Accuracy of Real-Time Network Water Quality Models

Developed by the University of Cincinnati

Prepared for the
National Institute of Hometown Security
368 N. Hwy 27
Somerset, KY 42503

December 2013

This research was funded through funds provided by the Department of Homeland Security, administered by the National Institute for Hometown Security Kentucky Critical Infrastructure Protection program, under OTA # HSHQDC-07-3-00005, Subcontract # 02-10-UK

Contents

1 Introduction **3**

2 Tracer Field Study Description **3**
2.1 Conductivity Monitoring 4
2.2 Injection Protocol 8

3 Real-Time Simulation of Tracer Movement **10**
3.1 Tracer Data Processing 10
3.2 Real-Time Water Quality Model 11
3.3 Accuracy Metrics 15
3.4 Summary Results 17
3.5 Observed and Simulated Tracer Signals 20
3.5.1 Region A 20
3.5.2 Region B 21
3.5.3 Region C 24
3.5.4 Region D 27
3.5.5 Region E 27
3.5.6 Region F 27

4 Acknowledgements **33**

A Appendix A: Prediction accuracy of chloride levels based on measured specific conductance **39**

B Appendix B: NSF Grade Liquid Calcium Chloride Product Data Sheet **41**

References **42**

1 Introduction

This report describes a field scale evaluation of a real-time hydraulic and water quality model of the Northern Kentucky Water District (NKWD) water distribution system. The data used include both those routinely collected through the District's Supervisory Control And Data Acquisition (SCADA) system, as well as field test data from injecting and tracking a series of salt pulses in a large section of the distribution system during November, 2012.

Tracer tests are the preferred method for calibration and validation of network hydraulic and water quality models, such as models to predict chlorine residual and trihalomethane formation. In addition, the transport of salt pulses may mimic the temporal signatures of contaminants intentionally introduced into a distribution system; thus these data may be used to evaluate whether real-time data analytics and models can enhance water security though support of event detection and emergency response.

The field activities included applying a calcium chloride ($CaCl_2$) solution as a series of pulses over a 12 hour period. The movement of the $CaCl_2$ pulses was observed by 38 continuous specific conductance meters located in the distribution system to provide information about the passage of the pulses at high spatial and temporal resolution. Operational data from the SCADA system was used to drive a real-time network hydraulic model, which underlay the real-time water quality model; the development and accuracy of the real-time hydraulic model was described previously (CitiLogics, 2013).

The field experiment described here is one of a few distribution system water quality studies that attempted to follow a large volume of finished water through an extensive portion of the distribution system. Tracer data provide unique information about processes that affect water quality in the distribution system, including water velocity, junction mixing, and flow path-dependent effects, and can rigorously evaluate the real-time network hydraulic and water quality models. It is the first such study to specifically use real-time modeling to drive the tracer simulations, and thus evaluate the fidelity of real-time simulation data processing techniques. The study design represents a severe test of model accuracy, as 24 of 38 monitors were located on small diameter distribution mains (17) or dead-end mains (7); thus the test is not evaluating only the ability of a real-time model to predict movement through the transmission main infrastructure, but is also evaluating the accuracy at a neighborhood scale.

The organization of this report is as follows. Section 2 describes the field test design and protocol, including details of the brine addition and field specific conductance monitoring. Section 3 outlines the approach to tracer data analysis and real-time simulations, and summarizes the accuracy of simulated tracer signals versus observed. Appendix A provides the supporting analysis and calculations for monitoring chloride levels using specific conductance monitors, as well as a fact sheet on the NSF food grade tracer which was used.

2 Tracer Field Study Description

Northern Kentucky Water District (NKWD) serves approximately 81,000 customer accounts, or nearly 300,000 people in Campbell and Kenton Counties, portions of Boone, Grant and Pendleton Counties, and the Greater Cincinnati Northern Kentucky International Airport.

It covers over 300 mi^2 of total service area through 1,282 miles of pipe. The three water treatment plants: Fort Thomas Treatment Plant (FTTP), Taylor Mill Treatment Plant (TMTP), and Memorial Parkway Treatment Plant (MPTP), have a combined capacity of 64 Million Gallons Per Day (MGD), and supply water through 16 pump stations and 20 storage tanks. A overview map of the portion of the NKWD service area that comprises the tracer study area is shown in figure 1. The area is divided into six regions labelled A to F, going from north to south, and monitoring at 38 locations was distributed throughout the study area, but also concentrated in specific regions to gather information about spatial variation in tracer transport. The original study design included 46 monitors and all were installed, but 8 monitors malfunctioned without storing any data. Flow to the study area originates at the south treatment plant, flowing north and south through transmission lines within a single pressure zone, before descending through regulating valves to a lower zone in the north containing monitoring sites A and B. The north treatment plant was not operating its high service pumps during the test, so that all of the flow to the monitors would be tagged by the brine pulses. This "un-split" mode is one of the normal operating modes for the utility, although the system is also operated in a split mode that requires the north plant to deliver water to regions A-D through its high service pumps. A more detailed description of the study area hydraulic behavior is available in CitiLogics (2013).

The tracer experiments included a calcium chloride tracer test applied as a series of four pulses over a 14 hour period. The injection pulses were designed to produce a specific conductance of 1000 μS/cm, more than twice the background level of approximately 350 μS/cm, though the peak conductivity achieved was somewhat less than the design value.

All sampling locations were located in the field at fire hydrants using standard hydrant adaptors, and a continuous flow rate of approximately 1.0 GPM was maintained to reduce residence time in the hydrant barrel to approximately 15 minutes. Each sampling location included a continuous conductivity monitor housed in a secure container (see Figure 2). Specific conductance data were downloaded from data loggers periodically. The discharge line from each hydrant was positioned to ensure that it drained to a sewer (if present) or to an area that allowed infiltration.

2.1 Conductivity Monitoring

The objectives of the monitoring location selection process were to identify locations that represented the range of hydraulic residence times while being spatially diverse. A secondary objective was to concentrate monitors in one or more more densely populated regions, so that variability over small areas could be assessed. Using the distribution system network model provided by the utility, water age and tracer simulations were performed using EPANET (Rossman, 2000), to provide input for monitor location.

Monitor placement was performed manually to locate a total of 46 conductivity sensors. While the overall intent was to place sensors to provide representative sampling locations with respect to spatial distribution and water age, there were additional key locations that were identified as important regardless of the underlying hydraulic characteristics. Specifically, one monitoring station was placed downstream of the injection location, and six monitoring stations were placed on the influent/effluent lines of the storage tanks within the study region. For the remaining 39 locations, the intent was to manually identify monitoring

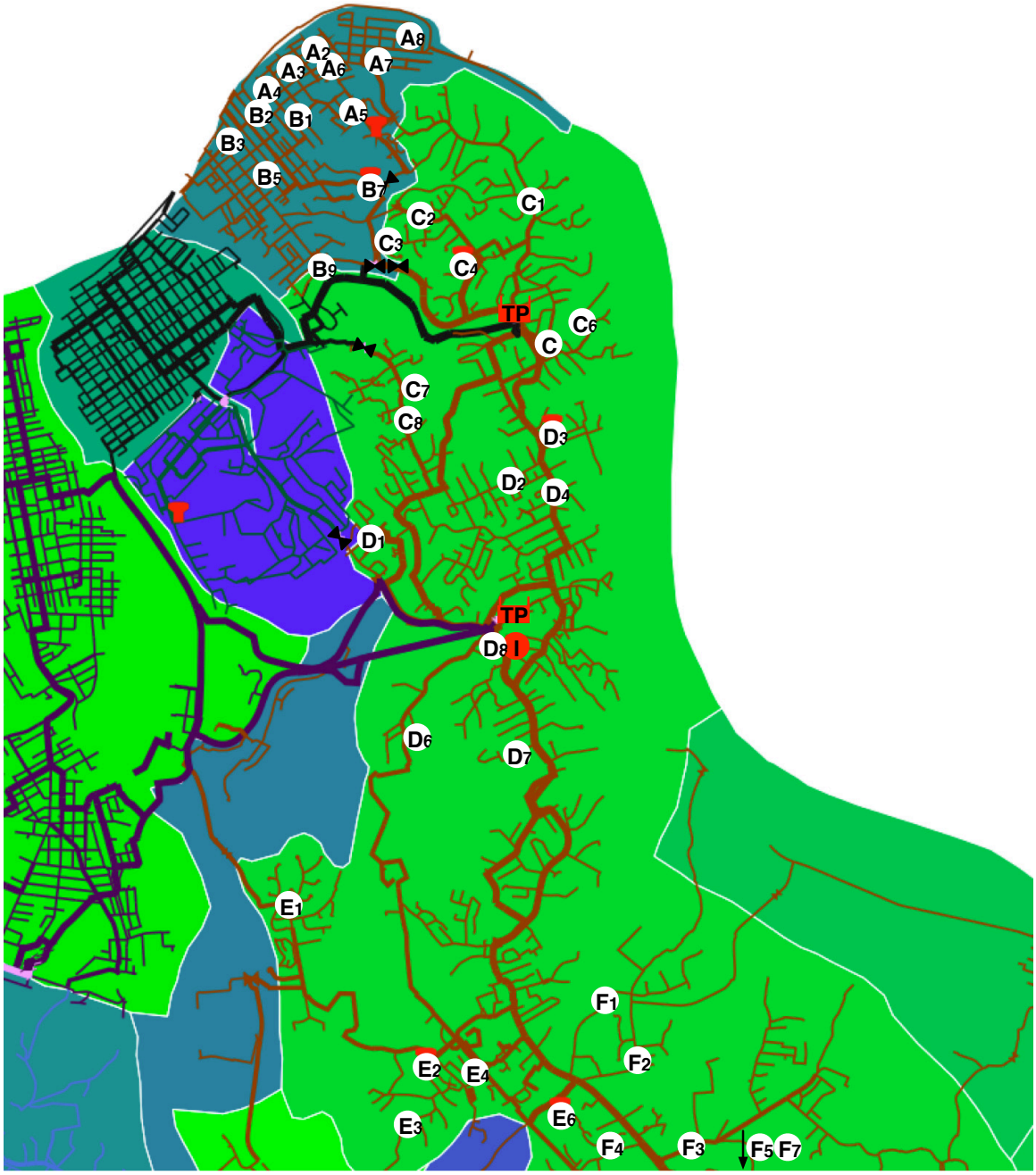

Figure 1: Tracer study area with 38 monitoring locations. Tracer monitoring stations are located within 6 regions labelled A though F and individually numbered. The different colored regions identify district metered areas used for real-time aggregate demand estimates. Brine injection occurred at the red circle labelled "I" adjacent to the south treatment plant.

Figure 2: Continuous conductivity analyzer (typical).

station locations to represent the distribution of water age as well as intensely sample a more populated "grid" within the network. Thirty of the remaining 39 monitoring stations were placed to have six monitoring stations in each of the water age quintile ranges using visual inspection to spatially distribute the monitoring stations. The remaining nine sensor stations were placed in the more dense, gridded region in the north of the system (regions A and B); three locations were intended to capture the influent water quality into this region and the other six locations selected to capture the potential variability in hydraulic or transport characteristics.

In achieving these design objectives, the monitoring locations represent a strong test of real-time water quality modeling accuracy. Difficult locations off of transmission mains in regions with small or localized demands were not discouraged. Table 1 shows the monitoring locations and pipe diameters for each of the 38 monitors that produced data for the analysis. In addition, the table notes identify whether the monitor is located on a dead-end main or a storage tank. Of 38 monitors, 24 are located on pipes of 8 in. diameter or less, and 7 are located on dead-end mains.

The electrical conductivity (EC) signals were measured on 1 minute intervals and logged continuously during the study period at each location. Each monitoring device consists of a conductivity sensor, monitor with display, data logger, battery, and flow-through piping, as shown in figure 3. (A 9v Lithium battery and data logger are within the monitor enclosure in these units; other setups used the same conductivity sensor and monitor but with a different physical configuration.) The conductivity sensor is a 4-Electrode conductivity

Table 1: Characteristics of selected monitoring locations.

Location	Pipe Diameter (in)	Note
A2	6	
A3	6	
A4	8	
A5	6	
A6	4	
A7	8	
A8	8	
B1	6	
B2	8	
B3	10	
B5	4	
B7	12	Tank
B9	8	
C1	8	
C2	12	
C3	8	Dead-end
C4	12	Tank
C6	6	Dead-end
C7	6	Dead-end
C8	12	
D1	12	
D2	8	
D3	8	Tank
D4	16	
D6	6	Dead-end
D7	6	Dead-end
D8	16	
E1	12	
E2	12	Tank
E3	8	Dead-end
E4	8	Dead-end
E6	20	Tank
F1	8	
F2	6	
F3	12	
F4	8	
F5	12	
F7	8	

monitor manufactured by Analytical Technology, Inc. (ATI model Q45C4), and can be used to measure specific conductance in the range of 0 to 2,000 μS/cm with an output voltage ranging from 0 to 2.5 V. All monitors were calibrated against 1000 μS/cm standards, and tested for variability between the devices by measuring three different lab samples.

Figure 3: Continuous conductivity analyzer components (typical).

2.2 Injection Protocol

The field activities included injecting a calcium chloride ($CaCl_2$) solution as a series of pulses at one location over a 14 hour period. Multiple pulses were used to provide more information about network flow dynamics, compared to using a single pulse. The high service pump station at the south treatment plant was selected in consultation with utility staff, based on safety, security, space, and flow control. A series of three valve changes were made prior to the test in order to "un-split" the main pressure zone served by the south treatment plant, and thus expand the service area of this plant and thus the region affected by the brine pulses. An NSF $CaCl_2$ solution was added to treatment plant finished water, producing a series of brine pulses of between 1-2 hours duration. The pulse injection rate was selected to produce a detectable increase in the specific conductance above the background (approximately 350 μS/cm), and yet maintain a significant safety factor when compared to the maximum allowable $CaCl_2$ increase based on applicable federal and state standards on Chloride. In between pulses the $CaCl_2$ feed was discontinued. For two days prior to the start of brine addition, and for a week afterward, the specific conductance was recorded at

the monitoring locations.

The U.S. EPA secondary standard on chloride is 250 mg/L. Based on historical data collected from NKWD over several years, the range in chloride concentrations of finished water was between 16 and 66 mg/L (given an analysis of finished water data collected between January, 2010 and August, 2011). More recent data from the past year for the distribution system and treatment plants indicated a similar range in chloride concentrations. Assuming a conservative background chloride concentration of 41 mg/L during the time period of the test, the applicable standards limit a chloride concentration increase to $(250 - 41) = 209$ mg/L. There was no applicable Federal or primacy agency standards for calcium, and thus it was regulated based on $CaCO_3$ solubility.

Food grade $CaCl_2$ was obtained in totes as a pre-mixed 33% (by weight) solution[1]. Assuming a specific gravity of 1.322 @ 60° F, the 33% solution equates to 436.26×10^3 mg $CaCl_2$/L, or 278.71×10^3 mg Cl^-/L.

It was essential to place reliable controls on the volumetric flow rate of the $CaCl_2$ solution injection pump, such that the chloride concentration was within the above regulatory limits. The maximum $CaCl_2$ injection flow rate of the food grade stock solution can be calculated from a mass balance at the injection site,

$$Q_{CaCl_2}^{max} = \frac{209\ mg\ Cl^-/L}{278.71 \times 10^3\ mg\ Cl^-/L} \times Q_{Prod} = (0.75 \times 10^{-3}) \times Q_{Prod}, \tag{1}$$

where $Q_{CaCl_2}^{max}$ is the maximum allowable flow rate of the NSF food grade $CaCl_2$ solution, and Q_{Prod} is the production flow rate (in the force main receiving the injection), with both flow rates expressed in the same units. Knowing the production flow rate Q_{Prod}, obtained from the SCADA system at the time of the injection, equation (1) was used to calculate the maximum injection flow rate for regulatory purposes. The adopted test protocol limited the maximum addition to 80% of this value.

For an effective tracer test, the injection of brine must create a measurable increase in specific conductance above background. The utility reported that the specific conductance in the distribution system varied between 248 and 637 μS/cm. The impact on the specific conductance can be estimated from the relationship between total dissolved solids (TDS, mg/L) and specific conductance (EC, μS/cm),

$$TDS = k_e EC, \tag{2}$$

or,

$$\Delta EC = \frac{\Delta TDS}{k_e}, \tag{3}$$

where the correlation factor $0.5 < k_e < 0.8$. Assuming the maximum $CaCl_2$ injection flow rate from equation (1), the resulting increase in total dissolved solids, ΔTDS, can be calculated,

$$\Delta TDS = 0.8 \times (0.75 \times 10^{-3}) \times (436.26 \times 10^3\ mg/L) = 261.76\ mg/L. \tag{4}$$

[1]Tetra Chemicals^{TM}NFS grade calcium chloride – see specification sheet in Appendix A.

Assuming a "worst case" $k_e = 0.8$, a corresponding increase in specific conductance was estimated,

$$\Delta EC = \frac{261.76}{0.8} = 327.2\ \mu S/cm \tag{5}$$

which is $> 80\%$ increase over background; this increase is significant given the accuracy of the specific conductance monitors that will be used.

To determine the appropriate calcium chloride dose on the day of the test, the background concentration of chloride was required on the day of the tracer test. To do so, a relationship between historic chloride and specific conductance was developed for water samples collected in finished water over the previous year. A summary of the analysis performed is attached as Appendix A. The analysis shows that based on a measured specific conductance, it is possible to estimate a range of chloride concentration. A very conservative approach was to assume the high end of this range as the current chloride concentration.

3 Real-Time Simulation of Tracer Movement

3.1 Tracer Data Processing

The specific conductance monitors typically produced data with a noise level that was well below the signal difference introduced with the tracer pulse injections. Each raw data stream was visually inspected for obvious anomalies, including sudden and persistent changes in conductivity to levels that were significantly below background or above the maximum conductivity injection peak, or sudden and persistent significant increases in noise levels that indicated a sensor instability. Based on such visual inspection, specific data ranges were removed from the following monitor locations: A7 (data before 11/19/12 00:00); B5 (data before 11/19/12 06:00); D2 (data before 11/19/12 13:20 and after 11/20/12 07:30); D4 (data before 11/19/12 11:50 and after 11/19/12 22:50); and F4 (data after 11/20/12 18:00).

Beyond the visual inspection, the tracer data was processed prior to analysis by ranging, interpolation, and smoothing. Specifically, each data stream was processed first by passing through a ranging filter that excluded points below 200 and above 1000 μS/cm. These data were then resampled on a common clock at a rate of once per minute, using linear interpolation between adjacent data points. Finally, the ranged and resampled/interpolated data was passed through a moving average filter with a 14 minute averaging window (7 minutes before and after each resampled point). This processing did not affect the underlying signal at times scales of interest, while simplifying subsequent data analysis procedures, and removing small anomalies that could simply the visual comparison between observed and simulated time series. An example of the results from these data processing steps is shown in Figure 4 for a location with a higher than normal level of sensor noise.

A final data processing step adjusted the time stamps of all tracer data streams to account for residence time in the hydrant barrel. The residence time will vary from one hydrant to the next, due to variations in depth of main and sample flow rate (which was set to approximately 1 gpm at each location). The hydrant nearest the injection site was used to measure the approximate delay due to the hydrant barrel residence time; at that location, the conductivity pulse arrived approximately 14 minutes after the start of the injection. Each

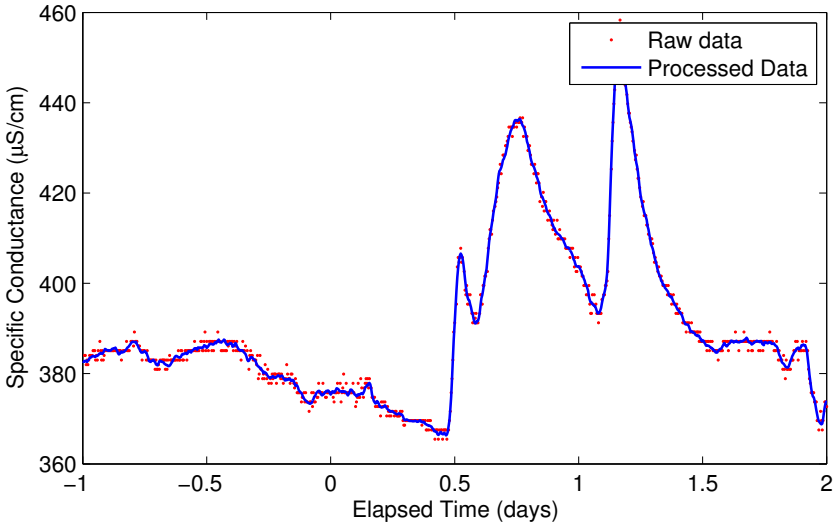

Figure 4: Illustrative results from data processing applied to raw tracer data.

of the tracer data streams was adjusted backward by 14 minutes, to better approximate the conductivity values in the main, and compare to tracer simulation values that would not include the hydrant barrel residence times.

3.2 Real-Time Water Quality Model

The tracer simulations presented below were developed using CitiLogics Polaris^{TM} software, a real-time data analytics environment based on the EPANET-RTX real-time extension for the EPANET programmer's toolkit (USEPA, 2013). Due to practical constraints, simulation results were created after the tracer event, as the real-time Polaris software was not installed at the water utility, and the conductivity monitors were not configured to collect data in a real-time database. Subsequent to the tracer test event, the water utility SCADA historian database was virtualized and brought off-site for connection to EPANET-RTX objects, through the Polaris interface. The tracer data was stored in a MySQL database using EPANET-RTX timeseries objects that have a persistent data store. Polaris was also used to construct the EPANET-RTX timeseries pipelines that performed the conductivity data processing steps described in section 3.1.

The data access, transformation, and simulations were conducted as they would have been in a real-time computation; the results are identical to those which would have been computed in real-time, automatically and without the intervention of an analyst, had that been a possibility during the conduct of the tracer test. The hydraulic behavior is defined by status, setting, and head boundary conditions, and district metered area demand computations, as defined for the real-time hydraulic model and described in CitiLogics (2013). No additional adjustments to assumptions or model parameters were made. Beyond those real-time hydraulic calculations, the real-time tracer simulation only requires initial and boundary conditions, as the tracer is non-reactive. The only relevant simulation parameter is the Epanet water quality time step, which was 15 seconds.

The boundary condition at the site of injection (see Figure 1) was constructed from continuous conductivity monitor observations, snapshot observations downstream of the injection location, and records of the start and stop time of the injection pump; the resulting tracer concentration is shown in Figure 5. Beyond the characteristics of the four salt pulses that were injected, the background conductivity at the injection site averaged 365.5 μS/cm, and was stable with a standard deviation of 4.2 μS/cm. The positive displacement injection pump discharged into a force main on the suction side of three parallel high service pumps. Coordination with operations staff ensured that the high service pumping status did not change while tracer was being injected. Thus, once the injection pump was turned on and the controller was set, the resulting conductivity was relatively stable. The changes in conductivity at the start of the first pulse in Figure 5 were due to experimentation with the injection rate. The injection was stopped for 5 minutes after 14 minutes had elapsed, because the first conductivity pulse had not been seen at a nearby downstream hydrant, creating uncertainty about the initial pulse characteristics. This uncertainty was resolved, and the pulse arrival time delay (estimated to be between 14 and 19 minutes) was attributed to residence time in the hydrant barrel.

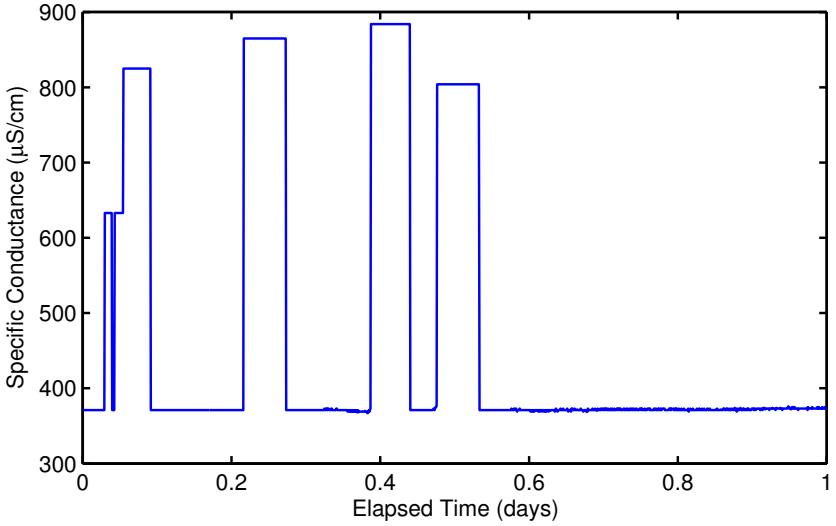

Figure 5: Conductivity boundary condition at the injection site used for tracer simulations.

One additional boundary condition was required at the north treatment plant (see Figure 1). Under the operating conditions during the tracer test period, this treatment plant operated part of the day, and all plant production flowed by gravity from the clearwell to lower pressure zones located to the west and southwest. These latter pressure zones were excluded from the tracer test monitoring area, because only a small fraction of their demand was satisfied via a regulator from the southern treatment plant that carried the tracer injection. Thus, in principal, the conductivity at the north treatment plant was not important for simulating tracer movement within the study area, because it only fed the lower pressure zones which were excluded from monitoring. As a result, a conductivity monitor was not placed at the north treatment plant boundary, and the initial plan was to arbitrarily set that tracer boundary condition to zero.

Under some operating conditions, water is delivered from the north treatment plant through high service pumps to the upper zone that interacts with the study area. Under the operating conditions in effect during the test, these high service pumps were off, and all demand in the study area was satisfied by the south treatment plant high service pumping. The north plant high service pumps were used, however, up until the start of the tracer injection. Thus water from the north treatment plant is expected to be within the study area, at the start of the tracer injection. There were two methods of dealing with this issue. The first would be to ignore it and start the simulation with the tracer addition, and represent as accurately as possible the initial conditions at the start of the test. The other would be to model the conductivity at the north treatment plant, and back up the simulation start to include a time period prior to the tracer addition, during which the background conductivity from both plants would be represented as tracer boundary conditions. This latter approach was chosen because of evidence that the background conductivity at the two plants were significantly different.

Given the lack of a monitor at the north treatment plant, two nearby monitors were used to develop an assumption of the north plant background conductivity. Figure 6 (top) shows a nearby storage tank level along with the status of the north plant high service pumps (if the status is 1, one pump is running; if 0, all pumps are off). The time scale is in days relative to the start of the tracer injection. More than three days before the start, the north plant was running continuously and the high service pumps delivered water to a significant portion of the study area. During this operation mode, the record shows that the nearby storage tank filled only when the north plant high service pumps were on. The time period from days -4 to -3 in Figure 6 (top) is indicative of this behavior. Through day -2, this tank was still filled only when the north plant high service pumps were on. A reasonable assumption, due to spatial proximity and pumping operations, is that the conductivity within this nearby storage tank represents an integrated measure of the background conductivity at the north treatment plant. Figure 6 (middle) shows the specific conductance measured at this tank along with its level, for a time period prior to the start of the test. The conductivity signal prior to day -2 occurs at the end of a prolonged drain cycle, and averages 445 μS/cm; this is the best estimate given the data of the conductivity within the tank when it is being filled from high service pumping at the north plant.

The conductivity monitor nearest to the north treatment plant can also be used to infer the background conductivity (see Figure 1; the monitor is to the southeast of the treatment plant). Data from this monitor was lost soon after the salt injection, but valid data was harvested for the preceding two days, as shown in Figure 6 (bottom). During the 2 days preceding the test, the area that includes this monitor was fed primarily by the south treatment plant. There are clear signatures, however, of higher conductivity pulses entering from the north treatment plant, associated with high service pumping activity. Thus these data provide additional evidence that the background conductivity at the north plant was significantly higher than the 365 μS/cmat the south boundary. The conductivity pulses at the nearby monitor are more likely, however, to be influenced by mixing with the south plant water. For that reason, the north plant boundary condition was set at 445 μS/cm – the value measured in the nearby storage tank prior to the change in operations.

The tracer simulation was divided into two parts: an initial period 53 hours prior to the test start at 11/19/2013 08:00; and a simulation period of 9 days duration. The start

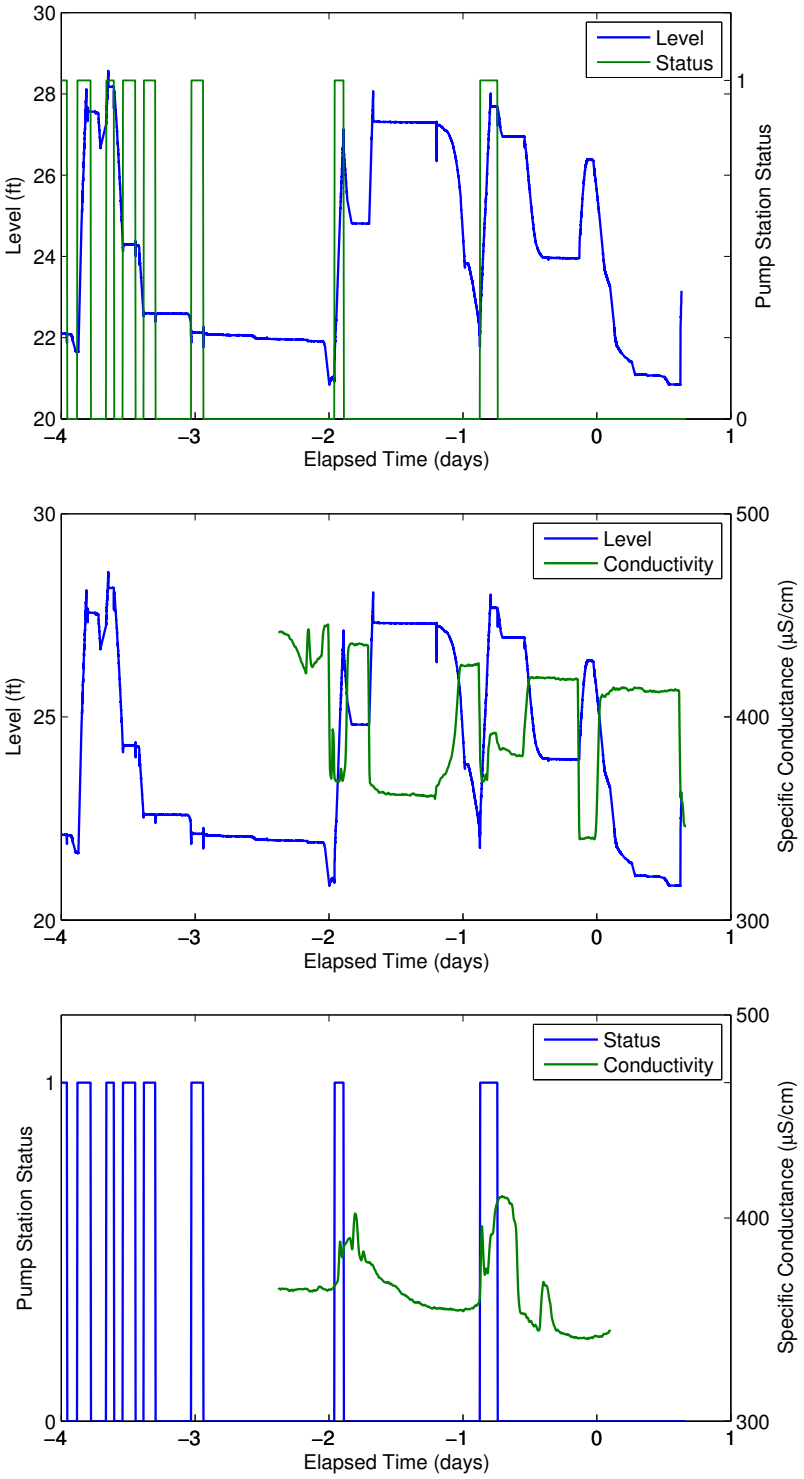

Figure 6: Evidence used to determine elevated background conductivity at north treatment plant. Water level at nearby storage tank and north plant high service pump status (top); same water level and conductivity on inlet/outlet line (middle); and high service pump status and nearby conductivity measure (bottom).

of the initial period was determined based on data availability. Most monitors were set to turn on at 11/16/2013 23:00, or 57 hours prior to the start. Delaying the start of the initial period by 4 hours allowed more monitors to be used to specify initial conditions, because of various data problems. Initial conditions were specified using data from the conductivity monitors distributed throughout the study area, at the start of the initial period. A nearest neighbor spatial interpolation was used to distribute those data to each network node. Initial conditions for storage tanks within the study area were treated separately. For each tank its observed conductivity was plotted along with its level, and the initial conductivity was estimated from that observed during a suitable drain period. At the start of the simulation period, all simulated tank levels were reset to observed levels. Thus at the start of the simulation period, the water quality initial conditions reflect the propagated background conductivity from both treatment plants over a 53 hour period, and the hydraulic initial conditions reflect measurements from SCADA.

3.3 Accuracy Metrics

It is difficult to assess the quantitative similarity between two time series. This is especially true of the simulated versus observed tracer time series from the current field study – the four tracer pulses were designed to produce unusual fluctuations in tracer signals, and time-shift errors between simulated and observed series could produce large disparities in traditional "goodness of fit" metrics, such as the root-mean-squared-error (RMSE). More to the point, the RMSE is not expected to discriminate between locations where the signals are time-shifted, and those where the simulated signal bears little qualitative resemblance to the measurements. One alternative approach would use the Pearson correlation coefficients at different time lags to measure the goodness of fit, as well as the lag that produces the largest correlation between simulated and observed. Another is to use dynamic time warping (DTW) to compare the paired series, which allows for variable and non-linear time shifts. While both of these approaches are expected to improve upon a simple RMSE (or similar metrics), we opted for an approach that compares the quantiles of the conductivity area above background, or CAAB. This approach relies on traditional statistical concepts, and naturally accommodates variable time shifting behavior, while being simpler than either the lag-correlation or DTW methods.

Figure 7 illustrates the quantified pulse characteristics for one particular tracer time series. Time is shown as elapsed time relative to the start of the tracer test at 08:00 on November 19, 2012. The first analysis step is to subtract the conductivity background individually for each series, to expose the conductivity pulse signals as the key features to be measured. The background exhibits some degree of variability, so background subtraction relies on an operational definition. We estimate the background for each series as the average conductivity during the 24 hours immediately preceding the tracer injection (either measured or simulated), and subtract this from the entire series. The series in Figure 7 shows the conductivity signal after the background is subtracted. In cases where the data record did not extend the full 24 hours prior to the injection, the background was estimated from what data was available. If no measured data existed prior to the start of tracer injection, the background of the observed signal was assumed equal to the simulated series background.

The key characteristics of the tracer pulses were estimated from integrating the area

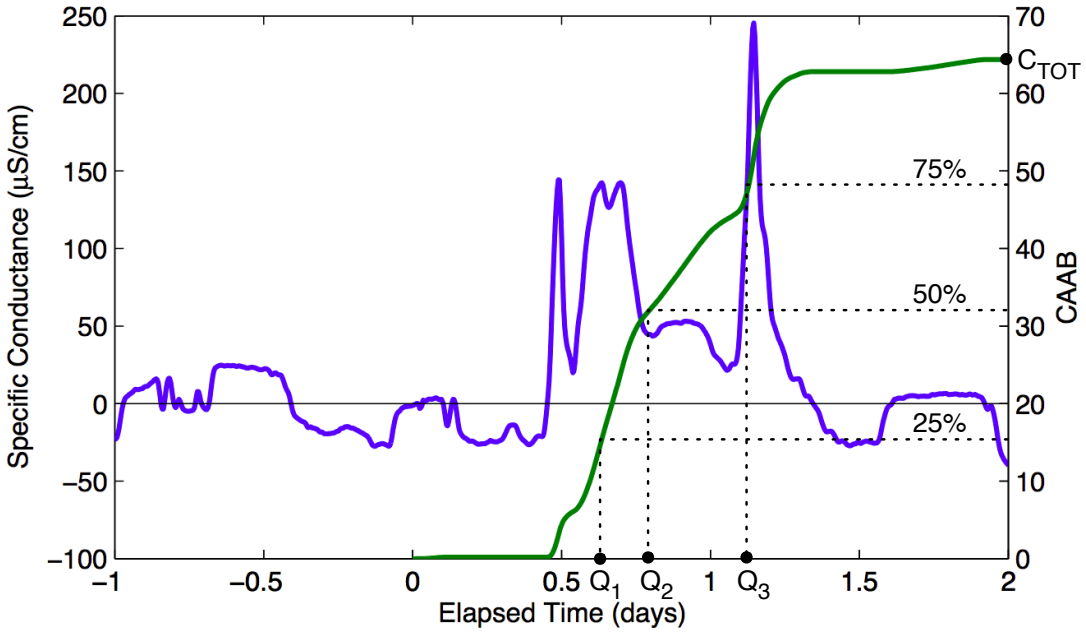

Figure 7: Tracer timeseries characteristics used for error analysis. Quartiles Q_1, Q_2, and Q_3 mark the times when 25, 50, and 75 percent of the tracer pulse signature (not necessarily tracer mass) has passed the sensor, allowing use of the median (50th percentile) and inter-quartile range ($Q_3 - Q_1$) as comparative characteristics. The conductivity area above background (CAAB) at the two day mark provides an integrated measure of signal strength and temporal signature.

under the conductivity signal (numerically), above the background, as a function of the elapsed time. This area is plotted in Figure 7 on the right axis. Note the integral was computed between the limits of 0 and 2 days elapsed time – an operational assumption reflecting that the tracer pulse signal had passed all of the monitor locations within the first 2 days after injection. Given the CAAB versus elapsed time curve, the first, second, and third quantiles of the area (Q_1, Q_2, and Q_3, each with units of time) are identified and used as key characteristics of the simulated and observed pulse signals at each location. More specifically, we compare the simulated and observed median, Q_2, or the time for 50% of the pulse area to pass the monitor, and the simulated and observed inter-quartile range, $IQR = Q_3 - Q_1$, as a measure of the time spread of the pulse. Since the simulated and observed medians and IQR could match exactly, even if the simulated pulse was attenuated or amplified relative to the observed, the total CAAB at 2 days elapsed time, labelled C_{TOT} in Figure 7, was also used to compare simulated and observed series.

3.4 Summary Results

The key characteristics of simulated and observed tracer time series are summarized in Table 2. These include the quantiles and total conductivity area above background (CAAB), as discussed above. The absolute errors in median and inter-quartile range of CAAB, and the percent error in total CAAB, are summarized in Table 3. Table 3 also gives the median and mean of each error statistic over all measurement locations; the median (mean) Q_2 error is 3.88 (4.47) hours, the median (mean) IQR error is 1.46 (2.57) hours, and the median (mean) C_{TOT} error is 37.11 (50.33) percent. These results suggest that errors affecting the average speed of the tracer pulses, as characterized by the Q_2 errors, dominate those that affect the dispersion of the pulses, as characterized by the IQR errors. The C_{TOT} errors usually reflect an attenuation of the observed series relative to the simulated (if these are computed using signed instead of absolute errors, the median C_{TOT} error is +13.83 %).

In an attempt to identify trends within the location error statistics, the three error statistics are plotted in Figure 8 as a function of pipe diameter at the measurement hydrant, after isolating those measurement locations at storage tanks and on dead-end mains. The horizontal and vertical lines on these plots indicate the median errors. The same error statistics are plotted in Figure 9 as a function of location region, by the same location labeling convention used in Figure 1. Neither of these analyses present visually compelling arguments for clear trends in the error statistics, of the sort that could help explain the sources of the simulation errors. It might be said that dead-end locations, for example, are poorer overall in simulation accuracy, as 5 of 7 have C_{TOT} errors that exceed the median, yet the same statement does not hold up for the quantile error statistics. In any case, the expectation that dead-end mains would have strikingly different, and poorer, overall simulation accuracy compared to looped mains, does not hold up to scrutiny in this case. Similarly, the regional categorization does not identify any one region with strikingly different overall simulation accuracy. While it could be argued, for example, that region A presents overall improved accuracy compared to B, especially in terms of the CAAB median, the results do not indicate a clear and consistent trend. Perhaps the lack of such trends reflects on the overall severe test posed by the tracer study design, with 24 of 37 measurement locations on dead-ends or mains with diameters less than or equal to 8 in. Just as likely, it may reflect that complexity

Table 2: Characteristics of simulated and observed tracer time series. All quantiles in days and C_{TOT} in (day$\times\mu$S/cm).

	Simulated				Observed			
Location	Q_1^{sim}	Q_2^{sim}	Q_3^{sim}	C_{TOT}^{sim}	Q_1^{obs}	Q_2^{obs}	Q_3^{obs}	C_{TOT}^{obs}
A2	0.615	0.753	0.833	67.946	0.632	0.788	1.122	64.096
A3	0.892	1.069	1.250	71.135	0.736	0.899	1.215	74.454
A4	0.569	0.767	1.073	71.277	0.660	0.823	1.118	72.018
A5	1.247	1.483	1.642	139.783	1.347	1.622	1.878	73.739
A6	0.667	0.757	1.167	75.481	0.646	0.785	1.115	54.429
A7	0.524	0.604	1.038	87.370	0.545	0.788	1.042	78.179
A8	0.913	1.045	1.476	78.415	0.660	0.833	1.135	70.000
B1	0.698	0.896	1.101	93.229	0.753	0.958	1.194	30.422
B2	0.490	0.597	1.035	66.101	0.726	0.847	1.156	65.670
B3	0.622	0.785	1.163	62.745	0.868	1.201	1.375	33.670
B5	0.486	0.618	1.021	65.624	0.618	0.983	1.306	122.286
B7	0.351	0.476	0.941	68.869	0.448	0.663	0.962	68.212
B9	0.389	0.503	0.941	78.368	0.521	0.788	1.010	83.297
C1	0.604	0.851	0.979	128.116	0.861	1.118	1.274	52.043
C2	1.097	1.177	1.243	83.718	1.399	1.538	1.771	40.529
C3	1.410	1.486	1.677	76.637	1.149	1.247	1.476	33.626
C4	0.674	0.837	1.108	77.546	0.642	0.792	0.979	48.172
C6	0.642	0.927	1.052	123.268	0.969	1.146	1.243	111.945
C7	0.542	0.826	0.889	116.072	0.663	0.872	1.052	78.424
C8	0.413	0.615	0.722	122.127	0.392	0.587	0.670	90.107
D1	0.924	1.069	1.337	133.188	0.601	0.799	0.986	62.177
D2	0.368	0.569	0.691	95.860	0.333	0.493	0.663	95.631
D3	0.688	0.847	1.003	34.971	0.229	0.431	1.028	13.403
D4	0.306	0.438	0.552	73.513	0.274	0.392	0.497	85.615
D6	0.559	0.892	0.920	60.895	0.611	0.705	0.819	128.969
D7	0.316	0.681	0.719	79.373	0.156	0.462	0.500	71.730
D8	0.274	0.431	0.507	83.598	0.274	0.424	0.503	78.251
E1	0.889	1.021	1.285	45.998	0.622	0.865	0.941	77.388
E2	0.632	0.806	1.222	52.170	0.472	0.656	1.038	87.035
E3	0.948	1.049	1.319	48.707	0.736	0.882	1.156	21.490
E4	1.559	1.681	1.774	65.256	0.465	0.531	1.052	34.392
E6	0.413	0.653	1.066	48.179	0.392	0.552	0.632	38.880
F1	0.531	0.622	0.802	65.859	0.344	0.444	0.639	41.987
F2	0.552	0.740	0.847	53.608	0.524	0.667	0.719	56.286
F3	0.455	0.507	0.705	46.768	0.434	0.462	0.493	35.008
F4	0.455	0.674	0.847	55.345	0.469	0.528	0.667	36.993
F5	0.500	0.677	0.767	62.123	0.479	0.639	0.701	53.725
F7	0.681	0.806	0.924	55.101	0.729	0.868	0.979	74.474

Table 3: Differences between simulated and observed tracer time series.

Location	$\lvert Q_2^{sim} - Q_2^{obs} \rvert$ (hr)	$\lvert IQR^{sim} - IQR^{obs} \rvert$ (hr)	$\frac{100 \times \lvert C_{TOT}^{sim} - C_{TOT}^{obs} \rvert}{C_{TOT}^{obs}}$ (%)
A2	0.83	6.50	6.01
A3	4.08	2.92	4.46
A4	1.33	1.08	1.03
A5	3.33	3.25	89.56
A6	0.67	0.75	38.68
A7	4.42	0.42	11.76
A8	5.08	2.08	12.02
B1	1.50	0.92	206.45
B2	6.00	2.75	0.66
B3	10.00	0.83	86.35
B5	8.75	3.67	46.34
B7	4.50	1.83	0.96
B9	6.83	1.50	5.92
C1	6.42	0.92	146.17
C2	8.67	5.42	106.56
C3	5.75	1.42	127.91
C4	1.08	2.33	60.98
C6	5.25	3.25	10.12
C7	1.08	1.00	48.01
C8	0.67	0.75	35.54
D1	6.50	0.67	114.21
D2	1.83	0.17	0.24
D3	10.00	11.58	160.92
D4	1.08	0.58	14.14
D6	4.50	3.67	52.78
D7	5.25	1.42	10.65
D8	0.17	0.08	6.83
E1	3.75	1.83	40.56
E2	3.58	0.58	40.06
E3	4.00	1.17	126.65
E4	27.58	8.92	89.74
E6	2.42	9.92	23.92
F1	4.25	0.58	56.85
F2	1.75	2.42	4.76
F3	1.08	4.58	33.59
F4	3.50	4.67	49.61
F5	0.92	1.08	15.63
F7	1.50	0.17	26.01
Median	3.88	1.46	37.11
Mean	4.47	2.57	50.33

of transport dynamics in looped networks can not be captured using simple concepts related to local pipe characteristics or geographic proximity.

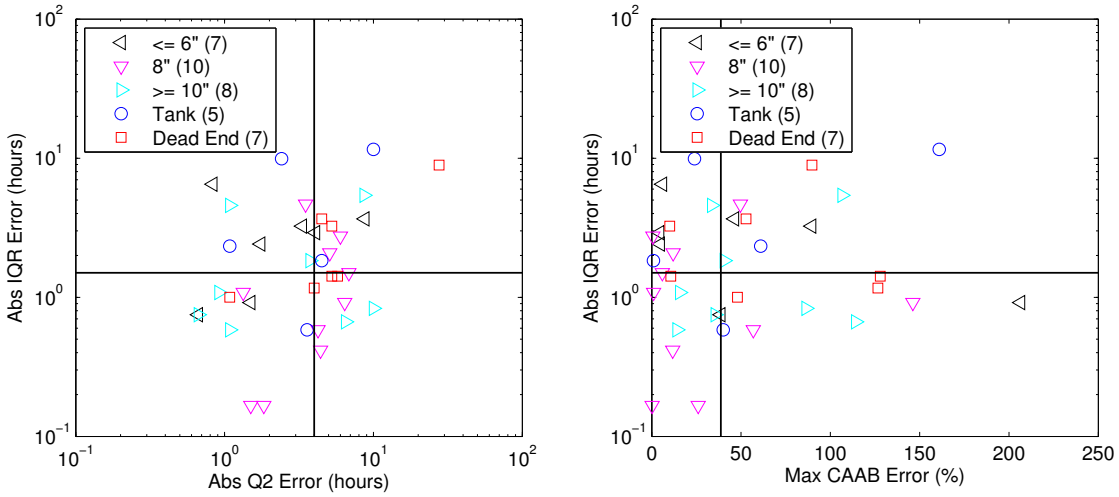

Figure 8: Comparison of tracer data and real-time simulations at 37 measurement locations, as a function of pipe diameter or location characteristic. Measurements at storage tanks and dead-end mains are isolated for comparison, and not included in the pipe diameter categories. Statistical comparison of observed and simulated conductivity pulses is focused on three pulse movement characteristics: absolute error in inter-quartile range (IQR), absolute error in median (Q_2), and percent error in the total conductivity area above background (CAAB), measured 2 days after tracer injection.

3.5 Observed and Simulated Tracer Signals

Time series data are presented here for the observed and simulated conductivity signals from 37 monitoring sites. The sites are grouped and plotted as regions A through F. The selection of these regions was based upon physical proximity for implementing the field study, and not necessarily because they span a pressure zone or demand metered area, or any other reason related to infrastructure or hydraulic behavior. Each location plot shows the simulated and observed conductivity signal above background, and includes information about the simulation accuracy metrics discussed in section 3.3. Inset plots are included that show the CAAB versus time for both simulated and observed series, allowing the conductivity pulse areas to be compared visually as they evolve over time. The CAAB quantiles are shown symbolically on each graph adjacent to the time axis, in the manner of a box plot. The outer box represents the first and third quartiles, the difference between them the inter-quartile range, and the line within the box the second quartile, or median.

3.5.1 Region A

Figures 10-12 shows the observed and simulated conductivity signals over a three day period. This region was densely monitored, providing an unusual spatial-temporal picture of tracer

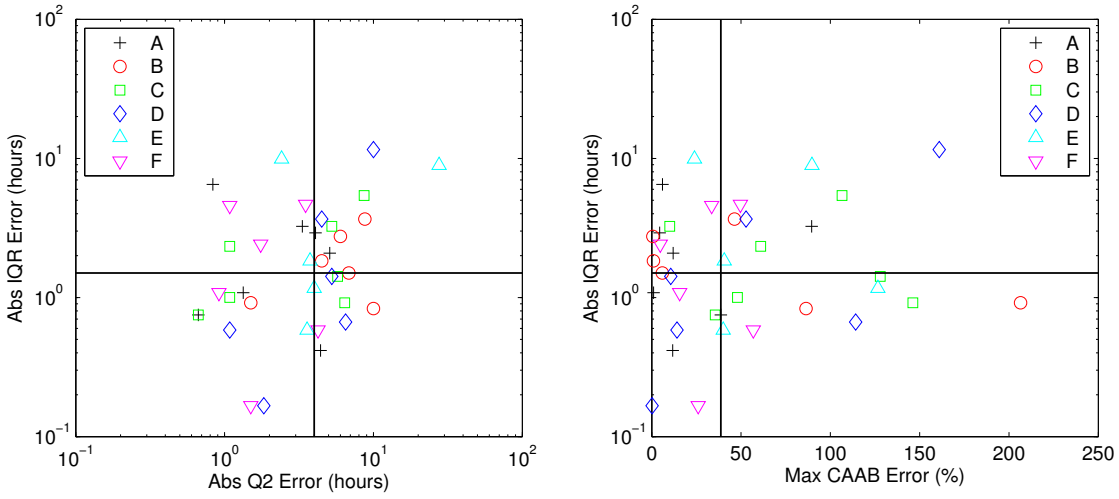

Figure 9: Comparison of tracer data and real-time simulations at 37 measurement locations, as a function of location region. Statistical comparison of observed and simulated conductivity pulses is focused on three pulse movement characteristics: absolute error in inter-quartile range (IQR), absolute error in median (Q_2), and percent error in the total conductivity area above background (CAAB), measured 2 days after tracer injection.

evolution in an older, more densely populated, urbanized area. No data is included for Location A1 because the conductivity unit malfunctioned. With the exception of A5, which was just outside of the "gridded" portion of Region A, the conductivity signals for Locations A2 – A8 showed similar conductivity signals. These results are interesting in light of the gridded pipe connectivity within this region. Although the observed signals for A2, A3, and A4 are very similar and these locations are within several blocks of each other, the simulated signal for A3 shows distinct characteristics that are not reflected in the data.

3.5.2 Region B

Figure 13-14 shows the locations and observed conductivity signals for the six monitoring stations in Region B, which were also located in a densely populated region of the distribution system. No data was collected from locations B4, B6, B8 and B10 due to monitor malfunction. While Locations B1 through B5 show similarity in the observed data, the degree of similarity is not as great as the monitors located within Region A. These observations also showed significant pulse attenuation compared to region A observations, and also compared to the simulated time series. Location B7 was located at a storage tank; the square pulses are generated by the tank drain and fill cycles. The signal from Location B9 was similar to locations in region A. The signal at B7 shows how visually similar observed and simulated series can exhibit large errors in the CAAB median, suggesting that perhaps quantile ranges are a better, and more stable, metric of simulation accuracy.

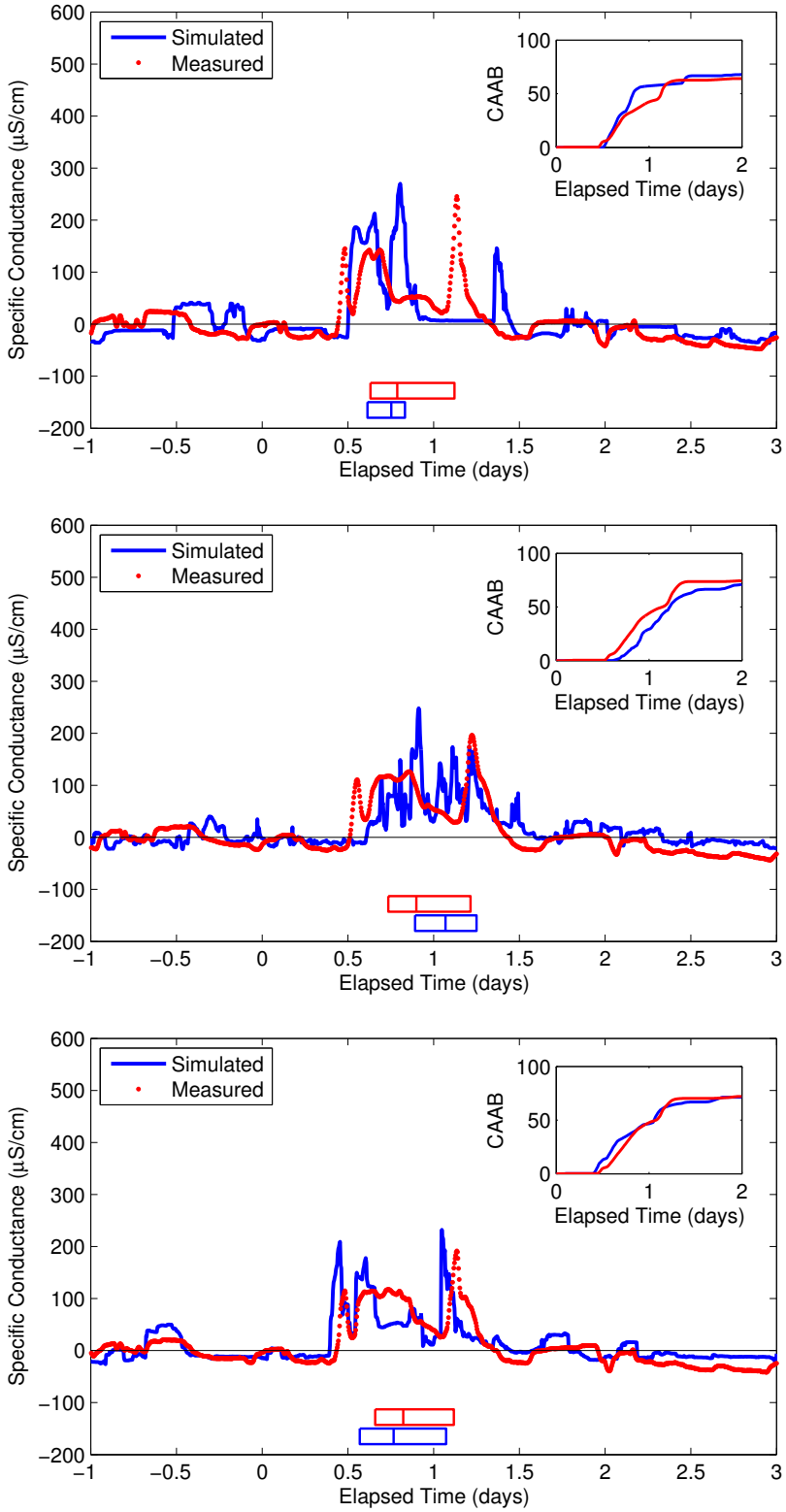

Figure 10: Observed and simulated tracer movement, locations A2 (top), A3 (middle), and A4 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

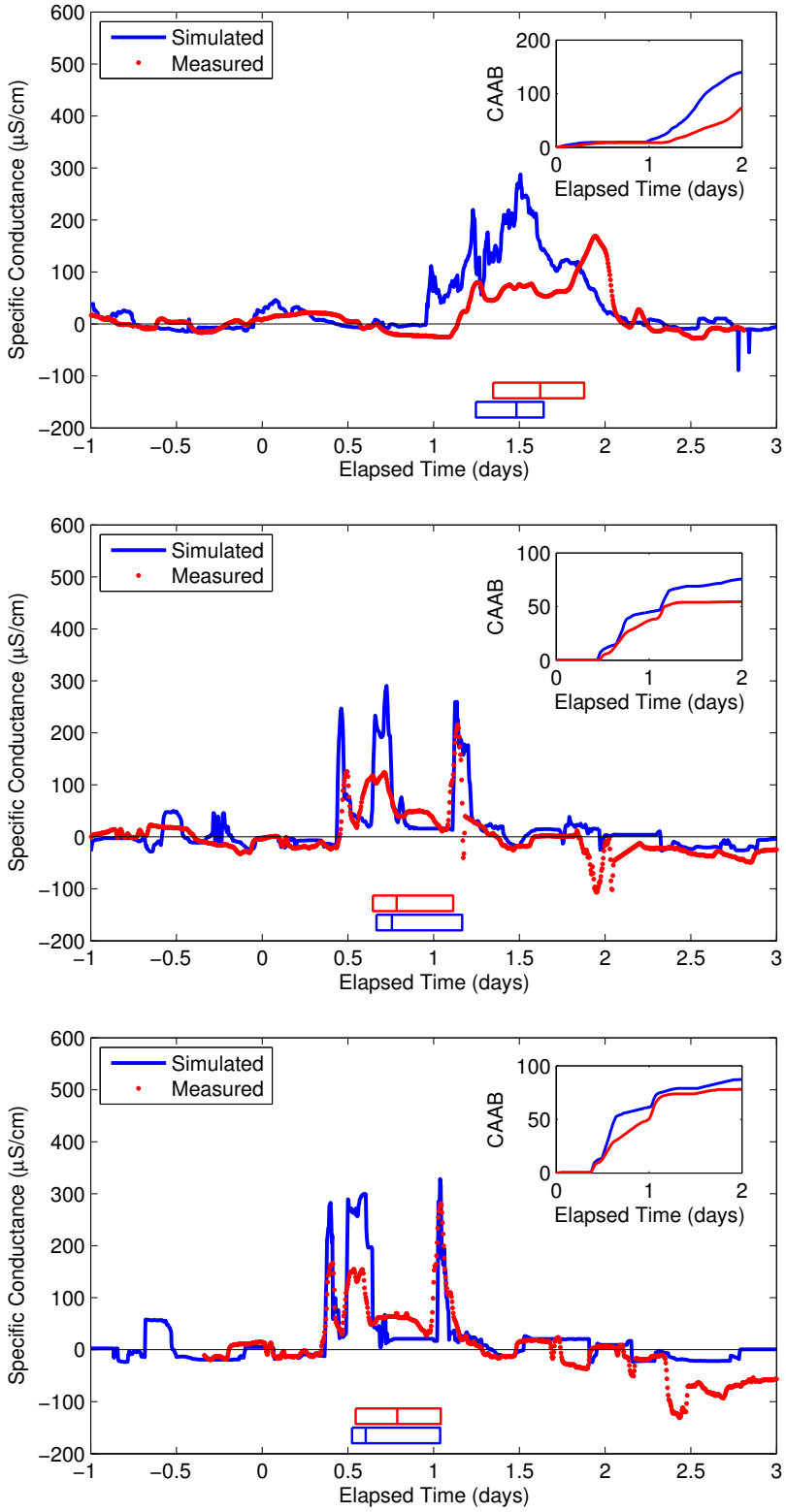

Figure 11: Observed and simulated tracer movement, locations A5 (top), A6 (middle), and A7 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

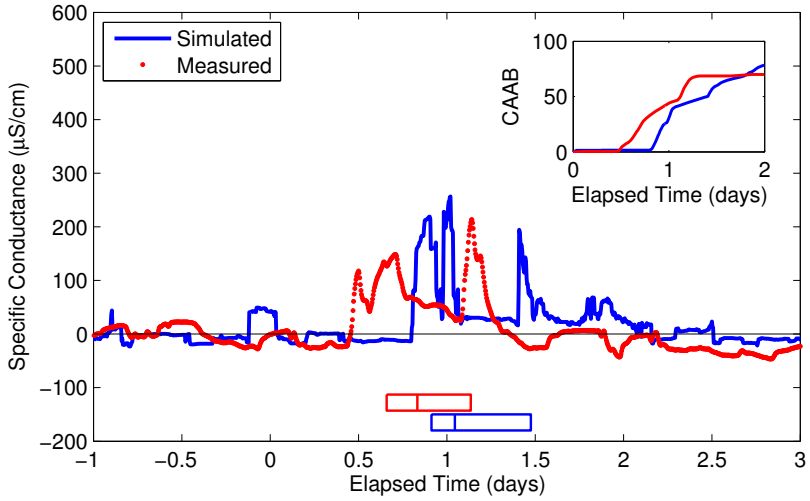

Figure 12: Observed and simulated tracer movement, locations A8. Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

3.5.3 Region C

Figure 15-17 presents the seven locations and conductivity signals associated with Region C, adjacent to and south of regions A and B. No data for Location C5 is presented due to monitor malfunction. Location C4 represents the conductivity signal adjacent to a storage tank. Locations C2 and C3 are under the influence of this tank as illustrated by the similarity in signal characteristics. Although the simulated conductivity signal at C4 is reasonable (e.g., $IQR = 2.33$ compared to the median IQR of 1.46 hrs.), the median CAAB at nearby signals at C2 and C3 are significantly too early and too late, respectively. Both these sites predict conductivity signal peaks transported from the injection source that do not appear in the data. Interesting, while location C3 is at the end of a long dead end main, C2 is within what could be called a dead end loop, due to the presence of a downstream regulator that is likely to be closed, according to utility personnel. Locations C7 and C8 present an interesting study on the impact of demands, and possibly transport mechanisms, within dead end mains. Location C8 is on a 12 in. main and C7 is just downstream on a 6 in. dead end. While C8 has visually and quantitatively good error characteristics, C7 is visually a much poorer fit to the signal, which has qualitatively different characteristics even, though is is just a short distance downstream. Presumably, these different characteristics are due to dead end demands as well as dispersion processes that may be dominant within the dead end. Finally, the comparison of C7 and C8 points out the challenges inherent in comparing two time series; while these locations have roughly equal quantitative error characteristics, the visual fit of the C8 simulated signal to the data is noticeably superior to that of location C7. What is happening in this case is the IQR is relatively good yet the times associated with passage of the first and third quantiles are significantly in error; it may be preferable to examine a

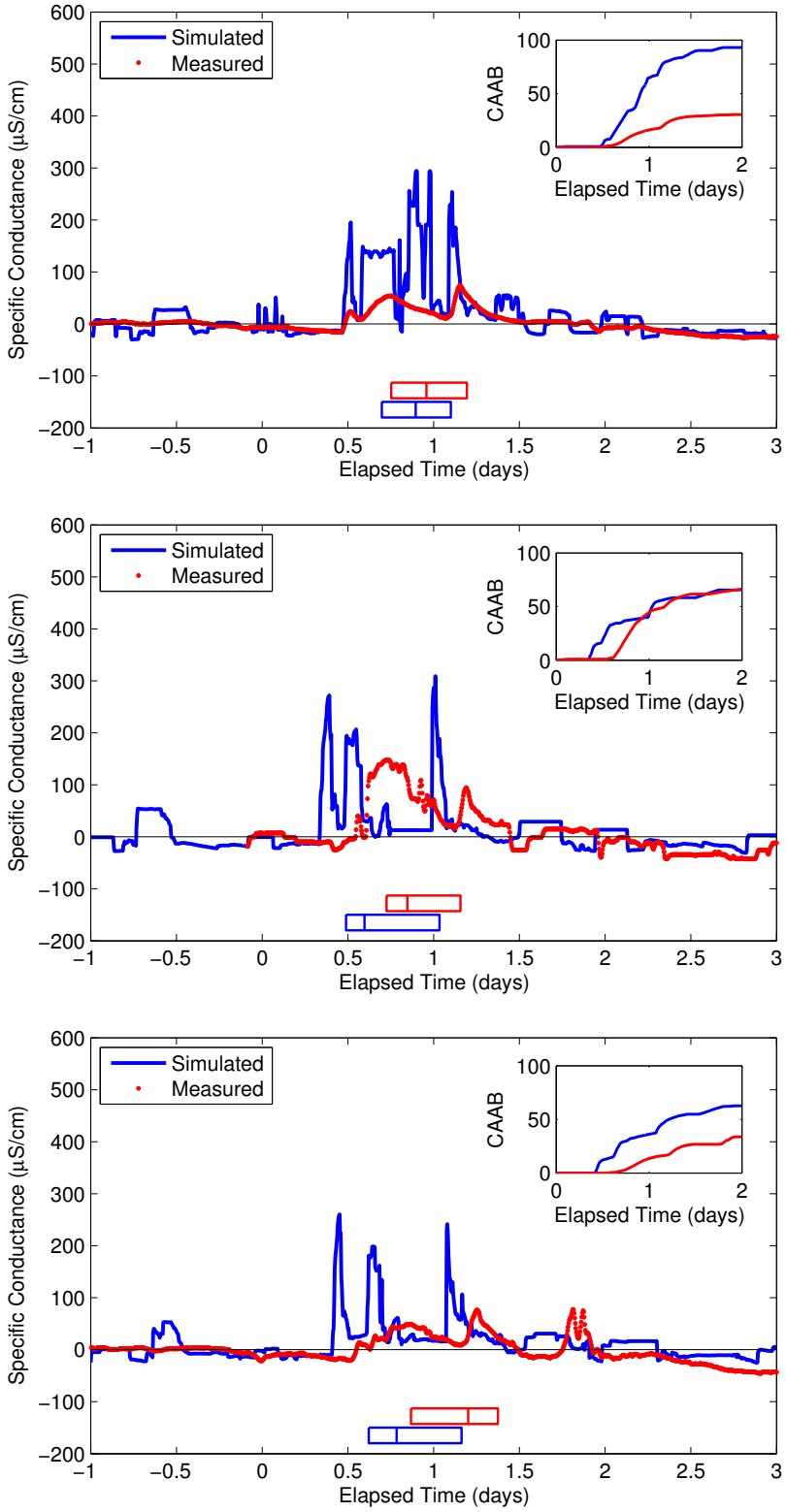

Figure 13: Observed and simulated tracer movement, locations B1 (top), B2 (middle), and B3 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

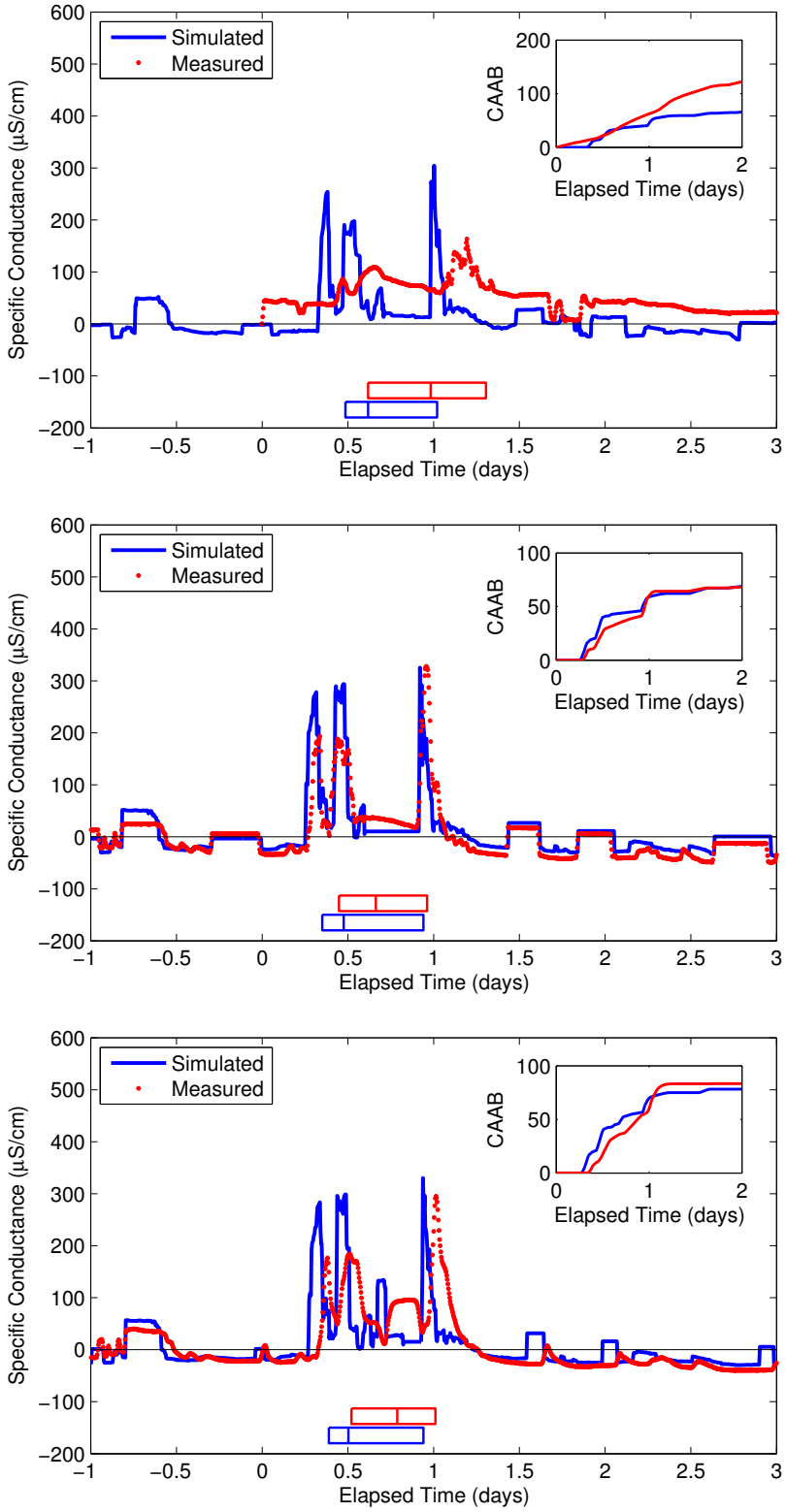

Figure 14: Observed and simulated tracer movement, locations B5 (top), B7 (middle), and B9 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

metric focussed more on these particular times, compared to using the IQR.

3.5.4 Region D

Figure 18-fig:D3 shows the conductivity signals for seven monitoring stations within Region D. Region D is south of Region A and includes the injection location, D8, at the south treatment plant in Figure 1. The monitor at location D8 started recording data after the first salt pulse, because it was moved to replace a malfunctioning unit. The close proximity of D8 to the injection boundary results in expected small simulation errors. Monitor D1 is located on the boundary of two pressure zones with flow governed by a regulating valve. The significant delay in the simulated median CAAB time compared to observed is a good indication that flow through the regulating valve is significantly greater than simulated, and thus that the downstream valve setting is in error. Locations D2, D3, and D4 are in close proximity to each other, with D3 located at a storage tank. Flow to this tank is actively controlled via a solenoid operated valve on the inlet/outlet line; the status of this valve is modeled explicitly by the real-time hydraulic model. When filling or draining, this tank level changes by 5-10 feet over short intervals of approximately 2 hours. Errors between simulated and observed conductivity at this tank are due to errors in pulse arrival times relative to when the tank is filling. Observations indicate that each conductivity pulse was transported past the tank while it was draining, whereas the simulation suggests that pulse arrival coincides with a tank fill period about 15 hours after the first injection. Location D4 is on a 16 in. main leading from the injection site to the tank, and is a good indication of the simulation error along one of the largest mains in the study area. Location D2, however, is on an 8 in. distribution main off of the 16 in. Locations D6 and D7 are both on 6 in. dead-end mains; D6 branches off of a 12 in. main, and D7 branches off of a 16 in. main – both leading from the injection site. The errors at both these locations indicate the simulation is too slow by about 5 hours, but those delays could be due to velocities in transmission mains, or within the dead-end pipes. It is interesting that the observed conductivity signal at D7 does not exhibit the pulse attenuation and dispersion observed at D6, and other dead-end signals.

3.5.5 Region E

Figure 21-22 shows the simulated and observed conductivity signals at the five monitors in Region E. Location E5 was omitted due to monitor malfunction. Monitors E2 and E6 are located at storage tanks. Location E1 is on a pressure zone boundary and under the influence of both the tank at E2, and a downstream regulating valve. Locations E3 and E4 are both on dead-end mains in between the two storage tanks at E2 and E6. The simulation errors at E3 and E4 suggest again how simulation results may be strongly affected by highly localized demand characteristics within small diameter pipes and especially within dead-end segments.

3.5.6 Region F

Figure 23-24 shows the observed and simulated conductivity signals for the six monitors in Region F. Locations F1, F2, and F3 are on 8, 6, and 12 in. distribution mains, respectively,

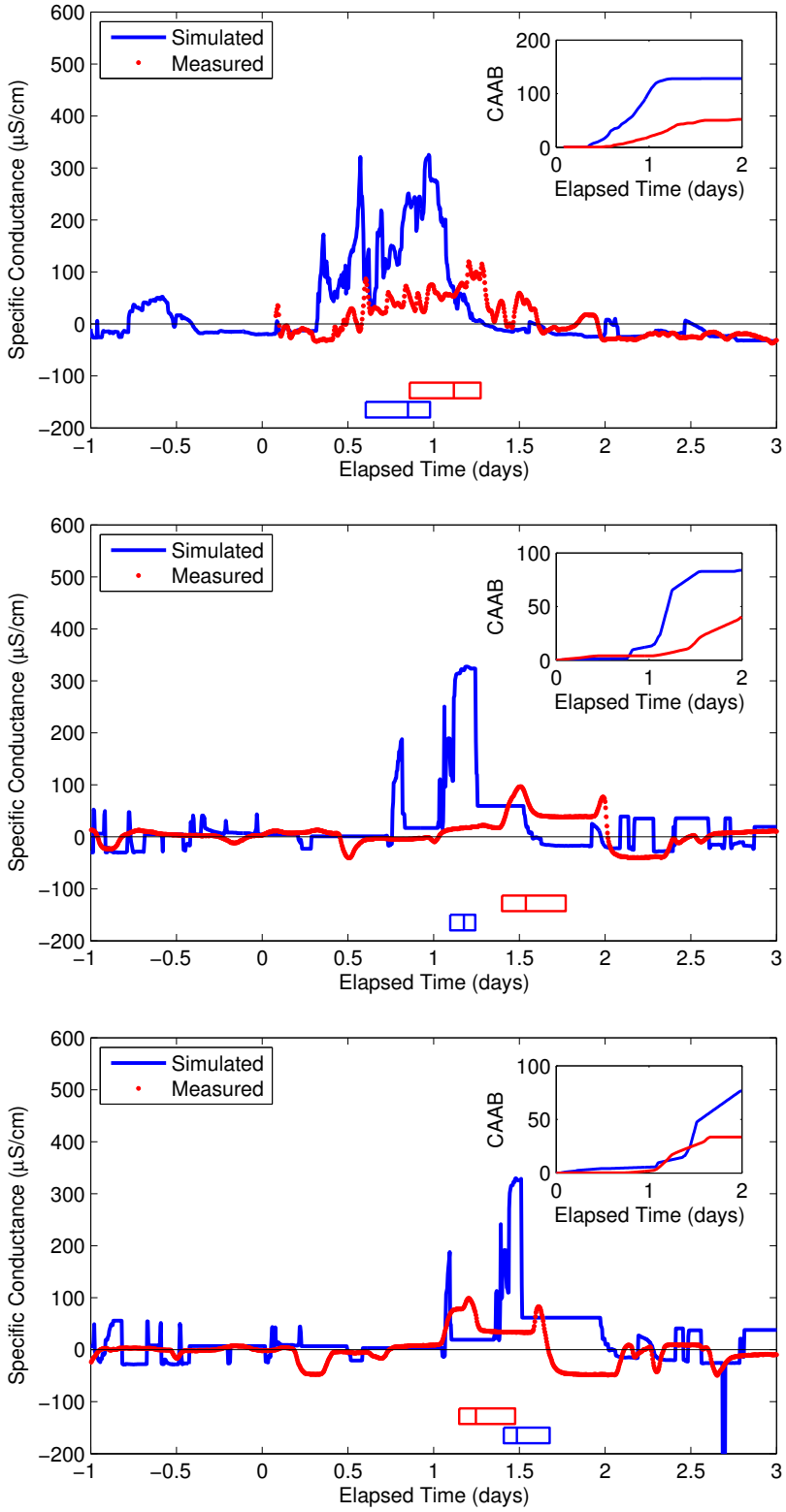

Figure 15: Observed and simulated tracer movement, locations C1 (top), C2 (middle), and C3 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

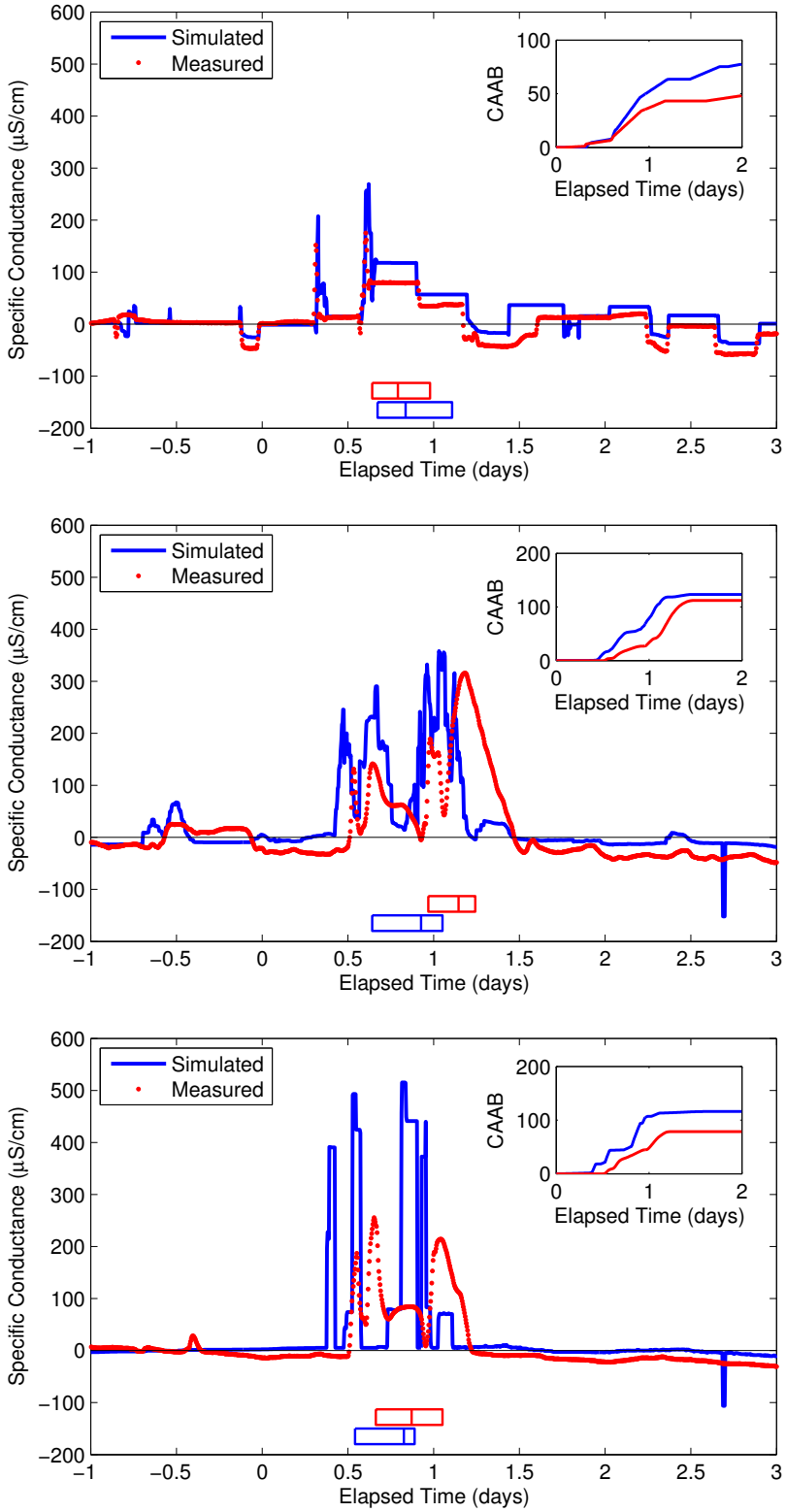

Figure 16: Observed and simulated tracer movement, locations C4 (top), C6 (middle), and C7 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series. 29

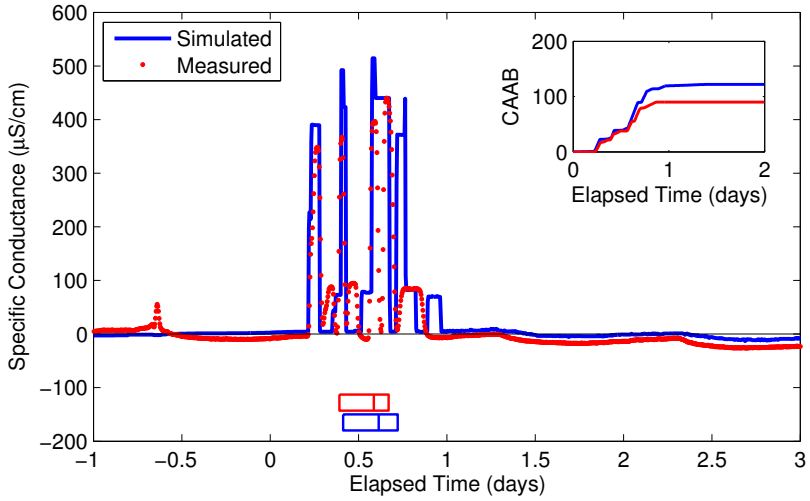

Figure 17: Observed and simulated tracer movement, locations C8. Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

that branch off of a 20 in. transmission main leading from the injection site to the south. None of these locations are on what could be called dead-end mains; they each have significant demand downstream, even if the network structure is mostly branched. Locations F1 and F2 observe three or four distinct pulses, and the simulations are slow compared to the observations. Location F3, however, observes a single pulse, even though it is downstream of F1 and F2 and off the same main. Obviously the flow is being affected downstream of F1 and F2, and in a manner that is contrary to the simulation. A booster station is located nearby and to the south of F2, off of the same 20 in. transmission main, that serves the area to the south (this area was outside of the study region). It is interesting that the real-time hydraulic simulation to the south of F3 was noticeably less accurate than other regions, including the operation of the booster station and the cycling behavior of three tanks that it serves(CitiLogics, 2013). Briefly, these three tanks are used more in terms of the depth of the fill/drain cycles, compared to the real-time hydraulic simulation behavior. Thus one explanation for the tracer simulation errors at F3 is that increased draw on the south tanks when the booster pump station is off, prevents some of the tracer pulse mass from being transported to the south, between Locations F1, F2, and Location F3. These same hydraulic errors likely affect the other locations in Region F, especially F4, although the errors in simulated signals at F5 and F7 are reasonable. The general change in character of the observed pulses at location F7 may be due, in part, to the impacts on flow from booster pumping and demands to the south.

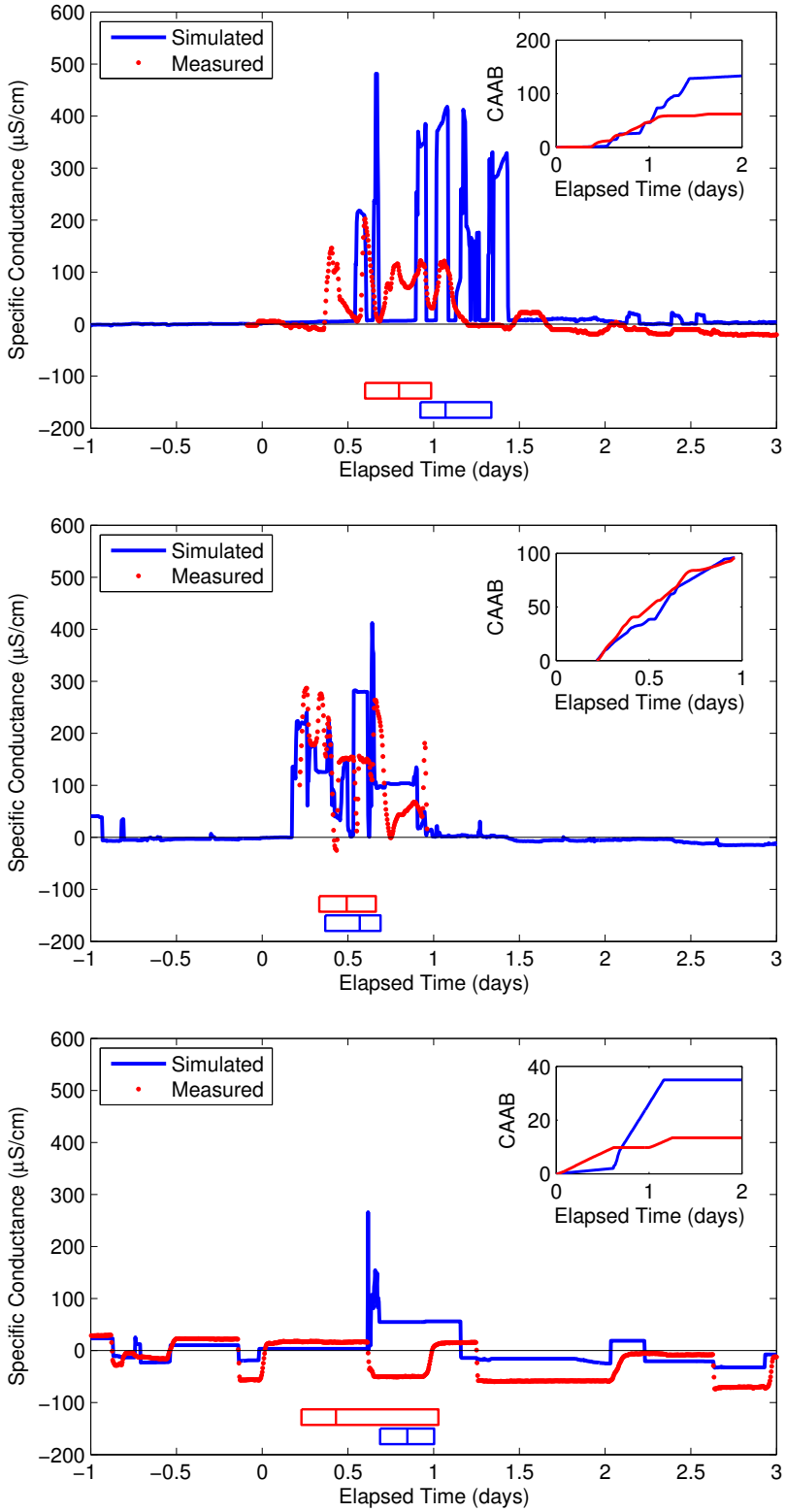

Figure 18: Observed and simulated tracer movement, locations D1 (top), D2 (middle), and D3 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

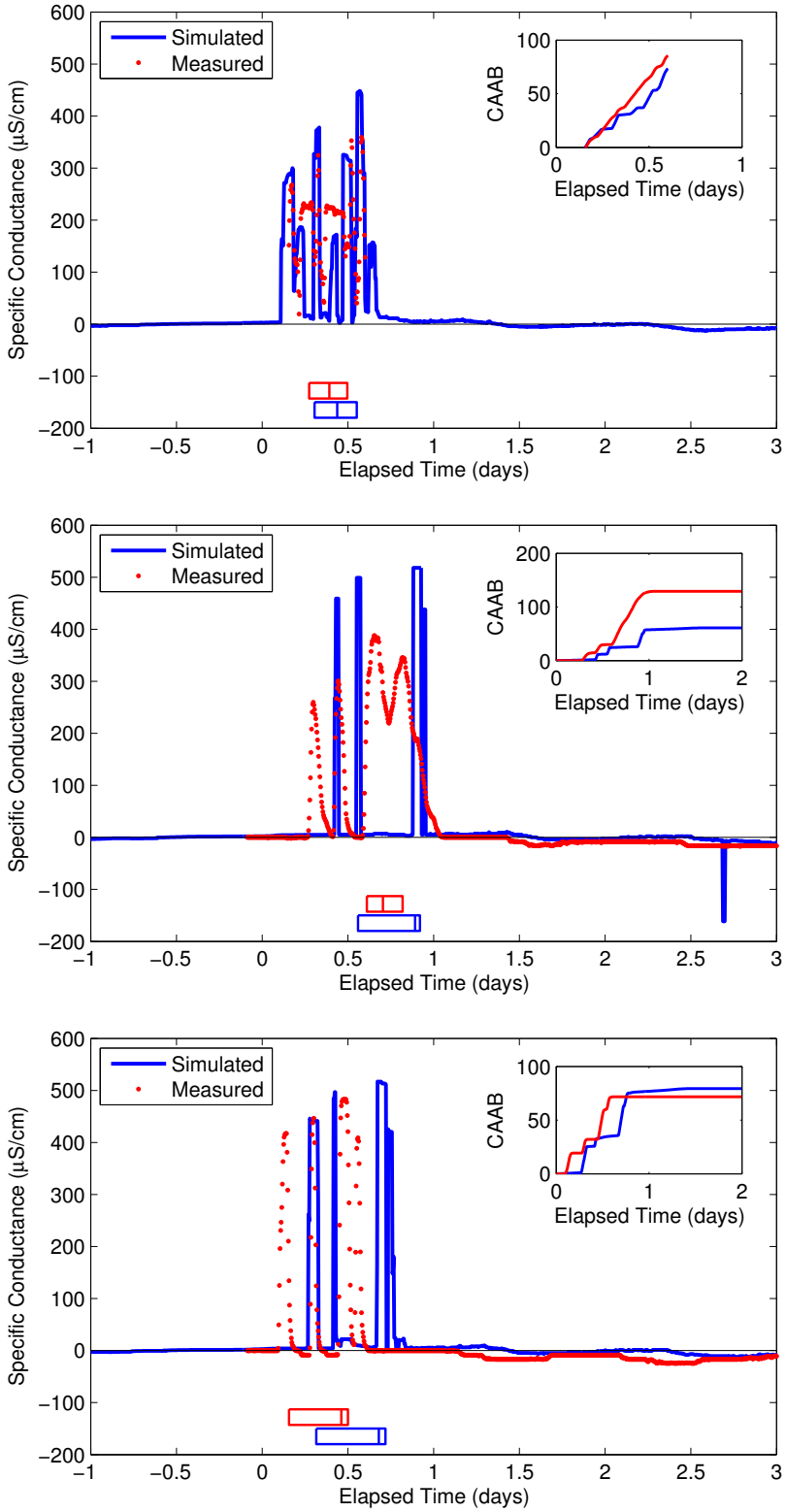

Figure 19: Observed and simulated tracer movement, locations D4 (top), D6 (middle), and D7 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

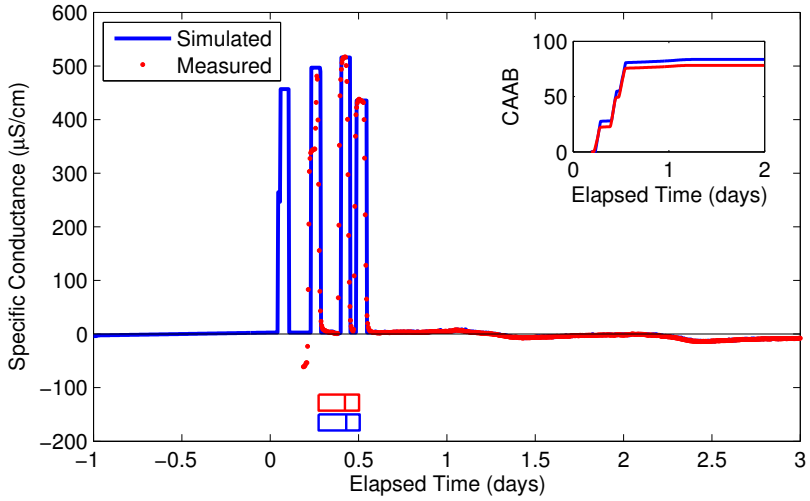

Figure 20: Observed and simulated tracer movement, locations D8. Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

4 Acknowledgements

The authors wish to thank a variety of individuals for their valuable efforts, without which the study could not have taken place. Sam Hatchett of CitiLogics, LLC contributed original material to this report as well as computer software and analysis expertise that were crucial to these results. We are especially grateful to Robert Janke of the USEPA National Homeland Security Research Center (NHSRC), who has provided steady encouragement and advice, and carried forward a vision of real-time modeling, without which the project would have stalled. John Hall of the USEPA/NHSRC helped coordinate the engagement with the NKWD. Lew Rossman and Michael Tryby of the USEPA National Risk Management Research Laboratory provided valuable guidance and technical support for the real-time model development, and the development of Epanet-RTX. Steve Allgeier, of the USEPA Office of Water, and Terra Haxton, from USEPA/NHSRC contributed valuable suggestions as members of the Epanet-RTX project committee. Ernesto Arandia of the University of Cincinnati, and Tom Taxon of Argonne National Lab, participated on the original Epanet-RTX design team, and helped to craft the software architecture that supported this study. Dick Males and Walter Grayman, both private consultants in Cincinnati, OH, gave valuable advice as part of the Epanet-RTX design review committee. Kevin Morley of the American Water Works Association, along with Rob Janke, helped to fund and organize water utility workshops in 2010 that provided valuable industry feedback about the functionality that water utilities would value from a real-time modeling system. Rob Janke, John Hall, Jeff Szabo, and Terra Haxton assisted with setting up and locating field water quality sampling stations and making sure the monitoring devices were working properly, and with Srinivas Panguluri and others at Shaw Environmental, provided support for calibration of monitoring devices. Joe Goodin from

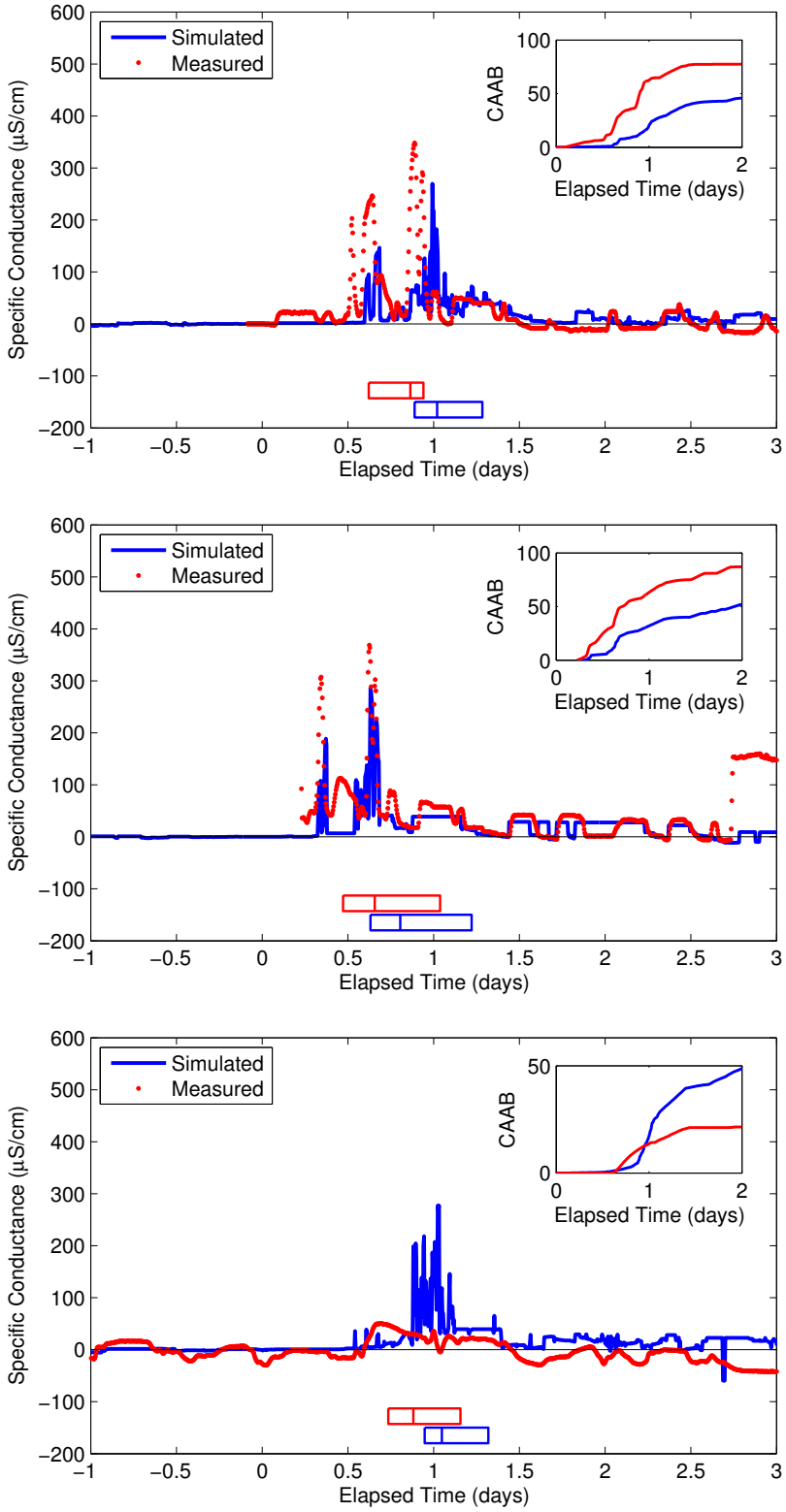

Figure 21: Observed and simulated tracer movement, locations E1 (top), E2 (middle), and E3 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

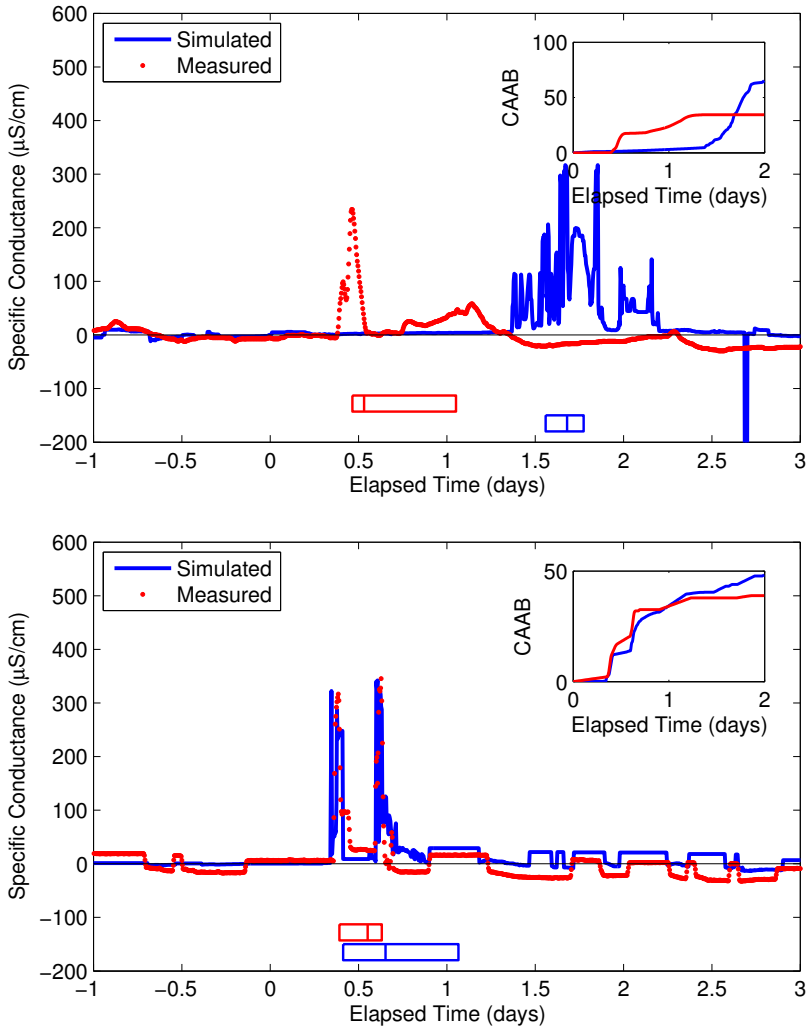

Figure 22: Observed and simulated tracer movement, locations E4 (top) and E6 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

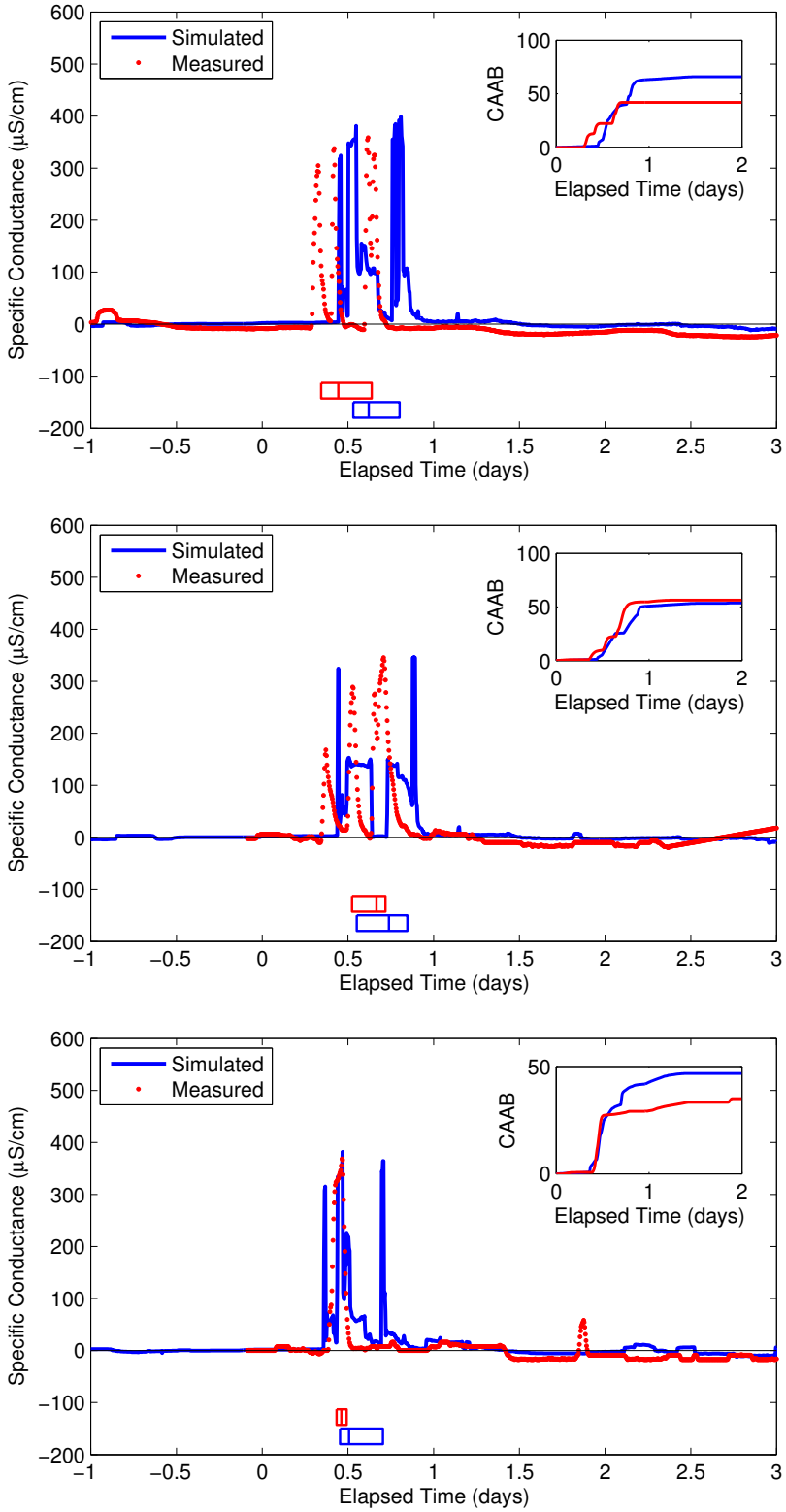

Figure 23: Observed and simulated tracer movement, locations F1 (top), F2 (middle), and F3 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

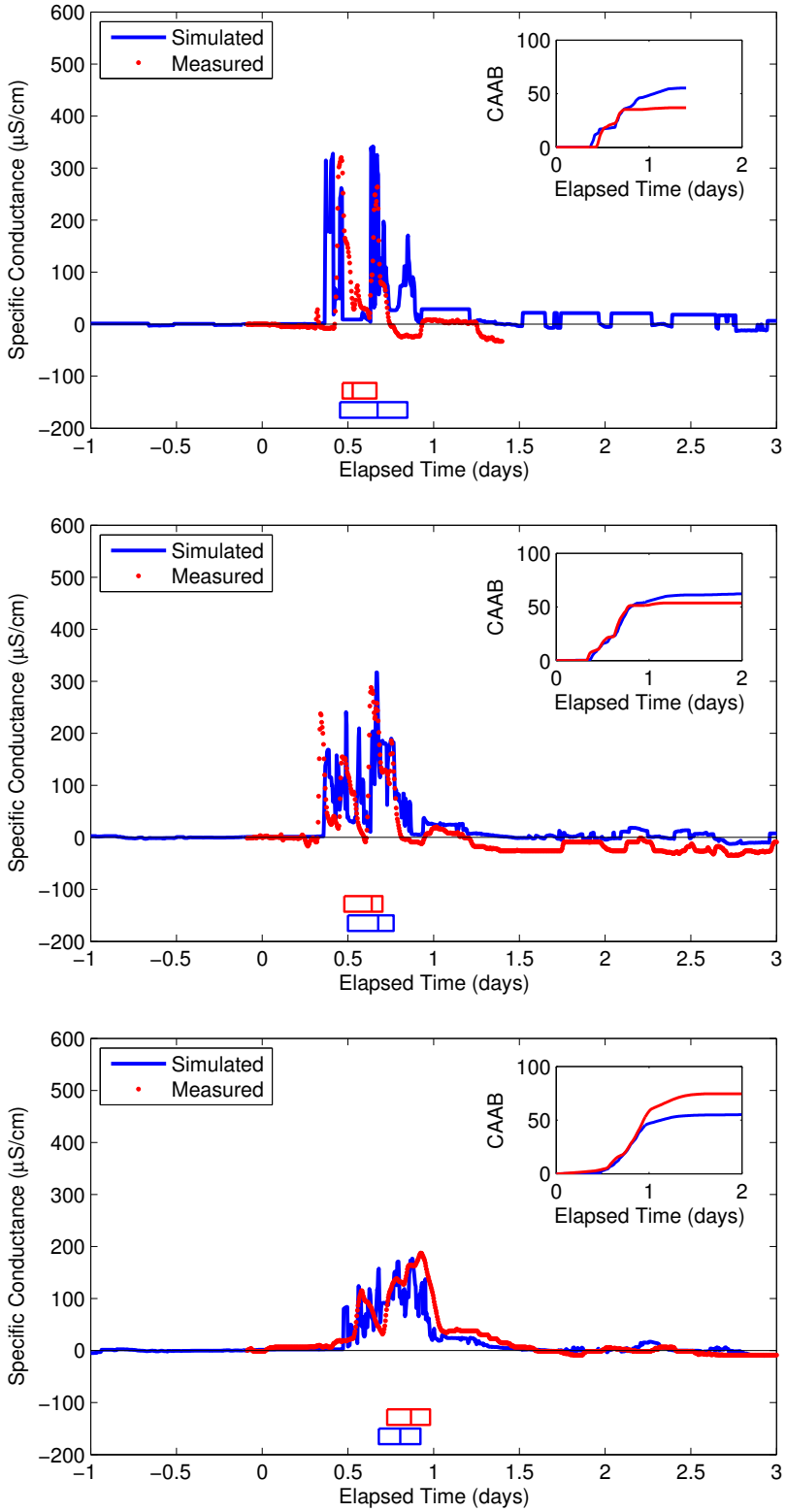

Figure 24: Observed and simulated tracer movement, locations F4 (top), F5 (middle), and F7 (bottom). Inset figure shows the conductivity area above background over a two day period commencing with the start of tracer injection, for both simulated and observed series. The symbols along the x-axis show the inter-quartile ranges (outer box boundaries) and medians (lines within boxes) for each series.

the University of Kentucky also assisted with setting up the water quality sampling stations. Morris Maslia from the Centers for Disease Control was instrumental in helping to secure necessary field equipment. Last but not least, individuals from the Northern Kentucky Water District have provided many hours of their time, and full access to information, that was critical for making this project possible. These included Operations Supervisors Amy Matracia and Bill Wulfeck, Information Technology manager William Stewart, and especially Engineering Supervisor Amy Kramer, who led the NKWD support effort. The active support of senior NKWD management was also vitally important, including Vice President of Engineering, Richard Harrison, and CEO Ron Lovan. Finally, we wish to thank Lindell Ormsbee for his guidance as Principal Investigator of the grant under which this work was, in part, undertaken.

Funding for this research was provided by the U.S. Department of Homeland Security, Science and Technology Directorate, through a technology development and deployment program managed by The National Institute for Hometown Security, under an Other Transactions Agreement, OTA #HSHQDC07300005, Subcontract #0210UK. This support is greatly appreciated. Additional financial support was also provided by the U.S. Environmental Protection Agency's (EPA) Office of Research and Development, under Work Assignment WSD 2-29 "Field Demonstration of a Real-Time Water Infrastructure Monitoring and Data Fusion Technology to Improve Operations and Enhance Security of Water Systems." The views expressed in this report are those of the authors, however, and do not necessarily reflect the views or policies of EPA. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. Finally, crucial in-kind support was provided by CitiLogics, LLC, Newport, KY, including the use of proprietary software and specialized technical support.

A Appendix A: Prediction accuracy of chloride levels based on measured specific conductance

The purpose of this analysis was to estimate the accuracy of chloride level estimates based on specific conductance. Paired chloride and specific conductance measurements collected from NKWD finished water (Fort Thomas Treatment Plant) between 9/7/11 and 8/22/12 were used in this analysis. A simple linear regression model was applied to the data set, as shown in Figure 25. The prediction interval at a 99% confidence level was calculated. This

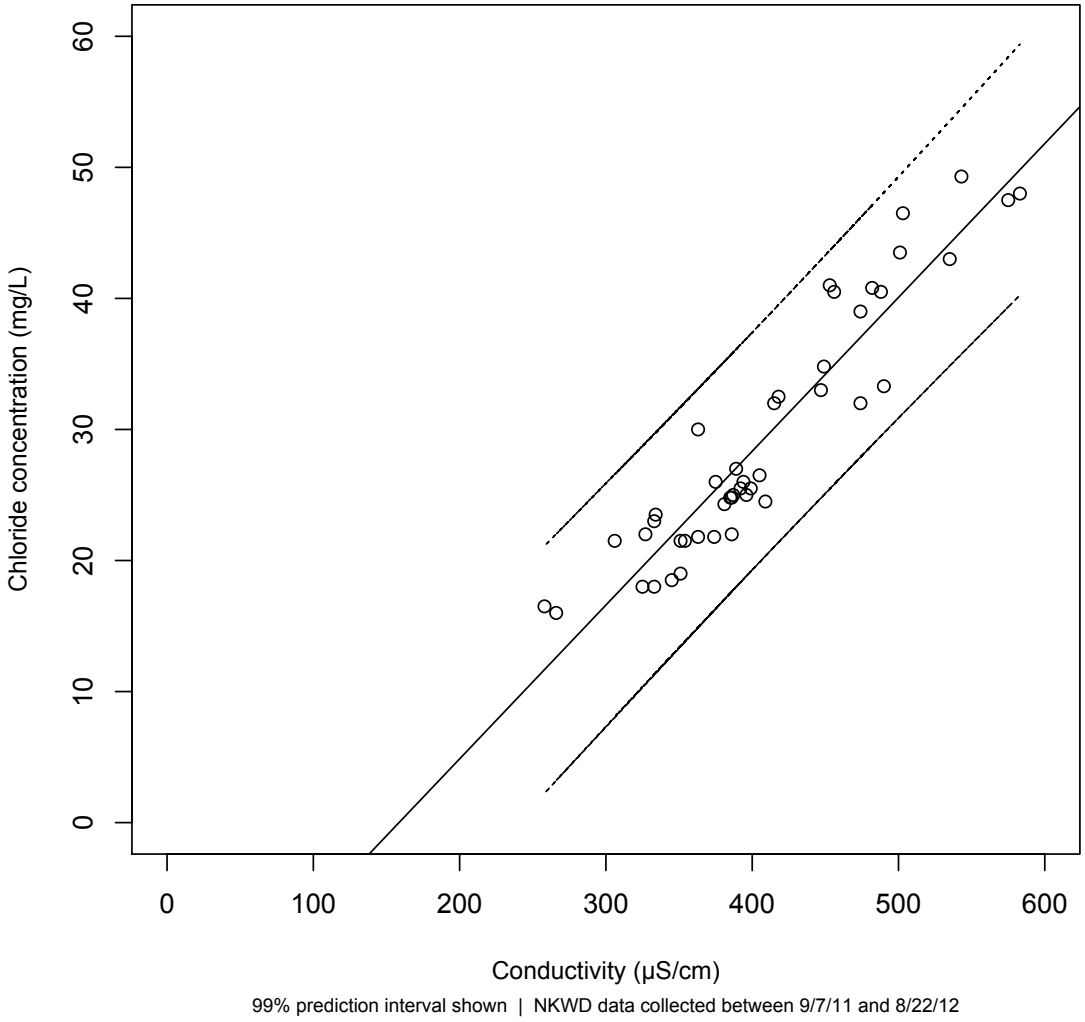

Figure 25: Relationship between specific conductance and chloride measurements

provides the ability to determine a range for chloride concentration based on a given specific conductance, and based on this data set. The following equation represents the lower bound ($[Cl^-]_{lb}$)on the prediction interval:

$$[Cl^-]_{lb} = 0.117EC - 27.5, \tag{6}$$

The upper bound ($[Cl^-]_{ub}$) is found by:

$$[Cl^-]_{ub} = 0.118EC - 9.68, \tag{7}$$

Using these two equations for a measured specific conductance value provides the 99% prediction interval on chloride concentration in that sample. For example, given a measured specific conductance of approximately 400 $\mu S/cm$, we can be state that the measured chloride concentration will fall between 19 and 37 mg/L, with 99% confidence.

B Appendix B: NSF Grade Liquid Calcium Chloride Product Data Sheet

NSF GRADE LIQUID CALCIUM CHLORIDE

Product Data Sheet

TETRA CHEMICALS

General Description

NSF Grade liquid calcium chloride ($CaCl_2$) is an odorless, slightly alkaline, colorless fluid with a typical concentration of 28 to 38 percent.

Applications

TETRA NSF Grade liquid calcium chloride is used for water treatment and complies with ANSI/NSF 60.

Availability

NSF Grade liquid calcium chloride is available from select plant and terminal locations throughout North America. For the location nearest you, refer to the plant and terminal map available on our website (www.tetrachemicals.com) or contact your TETRA sales or customer service representative.

Safety and Handling

Calcium chloride liquid is a strong salt solution. Protective clothing, rubber gloves and eye protection are recommended. Rubber safety boots should also be worn in work areas, since calcium chloride can damage leather. This product should be handled in areas with proper ventilation. Before using this product, refer to the MSDS (available on the Company's website) for complete safety and handling guidelines. For proper disposal guidelines for calcium chloride wastes, consult the appropriate local regulatory authorities.

PHYSICAL PROPERTIES	
Appearance	colorless liquid
Odor	None
Assay	28 to 38% by weight $CaCl_2$
Crystallization Temperature	-38°F (-39°C) to 42.1°F (5.6°C)
Specific Gravity @ 68°F (20°C)	1.264 to 1.3785
Bulk Density	10.53 to 11.49 lb/gal

CHEMICAL PROPERTIES	
Chemical	$CaCl_2$
pH	Slightly alkaline
Impurities (on 100% $CaCl_2$ basis)	
Alkali Chlorides	< 0.1% by weight
Magnesium (as $MgCl_2$)	< 0.1% by weight
Other Impurities (not H_2O)	< 1.0% by weight

TETRA Chemicals
25025 Interstate 45 North, Suite 600
The Woodlands, Texas 77380
Phone: 281.367.1983
Customer Service: 800.327.7817
Fax: 281.298.7150
www.tetrachemicals.com

Because use conditions and applicable laws may differ from one location to another and may change with time, Customer is responsible for determining whether products and the information in this document are appropriate for Customer's use and for ensuring that Customer's workplace and disposal practices are in compliance with applicable laws and other governmental enactments. Seller assumes no obligation or liability for the information in this document. NO WARRANTIES ARE GIVEN; ALL IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE ARE EXPRESSLY EXCLUDED. Further, nothing contained herein shall be taken as a recommendation to manufacture or use any of the herein described materials or processes in violation of existing or future patents.

Copyright © 2004 TETRA Technologies, Inc. All rights reserved.
TETRA and the TETRA logo are registered trademarks of TETRA Technologies, Inc.

January 12, 2005
NSFGrade

References

CitiLogics (2013). Case study of real-time network hydraulic modeling: Data transformation, model calibration, and simulation accuracy. Technical report, wsd 2-29, field demonstration of a real-time water infrastructure monitoring and data fusion technology to improve operations and enhance security of water systems, CitiLogics, LLC, 615 Madison Ave, Covington, KY.

Rossman, L. A. (2000). Epanet 2 user's manual. Technical report, U.S. Environmental Protection Agency.

USEPA (2013). Epanet-rtx real-time extension for the epanet toolkit. Open source software object library, U.S. Environmental Protection Agency. Available from http://openwateranalytics.github.io/epanet-rtx/.