

# An Online Algorithm For A Low Cost Real-Time Sensor-Based Fenceline Leak Detection System

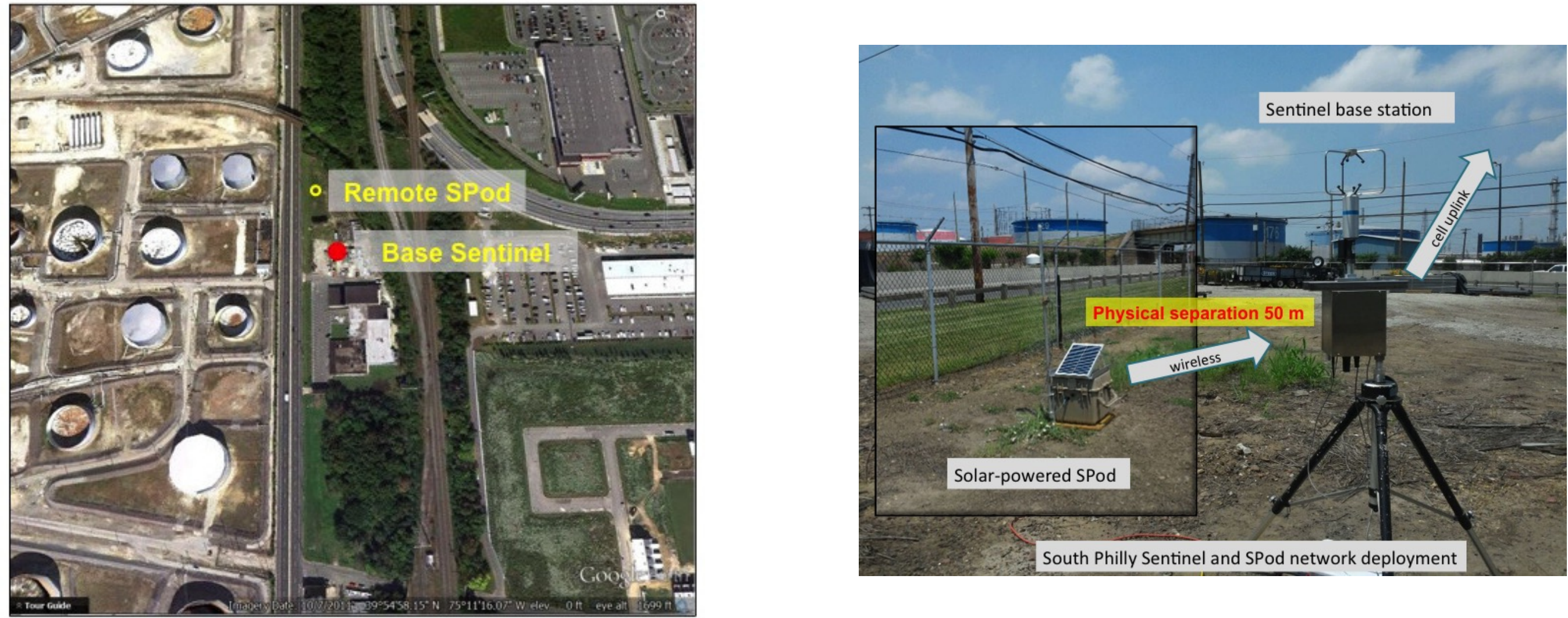
Halley Brantley, Yichen Si, & Dendi Suhubdy

North Carolina State University

## Introduction

Benzene is an air pollutant and carcinogen that is regulated by the U.S. Environmental Protection Agency (EPA) and often emitted by petroleum refineries as a result of leaking equipment and wastewater treatment [1]. Cost-effective and real-time methods for monitoring hazardous air pollutants emitted from industrial facilities could lead to faster leak detection, repair, and reduced emissions. One monitoring method currently under development by the EPA’s Office of Research and Development is the Sensor NeTwork INtelligent Emission Locator (SENTINEL) that combines time-resolved measurements and inverse models of various forms to improve source understanding on a variety of spatial scales. We present here an algorithm for locating the source of emissions using 1-sec benzene sensor and wind data.

## Data Description



- Solar-powered ”sensor pod” (SPod) controlled by a low cost and low power Arduino UNO computer
- Wired SENTINEL base station containing an Intel Atom™ computer (Santa Clara, CA, USA) and a model 81000V 3-D Ultrasonic Anemometer (R.M. Young, Inc., Traverse City, MI, USA) for 1 Hz wind measurements.
- Both systems log time-synchronized 1 Hz data from a custom EPA-developed sensor board containing a 10.6 eV passive photoionization detector [PID] (blue label piD-TECH®, Baseline-Mocon Inc. Lyons, CO, USA) used to detect VOCs (benzene).
- Deployed near a refinery in South Philadelphia from July, 2014 through February, 2015. Data from November, 2014 is analyzed here as an illustration.
- PID sensor response and stability are affected by temperature, humidity, lamp deposition and other variables resulting in unwanted baseline drift that complicates data analysis.
- The PID sensor output ranges from 0.2 to 5 volts.

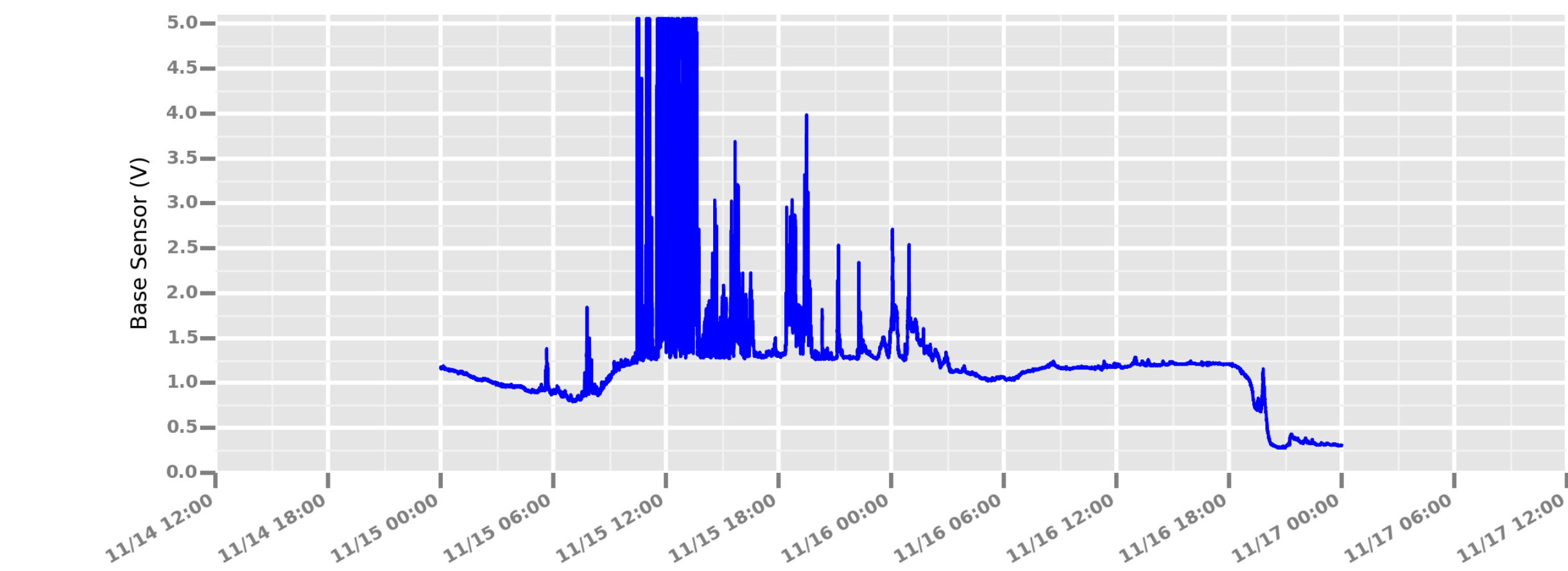


Figure 1: Example Base Sensor Signal

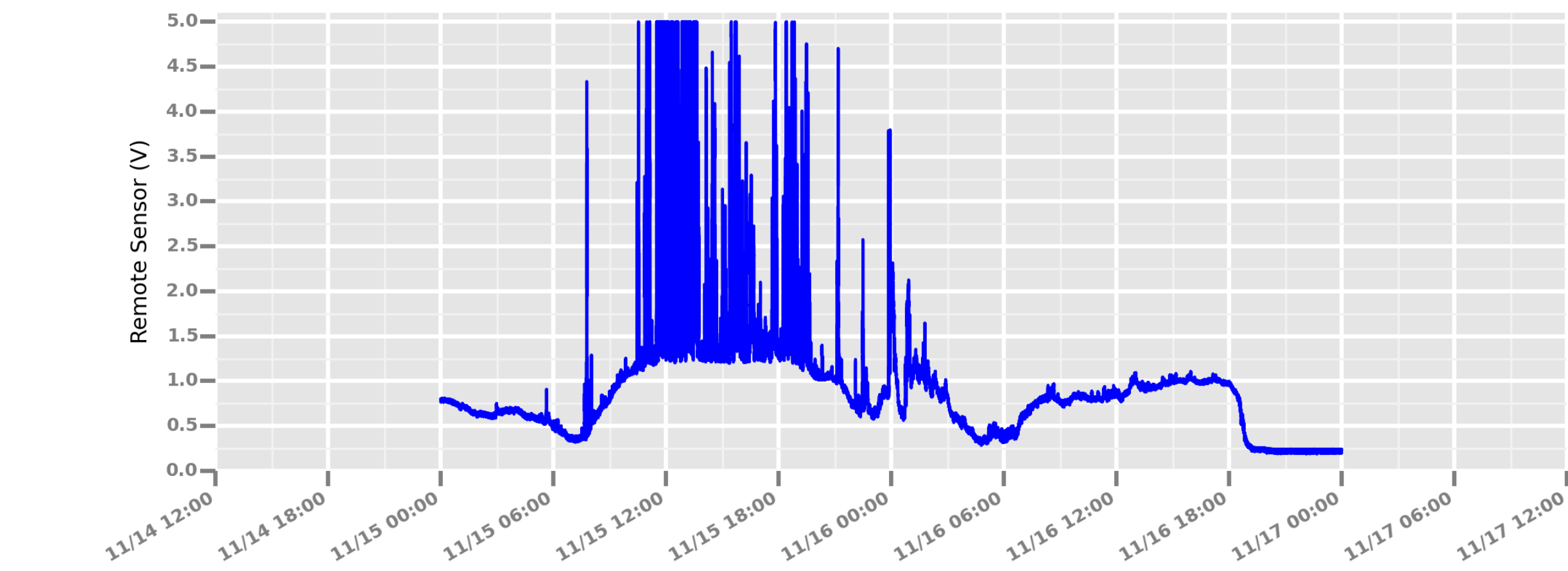


Figure 2: Example SPOD Sensor Signal

## Simulated Data

- Time series for one day were simulated for each of the sensors.
- Concentration was simulated at 0.05 Hz using the QUIC Dispersion Modeling System [2] and 5 min averaged wind measurements and up-sampled to 1 Hz by repeating the previous value.
- Each hour of the stochastic baseline was simulated as 1D Brownian bridge with each step drawn from a standard normal.
- An offset of 1.5 was added to the simulated baseline, to imitate the measured values
- After the simulated baseline and concentration were added, values above above 5 were censored to 5 and values below 0.25 were replace with random draws from a  $N(0.25, 0.05)$

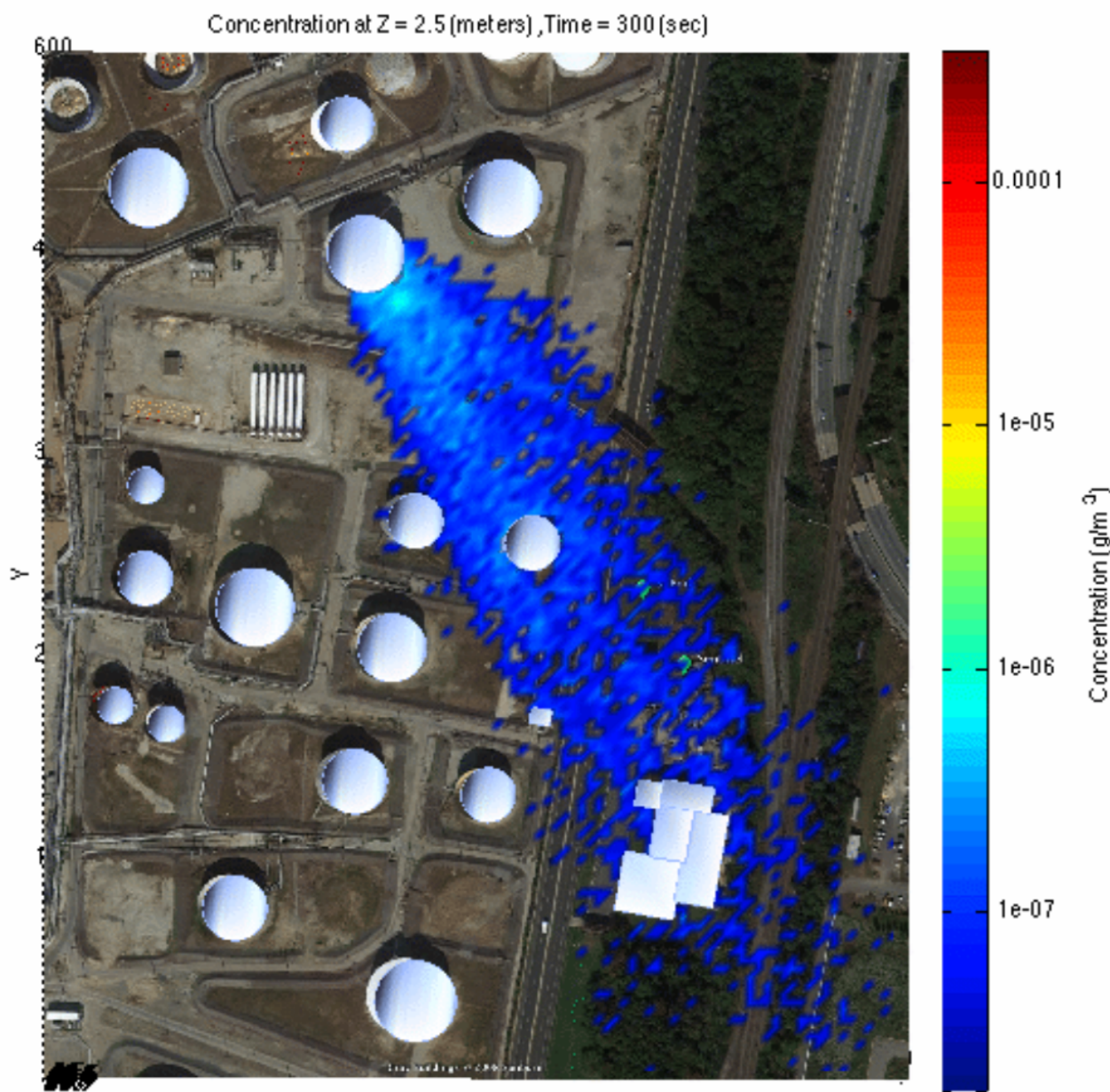


Figure 3: QUIC Plume Visualization

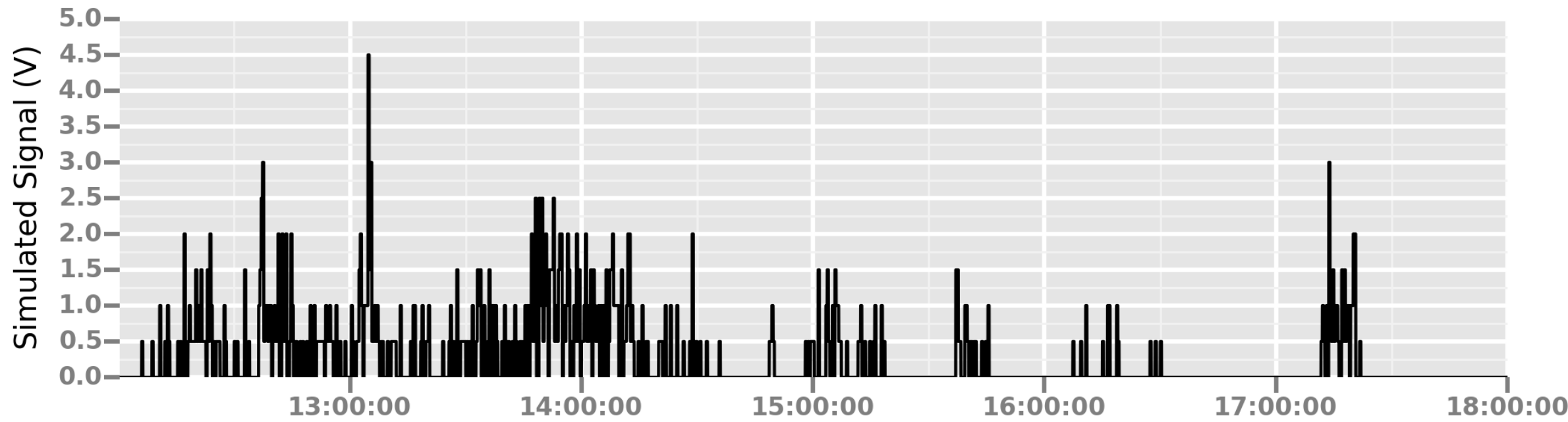


Figure 4: Base Sensor Simulated Concentration from QUIC

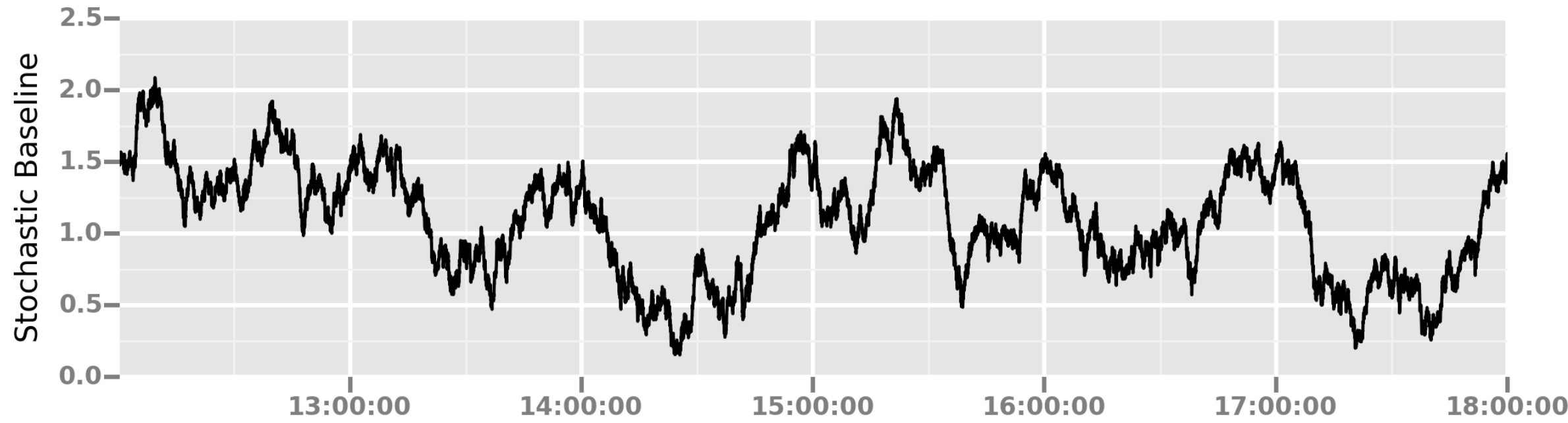


Figure 5: Base Sensor Simulated Stochastic Baseline

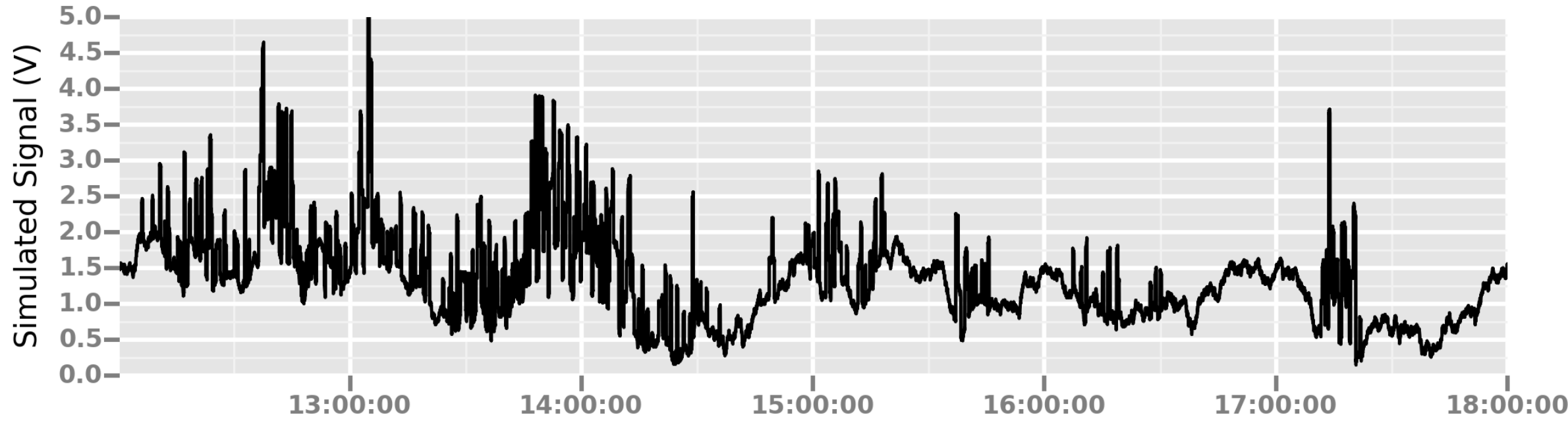


Figure 6: Base Sensor Simulated Data

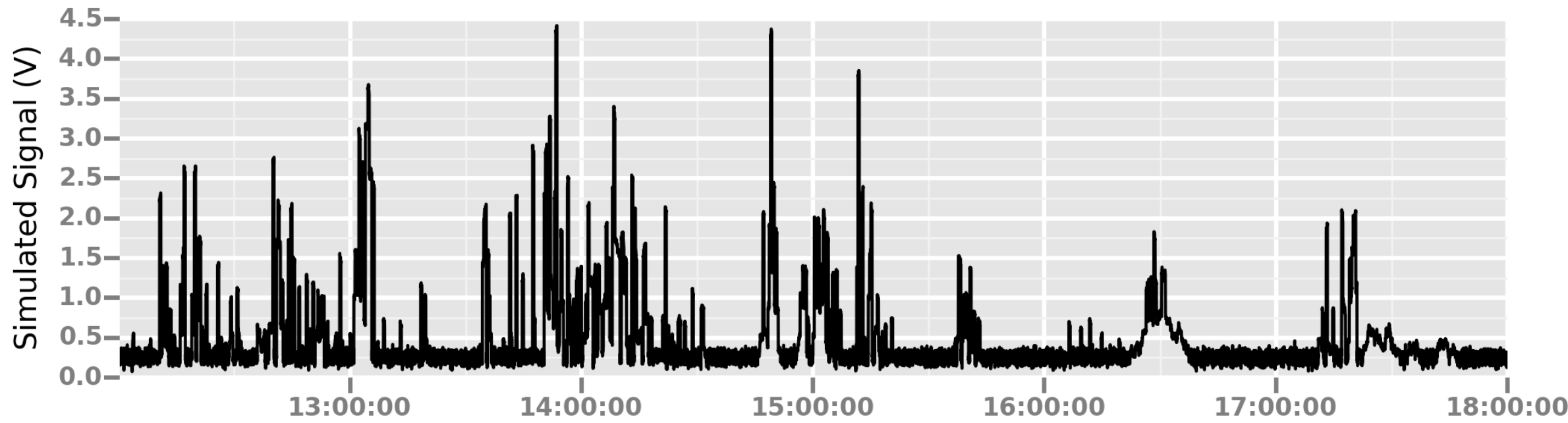


Figure 7: SPOD Sensor Simulated Data

## Algorithm Description

Two different methods of separating the concentration signal from the baseline were investigated and applied to 4 hour slices of data:

### Quantile Regression

Quantile regression was used to separate the baseline drift from the concentration signal by first smoothing the sensor measurements using a 40 sec moving average and then modeling the 1st quantile of the data as a smooth function of time using cubic splines with 5 degrees of freedom. After the baseline was subtracted the voltage was converted into a boolean variable with 1 sec values above 0.2 set to 1 and otherwise set to 0.

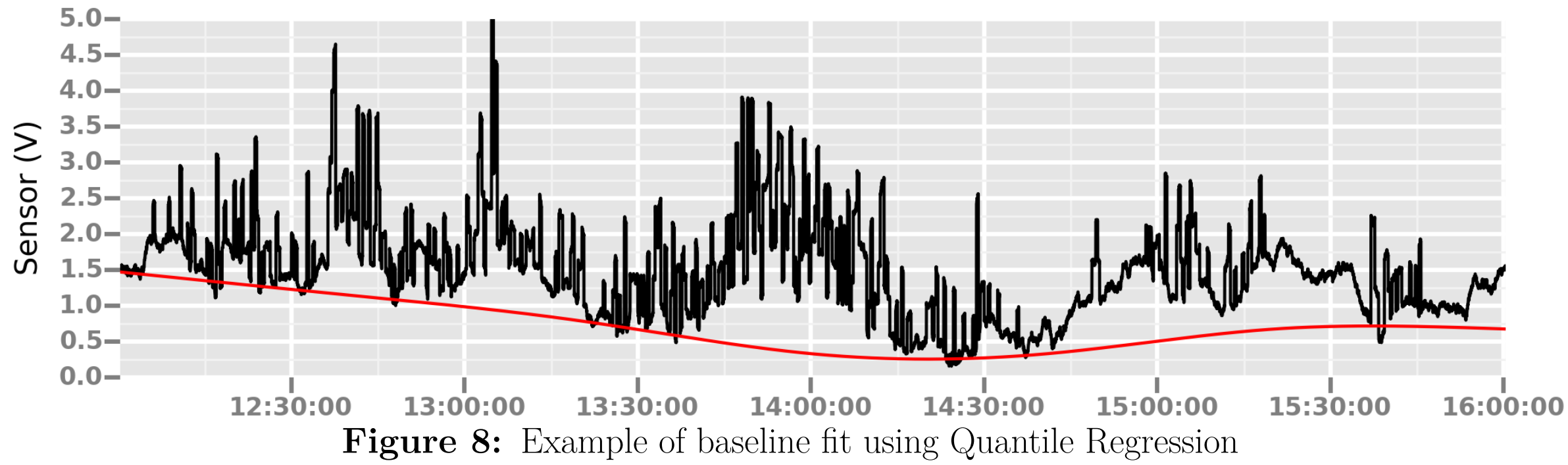


Figure 8: Example of baseline fit using Quantile Regression

### Butterworth Bandpass Filter

The measured data consists of a low frequency baseline drift component, a high-frequency noise component and the concentration signal which appears to occur at a frequency in between the drift and the noise. To separate out the concentration signal a Butterworth bandpass filter was applied using the scipy.signal toolbox with a frequency window of (0.01, 0.1) and order 2. After the filter was applied the voltage was converted into a boolean variable with 1 sec values above 0.08 set to 1.

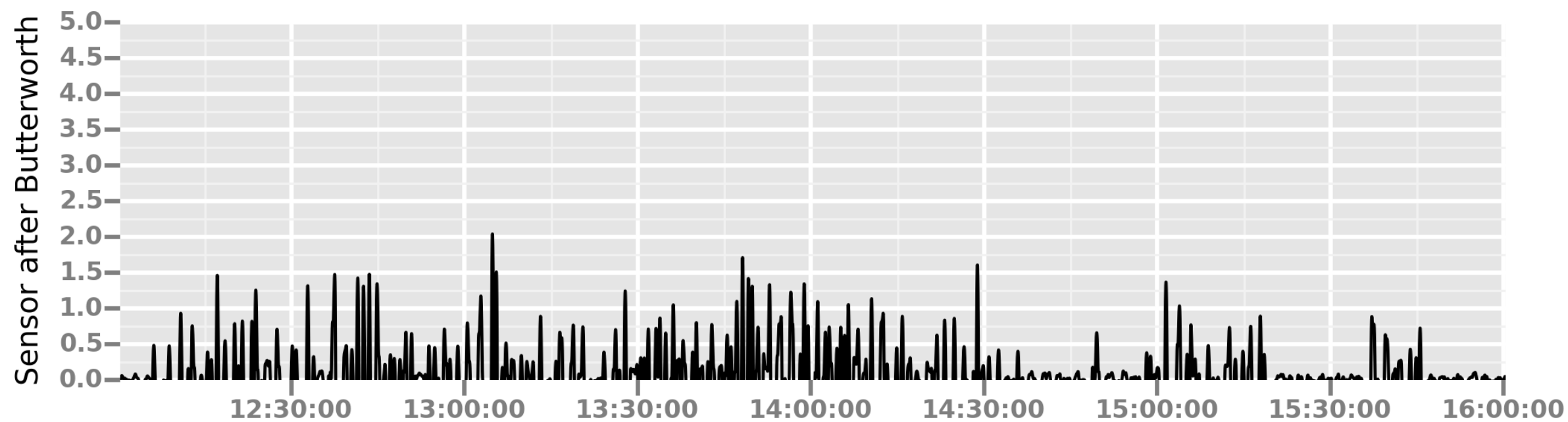


Figure 9: Example of sensor data after applying the bandpass filter

## Results

### Simulated Data

- True signal was defined as boolean variable with value 1 if the 1 sec simulated signal > 0.1 V, 0 otherwise.
- Each five minute interval was categorized into ”signal present” if the mean value of the true signal was > 0.017, i.e. more than 5 sec of true signal in the interval.
- Both the band pass filter method and the quantile regression method were able to correctly identify 80% of the 5 min periods with true signal.
- The band pass filter method had false positive rate of 4 % while the quantile regression method had a rate of 9 %.

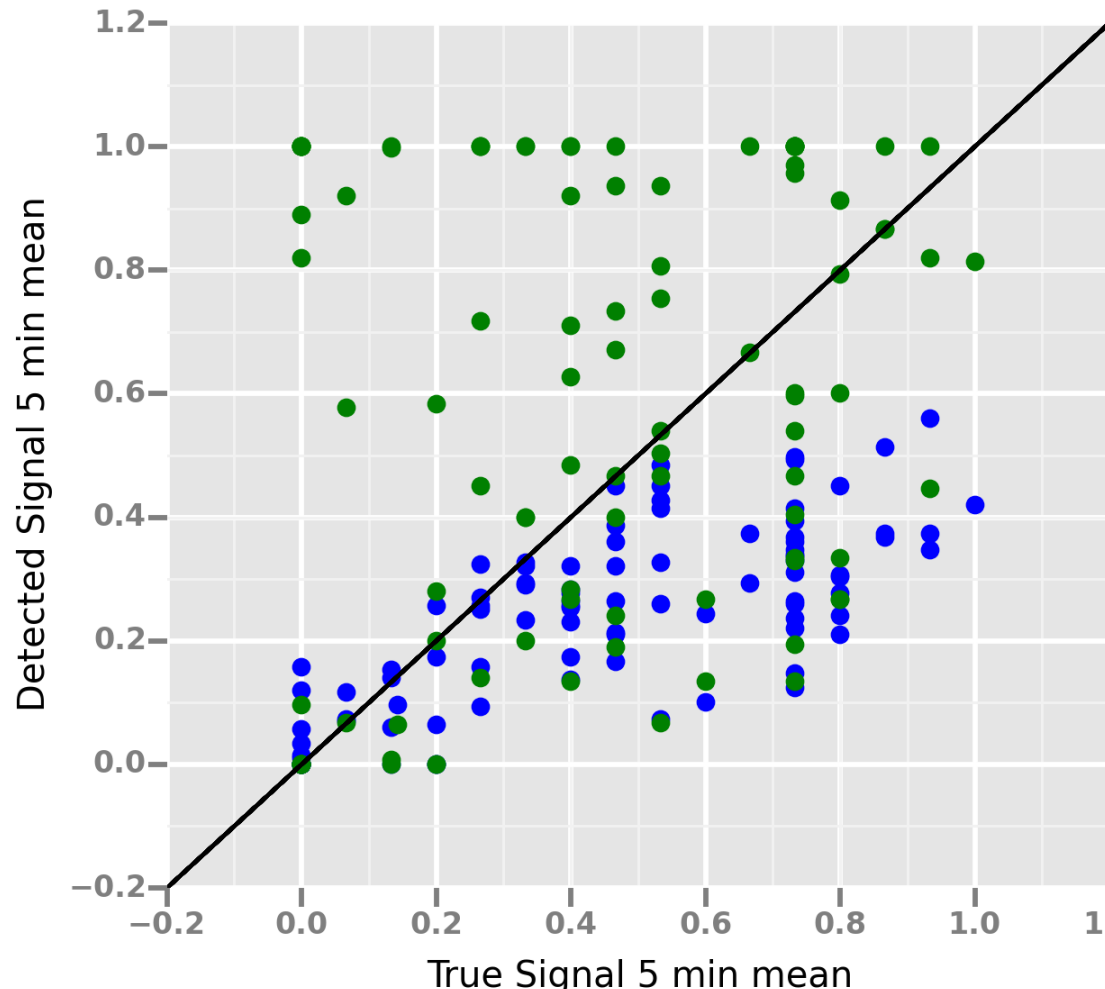


Figure 10: True versus detected signal, green represents quantile regression and blue represents the band-pass filter.

### Application

- The band pass filter algorithm was applied to the sensor data from 11/1/14 to 11/30/14
- The resulting 5 min filtered and averaged data set was divided into (1) time periods when signal was detected at the base sensor, (2) time periods when signal was detected at the SPOD sensor (3) time periods when no signal was detected.
- The windroses below show the distribution of 5 min wind vectors in each of the categories. The source of the emissions seems to be from a tank northeast of the sensors and is likely intermittent.

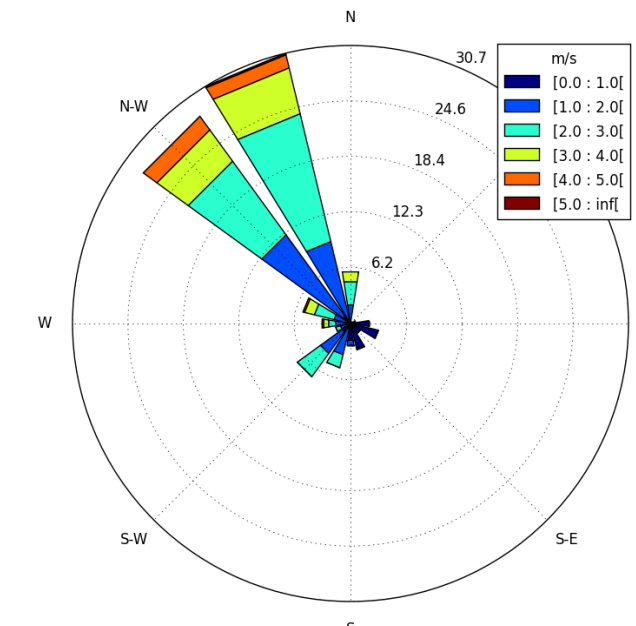


Figure 11: Signal detected at the Base sensor.

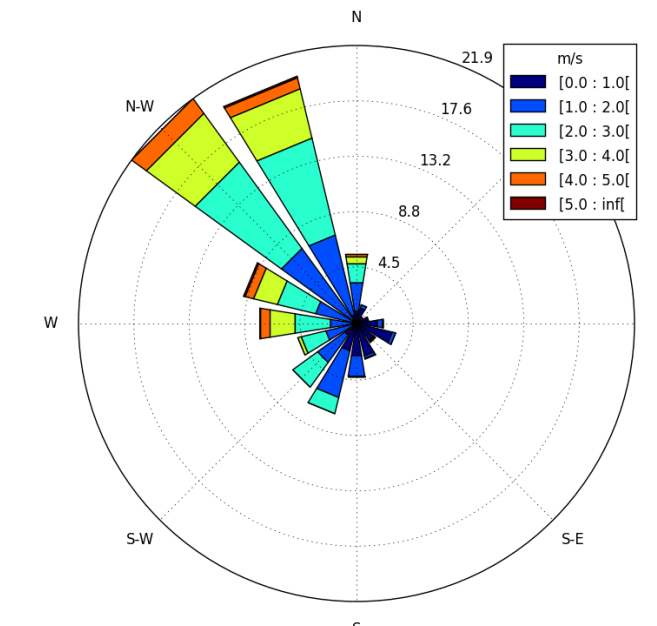


Figure 12: Signal detected at the Remote sensor.

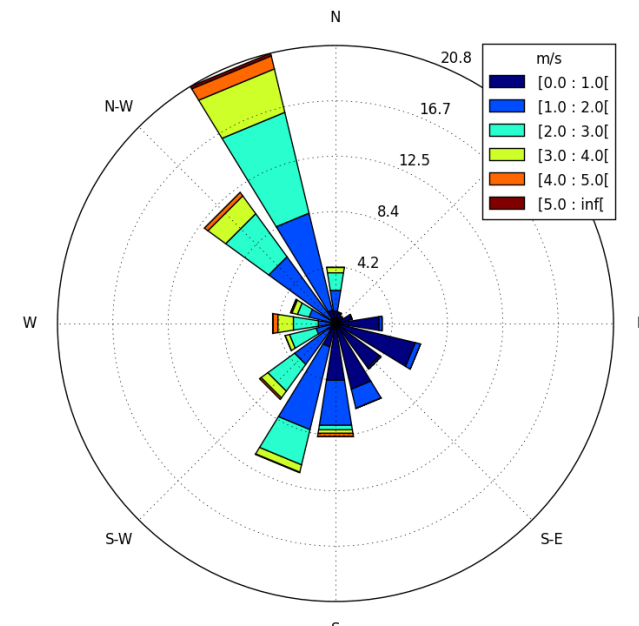


Figure 13: No signal detected.

[1] EPA. *Petroleum Refinery Sector Risk and Technology Review and New Source Performance Standards*. 2014. URL: <https://www.federalregister.gov/articles/2014/06/30/2014-12167/petroleum-refinery-sector-risk-and-technology-review-and-new-source-performance-standards>. [2] E. R. Pardyjak and M. J. Brown. *Quick Urban & Industrial Complex (QUIC) Dispersion Modeling System*.