SAND2007-6559P

Enhanced Water Quality Event Monitoring and Identification

Sandia National Laboratories
Presenter: Sean A. McKenna
samcken@sandia.gov
Local phone: 9381 0975

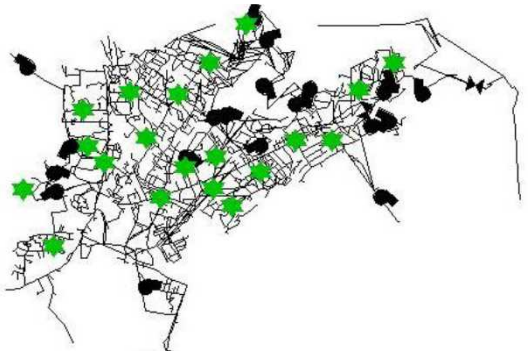
Sandia National Laboratories

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

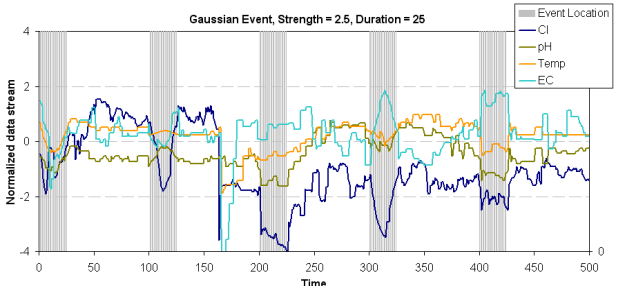
Outline

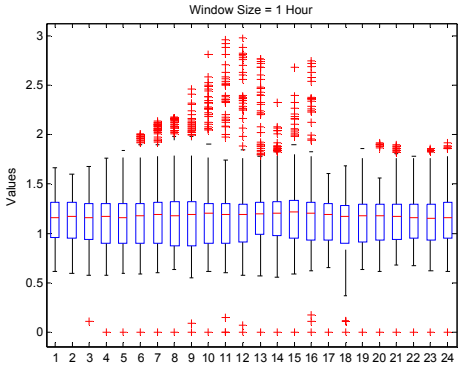
- Background
- Identified Application Areas
- Technical Approaches
- Tasks and Subtask Breakdown
- Costs and Schedule

Background



- Sandia National Laboratories has worked for the past 5 years to develop security solutions for water distribution networks
  - Focus on three areas
    - Event detection from surrogate parameter monitoring
    - Automatic source location identification
    - Optimal determination of monitoring locations





- PUB water distribution network
  - Complex network with on-line and grab sample monitoring
  - International recognition for utility operations

Sandia National Laboratories

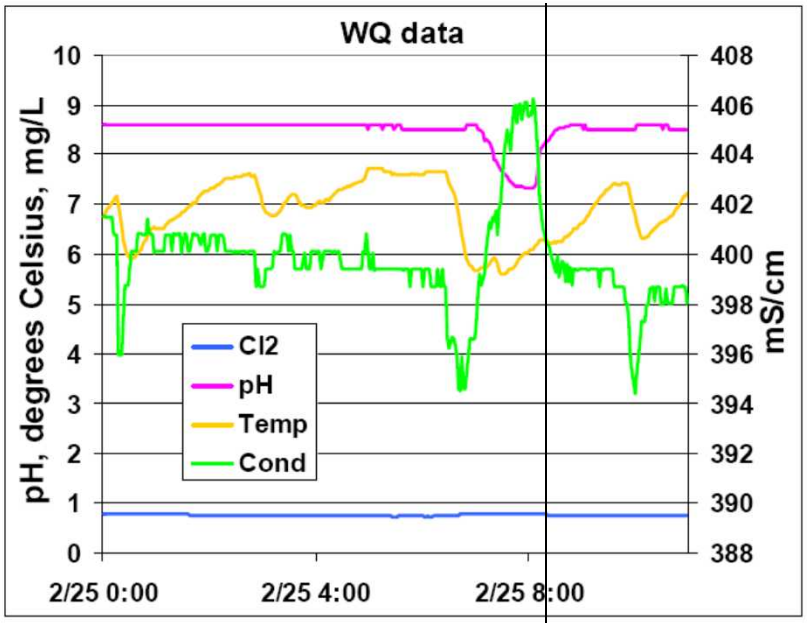
Application Areas

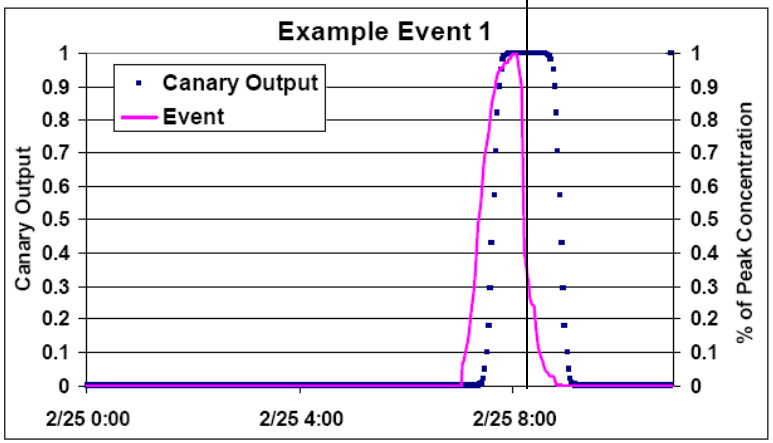
- Improved network monitoring and event identification
  - In-line real time monitoring of surrogate parameters in the upper network (detection and identification)
  - Combine grab sample, surrogate monitoring and customer inquiry data to identify water quality event source location
- Improved monitoring of the system provides dual-use benefits
  - Greater security in rapid detection of anomalous water quality
  - More efficient operations – identify and mitigate any operational causes of anomalous water quality

CANARY

- "Event" Detection software
  - "Event" is anomalous water quality for a specified length of time
  - Water quality signals are input and Prob(event) is output
- Works in two modes
  - Comparing new measurements to recent measurements (moving window) to find outliers
  - Pattern recognition from a library of patterns

CANARY Result





Example application of CANARY in a US utility

Four water quality signals

Results show unknown % of peak contamination and P(event) as identified by CANARY

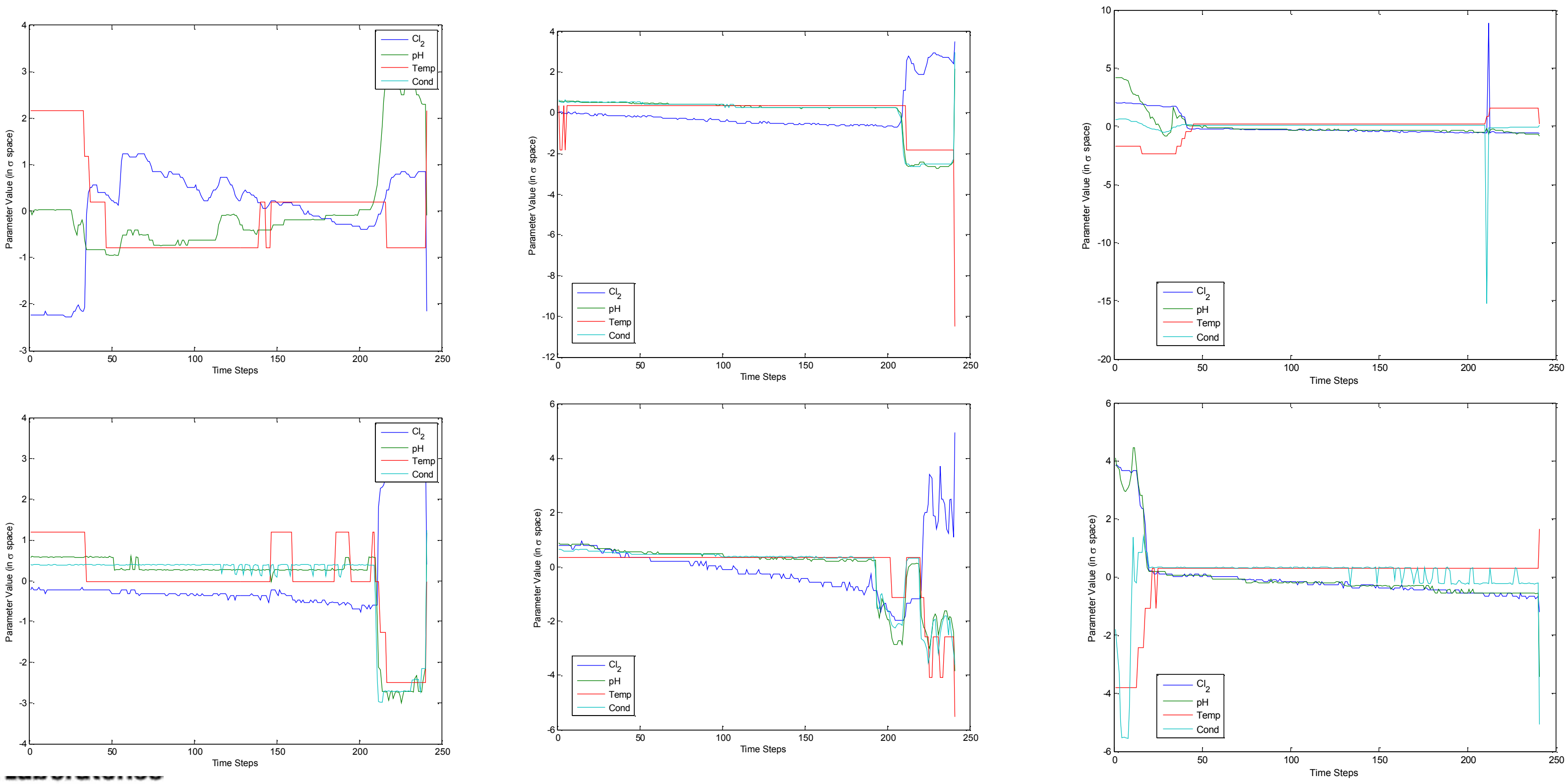
Note the long stretch of zero probability prior to the event

Lag time in detection is adjustable

Laboratories

CANARY Pattern Recognition

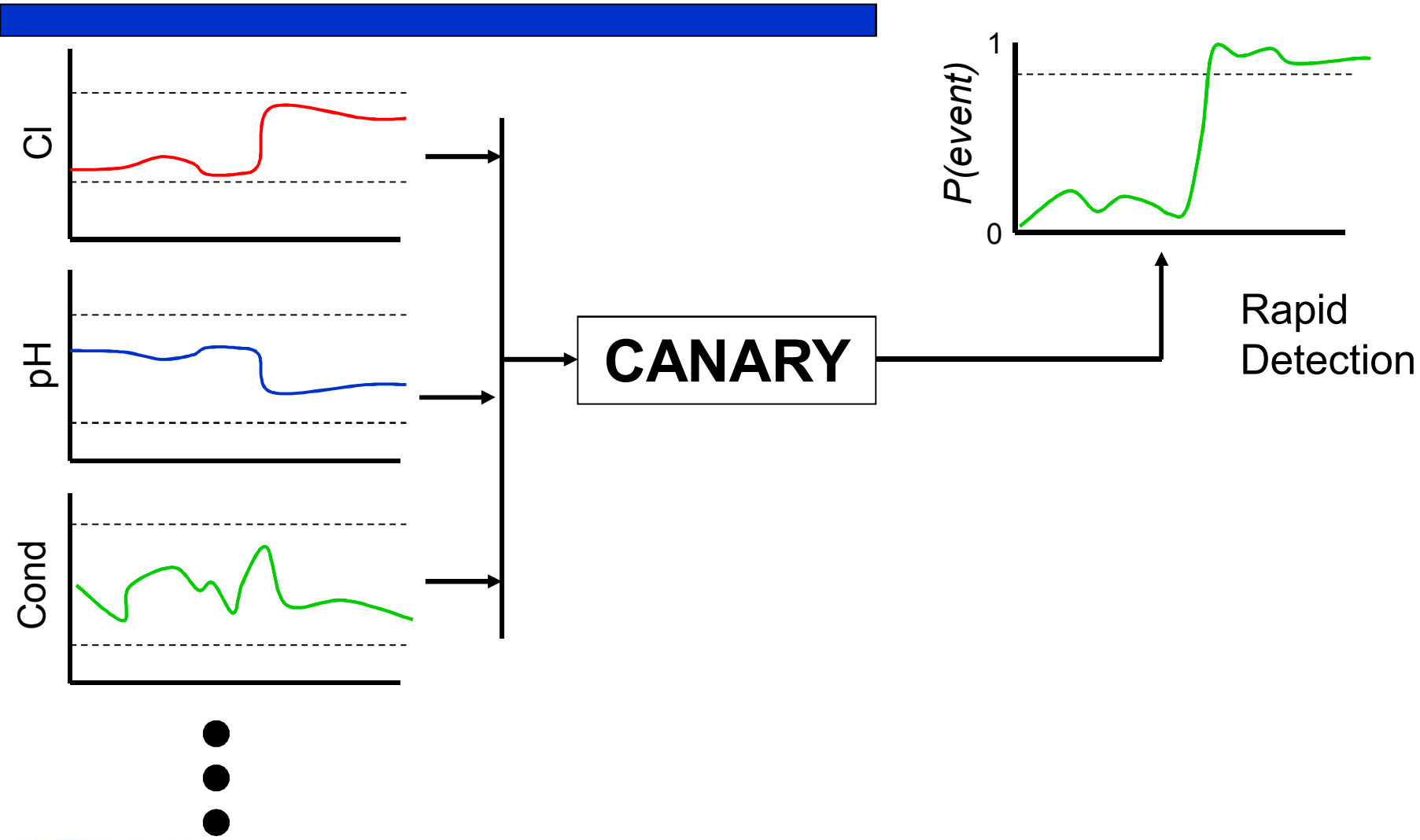
- Acceptable patterns pulled from training data
  - Patterns caused by mixing of source waters



PUB Application

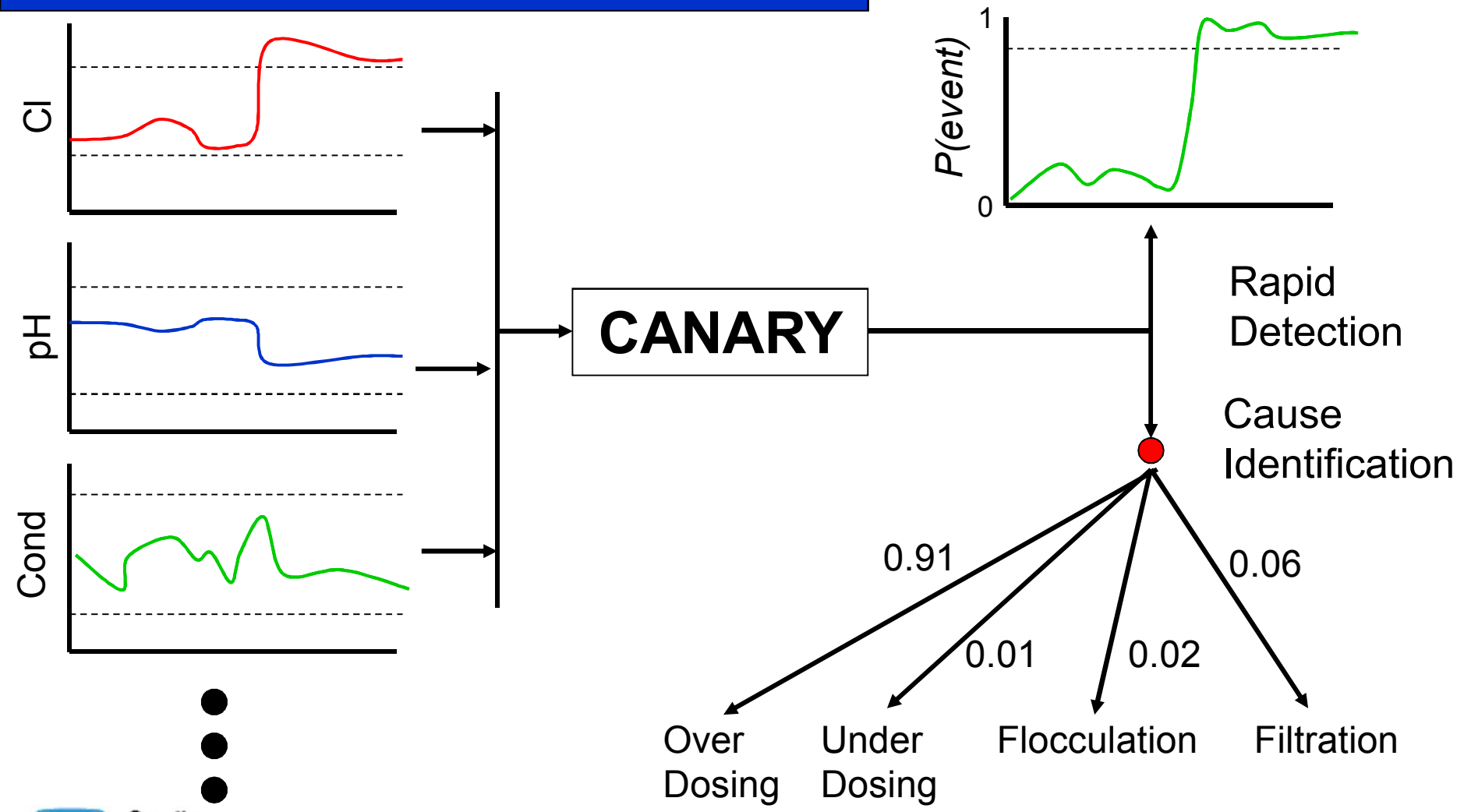
- Use CANARY for more efficient monitoring of source plant variations including intermediate water quality monitoring
  - Faster identification of variations in water quality due to treatment plant operations
  - Train CANARY to recognize water quality patterns related to specific treatment plant operations
  - Recognize bad water before it leaves the plant

PUB Application



Sandia National Laboratories

PUB Application



Identification template matching based on historical data

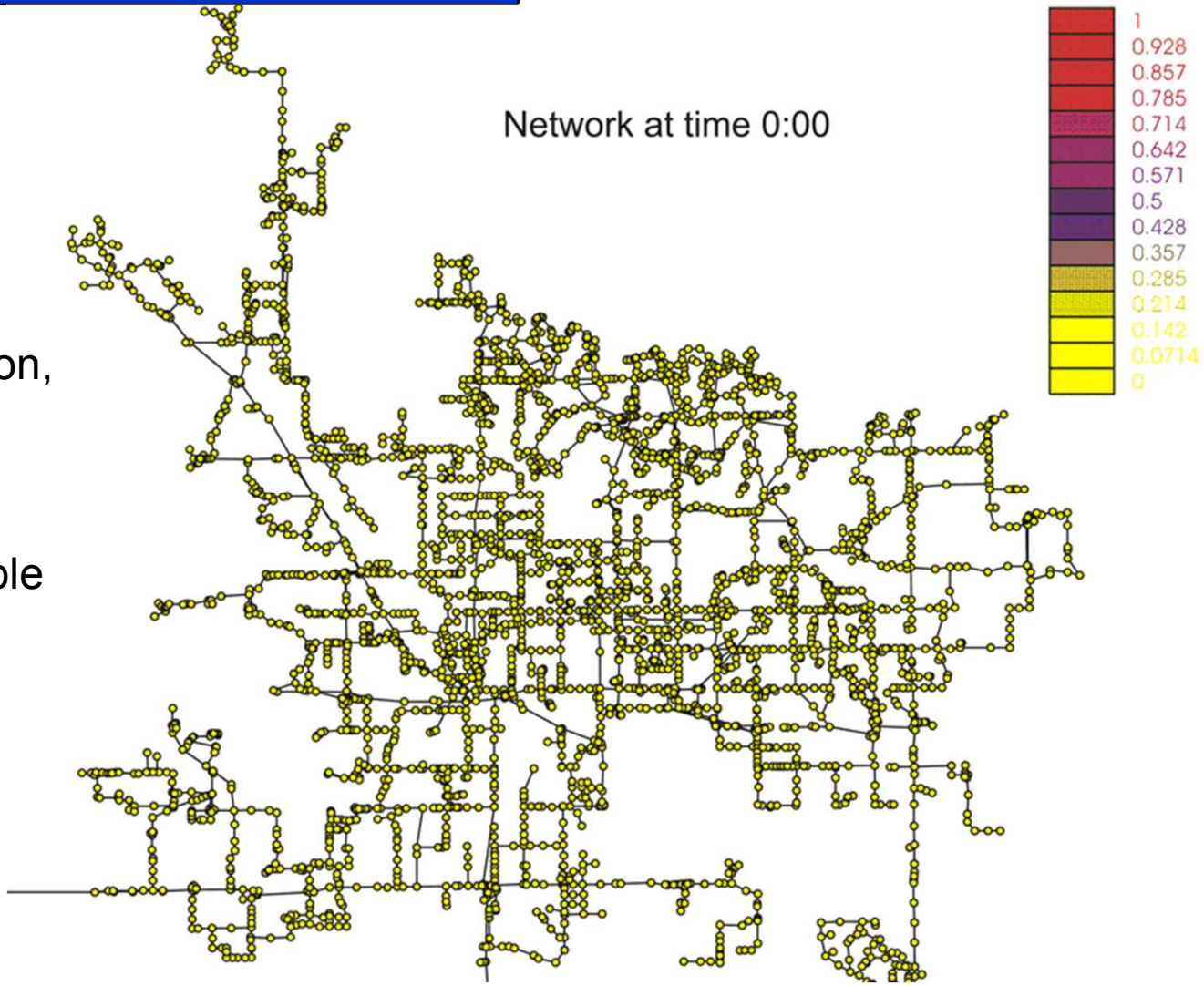
Sandia National Laboratories

Source Location Inversion

- Given some indication of an event in the system, can we determine the source location for the event?
  - Do this in real time using on-line monitoring
  - Do this on complex networks with minimum of number of sensors
  - Accurate determination of source location

Contamination Simulation

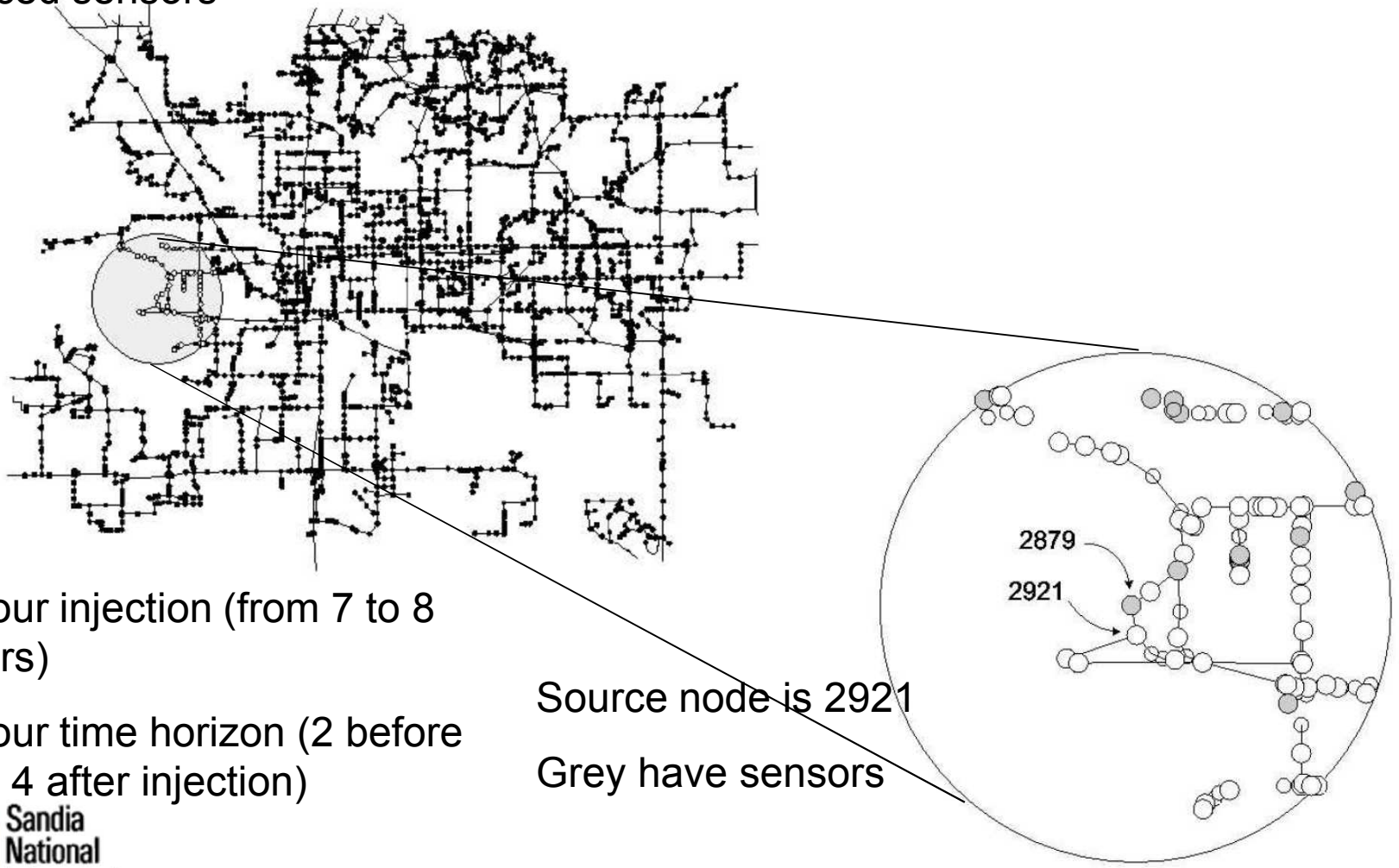
Unknown source location, strength and timing as well as unknown fine-scale demand patterns make it nearly impossible to predict results of contamination event



Sandia National Laboratories

Example (Large Network)

Approximately 3500 nodes, 350 randomly placed sensors



1 hour injection (from 7 to 8 hours)

6 hour time horizon (2 before and 4 after injection)

Source node is 2921

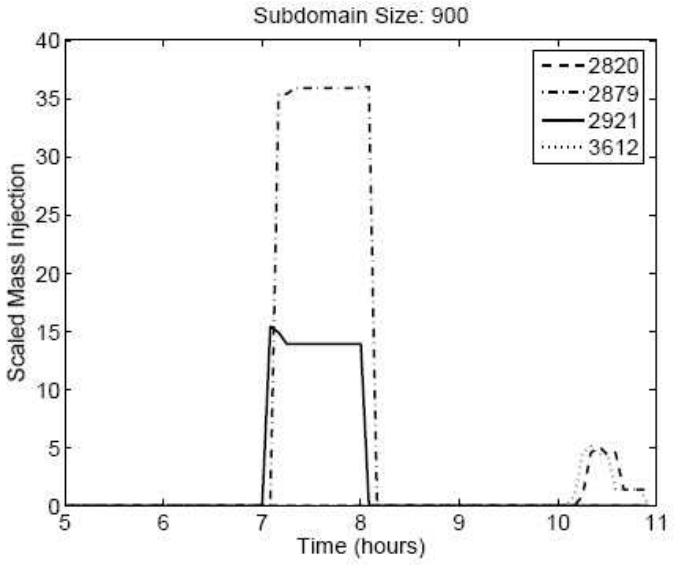
Grey have sensors

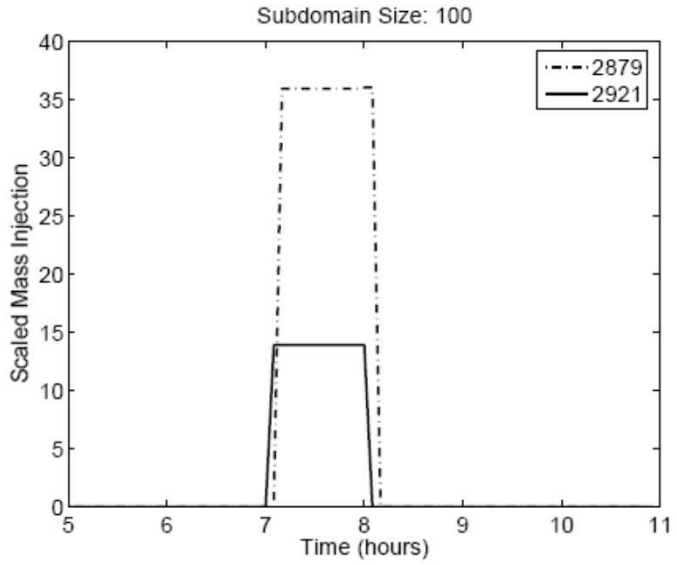
Sandia National Laboratories

(*From: Laird, Biegler, and van Bloemen Waanders, 2006*)

Large Network: Results

- Automatically break large network into smaller subdomains around sensor locations with contamination
- Solution determines how much mass came from every node in the subdomain and what time it was injected





5 minute time steps, real-time solutions on a 1.8GHz Pentium 4 computer

(*From: Laird, Biegler, and van Bloemen Waanders, 2006*)

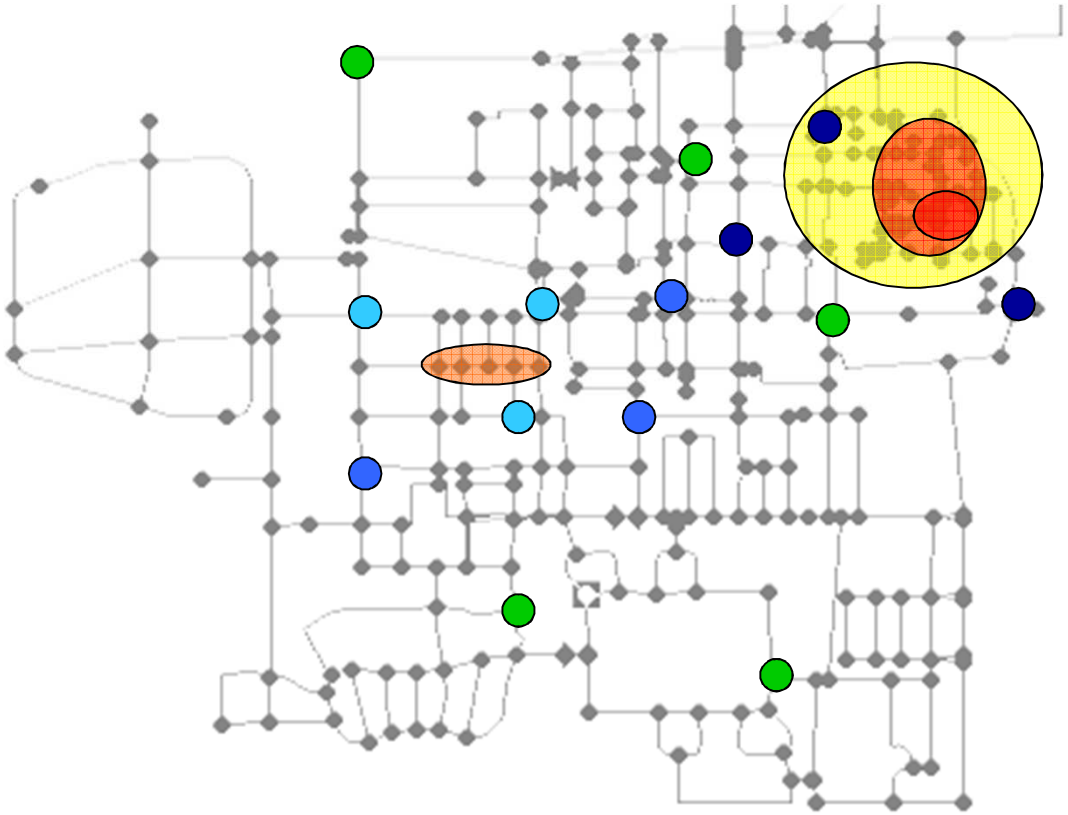
Sandia National Laboratories

Implementation at PUB

- Previous work with source location inversion has used continuous response of a sensor or CANARY
  - Apply to surrogate parameter monitoring in the upper network, and add other data sources from lower network
    - Customer inquiries (already recorded by PUB)
    - Grab Sample Data (weekly sampling plus follow up samples to inquiries)

Implementation at PUB

Simple schematic example of the process using a portion of an example network (not the PUB network)



- Surrogate Monitoring
- Consumer Calls
- Initial Grab Samples
- Round 2 Grab Samples
- Round 3 Grab Samples

Sandia National Laboratories

Task 1: Process Monitoring

- Add CANARY to surrogate parameter monitoring system with goals of faster alarms and identification of alarm source
  - PUB Subtasks
    - P1: Provide several years of training data with alarm periods identified
    - P2: Provide understanding & characterization of processes and events they cause
    - P3: Connection of CANARY to SCADA database (real time)
  - Sandia Subtasks
    - S1: Reevaluate template matching algorithm for use on PUB surrogates
    - S2: Populate pattern library using historical data from PUB
  - Joint Subtasks
    - J1: Implement pattern matching in real time operations (start with a few sites and then expand)
    - J2: Test and adjust the operation of CANARY for optimal performance

Task 2: Source Location

- Develop source location identification approach for water quality events using surrogate parameter monitoring, customer complaints and grab sample data
  - PUB Subtasks
    - P1:Run all water quality escalation scenarios under consideration using SynerGEE and store in a database that can be made accessible (big job)
    - P2: Provide information on previous events (complaints, suspected source, grab sample data, etc.)
  - Sandia Subtasks
    - S1: Incorporate surrogate monitoring data
    - S2: How to use complaint data in source location identification?
    - S3: How to use grab sample data in source location identification?
  - Joint Subtasks
    - J1: Testing of approach on historic events

Sandia National Laboratories

What is needed in terms of a software "tool" for this task?

Sandia Budget

Task / Subtask	Task/Subtask Name	Sandia Costs (k$ US)	Sandia Costs (k$ Sing)
1	**Improved Process Monitoring with CANARY**	**225**	**337.5**
S1.1	Extend Template Matching Algorithm	75	112.5
S1.2	Populate Template Library	40	60
J1.1	Real-Time Implementation of Pattern Matching	50	75
J1.2	Optimize CANRY Performance on PUB network	60	90
2	**Source Location Identification**	**300**	**450**
S2.1	Develop approach for joint use of surrogate and grab sample data	85	127.5
S2.2	Conditioning scenarios on complaint data	60	90
S2.3	Conditioning scenarios on grab sample data	60	90
J2.1	Testing on historic events	60	90
J2.2	Data worth evaluation	35	52.5
Mgmt/ Travel	**Project management duties outside of tasks and travel to and from US for project staff**	**40**	**60**
Totals		**565**	**847.5**

Sandia National Laboratories

Proposed Schedule

- Task 1: Duration of 6 months
- Task 2: Duration of 12 months
  - Quarterly progress reports for both tasks

Task 1: Process Monitoring

Task 2: Source Identification

Sandia National Laboratories