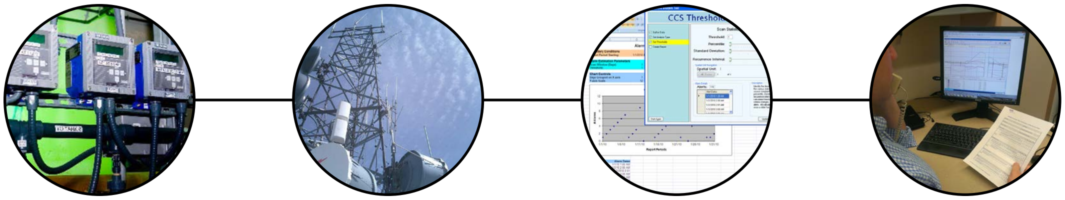
United States
Environmental Protection
Agency

Exploratory Analysis of Time-Series Data to Prepare for Real-Time Online Water Quality Monitoring



Office of Water (AWBERC, MC-140) EPA 817-B-16-004 November 2016

Disclaimer

The Water Security Division of the Office of Ground Water and Drinking Water has reviewed and approved this document for publication. This document does not impose legally binding requirements on any party. The information in this document is intended solely to recommend or suggest and does not imply any requirements. Neither the U.S. Government nor any of its employees, contractors, or their employees make any warranty, expressed or implied, or assume any legal liability or responsibility for any third party's use of any information, product, or process discussed in this document, or represents that its use by such party would not infringe on privately owned rights. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Questions concerning this document should be addressed to WQ_SRS@epa.gov or:

Steve Allgeier
EPA Water Security Division
26 West Martin Luther King Drive
Mail Code 140
Cincinnati, OH 45268
(513) 569-7131
Allgeier.Steve@epa.gov

Acknowledgements

The document was developed by EPA's Water Security Division, with additional support provided under EPA contract EP-C-15-012. Primary authors of this document are:

- Steve Allgeier, EPA, Water Security Division
- Katie Umberg, CSRA

Peer review of this document was provided by the following individuals:

- Brandon Grissom, San Francisco Public Utilities Commission
- Yeongho Lee, Greater Cincinnati Water Works
- David Travers, EPA, Water Security Division
- Matt Umberg, EPA, Water Security Division

Table of Contents

List of Figures

List of Tables

Section 1: Introduction

Online Water Quality Monitoring[1] (OWQM) analyzes ***real-time*** data collected from ***monitoring locations*** in a source water or distribution system to detect unusual water quality and generate information useful to utility operations. This data is reviewed by utility personnel, both as part of regular data review and in response to automated OWQM ***alerts***, as described in the *OWQM Primer*.

Personnel responsible for reviewing ***OWQM data*** must be able to differentiate between normal patterns and anomalous conditions requiring investigation and corrective action. This task is often not straightforward, as water quality can vary significantly even under normal conditions.

Figure 1-1 presents an example of data from a distribution system monitoring location with variable water quality. Even experienced drinking water utility personnel may not be able to determine whether or not the decrease in chlorine beginning on 5/24 is indicative of unusual water quality.

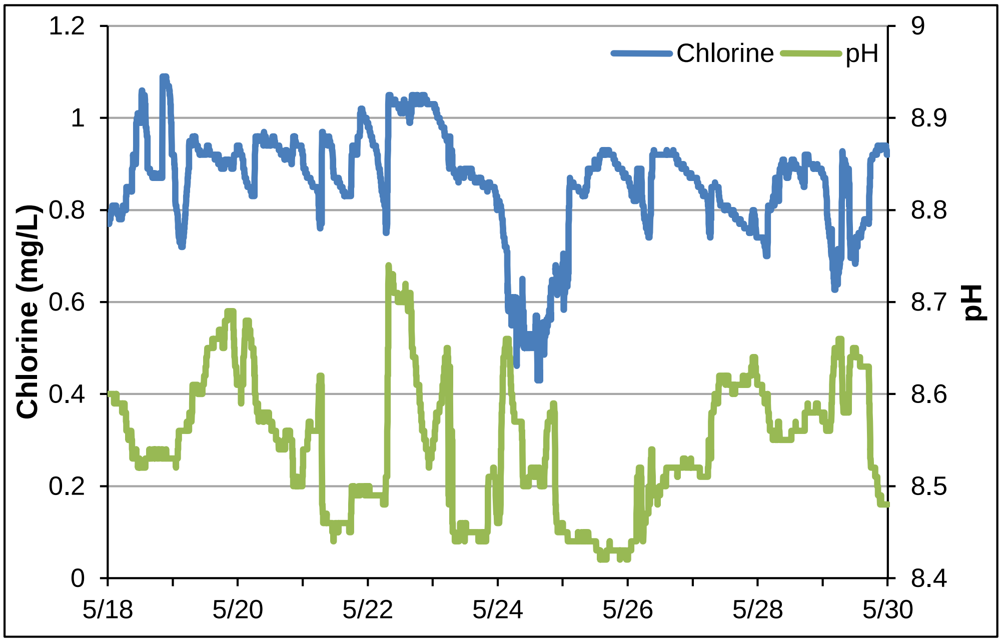


Figure 1-1. Example of OWQM Data Variability

This document describes an ***exploratory analysis*** of OWQM data collected from a drinking water distribution system. An exploratory analysis uses historical data (data that has already been collected and stored) to methodically characterize normal variability and identify the factors that impact water quality at each monitoring location. The intended audience for this document is drinking water professionals who will coordinate and manage an exploratory analysis.

While this document focuses on analysis of distribution system water quality data, the overarching analysis framework can be applied to other ***time-series data***, such as source water quality data.

The remainder of the document covers the following topics:

- **Section 2** describes preparatory steps for an exploratory analysis.
- **Section 3** provides guidance on conducting an exploratory analysis.
- **Section 4** demonstrates how the findings of an exploratory analysis can be used to develop materials to train and support utility personnel responsible for reviewing OWQM data.

---

[1] Words in bold italic font are terms defined in the Glossary at the end of this document.

Section 2: Prepare for Exploratory Analysis

This section describes the following steps to prepare for an effective and efficient exploratory analysis:

1. Establish objectives for the analysis (Section 2.1)
2. Identify personnel who will support the analysis (Section 2.2)
3. Identify analysis tools (Section 2.3)
4. Develop evaluation datasets (Section 2.4)

2.1 Establish Objectives for the Exploratory Analysis

Objectives define the outcome a utility expects to achieve through an exploratory analysis. They serve to focus the analysis and ensure that the results are useful to a utility. These objectives may be informed by the overarching design goals established for the OWQM component, as described in the *OWQM Primer*.

Common objectives for an exploratory analysis of distribution system water quality data include:

- Characterize typical water quality and variability at each monitoring location
- Establish and verify system flow paths and the resulting ***hydraulic connectivity*** among monitoring locations
- Identify the manner in which system operations and source water quality impact distribution system water quality
- Develop tools and resources to support real-time management of distribution system water quality
- Develop tools and resources to support the investigation of OWQM alerts

2.2 Identify Personnel who will Support the Analysis

Accurate interpretation of OWQM data often requires consideration of a variety of factors such as treatment and distribution operations, distribution system flow paths, and ***water quality instrument*** performance. Thus, it is valuable to engage utility personnel with a variety of expertise in the exploratory analysis, such as those described in **Table 2-1**.

Table 2-1. Roles that Commonly Support an Exploratory Analysis

Role	Support Provided to an Exploratory Analysis
Water quality specialist	Evaluate and characterize typical system water quality
Instrument technician	Document instrument issues and maintenance
Distribution system operator	Describe distribution system operations and flow paths
Distribution system maintenance supervisor	Provide information about distribution system maintenance activities and upsets
Distribution system modeler	Conduct simulations to better understand system flow paths and water quality
Data analyst	Use statistics, visualization techniques, and data analysis tools to analyze data
IT specialist	Provide access to data, and assist in the identification of data quality issues due to IT or communications problems

Individuals assuming these roles may support the analysis at different levels, with some involved during all steps of the analysis and others consulted only as needed. In some cases, an individual may assume responsibility for multiple roles during an exploratory analysis. A kickoff meeting can provide the opportunity to convey the project objectives, approach, and schedule, as well as each member's responsibilities and expected level of effort.

Given the variety of roles that may be involved in an exploratory analysis, it is useful to appoint an analysis coordinator responsible for:

- Identifying utility personnel with the right skills and knowledge to support the analysis
- Managing analysis activities and coordinating efforts among team members
- Coordinating with those responsible for overall design and operation of the OWQM component to ensure analysis efforts are aligned with OWQM design goals

2.3 Identify Analysis Tools

An exploratory analysis relies on tool(s) for accessing, analyzing, and viewing the evaluation datasets. These tools should provide analysts with the ability to:

- View the raw data
- Run queries and conduct basic statistical analyses
- Create visual representations of the data including time-series plots, histograms, and box-and-whisker plots
- Export and print data, analysis results, and figures

Time-series plots form the basis of many analysis activities, thus the ability to effectively plot data is particularly important for an exploratory analysis. Ideally, analysts should be able to create custom plots in which they can:

- Change the y-axis scale (for example, change the maximum value of chlorine data displayed from 1.5 to 2 mg/L), which is important if data values change significantly over time
- Change the x-axis scale to view data at different time scales (for example, display one month, one week, or one day of data), which supports the analysis of short-term variability and long-term trends
- View multiple parameters on a single plot, which allows users to investigate if and how parameters change together

An efficient way for analysts to access OWQM data is through an information management system, such as a supervisory control and data acquisition system, with a ***user interface*** designed to facilitate data review. However, access to multiple information management systems may be necessary if data needed for the exploratory analysis is stored in separate systems (for example, if some data is stored in a laboratory information management system). While it may take some effort to develop or configure system(s) to support an exploratory analysis, it can significantly decrease the overall amount of time required to conduct the exploratory analysis.

If the information management system(s) cannot be directly used to support the exploratory analysis due to constraints such as the inability to modify the user interface or limited access to the system due to cybersecurity policies, other methods can be used. For example, reviewers could be provided with the raw data, which could then be analyzed in external data analysis programs. This manual approach could be made more efficient by assigning one person the responsibility for extracting data and generating predefined plots, tables, and other outputs for use by all analysts.

2.4 Develop Evaluation Datasets

This section describes development of ***evaluation dataset(s)***, which contain the data that will be used for the exploratory analysis. The process of developing datasets includes the following activities:

1. Identify data available for analysis (Section 2.4.1)
2. Select an evaluation period (Section 2.4.2)
3. Assemble evaluation dataset(s) (Section 2.4.3)

2.4.1 Identify Data Available for Analysis

During this activity, ***datastreams*** that have been collected and are available for analysis are identified. The primary datastreams of interest are the OWQM data generated at distribution system monitoring locations, which are determined by the design of the OWQM component, specifically the monitoring locations and parameters.

A variety of supplemental datastreams, such as those described in **Table 2-2**, can be useful when considering the causes of changes in water quality. However, for these datastreams to be considered in the exploratory analysis, the analyst must have access to the data in a format that can be extracted.

Table 2-2. Types of Supplemental Information that can Support an Exploratory Analysis

Information Type	Example Datastreams	Value to an Exploratory Analysis
Instrument performance indicators	Instrument error codes, remote diagnostics, data quality indicators	Used to tag data of suspect quality
Communications system status indicators	Communications status between instruments and the information management system	Used to tag data of suspect quality or explain missing data
Instrument maintenance records	Log of instrument maintenance activities and record of known instrument issues	Used to tag data of suspect quality or explain missing data
Grab sample results	Results from field measurements of chlorine residual and pH	Used to determine if OWQM data is accurate
Treatment process operational data	Treatment process settings, chemical doses, process effluent monitoring data	Supports the investigation into the cause of a water quality change and identification of finished water source(s) supplying each monitoring location; treatment process settings impact finished water quality, which in turn is reflected in distribution system water quality
Distribution system operational data	Tank levels, pump status, valve status, system flow and pressure data	Supports the investigation into the cause of a water quality change and the impact of operations on water flow paths; changes in operations often cause changes in system flow paths and corresponding changes in the water quality measured at a monitoring location
Modeling results	System hydraulics and flow paths	Provides insight into flow paths between monitoring locations
Record of distribution system work	System maintenance records and work orders	Supports the investigation into the cause of a water quality change; system work or upsets can impact system flow paths and water quality
Customer complaints	Customer complaints about water quality issues, such as taste, odor, or appearance	Supports the investigation into the cause of a water quality change; some water quality changes impact water aesthetics which can be detected and reported by customers
Calendar of regional events	Date and location of large community events	Supports the investigation into the cause of a water quality change; large events which significantly alter water demand in a specific area can impact system flow paths and water quality

2.4.2 Select an Evaluation Period

An ***evaluation period*** is defined by the start and end dates between which data is compiled and used for an exploratory analysis. The duration of the evaluation period used for an exploratory analysis should be long enough to capture different operational and seasonal conditions.

Figure 2-1 illustrates the importance of analyzing OWQM data over a sufficiently long evaluation period. The figure shows two plots displaying data from the same monitoring location at different times during the year. It is evident that both the average values and variability of chlorine residual and conductivity data are significantly different in February and August.

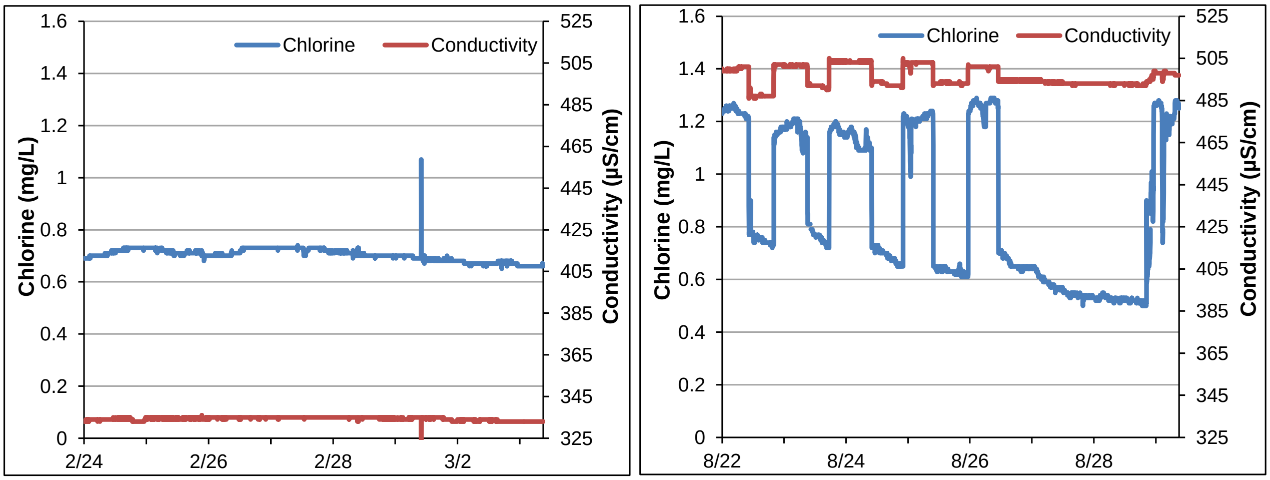


Figure 2-1. Example of Water Quality at a Monitoring Location during Different Time Periods

The recommended duration of an evaluation period is one year, as it shows water quality across all seasons and usually includes all operating conditions. It may be possible to use a shorter time period if (1) finished water quality entering the distribution system varies by less than ~10% from the average value over the course of a year, and (2) there are no significant changes in the configuration or operation of the distribution system over the course of a year.

Once the duration for the evaluation datasets has been determined, the precise start and end date of the evaluation period (for example, May 1, 2016 through April 30, 2017) can be chosen. The most recent data is typically selected for an exploratory analysis based on the assumption that it is most representative of current system conditions including source water quality, utility facilities in use, demand patterns, and pumping schedules.

2.4.3 Assemble Evaluation Datasets

Dataset(s) are assembled that contain data for all datastreams during the selected evaluation period and in a format compatible with the data analysis tool(s) that will be used. This may require data transformations such as:

- **Converting data to acceptable file format(s)**. Common formats include database files, spreadsheets, and formatted text files. Data may need to be stored in a single file, or be separated into multiple files, such as a file for each monitoring location.
- **Converting values to consistent measurements**. For example, converting all temperatures to Fahrenheit or all chlorine residual measurements to mg/L.
- **Ensuring valid values**. Some analysis tools cannot accept certain inputs such as blank cells, non-numeric values, or negative values.
- **Enforcing a consistent time interval**. The datastreams selected for evaluation may use different timesteps and may be missing values, which can create a problem for data analysis tools that require data to be reported for all datastreams at a constant, consistent timestep (such as every 5 minutes). A consistent time interval across datastreams may be achieved by averaging values for datastreams that are measured more frequently than the desired timestep or interpolating values for datastreams that are measured less frequently than the desired timestep.

SATISFYING DATA QUALITY OBJECTIVES

Data quality objectives are metrics or criteria that establish the quality and quantity of data needed to support its intended use. It is assumed that data quality objectives, particularly data ***accuracy*** and ***completeness***, were developed before OWQM data collection was initiated and that the data to be used for the exploratory analysis meets these objectives.

Data that does not meet the data quality objectives should be tagged and excluded from analysis since it may not represent true water quality. Data may be screened to remove inaccurate values, such as those:

- That fall outside of an instrument's meaningful measurement range, such as negative values
- That are concurrent with an instrument or communications system malfunction
- During periods where data values are not changing or are changing at a highly improbable rate or magnitude
- That have been flagged as invalid by the information management system or system operators

Though not included in the primary analyses, review of tagged data can provide valuable information. For example, it is useful to note data patterns or parameter values associated with specific instrument issues so that a malfunctioning instrument can be easily identified during subsequent data review.

Section 3: Conduct Exploratory Analysis

This section describes the process of conducting an exploratory analysis. A fundamental objective of the analysis is to characterize normal water quality variability at monitoring locations. The following factors determine water quality at a monitoring location at any given time:

- The quality of ***finished water*** entering the distribution system. ***Finished water sources*** include treatment plants, groundwater wells, and interconnections with other distribution systems.
- Flow paths of water throughout the distribution system, which determine the finished water source(s) as well as ***storage facility(ies)*** that supply water to a monitoring location. Flow paths are largely determined by system operations and water demand.
- Changes in water quality as it flows through the system, which can be caused by:
    - Degradation of disinfectant residual due to increasing water age
    - Chemical reactions, such as nitrification and corrosion
    - Biological activity, including interaction with biofilm and utilization of assimilable organic carbon
    - The introduction of foreign substances through cross connections, infiltration, and other sources of contamination

Often, these factors result in more than one ***water type*** being observed at a monitoring location. Each water type has a unique combination of parameter values that is largely determined by its finished water source(s), flow path, and water age.

An exploratory analysis consists of the following three activities:

1. Characterize typical water quality at each monitoring location (Section 3.1)
2. Characterize flow paths to monitoring locations (Section 3.2)
3. Identify and investigate anomalies in OWQM data (Section 3.3)

3.1 Characterize Typical Water Quality at Each Monitoring Location

This section describes the first activity in an exploratory analysis in which data from each monitoring location is analyzed to characterize typical water quality. Two approaches are described:

1. Analysis of time-series plots to understand typical water quality (Section 3.1.1)
2. Statistical analysis to define the range of typical values for monitored parameters (Section 3.1.2)

It is most effective to begin the analysis with finished water sources, move to distribution system monitoring locations near these sources, and then progress farther out into the distribution system. By analyzing monitoring locations in this order, knowledge gained from the analysis of upstream monitoring locations can be applied to downstream locations.

HELPFUL HINT

Both individual review and group discussion can be useful during an exploratory analysis. For example, individual analysts can complete specific tasks independently (such as reviewing time-series plots for a single monitoring location) and then discuss observations and conclusions as a group.

3.1.1 Analyzing Time-series Plots to Understand Typical Water Quality

Viewing time-series plots is an effective way to get an initial sense of the range of water quality values typically seen at a monitoring location, as well as the frequency and patterns in water quality. Changes in the water type at a monitoring location can often be identified through a visual analysis of time-series plots, as illustrated in **Figure 3-1**, which shows conductivity and chlorine data from four monitoring locations. Notations on the conductivity datastream in each plot indicate the number of water types observed at the location. Example observations from an analysis of these plots are provided in **Table 3-1**.

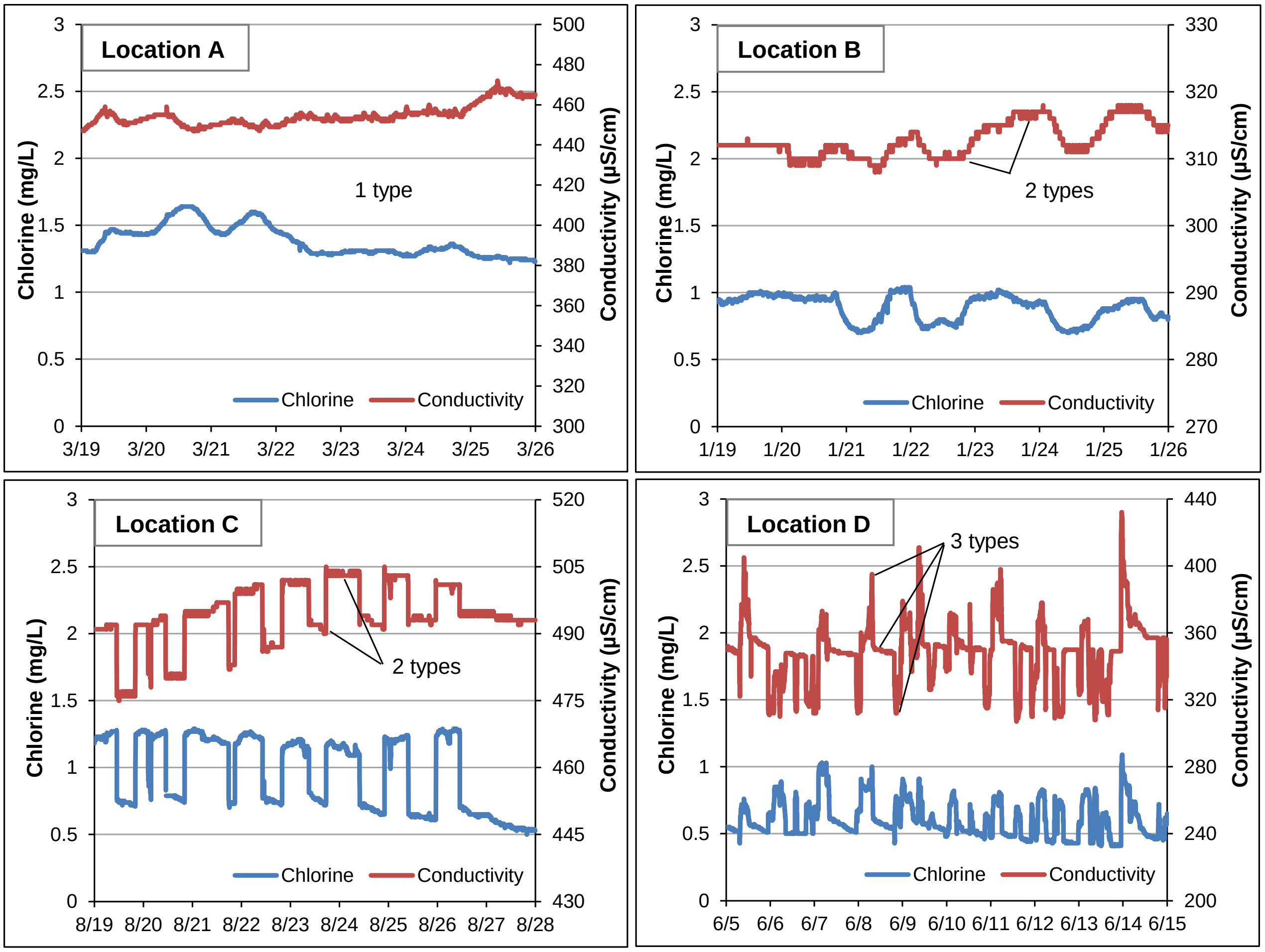


Figure 3-1. Example Time-series Plots of OWQM Data at Four Monitoring Locations

Table 3-1. Example Observations from Analysis of the Time-series Plots in Figure 3-1

Monitoring Location	Number of Water Types	Frequency of Change	Transition Between Water Types	Notes
Location A	1	N/A	N/A	Minor fluctuations due to variability in finished water quality. Changes in chlorine and conductivity values do not occur in unison.
Location B	2	0-1 times per day	Smooth with no clear transition point	One water type has higher chlorine and conductivity than the other
Location C	2	0-3 times per day	Abrupt shift in water quality within ~10 minutes	One water type has higher chlorine and conductivity than the other
Location D	3	1-8 times per day	Abrupt shift in water quality within ~10 minutes	The water type with the middle conductivity and lowest chlorine values is most common, with changes to other types lasting for only a few hours

This initial activity of the exploratory analysis focuses on identifying the number of water types observed at each monitoring location. Identification of the specific origins of these water types will be considered during the analysis activity described in Section 3.2.

3.1.2 Statistical Analysis of Typical Water Quality Parameter Values

In addition to time-series analysis of water quality variability, statistical calculations can provide precise quantification of each parameter's typical range of measured values. **Table 3-2** shows the results of common statistical calculations for Location C (from Figure 3-1).

Table 3-2. Example Statistical Summary of OWQM Data for Location C

Statistic	Chlorine (mg/L)	Conductivity (µS/cm)
5th percentile	0.55	480
10th percentile	0.61	483
25th percentile	0.69	492
50th percentile	1.10	494
75th percentile	1.22	501
90th percentile	1.26	503
95th percentile	1.27	503
Average	1.17	497
Standard deviation	0.28	6.8

While the statistics listed in Table 3-2 provide the precise values of the percentiles, average, and standard deviation for each datastream, they do not indicate how the values vary within the dataset. Other analysis methods can provide more insight into the distribution of data, such as the two example plots shown in **Figure 3-2**.

The plot on the left is a histogram; the bars show the number of times each chlorine value occurred in the dataset. The presence of two water types can be clearly seen as two spikes in the frequency of chlorine values, one from 0.72 to 0.76 mg/L and the other from 1.20 to 1.28 mg/L.

In the plot on the right of Figure 3-2**,** each point shows the chlorine and conductivity values from one timestep. The two water types observed at this location form distinct "cluster**s**" on this plot, where data points are relatively dense.

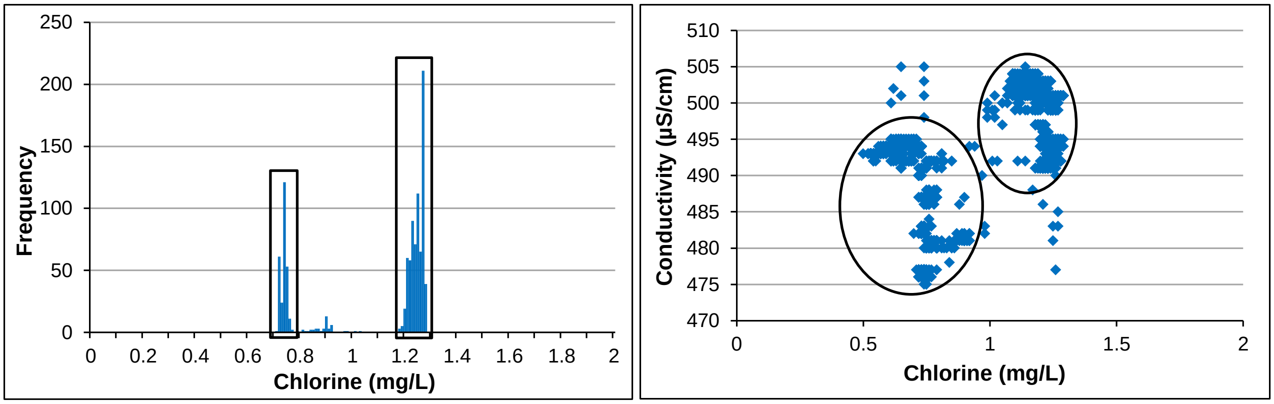


Figure 3-2. Statistical Plots Showing Two Water Types for Location C (from Figure 3-1)

While the recommended evaluation period is one year, the statistical analysis or visual representation of different water types may need to be much shorter. For example the two plots in Figure 3-2 are based on only 10 days of data. When longer time periods are plotted, distinct water types may not be evident because variability in operations and finished water quality over time can result in a wide range of values for each water type, creating significant overlap in the plots. Thus, when analyzing a dataset for water type clusters, the evaluation dataset should be considered in sufficiently short time periods such that water quality at each finished water source, as well as general system operations, remain fairly constant. For example, data can be analyzed separately for each season or operating condition.

3.2 Characterize Flow Paths to Monitoring Locations

To understand the cause of normal water quality variability, it is necessary to determine flow paths between finished water sources, storage facilities, and monitoring locations. This section describes a systematic process for determining flow paths by using:

1. Distribution system maps, operator knowledge, and other resources to identify general flow paths among monitoring locations (Section 3.2.1)
2. OWQM data to confirm and refine the understanding of flow paths among monitoring locations (Section 3.2.2)
3. Operational data to gain an understanding of the operating conditions under which each monitoring location receives water along different flow paths (Section 3.2.3)

3.2.1 Identify General Flow Paths

The first step in characterizing flow paths is to develop a general understanding of likely flow paths using resources such as:

- Operator experience: System operators and engineers often have a general understanding of how water flows through the distribution system under different operating conditions.
- System maps: A distribution system map can be useful for visualizing the spatial and hydraulic relationship among zones and facilities in the system.
- Distribution system model: Model simulations can be used to study system flow paths under different operating conditions.
- Tracer studies: If tracer studies have been performed, they can be used to verify flow paths and the accuracy of model simulations.

It is useful to first identify the general paths water takes through the distribution system, such as between pressure zones or along major transmission and distribution mains. **Figure 3-3** shows an example of a system map annotated to document flow paths. The image on the left shows the distribution system's major pressure zones, color-coded to indicate the finished water source supplying the zone. Arrows indicate general flow directions in the system.

The image on the right in Figure 3-3 is an enlarged view of Detail A that shows greater detail in the flow paths. In Detail A, an arrow on each distribution main, color coded to indicate the finished water source or storage facility, shows the flow direction. In some cases, multiple arrows flow to a single monitoring location, indicating that the location can receive water from different finished water sources or storage facilities, and thus observe different water types.

The distribution system and monitoring locations depicted in Figure 3-3 will serve as the basis for subsequent examples in this section.

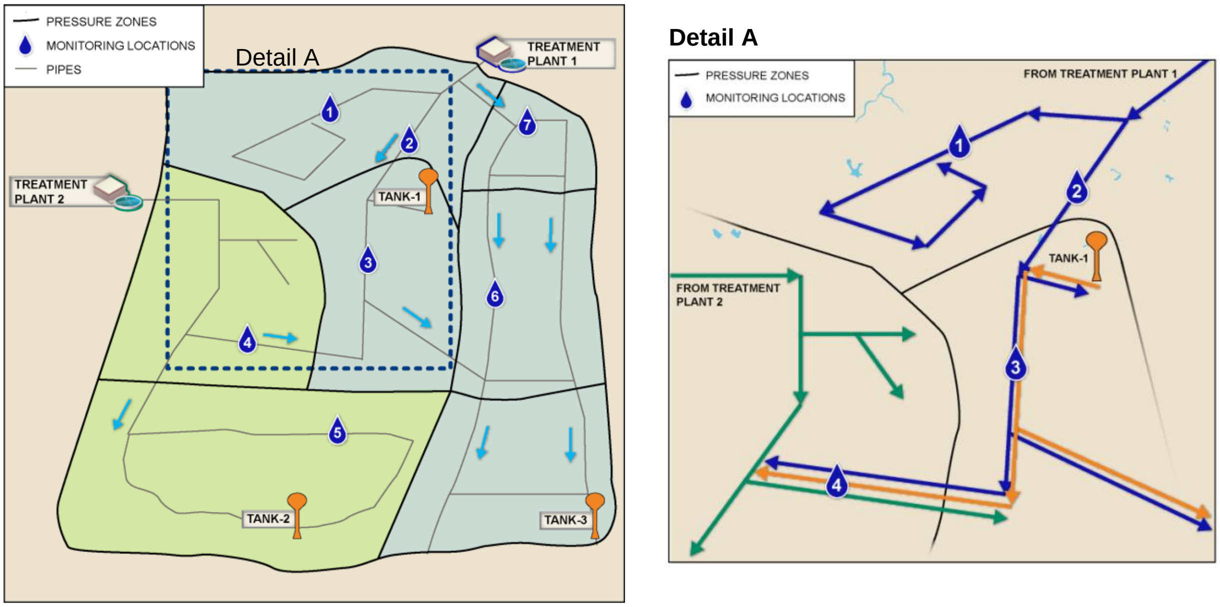


Figure 3-3. Example of a Visual Summary of System Flow Paths

It may become apparent that the number of water types observed at a monitoring location, determined during the analysis of time-series data (Section 3.1), does not match the number of flow paths identified during this step. For example, a location's time-series data may clearly indicate that the water quality is alternating between two water types, while the analysis of flow paths identifies only a single path leading to the location. In such cases, further analysis is necessary to determine if (1) water variability along a single flow path led to the erroneous conclusion that there were two water types, or (2) there is second flow path to the monitoring station that correlates to the second water type. The analysis described in the next section should help to resolve such apparent inconsistencies.

3.2.2 Verify Flow Paths using OWQM Data

In this step, the flow paths identified using the process described in Section 3.2.1 are verified by comparing OWQM data from hydraulically-connected monitoring locations. If two monitored locations are truly hydraulically-connected, then there should be similarities in the water quality observed at these locations, although the impact of water age and the potential for mixing with other sources must also be taken into account. The examples in this section show comparisons between data at monitoring locations and their finished water source(s), although similar analyses could be conducted between any two hydraulically-connected locations.

Analysis of conservative parameters such as TOC, conductivity, pH, and temperature, which are generally stable in a distribution system, can be useful for verifying the finished water source(s) that supply water to a monitored location. **Figure 3-4** presents an example in which conductivity data from Location-4 and Plant-2 (from Figure 3-3) is viewed on a single time-series plot. The fact that conductivity values at Location-4 mirror the fluctuations seen at the plant, with a delay due to the time it takes for water to travel between the locations, confirms that Plant-2 is the primary finished water source for Location-4 during this time period.

On this plot, there are several instances of rapid changes in conductivity at Location-4 that are not observed in the data for Plant-2. These changes, which have been circled in black, indicate a change in water type. Based on the flow paths shown in Figure 3-3, these changes are likely due to either (1) water from Plant-2 traveling through Tank-1 before reaching Location-4, or (2) water from Plant-1 supplying Location-4.

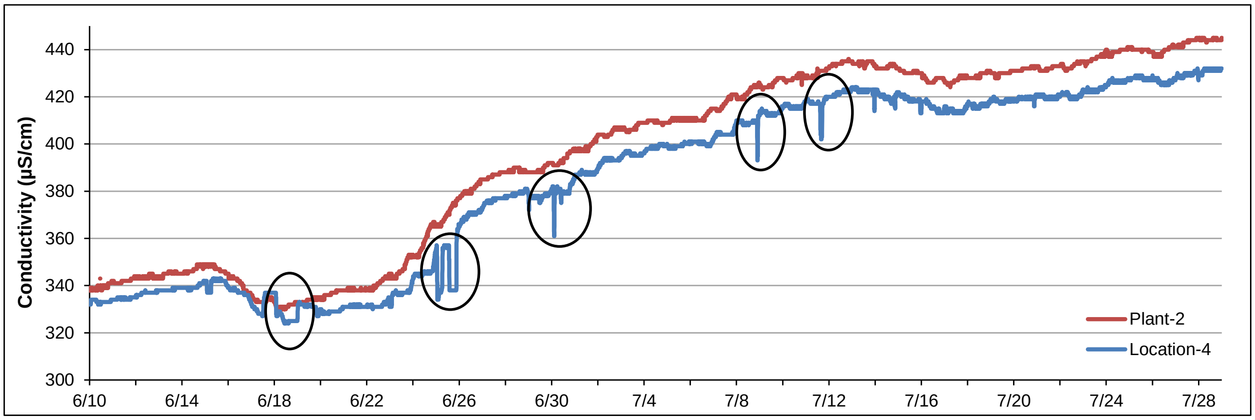


Figure 3-4. Conductivity Data from Location-4 and One of its Upstream Finished Water Sources

Figure 3-5 demonstrates another method for investigating hydraulic connectivity between finished water sources and monitoring locations. This ***box-and-whisker plot*** shows the range of pH values for Plant-1, Plant-2, Location-3, and Location-4 (from Figure 3-3). The boxes for the two plants, which represent the ***interquartile range*** (i.e., the middle 50% of data values), are highlighted by the two shaded bands that extend the width of the plot.

The fact that the boxes for the two plants do not overlap illustrates that the plants have significantly different typical pH values. Thus, analysts can use pH to identify the original source of finished water reaching the monitoring locations. Since the interquartile range of pH values for Location-3 falls within that of Plant-1, it could be concluded that this location receives water exclusively from Plant-1. However, since the interquartile range for pH for Location-4 overlaps the ranges from both Plant-1 and Plant-2, it could be concluded that this location receives water from both treatment plants.

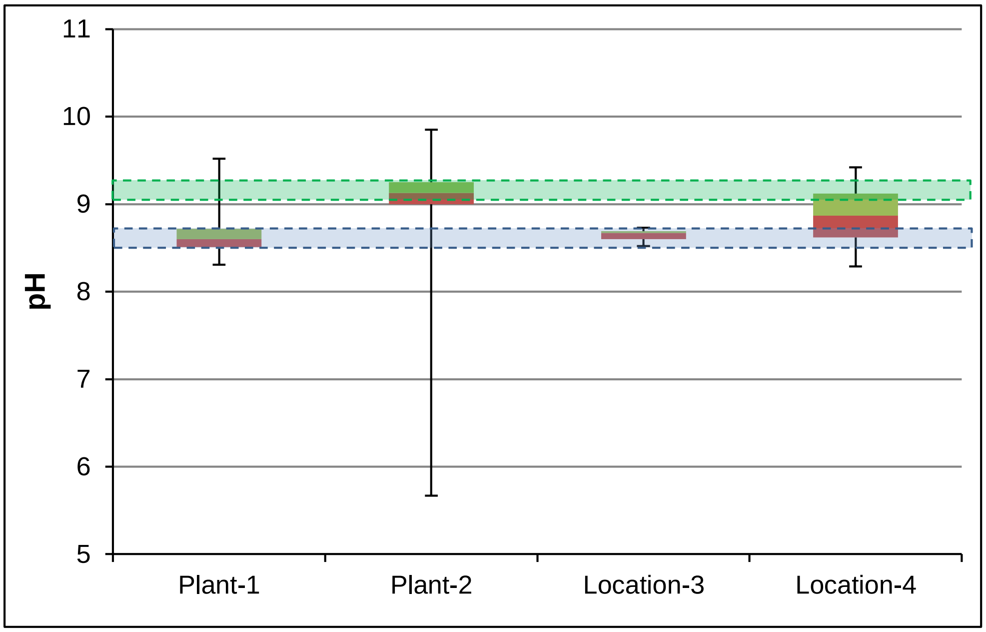


Figure 3-5. Box-and-whisker Plot Comparing Water Quality among Monitoring Locations and Finished Water Sources

Non-conservative parameters that change with water age (such as disinfectant residual) can also be used to verify water flow paths. It would be expected that disinfectant residual concentrations at a monitoring location would be lower than those at upstream locations, and higher than those at downstream locations.

3.2.3 Investigate Causes of Changes in Water Types Using Operational Data

In this step, distribution system operational data is analyzed to identify changes in system operations that alter flow paths to monitoring locations, resulting in a change in the water type observed at a monitoring location. This step of the analysis helps to establish the cause of observed changes in water quality.

The effect of ***operational changes*** on water type observed at a monitoring location can be investigated by viewing time-series plots of water quality and operational data together. **Figure 3-6** provides an example in which periodic changes in water type can be attributed to changes in system pumping. At this location, a shift in water quality occurs approximately 20 minutes after the status of either pump changes (the higher value for pump status indicates that the pump is on).

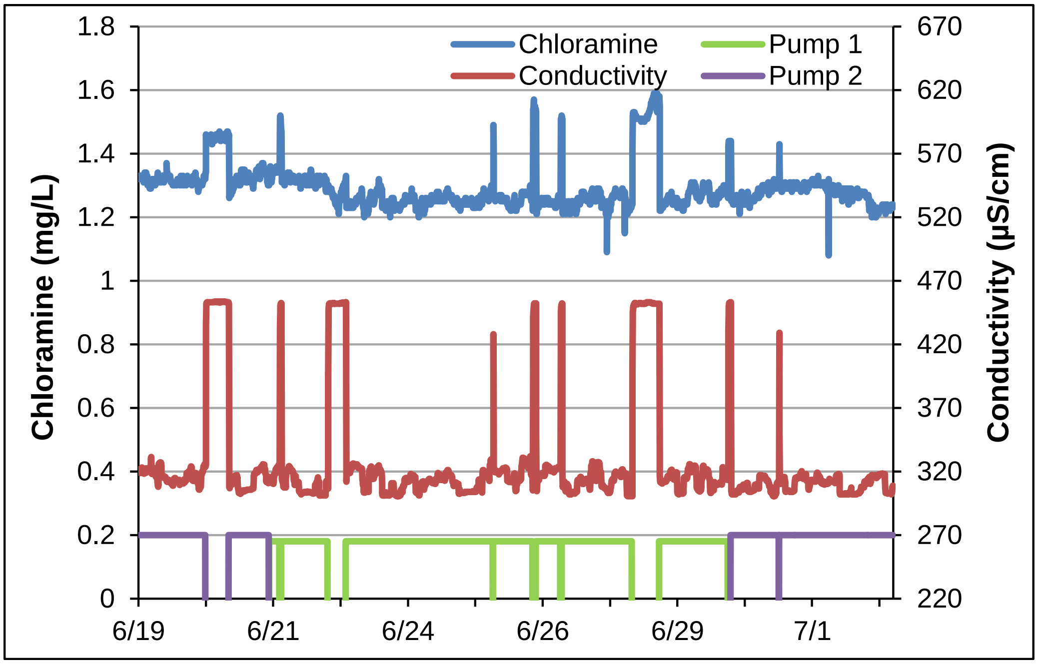


Figure 3-6. Example of the Effect of Pump Operation on Water Type Observed at a Monitoring Location

Figure 3-7 provides another example in which a change in water type is observed approximately one hour after a system valve was temporarily opened (the higher value for "valve status" indicates an open valve).

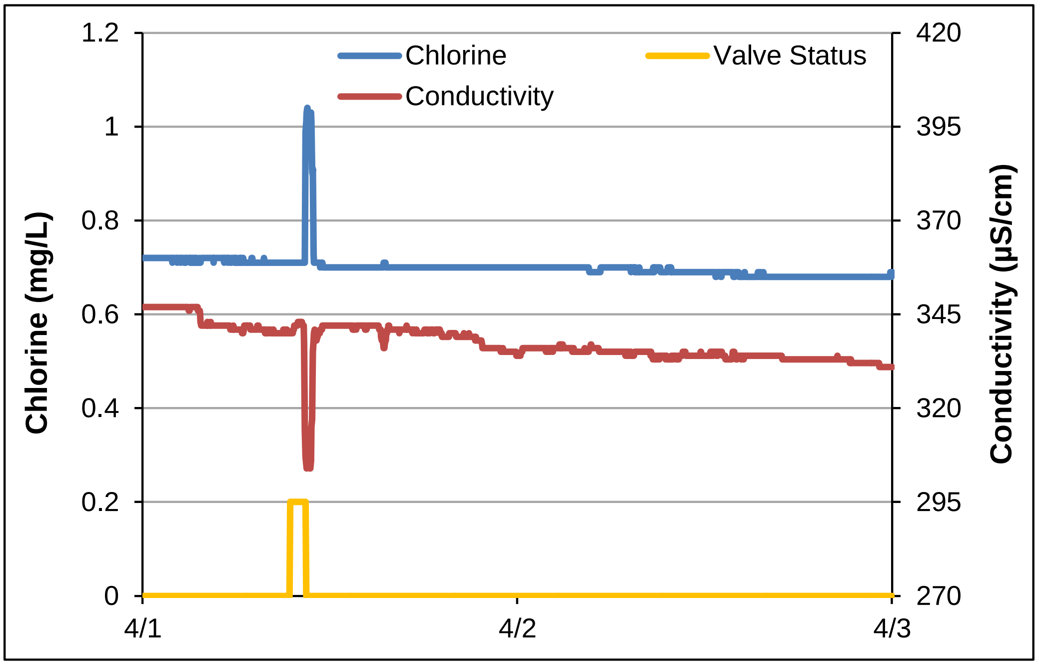


Figure 3-7. Example of the Effect of Valve Status on Water Type Observed at a Monitoring Location

Figures 3-6 and 3-7 show how analysis of time-series data can be used to identify operational changes that impact water quality. An alternative approach is to first identify significant changes in system operations and then review data for downstream monitoring locations during a time frame consistent with the travel time between the location of the operational change and the monitoring location. If the locations are hydraulically connected, the operational change may result in a change in water type observed at the monitoring location.

Table 3-3 summarizes the hydraulic connectivity for Location-3 in Figure 3-3, as determined and verified through the analyses presented in this section. Specifically, the table shows the finished water source(s) and storage facility(ies) that supply water to the area monitored by Location-3, monitoring locations upstream of Location-3, and monitoring locations downstream of Location-3. An example summary for a monitoring location with more complex hydraulic connectivity is provided in Appendix A.

Table 3-3. Example Summary of Hydraulic Connectivity for Location-3

Connectivity Type	Connected Monitoring Location	Comments
Upstream Water Source or Storage Facility	Tank-1	If Tank-1 is draining, water flows to Location-3 from this tank
	Plant-1	When Tank-1 is not draining, Location-3 receives water directly from Plant-1
Upstream Monitoring Locations	Tank-1	The effluent of this tank is an upstream monitoring location for Location-3 only if water is flowing to Location-3 from Tank-1
	Location-2	Location-2 is upstream of Location-3 only when water is flowing to Location-3 directly from Plant-1
Downstream Monitoring Location	Location-4	A downstream location for Location-3 unless Location-4 is receiving water solely from Plant-2

3.3 Identify and Investigate Anomalies in OWQM Data

This section describes the final activity of an exploratory analysis: identifying and investigating ***anomalies*** in the evaluation datasets. An anomaly is an instance where data is different than expected at a monitoring location. Anomalies can be reflected in one or more parameters and can last for time periods ranging from minutes to days. A data anomaly is not necessarily indicative of abnormal water quality. For example, an anomaly could be caused by an instrument malfunction that results in inaccurate data. This step of an exploratory analysis will provide a foundation for the investigation of anomalies during real-time OWQM operations.

Figure 3-8 shows time-series plots of chlorine residual at three monitoring locations (from Figure 3-3). At all three locations, the drop in chlorine residual was determined to be a deviation from typical values and thus an anomaly.

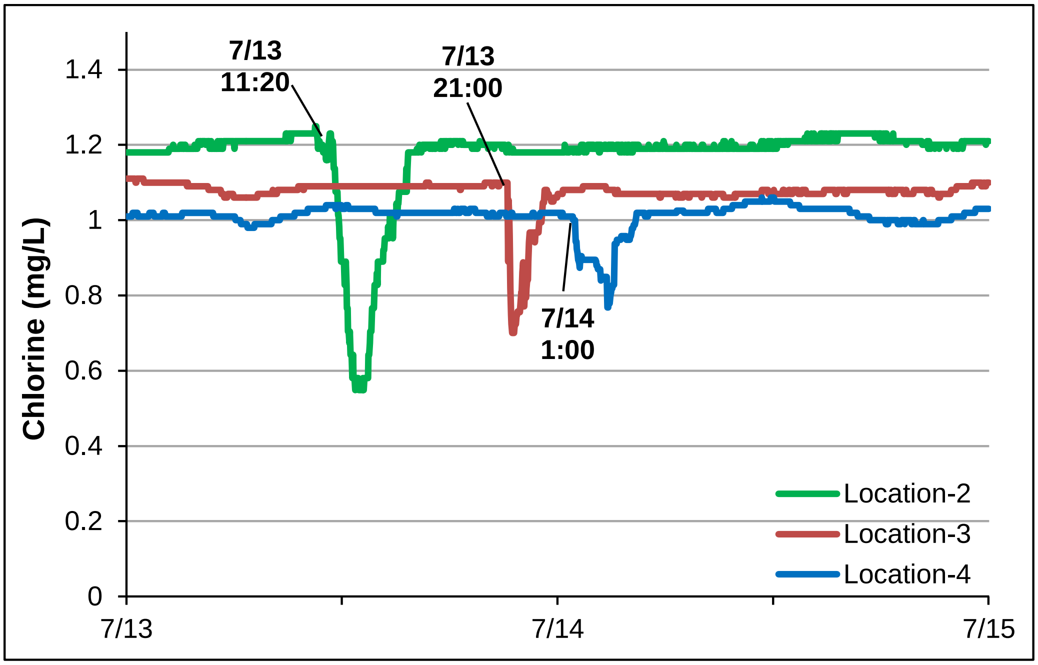


Figure 3-8. Example Anomalies in OWQM Data

An ***anomaly cluster*** is a group of anomalies from different monitoring locations that occur during the same time period and which exhibit similar water quality changes. The three anomalies shown in Figure 3-8 would be considered a cluster since they all reflect a decrease in chlorine and occur during the same time period.

This section describes the following process for identifying and investigating anomalies:

1. Identify anomalies and clusters in the evaluation datasets (Section 3.3.1)
2. Investigate the cause of the anomalies (Section 3.3.2)

3.3.1 Identify Anomalies and Clusters

After completing the activities described in Section 3.1 and 3.2, analysts should have a good sense of typical water quality at each monitoring location and thus be able to identify anomalies in the data. At this point, time-series plots of the evaluation dataset for each monitoring location should be reviewed again to identify anomalies. Some judgment is necessary to distinguish between anomalies and normal variability. However, when in doubt it is better to identify all suspected anomalies at this stage of the analysis. It is generally most effective to review data in two-week intervals for monitoring locations with a high degree of variability and four-week intervals for monitoring locations with relatively stable water quality.

Details of each anomaly should be documented, noting the monitoring location, start date and time, end date and time, and parameters that were anomalous. Documentation of anomalies may also include time-series plots that show the details of each anomaly. The list of anomalies from all monitoring locations will be referred to as the ***anomaly list***.

Table 3-4 provides an example anomaly list that includes the anomalies shown in Figure 3-8, as well as additional examples from the monitoring locations shown in Figure 3-3. While only chlorine was shown in Figure 3-8, this table reflects the fact that a change in other parameters was observed at these three locations. Each anomaly is assigned an identifier that combines the location and start date of the anomaly.

Table 3-4. Example Anomaly List

Anomaly ID	Monitoring Location	Anomaly Start Time	Anomaly End Time	Anomalous Parameters	Notes
Location-1, 4/18/2016-2250	Location-1	4/18/2016 22:50	4/19/2016 9:00	pH	Rapid decrease in pH from 7.8 to 7.2
Location-2, 4/19/2016-800	Location-2	4/19/2016 8:00	4/19/2016 13:00	pH	Gradual decrease in pH from 7.9 to 7.3
Location-2, 7/13/2016-1120	Location-2	7/13/2016 11:20	7/13/2016 23:15	Chlorine, conductivity, turbidity	Rapid decrease in chlorine by 0.6 mg/L Rapid increase in conductivity by 100 µS/cm Rapid increase in turbidity from 0.2 NTU to 1.8 NTU
Location-3, 3/1/2016-1925	Location-3	3/1/2016 19:25	3/1/2016 21:14	Chlorine, conductivity, pH	Short but significant change in listed parameters
Location-3, 4/19/2016-1120	Location-3	4/19/2016 11:20	4/19/2016 17:00	pH	Unusual decrease in pH by 0.2 units
Location-3, 7/14/2016-2100	Location-3	7/13/2016 21:00	7/14/2016 0:20	Chlorine, conductivity	Rapid decrease in chlorine by 0.4 mg/L Rapid increase in conductivity by 70 µS/cm
Location-4, 7/14/2016-100	Location-4	7/14/2016 1:00	7/14/2016 7:50	Chlorine, conductivity	Rapid decrease in chlorine by 0.2 mg/L Rapid increase in conductivity by 30 µS/cm

Once an anomaly list has been developed, it should be methodically reviewed to identify clusters of anomalies as follows:

1. Check the anomaly list for anomalies with similar changes in water quality that occur within the same time period.
2. Check data from both upstream and downstream monitoring locations without a related entry in the anomaly list to see if a similar water quality change occurred but was not previously noted as an anomaly. Viewing the same parameter(s) from multiple locations on the same time-series plot (as illustrated in Figure 3-8) can be effective for identifying clusters.

It is most efficient to begin with anomalies at monitoring locations closest to the primary finished water source(s) and proceed out, following typical hydraulic flow paths. In some cases, anomalous water quality may not be discernable at locations known to be hydraulically connected due to dilution and mixing that can occur with increasing travel time from the source of the anomaly. Also, it can be difficult to detect anomalies at locations with high water quality variability.

HELPFUL HINT

The magnitude and the visual “shape” of water quality changes can vary across monitoring locations for the same cluster. Reasons for this include dispersion, mixing, chemical reactions, and water age.

In **Table 3-5**, the anomaly list shown in Table 3-4 has been reorganized, with clusters of anomalies grouped together. In this table, each anomaly or cluster is identified by the location(s) at which it was observed and the start date of the first anomaly in the cluster.

Table 3-5. Example List of Anomaly Clusters

Cluster Identifier	Anomalous Parameters	Monitoring Location(s)	Anomaly Start Time
Location-3, 3/1/2016	Chlorine, conductivity, pH	Location-3	3/1/2016 19:25
Location-1-2-3, 4/18/2016	pH	Location-1	4/18/2016 22:50
		Location-2	4/19/2016 8:00
		Location-3	4/19/2016 11:20
Location-2-3-4, 7/13/2016	Chlorine, conductivity, turbidity	Location-2	7/13/2016 11:20
		Location-3	7/14/2016 21:00
		Location-4	7/14/2016 5:20

The grouping of anomalies into clusters at this point is based solely on the fact that the anomalies had a similar change in water quality and occurred in the same time period at locations that are hydraulically connected. In the next section, each cluster will be investigated to determine whether or not the anomalies are truly related.

REFINE HYDRAULIC CONNECTIVITY BY ANALYZING ANOMALY CLUSTERS

The flow paths among monitored locations can be verified through analysis of anomaly clusters. If the anomalies in a cluster are indicative of a true water quality change, they should track the flow of anomalous water through the system and be observed at hydraulically-connected monitoring locations in the sequence established through flow path analysis. The time delay between the start (or peak) of an anomaly at hydraulically-connected monitoring locations can be used to estimate hydraulic travel times between the locations.

Each anomaly cluster represents flow under unique system operating conditions. Thus, the analysis of anomaly clusters may reveal a variety of flow paths and travel times. However, if an anomaly cluster reveals unexpected flow paths it may indicate the need to re-evaluate hydraulic connectivity, or the anomalies may be unrelated.

3.3.2 Investigate Causes of Anomalies

In this section, analysts will attempt to identify the cause of each anomaly or cluster using supplemental information resources, such as those listed in Table 2-2. The investigation may also involve personnel with experience and information that is not captured in recorded datastreams. For example, instrument technicians can provide more detail and insight into equipment performance than can be gained by simply reviewing maintenance logs.

Table 3-6 describes common causes of anomalous OWQM data and lists characteristics of an anomaly or cluster that may suggest its cause. The causes in this table are listed in order of the ease with which they can generally be identified. However, the ease of investigation will vary based on the personnel participating in the analysis and the supplemental information available. For example, while it can be easy to identify an anomaly caused by a change in system operations using operational data, it can be quite difficult to establish this causation if operational data is not readily available.

Table 3-6. Common Causes of Anomalies in OWQM Data

Cause of Anomaly	Description	Indicators that this is the Cause
Instrument issue	Inaccurate data values due to an instrument problem; not a true change in water quality	• An anomaly observed at only a single monitoring location • Water quality values or changes that are physically impossible or improbable • Instrument or system faults • Inadequate station pressure or flow • Problems with a station or instrument noted in maintenance logs
Change in finished water quality	A change in distribution system water quality due to a change in the quality of finished water entering the distribution system	• An anomaly cluster which includes all locations receiving water from one finished water source • Water quality changes evident in plant effluent water quality data
Distribution system activities	A change in system flow paths or water quality at a location caused by distribution system work such as flushing operations or pipe repair	• Unusual water types observed at monitoring locations • Records of distribution system work in the vicinity of the monitoring location at which the anomaly was detected • Abnormal system flow or pressure data
Change in system operations	A change in system flow paths caused by unusual system operations, such as changes in pumping or valving	• Unusual water types observed at monitoring locations • Changes in system operations, such as tank, pump, or valve operations
Unusual hydraulic conditions	A change in system flow paths due to unusual hydraulic conditions such as: • Main breaks • Changes in system demand, which could be caused by firefighting activities or large, community events • Low pressure • Pressure transients	• Unusual water types observed at monitoring locations • Abnormal system flow or pressure data
Water contamination	A water quality change due to accidental or intentional introduction of a foreign substance into a distribution system such as: • A cross connection • Contamination of a storage facility • Backflow through a hydrant or service connection • Entry of contaminated fluid or debris through an open pipe	• Unusual results from analysis of grab samples collected from the area during the time of the anomaly • Customer calls noting abnormal taste, odor, or appearance of water in the area • Significant decrease in disinfectant residual or increase in total organic carbon (TOC), ultraviolet (UV) absorbance, or turbidity

Table 3-7 identifies the cause of each cluster listed in Table 3-5 and provides a summary of the investigation to determine the cause.

Table 3-7. Causes of Example Clusters Listed in Table 3-5

Cluster Identifier	Cause of Anomaly	Summary of Investigation
Location-3, 3/1/2016	Instrument issue	An instrument issue was suspected since abnormal OWQM data was observed at only one location. It was confirmed that the shape and duration of the water quality changes seen were consistent with a loss of flow to the monitoring location.
Location-1-2-3, 4/18/2016	Change in finished water quality	• An instrument issue was ruled out because a similar water quality change (a decrease in pH) occurred at multiple locations. • A change in finished water quality was confirmed by reviewing effluent water quality data from Plant-1; a similar decrease in pH was observed in the water leaving the plant.
Location-2-3-4, 7/13/2016	Distribution system activities	• An instrument issue was ruled out because similar changes in chlorine and turbidity were observed at multiple locations. • A change in finished water quality was ruled out because no change in these water quality parameters was seen in the Plant-1 effluent. • When reviewing work logs, it was found that a maintenance activity to replace a large section of distribution main occurred the morning of 7/13. The location of the work was consistent with the monitoring locations that detected the anomaly. Thus, the probable cause is a foreign substance inadvertently introduced into the pipe, which exerted an excessive chlorine demand.

In general, anomalies that comprise a cluster are related to the same underlying cause, with changes in water quality evident at different monitoring locations as the slug of unusual water flows through a distribution system. Thus, the cause identified for a cluster of anomalies applies to each anomaly in the cluster. However, the grouping of anomalies into clusters in Section 3.3.1 was based solely on the fact that the anomalies had a similar change in water quality and occurred in the same time period. If an investigation reveals that some anomalies previously grouped into the same cluster are in fact not related, these anomalies should be documented as discrete, unrelated anomalies.

In some cases, no cause will be able to be identified based on the information available. This should be noted, along with any relevant investigative steps that were completed.

THOROUGHLY DOCUMENT THE FINDINGS OF AN EXPLORATORY ANALYSIS

It is important to clearly and completely document the knowledge gathered during an exploratory analysis. While this detailed documentation is generally not used during day-to-day operations, it is an important resource that:

- Preserves the knowledge developed over the course of the analysis
- Demonstrates the analysis methodology so that it can be reproduced if necessary (such as if a new monitoring location is added or if flow paths to existing locations change)
- Serves as a basis for development of training materials and investigation tools (discussed in Section 4)

Documentation should include any tables, figures, or summaries developed during the exploratory analysis, as well as notes that provide context for the analysis.

Section 4: Apply Findings

This section describes the practical application of results from an exploratory analysis, such as:

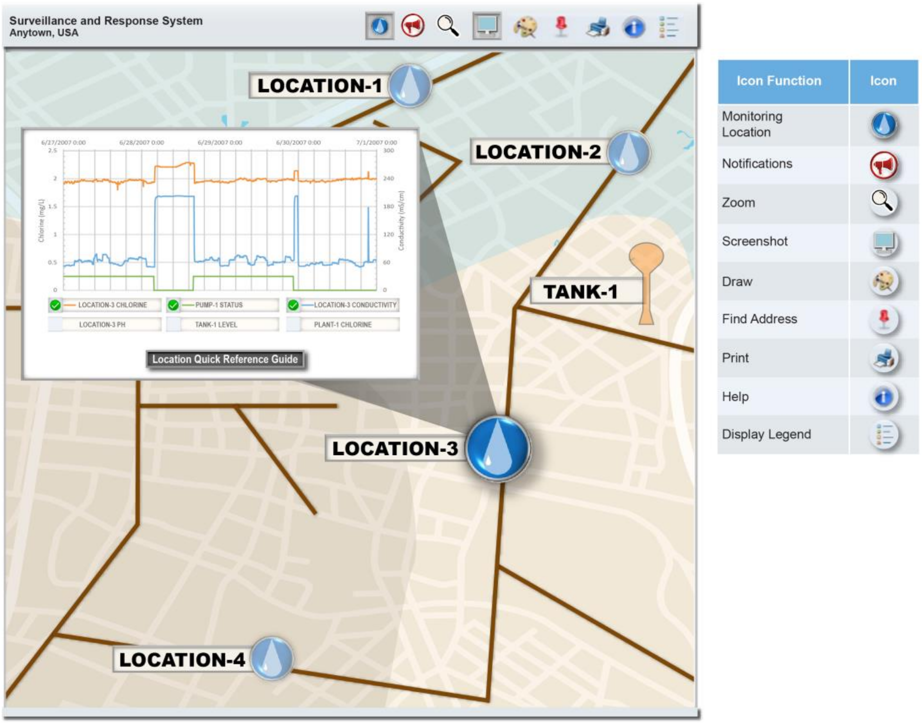
- Develop tools for data review and investigation (Section 4.1)
- Develop training materials (Section 4.2)
- Apply findings to configure anomaly detection systems (Section 4.3)

4.1 Develop Tools for Data Review and Investigation

During real-time operation, utility personnel review OWQM data regularly as part of standard operating procedures, as well as when a water quality anomaly has been detected. The results of an exploratory analysis can be used to develop tools to support these activities.

A user interface available through a utility information management system is usually the main tool investigators use to obtain the information necessary to carry out their OWQM responsibilities. Thus, a powerful way to facilitate data review is to develop clear and informative user interface screens. *Dashboard Design Guidance for Water Quality Surveillance and Response Systems* provides examples of effective user interfaces and guidance on establishing requirements for a dashboard.

Figure 4-1 shows an example user interface designed to facilitate review of OWQM data for the distribution system shown in Figure 3-3. The background figure shows the home screen, which displays monitoring locations and utility facilities on a distribution system map. When a user clicks on the icon for a monitoring location, a pop-up window displays the time-series data plots for that location. The plots are pre-populated with default datastreams, and users can select the time period and datastreams to display. The button titled "Location Quick Reference Guide" at the bottom of this window brings up a summary of the location's finished water sources, hydraulically-connected locations, and relevant data from system operations (not shown).



Icon Function	Icon
Monitoring Location	💧
Notifications	📢
Zoom	🔍
Screenshot	🖥
Draw	🎨
Find Address	📌
Print	🖨
Help	ℹ
Display Legend	☰

Figure 4-1. Example of a User Interface Screen for Viewing OWQM Data

Factsheets or quick reference guides can be effective resources to support real-time OWQM data review. They provide useful information and instructions to guide investigators through complex or less frequently-implemented tasks. Examples of guides that can be derived from the analyses described in this document include summaries of:

- System flow paths and hydraulic connectivity between monitoring locations, including typical travel times (often shown on a distribution system map or in a table)
- Typical water quality and variability at each finished water source and storage facility (often shown in tables or time-series plots)
- Data patterns associated with common OWQM instrument problems (often shown as time-series plots)

An example of a monitoring location quick reference guide that summarizes flow paths and hydraulic connectivity is provided in **Appendix A**.

4.2 Develop Training Materials

Personnel responsible for reviewing OWQM data during real-time operations will need training on the key results and findings from the exploratory analysis. Training materials that can be derived from the knowledge gained during the analysis include:

- Typical water quality at each monitoring location, finished water source, and storage facility
- System flow paths under different conditions, and the hydraulic connectivity among monitoring locations
- System operations that impact the finished water source or storage facility supplying each monitoring location
- Examples of OWQM data patterns indicative of common instrument problems
- Examples of anomalies and clusters that fall into each of the cause categories (e.g., Table 3-6)

Data plots and tabular summaries are an effective way to convey much of this information and can be developed from the evaluation datasets used during the exploratory analysis.

4.3 Apply Findings to Configure Anomaly Detection System(s)

Automated ***anomaly detection systems*** process real-time OWQM data and generate an alert if anomalous data is detected. A properly-configured anomaly detection system ensures that personnel are reliably notified of anomalous water quality without receiving too many ***invalid alerts*** (alerts not caused by abnormal water quality).

The evaluation datasets for each monitoring location are ideal for configuring and evaluating an anomaly detection system, as they capture water quality conditions across time and for different operating conditions. In addition, the anomalies identified and characterized during the exploratory analysis can be used to evaluate the system’s ability to detect real water quality incidents.

Some anomaly detection systems can be configured to reduce the number of invalid alerts based on observations gathered during an exploratory analysis. For example, logic could be developed to suppress an alert when a change in system operations causes a change in the water type flowing through a monitoring location.

IMPLEMENT A PLAN FOR UPDATING OWQM INVESTIGATION TOOLS

To ensure that investigation tools and anomaly detection systems are properly configured for current conditions, they should be reviewed periodically in light of the most recent OWQM data. Typical water quality patterns at individual monitoring locations can change over time for a variety of reasons such as expansion of the distribution system, changing demand patterns, or changes in operating conditions. A plan for maintaining these tools should be developed and an individual should be designated to oversee its implementation. Occasions when these tools should be reviewed include:

- At a pre-defined interval. Quarterly review can be effective so that seasonal changes can be considered.
- When system changes occur that could impact water quality or hydraulics. Examples include bringing a utility facility online or taking it out of service, changing pump and tank operations, or reconfiguring system valves.
- When information in the tools is identified as inaccurate. Over the course of real-time OWQM operations, reviewers may notice that information in tools is outdated or that an anomaly detection system is no longer achieving performance objectives.

Fortunately, subsequent analyses usually do not take as much time and effort as the initial exploratory analysis because the framework, tools, and knowledge gained through the initial analysis and real-time OWQM data review can be leveraged. Also, the investigation tools for only a subset of monitoring locations require updates at any given time.

Glossary

accuracy. The degree to which a measured value represents the true value.

alert. An indication from an SRS surveillance component that an anomaly has been detected in a datastream monitored by that component. Alerts may be visual or audible, and may initiate automatic notifications such as pager, text, or email messages.

anomaly. A deviation from an established baseline in a monitored datastream. Detection of an anomaly by an SRS surveillance component generates an alert.

anomaly cluster. A group of anomalies from different monitoring locations, with similar water quality changes and occurring during the same time period.

anomaly detection system. A data analysis tool designed to detect deviations from an established baseline. An anomaly detection system may take a variety of forms, ranging from complex computer algorithms to thresholds.

anomaly list. The list of all anomalies identified in an evaluation dataset.

box-and-whisker plot. A graphical representation of nonparametric statistics for a dataset. The bottom and top whiskers represent the 10th and 90th percentiles of the ranked data, respectively. The bottom and top of the box represent the 25th and 75th percentiles of the ranked data, respectively. The line inside the box represents the 50th percentile, or median, of the ranked data.

completeness. The percentage of data that is of sufficient quality to support its intended use.

data quality objectives. Qualitative and quantitative statements that clarify study objectives, define the appropriate types of data, and specify the tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions.

datastream. A time series of values for a unique parameter or set of parameters. Examples of datastreams useful during an exploratory analysis include water quality parameter values (e.g., chlorine, pH, turbidity), flow and pressure data, system operational data, and water quality instrument error codes.

evaluation dataset. The data collected from a monitoring location, along with supplemental data, that will be used for an exploratory analysis. An evaluation dataset is defined by the datastreams and time period that will be analyzed.

evaluation period. The time period of previously-collected data that will be used for the exploratory analysis.

exploratory analysis. An off-line data analysis conducted to characterize normal water quality variability and identify factors that impact water quality at monitoring locations.

finished water. Treated drinking water that leaves a treatment facility, groundwater well, or connection with another drinking water system and enters a distribution system.

finished water source. A location at which finished water enters a distribution system.

hydraulic connectivity. The hydraulic relationship between locations in a distribution system. Two locations are hydraulically connected if water flows from one to the other.

interquartile range. In a ranked dataset, the values between the 25th and 75th percentile values. On a box-and-whisker plot, this is designated by the box portion of each plot.

invalid alert. An alert from an SRS surveillance component that is not due to an occurrence of abnormal water quality.

monitoring location. A specific point in a water distribution system or source water where water quality data is collected, such as an OWQM station or a storage facility.

Online Water Quality Monitoring (OWQM). One of the surveillance components of an SRS. OWQM utilizes real-time data collected from monitoring locations that are deployed at strategic locations in a source water or distribution system. Monitored parameters can include common water quality parameters (e.g., chlorine residual, pH, specific conductance, and turbidity) and advanced parameters (e.g., total organic carbon and UV-Vis spectral data). Data from distribution system monitoring locations is transferred to a central location and analyzed for anomalies.

operational change. A change in the way a distribution system is operated, such as changes in pumping, storage facility operations, or valve configuration.

OWQM data. Real-time data generated at OWQM locations.

real-time. A mode of operation in which data describing the current state of a system is available in sufficient time for analysis and subsequent use to support assessment, control, and decision functions related to the monitored system.

storage facility. A structure in a distribution system where water is temporarily held, such as a tank or reservoir.

time-series data. A sequence of data points ordered by the time they were generated, generally collected at a regular interval.

user interface. A visually-oriented interface that allows a user to interact with an information management system, facilitating data access and analysis.

water quality instrument. A unit that includes one or more sensors, electronics, internal plumbing, displays, and software that is necessary to take a water quality measurement and generate data in a format that can be communicated, stored, and displayed. Some instruments also include diagnostic tools.

Water Quality Surveillance and Response System (SRS). A system that employs one or more surveillance components to monitor and manage source and distribution system water quality in real time. An SRS utilizes a variety of data analysis techniques to detect anomalies and generate alerts. Procedures guide the investigation of alerts and the response to validated water quality incidents that might impact operations, public health, or utility infrastructure.

water type. Water with a unique combination of parameter values, largely determined by the finished water source(s), flow path, and water age as water travels from the source(s) to a monitoring location.

Resources

Online Water Quality Monitoring Primer

This document provides an overview of the Online Water Quality Monitoring (OWQM) component of Water Quality Surveillance and Response System. The OWQM component analyzes real-time data collected from monitoring locations in a source water or distribution system to detect unusual water quality and generate information useful to utility operations. EPA 817-B-15-002A, May 2015.
https://www.epa.gov/sites/production/files/2015-06/documents/online_water_quality_monitoring_primer.pdf

Dashboard Design Guidance for Water Quality Surveillance and Response Systems

This document provides information about useful features and functions that can be incorporated into a dashboard designed to support operation of a Water Quality Surveillance and Response System. It also provides guidance on a systematic approach that can be used by utility managers and IT personnel to define requirements for a dashboard. EPA 817-B-15-007, November 2015.
https://www.epa.gov/sites/production/files/2015-12/documents/srs_dashboard_guidance_112015.pdf

Appendix A: Example of a Monitoring Location Quick Reference Guide

This appendix provides an example of a monitoring location quick reference guide that could be developed using the results of an exploratory analysis. Quick reference guides are intended to serve as a resource for personnel who will review data during real-time OWQM operations. It is not intended to be a comprehensive summary of the results of the exploratory analysis, which would include additional details and materials such as time-series plots, summaries of statistical analyses, and documentation of key observations.

This quick reference guide contains:

- A summary of the water types observed at the monitoring location, with general guidelines on the distribution system conditions under which each water type is observed
- Guidance on distinguishing between the water types based on time-series plots of data from the monitoring location
- A list of monitoring locations that are upstream and downstream of this location

This example guide contains only location-specific water quality information. General analysis principles that apply to all locations (such as the usefulness of each water quality parameter in investigating changes in water quality) would be conveyed through trainings or in other documentation.

Also, this example guide is applicable to a specific set of operating conditions. If distribution system operations change at different times of the year (e.g., different valve configurations, pumping schedules, tank cycling, and demand patterns), it may be necessary to develop a separate quick reference guide for each set of conditions.

Quick Reference Guide:
Woodrow Fire Station (Location-17)

Summary of Flow Paths for Location-17

Finished Water Source or Storage Facility	Travel Time between this Facility and Location-17	Distribution System Operating Conditions under which Location-17 Receives Water From this Facility	Supplemental Data to Determine Whether Location-17 is Receiving Water From this Facility
North Park Reservoir (supplied by the Main Treatment Plant)	0.5 – 1 hours	• Location-17 receives water from North Park Reservoir whenever the reservoir is draining	• North Park Reservoir outflow • North Park Reservoir effluent chlorine residual and pH
Main Treatment Plant	4.5 – 7 hours	• If the North Park Reservoir is not draining, water usually comes directly from this plant	• North Park Reservoir outflow • Main Treatment Plant effluent chlorine residual and pH
Bluebird Tank (supplied by the Main Treatment Plant)	1.5 – 3.5 hours	• If the North Park Reservoir is not draining, water may come from Bluebird Tank • Conditions under which this occurs: ○ Bluebird Tank is draining AND ○ Pumps 1 and 3 at Green Ridge Pump Station are turned off AND ○ Valve 08A is closed	• North Park Reservoir outflow • Bluebird Tank water level • Status of Green Ridge Pump Station Pumps 1 and 3 • Status of Valve 08A
Warner Street Treatment Plant	2 – 4.5 hours	• On rare occasions, Location-17 receives water from the Warner Street Treatment Plant • Conditions under which this occurs: ○ North Park Reservoir is not draining ○ Pumps 1 and 3 at Green Ridge Pump Station are turned off AND ○ Valve 08A is open AND ○ Demand in the Northwest pressure zone is high, such as during major sporting events	• North Park Reservoir outflow • Warner Street Treatment Plant effluent chlorine residual and pH • Status of Green Ridge Pump Station Pumps 1 and 3 • Status of Valve 08A • Calendar of regional events

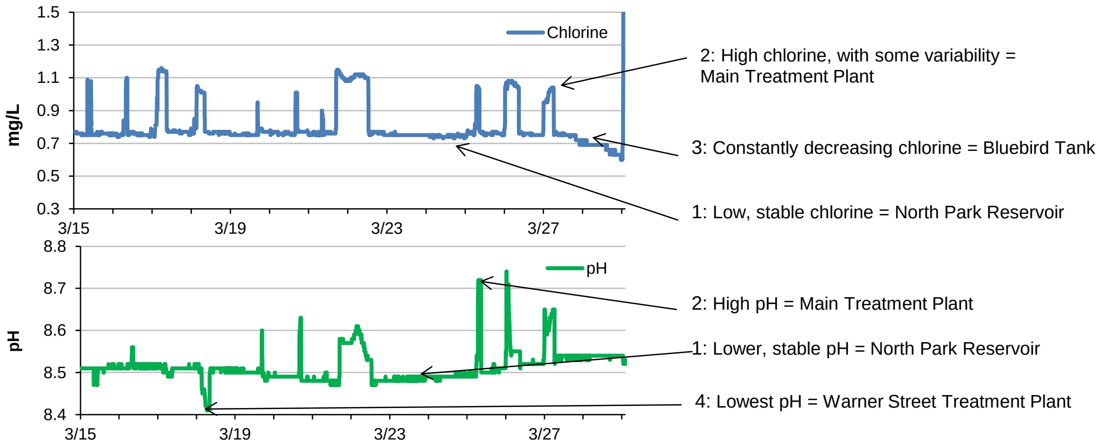
Guidelines for using Time-series Plots to Identify Water Types for Location-17

Water Type Descriptor*	Primary Flow Path	Visual Indicators of the Water Type
Stable with low chlorine (1)	From Main Treatment Plant via North Park Reservoir	• Very stable data for all parameters; there is little variability in the water quality that comes through this reservoir
Highest chlorine and pH (2)	Directly from Main Treatment Plant	• Minor water quality variability, which is caused by variability in the plant effluent water quality • Usually the highest chlorine and pH of all water types • Location-17's water quality tracks closely with this plant's effluent
Low, constantly decreasing chlorine (3)	From Main Treatment Plant via Bluebird Tank	• Chlorine values are constantly decreasing while Location-17 receives this water type, though other parameters are stable • Usually the lowest chlorine across the water types
Low pH (4)	From Warner Street Treatment Plant	• The lowest pH of the water types • Chlorine is usually slightly lower than Main Treatment Plant but higher than the other types • Location-17's conductivity tracks closely with this plant's effluent

* Number in parentheses corresponds to the trends illustrated in the time-series plots below

Plots Illustrating the Above Relationships

While these plots show only 2 weeks of data, the relationships illustrated held throughout the evaluation datasets.



2: High chlorine, with some variability = Main Treatment Plant

3: Constantly decreasing chlorine = Bluebird Tank

1: Low, stable chlorine = North Park Reservoir

2: High pH = Main Treatment Plant

1: Lower, stable pH = North Park Reservoir

4: Lowest pH = Warner Street Treatment Plant

Hydraulic Connectivity between Location-17 and Other Monitoring Points

Upstream Monitoring Locations, Ordered by Travel Time between Locations

Monitoring Location	Average Travel Time from this Monitoring Location to Location-17	Conditions Under which this Location is Hydraulically-Connected to Location-17 (and thus an upstream monitoring point)
University Hospital	30 – 45 minutes	Hydraulically-connected only when water is coming through North Park Reservoir; when they are hydraulically connected, the two locations have very similar water quality
North Park Reservoir	0.5 – 1 hours	Hydraulically-connected when water comes through this reservoir
Grab sample location 43	1 – 1.5 hours	Hydraulically-connected only when water is coming through Bluebird Tank or directly from the Main Treatment Plant
Bluebird Tank	1.5 – 2.5 hours	Hydraulically-connected only when water is coming through this tank
Warner Street Treatment Plant	3 – 5.5 hours	Hydraulically-connected only when water is coming from this treatment plant
Main Treatment Plant	4.5 – 7 hours	Hydraulically-connected only when water is coming from this treatment plant

Downstream Monitoring Locations, Ordered by Travel Time between Locations

Monitoring Point	Average Travel Time from Location-17 to this Monitoring Location	Conditions Under which this Location is Hydraulically-Connected to Location-17 (and thus a downstream monitoring point)
Grab sample location 45	3 – 4.5 hours	Sometimes hydraulically-connected, though the conditions under which it is are unclear
Grab sample location 78	5 – 7.5 hours	Always hydraulically-connected
Maple St. Fire Station	5 – 9 hours	Hydraulically-connected when Pinewood Pump Station is operating