supply volume (as defined either by physical or regulatory capacity) and total annual demand represents remaining available capacity to support community growth. If your remaining capacity is less than your community's projected or planned growth, a better understanding of your customer usage patterns can help determine the extent to which the capacity deficit might be closed through demand management efforts.

3. Manage peak demand

Some systems have sufficient capacity on average but struggle to meet demand during peak days. Understanding the distribution and causation of peak demand can help with operational planning as well as customer outreach and strategic pricing structures to help reduce discretionary use, such as lawn watering.

[2] ASR data 2009 - 2017

4. Address revenue sufficiency and stability

As mentioned above, a consistent downward trend in water use will eventually result in a revenue deficiency unless there are regular rate adjustments. Additionally, relying on revenue from peak usage to meet baseline costs can leave systems vulnerable to deficits in wet years.

B. SET ANALYSIS GOALS

Knowing what questions you need answered should guide the analyses you undertake. Our sample analysis is based on data from a municipal water system in eastern Massachusetts (see Appendix C for relevant information about the system and its community). This system has identified ongoing challenges with meeting peak demands, and concerns about revenue instability – especially loss of revenue in wet years. These themes inform the analyses in the remainder of this document. Figure 1 describes the logic (i.e., the underlying questions) that flows through the sequence of the analyses.

Figure 1. Diagram of Logic Flow for Sample Analyses

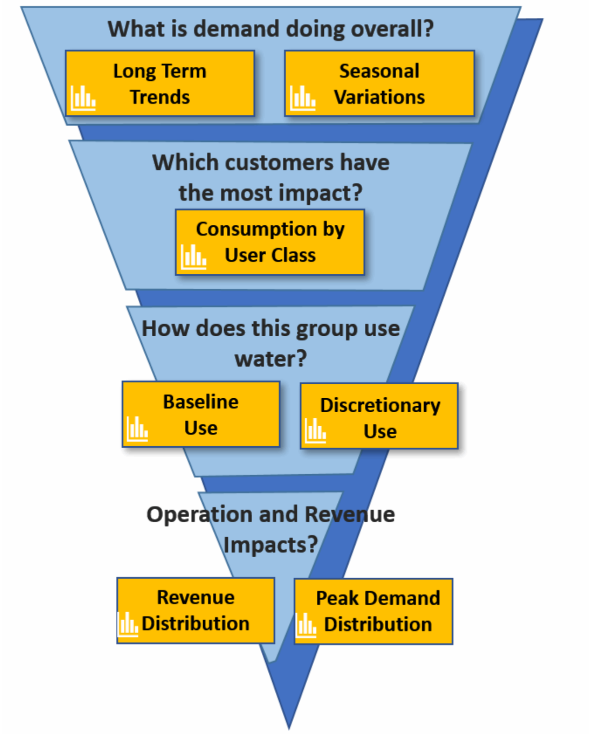


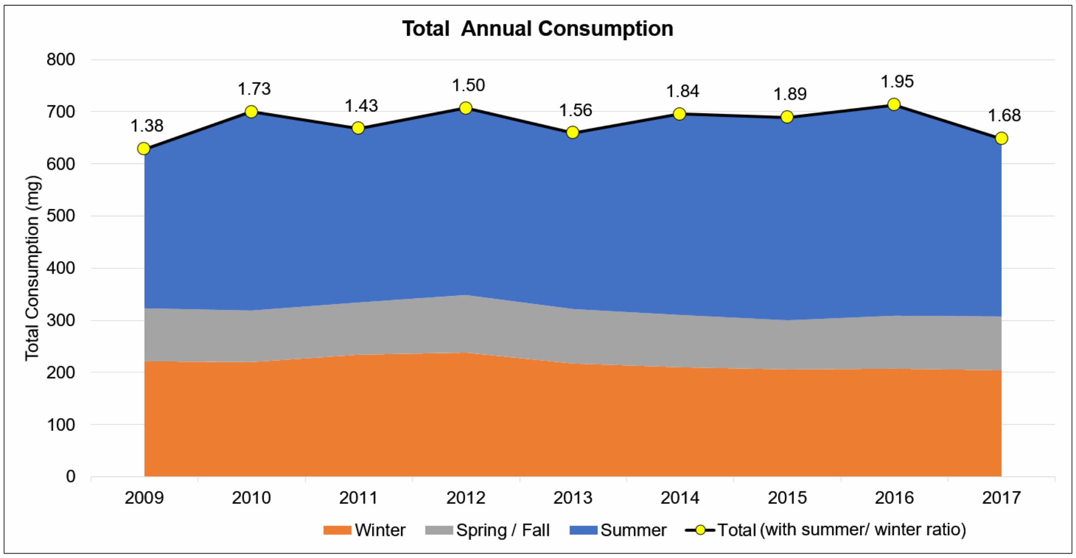
Figure 1 depicts a logic that starts by looking at a system-wide overview in a few different ways and subsequently goes into more detail based upon the findings. Although the following demonstration is only one way to approach your data, starting with broad overview questions and narrowing from there is appropriate for most circumstances.

Although only a subset of what is possible with good data management practices, the particular analyses we present would likely be useful for many systems. They will hopefully spur other ideas for how you may want to look at your data.

What Is Demand Doing Overall?

Reviewing long-term trends in water usage is a good place to start, as these trends will likely play a large role in future system needs. These trends are part of the "big picture", and after examining them we can hone in on more details. Figure 2 shows overall consumption for our sample dataset for a nine-year period, broken down by season[3]. This helps evaluate trends in overall consumption (represented by the upper most line) and seasonal consumption, shown by the colored bands.

Figure 2. Total Annual Water Usage



C. EXAMINE ANNUAL LONG-TERM CONSUMPTION TRENDS

We can see with Figure 2 that despite some fluctuation from year to year in summer use, the overall level of water use is quite stable for this system.

If this analysis had shown that our overall use was increasing, we would need to think about whether our supplies can support continued increases. If our supply is likely to be a constraining factor, we may be able to build in capacity for future growth through conservation. Our subsequent analyses could be set up to answer questions about consumption patterns, so we could target our conservation efforts to make the most difference. Even if our supplies are not constrained physically or by regulation, it might be more cost-effective to support future growth through conservation, rather than increasing our supply.

Alternatively, if this analysis had shown that our overall use was decreasing, we would need to develop a long-term rate strategy that allowed us to recover our full costs, even as demand declines. Our subsequent analyses could be set up to pursue detailed questions about how we are currently bringing in revenue and guide us toward sustainable rate structures.

[3] The seasonal designations used to create the colored bands in this figure were defined as follows: **Summer** = May through September; Winter = January, February, March, November, and December; **Spring/ Fall** = April and October. For different purposes, seasonal use may be defined differently.

D. EXAMINE SEASONAL VARIATION

Most water systems experience a measurable increase in water usage during the summer months. In Figure 2, we quantify seasonal variation with the "summer/winter ratio", as defined by the MassDEP on the Annual Statistical Report (ASR). The summer/winter ratio for the sample data set ranges from a high of 1.95 in 2016 (a record drought year) to a low of 1.38 in 2009 (a relatively wet year). This helps depict visually why the system may be struggling to meet peak demands and with revenue stability. Not only does water use increase substantially during most summers, creating operational burdens, but those increases fluctuate year to year. Because this system has tiered rates, the higher volumes sold during the summer contribute disproportionally to the overall revenue stream (from higher per-gallon charges). As summer volumes fluctuate, revenue fluctuates even more.

Next, we break down these patterns at a more detailed level and explore their implications for how we could operate our system and set our rates into the future.

Note that for the remainder of the analyses in this document, we use data from 2018. Much insight can be gained by looking in depth at a single year, as long as it does not represent a significant aberration. If time and resources permit, data derived from a 3-to-5-year average for each customer can provide a broader picture. Alternatively, for some questions it may work best to look separately at some extremes, such as a particularly dry year and a particularly wet year.

Which Customers Have the Most Impact?

E. ANALYZE USAGE BY CUSTOMER CLASS

We next look at how total usage is distributed across the various customer classes. Figure 3 is a stacked column chart that shows total monthly consumption broken down by customer class, for 2018. The total usage for a given month is represented by the top of the column, and the numbers above each column represent the percent of total system consumption attributed to that month. The colors depict the monthly usage for each customer class.

Figure 3. Monthly Consumption by Customer Class

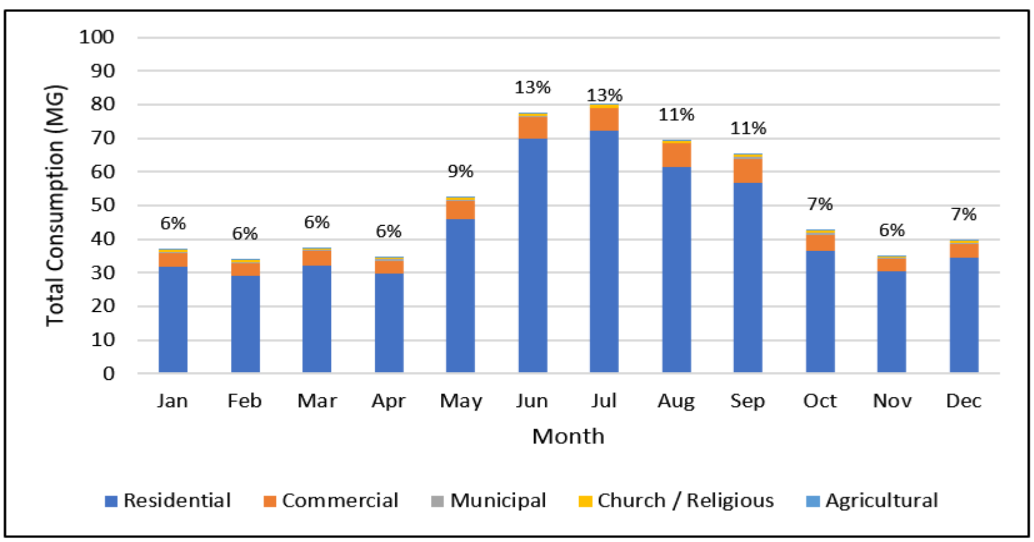
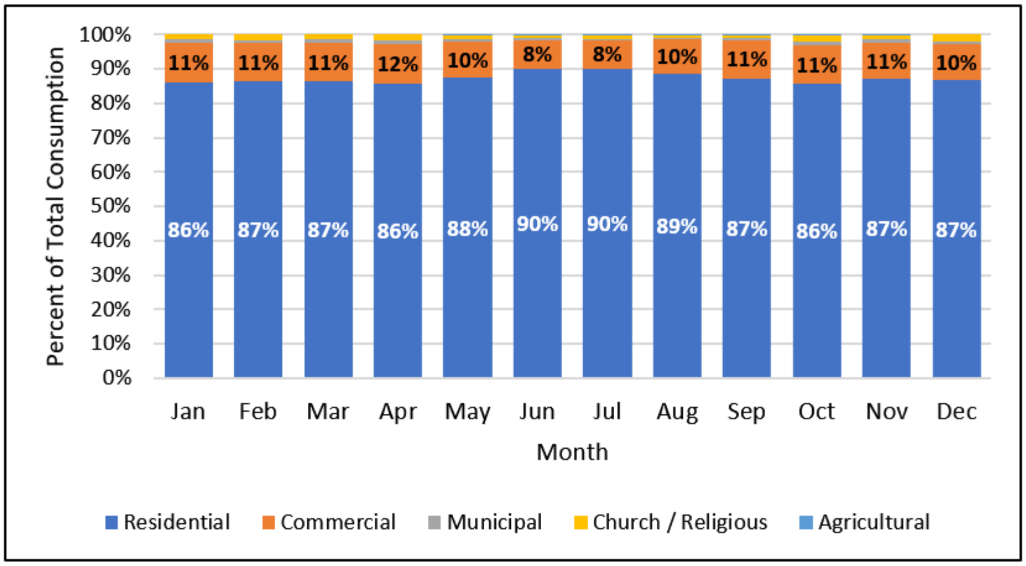


Figure 3 shows that the seasonal increase in usage is largely between May and September, which validates the seasonal designations used to derive the summer-to-winter ratio in Figure 2. (Checking assumptions is an important use of analysis.) Figure 3 also illustrates that the bulk of water usage is associated with the residential class, with the commercial class a distant second and the remaining classes quite nominal.

Figure 4, a 100% stacked column chart, quantifies the relative use by customer class, showing the percent contribution of use by each customer class, for each month. The percentages for residential and commercial use were added as data labels, to help with reading and interpretation.

Figure 4. Percent Consumption by Customer Class



Although Figure 3 showed that the bulk of usage was from the residential class, with commercial use a distant second, the enhanced precision of Figure 4 shows that in the summer months, the relative percent used by the residential class increases slightly. This indicates that not only does the residential class use the most water, but in the summer, the dominance of the residential class is even higher.

How Does the Most Impactful Customer Class Use Water?

Next we drill down into the most significant user class. How do residential customers use water? Is there much variability between customers in their overall use or their seasonal usage patterns? Is there a subset of customers largely responsible for the peak stresses on the system? If so, what portion of the customers is this?

F. EXAMINE RESIDENTIAL USE BY SUBCLASS

We first examine which residential subclasses are significant. Table 1 shows the number of accounts and subtotaled usage for each residential subclass. Using the seasonal designations discussed above, water use for the summer and winter seasons were totaled separately for the top four residential subclasses, and the remaining subclasses were pooled into a single total.

Table 1. Total Residential Water Usage By Subclass

	Number of Accounts		Summer Usage		Winter Usage	
Description	**Total**	**% of Total**	**Total (gal)**	**% of Total**	**Total (gal)**	**% of Total**
SINGLE FAM MDL-01	5,435	76.9%	247,668,188	80.8%	65,938,696	68.2%
CONDO NL MDL-00	855	12.1%	19,535,926	6.4%	10,354,261	10.7%
TWO FAMILY	171	2.4%	5,830,519	1.9%	3,025,843	3.1%
MOBILE HM	42	0.6%	4,346,520	1.4%	2,787,536	2.9%
OTHER	556	7.9%	28,990,966	9.5%	14,511,026	15.0%
TOTAL	7,066	100%	306,372,119	100%	96,617,362	100%

Table 1 shows that single family customers dominate residential use, both in terms of number of accounts and total usage. The fact that single family customers represent 76.9% of all accounts, but 80.8% of summer usage means not only is this the largest subclass, but the most impactful subclass on a per-account basis in the summer. The second most impactful subclass, condominiums, represents 12% of the accounts but only 6.4% of summer usage. The remainder of the residential subclasses use even less. Based on these findings and this system's concerns centering on peak usage capacity, Single Family Residential (SFR) users were selected for further evaluation. *Note that if this system was concerned with annual usage, it would probably be worth looking more closely at usage patterns within the condominium subclass.*

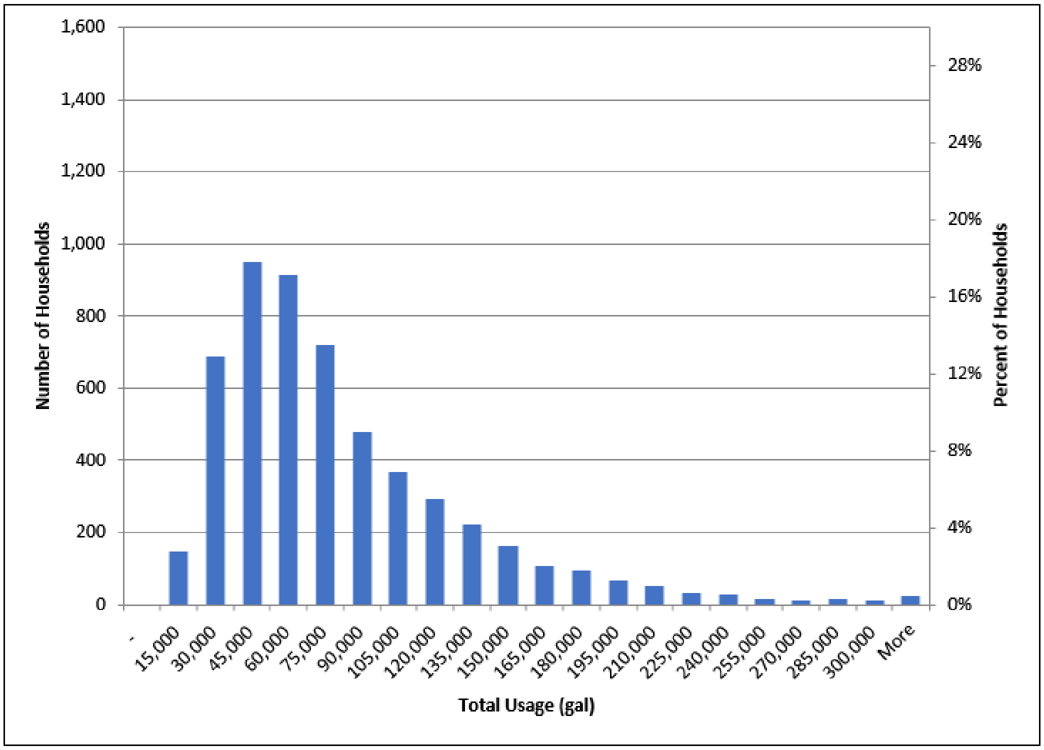
In preparing for the next step of the analysis, inactive or unoccupied accounts were eliminated from the data[4]. A year-round average of 20 gallons per day (gpd) – likely the lowest that would be expected for one person living in a house with water-efficient fixtures and no outdoor use – was used as a lower threshold for determining occupancy. In the sample data, 81 accounts had average usage below 20 gpd. These were eliminated from the analyses below. As noted above in Phase 2 Data Cleaning and Preparation, it can be worth investigating these accounts for coding errors, faulty meters, or other issues.

G. ANALYZE HOW SINGLE-FAMILY CUSTOMERS USE WATER

Figure 5 shows the total annual usage by customer for the single-family residential class in histogram format. A histogram presents data by determining how many individual data points fall within particular ranges of values (called "bins"). In Figure 5, each bin represents a range of total annual usage. The left-hand y-axis shows the *number* of single-family customers that fit within each range, while the right-hand y-axis shows the *percentage* of single-family customers within each range. The first column represents customers who used up to 15,000 gallons. The second column represents customers who used between 15,000 and 30,000 gallons. The third column represents customers who used between 30,000 and 45,000, etc. In other words, the value shown on the x-axis represents the *maximum value for that bin*.

[4] All analysis should be performed on a separate copy of the data. All original data should be preserved for record keeping purposes.

Figure 5. Total Annual Water Use by Single-Family Residential Customers

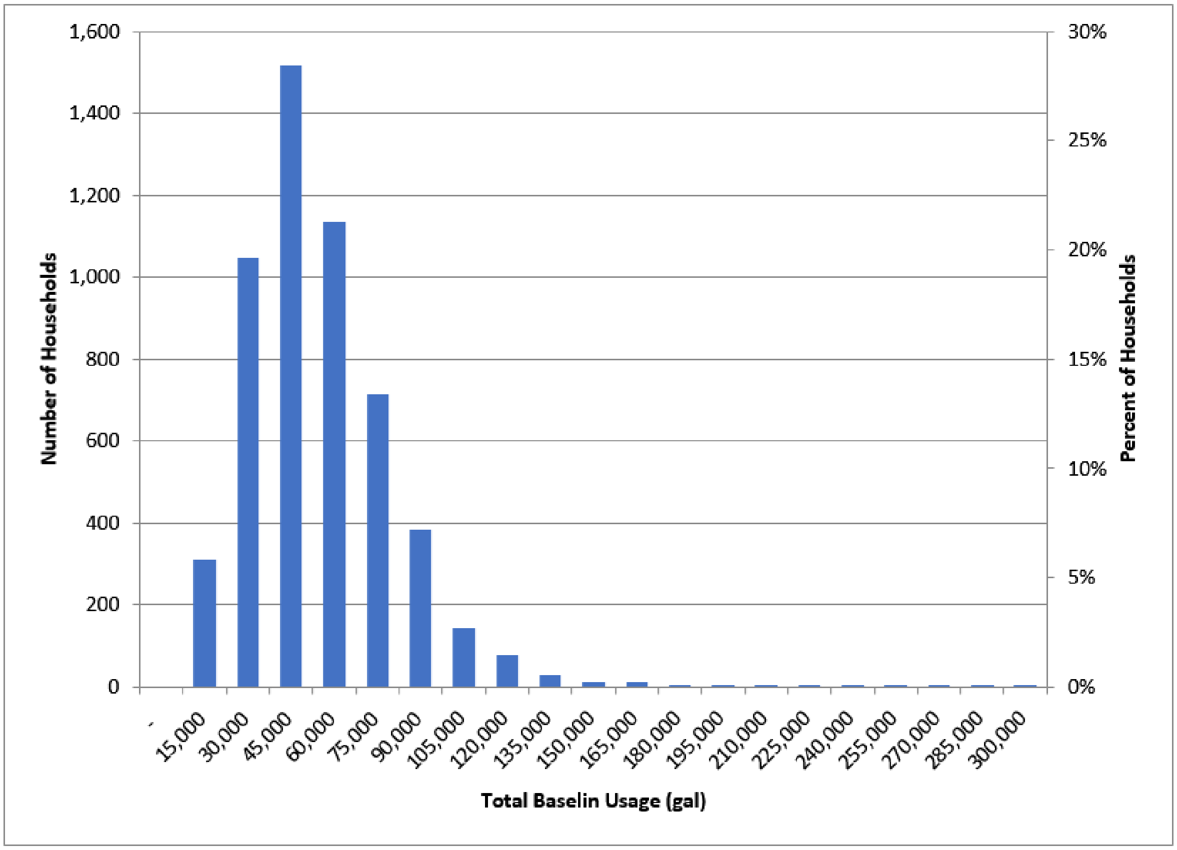


Histograms are a useful first cut to look at variability and tendencies within a group. In Figure 5, we see that variability of total annual use among single family customers ranges from less than 15,000 gallons to more than 300,000 gallons – more than a twenty-fold difference. We also see that the most typical annual usage, as shown by the tallest two bars, is between 30,000 and 60,000 gallons, and that in general the data are somewhat clustered between 15,000 and 75,000 gallons (the tallest four bars), suggesting that a large portion of customers have annual usage in this range.

Because this system has difficulty meeting peak demands, we next unpack the data to better understand what portion of the annual use shown in Figure 5 appears to be more discretionary, and therefore presumably more easily reduced. To do this, we break all customer usage into either "baseline use" or "discretionary use". Baseline use is intended to represent use for stable, essential needs like drinking, cooking, cleaning, and bathing. These are generally indoor uses. Alternatively, most outdoor uses, such as watering lawns and landscapes, filling swimming pools, washing cars and patios, etc. tend to be more discretionary (i.e., less essential) and more variable by month and year. In New England, residential winter-time usage is often a reasonable proxy for baseline use. While baseline use can be reduced through efficiency upgrades in appliances and some behavioral modification (such as taking shorter showers), in the short-term, baseline use tends to be more difficult to reduce than discretionary use.

Figure 6 is a histogram that shows the component of annual water use by single-family customers that we designated as "baseline" use. This was derived by looking at winter use for each customer (total use for the billing quarter for November, December, and January) and multiplying it by 4, to represent 12 months of use at that level. *Note that where this calculation resulted in a total that was greater than that customer's actual total annual use, the baseline use was capped at the total annual* *use. These represent customers whose year-round use is stable, or whose use actually declines in the summer. This was true for approximately 24% of single-family customers.*

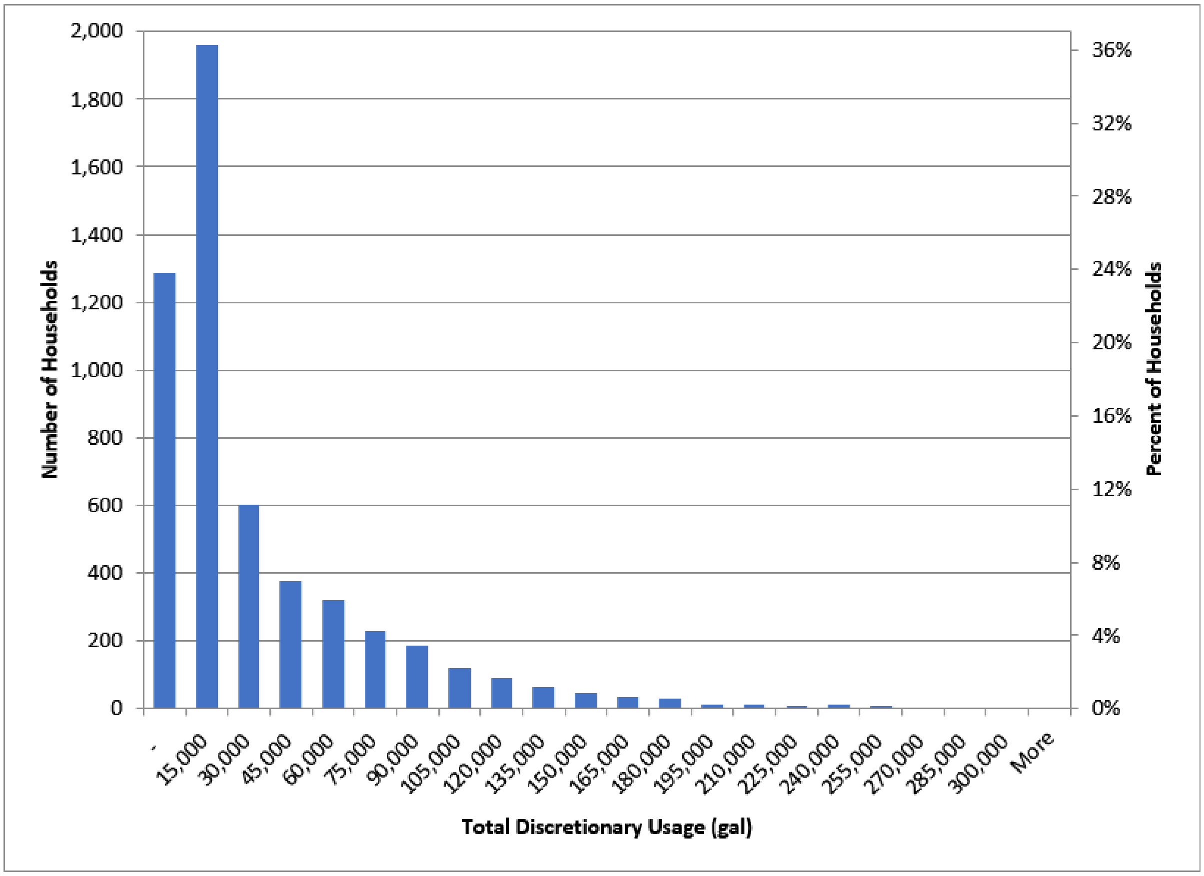
Figure 6. Annual Baseline Water Use by Single-Family Residential Customers



In Figure 6, we see that while there is still a range in baseline use, there is less variability over the customer base. Most customers use between 15,000 and 60,000 gallons, and the vast majority of customers use less than 90,000 gallons. *This type of baseline-use histogram can be very helpful for establishing volume breakpoints for pricing tiers intended to cover stable, essential use.*

To that end, we next look at discretionary use, which is the difference between baseline and total use, for each customer.

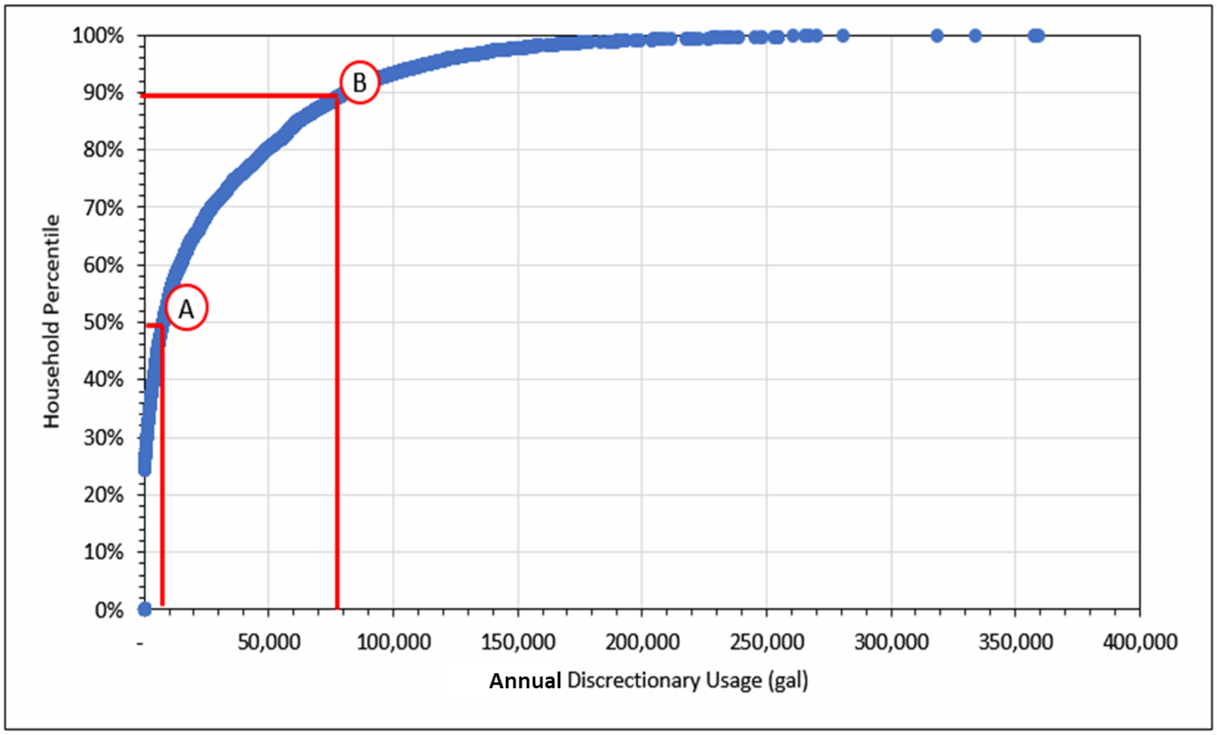
Figure 7. Annual Discretionary Water Use by Single-Family Residential Customers



In Figure 7, we see that discretionary use clusters at the very low levels and then spreads gradually into the higher use categories (i.e., the long tail to the right of the tallest bar). Recall that discretionary use is defined as total use over baseline, over the whole year. In this graph, the very first column represents zero-use. These are the 24% of customers we mentioned above, who showed no seasonal increase in use. In the next column are those whose discretionary use was greater than zero, but less than 15,000 gallons. These two columns together appear to represent about 60% of SFR customers. However, the rest of the customers vary quite widely in how much discretionary water they use, as shown by the gradual tapering of the histogram. To help us better understand this variability, let us look at discretionary use a different way.

Figure 8 is a cumulative distribution curve, showing annual household discretionary use along the x-axis, and household percentile rank on the y-axis. The graph in Figure 8 is created by ranking the annual discretionary use for each SFR household from lowest to highest and then calculating each household's *percentile rank*. Percentile rank refers to the percentage of all SFR customers that use *the same amount or less* discretionary water than that household. For example, the 50th percentile household, marked by the letter "A" in Figure 8, uses about 8,000 gallons of discretionary water over the whole year (as shown by dropping down to the x-axis from the letter "A"). Another way to say this is that 50% of all SFR households use 8,000 gallons or less of discretionary water. (The 50th percentile value is more commonly called the "median".)

Figure 8. Cumulative Distribution of Single-Family Residential Discretionary Usage



There is an almost ten-fold increase in use over the next 40% of customers, as we move up the y-axis from the 50th to the 90th percentile. This point, marked by the letter "B", shows that 90% of customers use 78,000 gallons or less of discretionary use. One of the most remarkable observations from this graph is that the top 10% of customers uses dramatically more discretionary water – some using even more than four times the 90th percentile value!

We have determined that a small number of customers have very high discretionary use, but how much of a difference do these few customers make when it comes to this system's total peak demand stresses? How responsible do we think these few customers are in this system's revenue fluctuations? We will turn our attention to these questions next.

What are the Operational and Revenue Impacts of the Top Users of Discretionary Water?

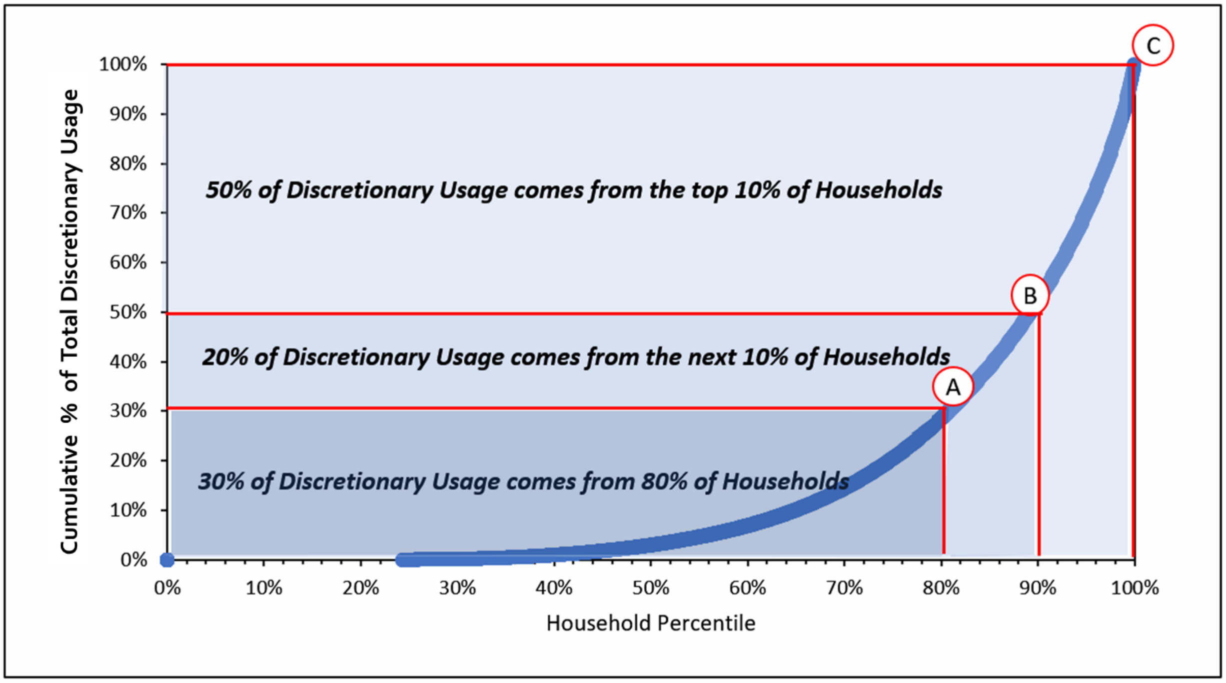
In the final section of this document, we will look at how the top discretionary water users are impacting the operations and revenue of the sample system. These final analyses will help communicate the key issues to this system's customers, water commissioners, and other community officials, and help inform how the system may want to tackle peak supply and revenue stability issues.

H. ANALYZE THE RELATIVE IMPACT ON PEAK DEMAND CAUSED BY TOP USERS

Figure 9 is another cumulative distribution curve, but rather than showing *volumes of discretionary use* along the x-axis, it shows the *percentile rank of customers based on their discretionary use* (i.e., the value that was shown on the y-axis in Figure 8). The y-axis in Figure 9 shows the *cumulative percent of all discretionary use* accounted for. This graphic presents usage data in a different way than we are used to seeing and thus may take more time for people to digest; however, it provides useful insights.

The x-axis shows customer's percentile rank for discretionary usage. Recall from Figure 8 that each percentile rank is associated with a total annual discretionary volume. As we move along the x-axis from left to right, the cumulative total of this underlying discretionary use increases. The y-axis shows what percent of all discretionary use that running total represents. For example, when we get to point "A" on the x-axis, we see that 80% of our customers have been accounted for (x-axis value), while only 30% of our total SFR discretionary demand has been accounted for (y-axis value). Moving further along the x-axis, we find that 90% of our customers account for 50% of all our discretionary demand (point "B"). Finally, we see very clearly what impact the top 10% of users have on peak demand – namely, they account for a full 50% of our discretionary demand (the lightest blue box in Figure 9)!

Figure 9. Cumulative Percent of Single-Family Residential Discretionary Usage by Customer Percentile



This has significance for how we might address this system's peak demand challenges. Figure 9 suggests there could be large operational savings possible from targeting the discretionary use of just the top 10% of customers. Perhaps the system could conduct targeted education and messaging to these users, or offer outdoor water use audits or weather-based irrigation interruption devices to these customers. The system could also potentially establish very high per-unit price increases for a tier of usage only the highest-use customers are likely to trip into. Outdoor watering restrictions could also be instated that allow the majority of customers to maintain their existing habits, while bringing the top users into more efficient irrigation routines.

Before any of these actions are taken, additional data analysis could be used to fine-tune the details. For example, the analysis in Figure 9 could be re-run, honing in on the months where peak demand stress poses the greatest problems. Target reductions in demand can be set for those months, based on operational constraints, and then examined in terms of their equivalent as a percentage of total discretionary water use during those months. This may help us identify, for example, that we need to reduce discretionary water use by 30% to avoid peak demand stresses. With further number crunching, we might then determine that we could achieve that if we pursued concerted conservation efforts that resulted in the top 25% of users reducing their discretionary use by an average of 20%. While arriving at such specific conclusions would take a few more steps than we demonstrate here, our point is to illustrate that Figure 9 is a helpful jumping off point in considering operational strategies to address peak capacity issues. Understanding consumption patterns in this way not only informs decision-making, but it increases the likelihood of targeted actions achieving their intended outcomes.

In our next and final analysis, we examine revenue stability.

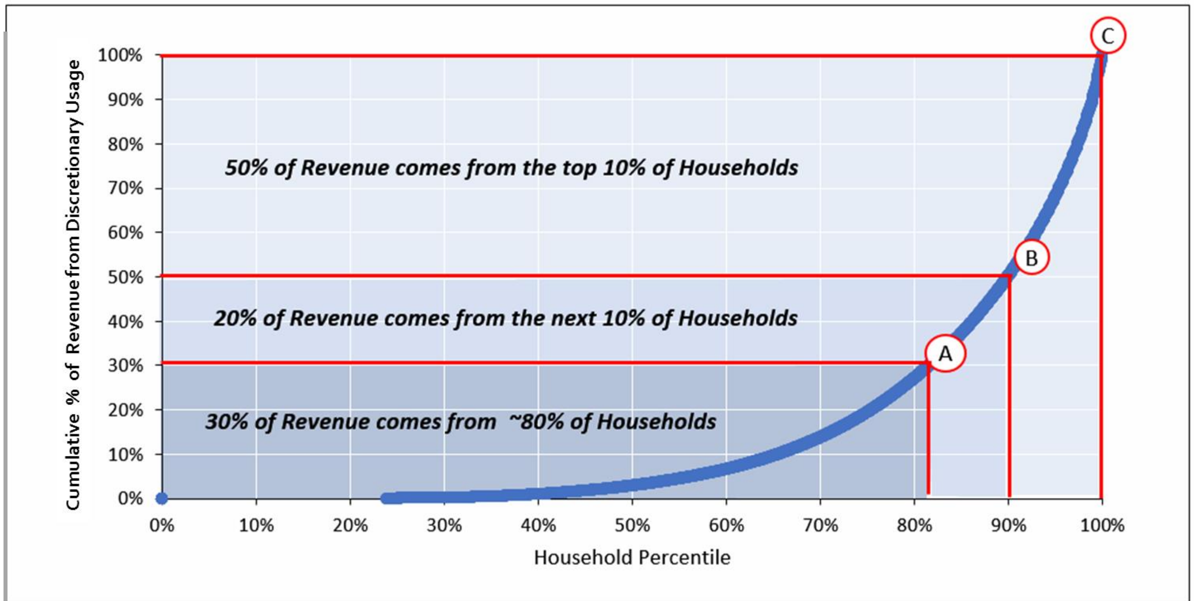
I. ANALYZE THE RELATIVE CONTRIBUTION OF THE TOP USERS ON SEASONAL REVENUE

If this system curtails demand to help address peak capacity limitations without changing its rate structure, it can expect a reduction in revenue. To ensure it can reliably cover its costs into the future, it will need to shift the revenue that is currently associated with usage it hopes to curtail to usage that it can rely on moving forward. Even if this system were not going to pursue demand reduction strategies, ensuring it can recover its costs with revenue collected from *stable, reliable use* would mean avoiding sudden shortfalls during wet years.

To help understand the revenue instability this system experiences, we will look at an analysis that is very comparable to the analysis depicted in Figure 9. However, in Figure 10, we replace the cumulative percent of *discretionary usage* with the cumulative percent of *revenue earned from discretionary usage.* For each single-family residential customer, we calculate the annual amount they were charged *for the portion of their use estimated to be discretionary.* This is done by first estimating what they spent on their baseline usage. This was derived by taking their total bill from the winter quarter (November-January) and multiplying it by four. As with usage, if this created an annual total greater than the actual total annual charges for that customer, than the estimated baseline charges were capped at the annual total. The revenue associated with discretionary use was then calculated as the difference between each customer's total annual charges and their estimated charges for baseline use. *Note that this exercise would need to be modified if this system had seasonal increases in rates, as "baseline" use in the summer would not cost the same amount as baseline use in the winter.*

Figure 10 looks very similar to Figure 9. Reading this graph in the same way we read Figure 9, we see that just over 80% of SFR customers are responsible for 30% of the revenue this system derives from discretionary use (point "A"). Another 10% of customers are responsible for the next 20% of revenue. Again, we see that the top 10% of discretionary water users account for a full 50% of the revenue associated with the sale of discretionary water. If this system had steeper tiers in place, we would likely see an even higher percentage of revenue associated with the top users.

Figure 10. Cumulative Percent of Revenue from Discretionary Usage by Customer Percentile



This type of analysis is a very useful starting point for analyzing vulnerabilities in cost recovery. For example, we see immediately that if this system successfully reigned in the top users to help meet peak demands, a full 50% of their revenue from seasonal use would take a large hit. Similarly, each year that these top customers use significantly less due to perhaps a wet summer, revenue takes a large hit. This graphic can help communicate to stakeholders, including water commissioners who oversee rate structures, the underlying causes of revenue instability.

As with the usage analysis, additional steps would be needed to help set tier breaks and prices to achieve more stable revenue streams. In fact, sophisticated rate models will generally be necessary to fully analyze the impact of rate proposals on overall revenue and on various customer user groups. But analyses such as those in Figure 10 can help create starting points for rate structure proposals. For example, based on the results in Figure 10, a system could establish as a criterion for a revised rate structure that costs would be covered in full from the volume of water currently sold to the bottom 90% of customers, and very aggressive pricing would be used to try to curtail the highest use (without compromising cost recovery). As with any major rate restructuring, it may have to happen in phases, and involve significant stakeholder engagement.

In places where major rate changes have been successfully adopted and accepted by the community, water systems have almost always first undertaken community-wide conversations laying out the long-term problems, using visually compelling graphics, such as those depicted here, and explored the range of options that could be pursued to address the problems. A wide variety of additional analyses and visual tools can help people characterize the impacts of different options on different user classes and stakeholder groups and weigh in on what feels like a fair and feasible solution for their particular community.

J. ANALYSIS SUMMARY

Analyzing water use data is best accomplished using a step-wise, drill-down process by which each step is guided by the conclusions of the one before it. This process will help identify your system's particular goals and challenges and help develop potential solutions. While some goals and challenges are apparent simply from experiencing the day-to-day management of the system, others may require taking a broader perspective by engaging with community leaders to discuss growth projections and goals. Even after such a review, some issues might only come into focus as you start to dig into the data and look at it from different high-level angles. In short, each step of analysis should be guided by the primary questions you are pursuing, which may become increasingly clear as you go.

In the example system used in this guide, difficulty meeting peak demands and experiencing revenue instability were known challenges going into the analysis. Early high-level analysis clarified that the bulk of both total usage and seasonal usage were from single-family residential (SFR) customers, which is typical for systems in largely suburban communities, including those throughout much of Massachusetts. From there, we pursued a greater understanding of what portion of SFR demand covered "baseline" use and what portion appeared to be "discretionary", as well as how variable each of these demand components are across the customer base. This led to the observation that more than half of the customer base appeared to have modest to no discretionary use, while the remaining customers showed high variability in discretionary use, including some extreme outliers.

Lastly, we examined the *impact* of these higher discretionary users on our total seasonal demands and the *relative contribution* of these higher discretionary users on our seasonal revenue streams. This led to the recognition that the very highest discretionary users accounted largely for the system's seasonal demand spikes, and that the discretionary demand of these same highest users accounted for most of the seasonal revenue. Both of these conclusions should help this system fine-tune strategies to address peak demand stresses and revenue stability issues that are grounded in the system's particular realities. In particular, conservation strategies targeting the top 10% of discretionary users will go a long way to reign in peak demands. To avoid sharp revenue losses that would otherwise be felt by curtailing the highest discretionary volumes, rate strategies should be pursued that would fully recover costs if discretionary use topped out at the existing 90th percentile. For discretionary use over the 90th percentile, sharp per-unit price increases could be used to create a financial reserve that could be drawn on during very wet years when revenue shortfalls may still occur.

The analyses demonstrated here are only a small subset of how this system could look at consumption and billing data to improve operational and financial sustainability. Even the strategies just described could benefit from fine-tuning with some additional analyses. We hope, however, that the flow and logic of this demonstration has sparked curiosity about what similar analyses would show for your own system, as well as ideas for any number of other analyses you could undertake with your billing and consumption data to pursue the challenges you deal with every day. We also hope some of the visual graphics shown here spark ideas for ways to effectively communicate these same challenges to your water commissioners, elected officials, and customers.

V. CONCLUSIONS

Managing modern water supply systems involves balancing a complicated array of goals, navigating many constraints, and being simultaneously able to plan at the 10,000-ft level and track the goings-on at a detailed operational level. Developing a deeper understanding of the daily and seasonal patterns and the long-term trends of your customers' water demands, and how they impact revenue will be useful for any system in terms of overall efficiency and optimization.

The Best Practices laid out in Section II *Data Collection and Management* and Section III *Data Assembly and Preparation* represent a new first step for some suppliers, and a way to review and fine-tune practices for others. For those in the former category, getting your data in order may take some time, but the hope is that the analyses shown in Section IV *Data Analysis and Interpretation* have shown the benefits of improving data collection and management. For those of you with strong data management practices, the hope is that Section IV spurred your thinking about useful approaches you could take to analyze your own data in pursuit of your current needs.

This document is intended to be one of several resources the Massachusetts Executive Office of Energy and Environmental Affairs and its agencies are developing to support Massachusetts water suppliers in advancing the long-term financial and operational sustainability of their systems. In particular, the MA Department of Conservation and Recreation's Office of Water Resources and the MA Department of Fish and Game's Division of Ecological Restoration have been working to respond to resource gaps brought to our attention by MA water suppliers concerning revenue stability and sufficiency, water conservation strategies, and data management. We are grateful to all MA suppliers who are working to help us understand the existing resource gaps and shape the efforts of the Commonwealth to address them.

APPENDIX A – WATER METER AND BILLING SOFTWARE

The majority of water suppliers in Massachusetts utilize some type of electronic meter reading system and integrated billing software. The meter reading device collects read data from customer water meters, which is then transferred or downloaded using a proprietary software package. The software package then translates meter reading data into a format usable by a second software application designed to prepare water bills from meter data. These software applications are the primary source of the water consumption data used in the analyses described in this document. The ability to extract usage data varies considerably between the various software packages. The features and capabilities of most commonly used software metering and billing software packages reported are summarized below.

	Name	Can the User Download Consumption Data?	Can the User Use Crystal Reports?	Data Export Format	Data Available in Export	Data History/ Logging/Storage	Software Update Process/History
Billing Software	CUSI	Yes	Yes	CSV, Excel, PDF	Bills, usage	Unlimited history management	Software automatically updates
	Munis	Yes	Yes	CSV, Excel	Unlimited user-defined fields	Old data is stored in an archive	Clients must download updates from Managed Internet Update (MIU)
	NDS	Yes	No	CSV, Excel	Shutoffs, collections, payment plans, inventory, liens, service history	New data overwrites old data, and old data is sent to an archive	NDS-staff assisted update procedures ("zero" downtime)
	SoftWrite	Yes	No	.FTL, .BIL, DTED, .HGT, .IMG .TIF	Bills, usage	Old data can be stored using the application's Archive function	Files need to be merged with the updated vesion using the application's Merge function
Water Meter Radio Software	Sensus	Yes	No	CSV, Excel	Meter reading	Data history stored on the tranceiver device	No updates needed
	Itron	Yes	Yes	CSV, Excel	Hourly interval meter readings	data remains on device for 40 days before being archived	Software updates only, no AMR updates. Updates take place on Windows 7 or 10
	Neptune	Yes	Yes	CSV, Excel	Maps & meter reading values	Old data is stored in an archive	Works with prior and current generations of meters
	Badger	Yes	Yes	Excel	Customizable reports and ability to select categories - any information that can be given by the meter	Historic consumption data can be accessed at any time	Software automatically updates

While both the meter reading and billing software contain customer water use data, the billing system is the best place to start. In many water systems, the person most familiar with this software is the billing clerk. The clerk will likely not have ever needed to download usage data and therefore may not be familiar with the steps required or options available. Consumption data sets are very large and difficult to manipulate in spreadsheet format. It is well worth the investment in time to determine the capabilities of the software in terms of exporting data, specifically relative to file types, included fields and layout. The goal is to determine the most efficient way to obtain usable data, proceeding in a step-wise fashion as follows:

1. Research software capabilities

Find out which version of the software is installed and contact the software provider's customer representative. Ask for instructions to download software and what options or features can help obtain the data in desired format. Not only is this the best first step, it can save considerable time and effort by finding out if the data can be obtained in usable format directly.

2. Review data

If the usage data is not in a usable format, first check to make sure all options have been explored. If the data cannot be directly downloaded in usable format, it must be manipulated in either Excel or Crystal Reports. Crystal Reports is a software package designed to create reports from a variety of data sources. This will likely require less effort than Excel. Many communities have this capability, or this service can be obtained by third-party providers.

APPENDIX B – MASSGIS

A Geographic Information System (GIS) can be described as a hybrid mapping / spreadsheet software platform. By combining these two features, data can be viewed and analyzed quickly and easily. This technology has been broadly adapted in Massachusetts and most communities have some level of GIS capability; if not, consult with the local regional planning agency. For a water provider, developing a water-related GIS is a worthwhile investment; however, it is important to note that this is not required to take advantage of this technology or the abundance of available data.

MassGIS is a state agency that maintains vast libraries of geospatial (GIS) data. Much of the intermediate and advanced customer data described in Section III *Phase 2: Planning and Preparation* are available through MassGIS at no cost. One of the datalayers available is the Standardized Assessors' Parcel database[5] for all Massachusetts communities. Parcels within the database each have a unique identifier and all attributes relate to that identifier. Each parcel is associated with a street address, which can be matched to the street addresses in the water customer database[6]. The parcel database also includes information on lot size, year built, number of living units, zoning district, among other data.

Land use codes are also available for each parcel; codes are based upon a guidance document published by the Department of Revenue's Division of Local Services[7]. Land use codes enable a water system to analyze each class of water customers in a very detailed fashion and can be used to identify the single-family residential customers quickly and easily.

While GIS is commonly thought of as a map-based application, all information from the attribute table can be readily exported to a spreadsheet format. Once developed, this spreadsheet can be used to supplement water usage data to perform all of the analyses described in this document.

---

[5] https://docs.digital.mass.gov/dataset/massgis-data-standardized-assessors-parcels

[6] Often, the street names found in a water system customer database will not exactly match the 'official' street names used by MassGIS. The most common reason for this is that streets are referred to differently and sometimes abbreviations are used locally (i.e., "Mass Ave" vs. "Massachusetts Avenue"). Some processing will be required to correctly match addresses across the two databases.

[7] https://www.mass.gov/files/documents/2016/08/wr/classificationcodebook.pdf

APPENDIX C – SAMPLE DATA SET

The sample dataset used to illustrate the analyses in Section IV *Phase 3: Data Analysis and Interpretation* came from a Massachusetts water system with the following characteristics:

Population:	24,000
Average Household Size:	2.5
Water Customers:	7,484
Residential Usage (RGPCPD):	63
UAW:	5%
Summer/Winter Ratio:	1.68
Total Annual Usage:	650 MG
Percent Residential:	95%
Read Frequency:	Monthly
Bill Frequency:	Quarterly