Evaluation of an Event Detection System to Identify Spill Events Related to Marcellus Shale Development

Joseph C. Wickline

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation


This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Evaluation of an event detection system to identify spill events related to Marcellus Shale development

Joseph C. Wickline

Thesis submitted
to the Benjamin M. Statler College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements for the degree of
Master of Science in
Civil and Environmental Engineering

Leslie Hopkinson, Ph.D., Chair
Antarpreet Singh Jutla, Ph.D.
Seungho Hong, Ph.D.

Department of Civil and Environmental Engineering

Morgantown, West Virginia
2018

Keywords: water quality, continuous water quality monitoring, Marcellus Shale, unconventional oil and gas, surface spills, event detection, data processing, early warning systems

Copyright 2018 Joseph C. Wickline
ABSTRACT

Evaluation of an event detection system to identify spill events related to Marcellus Shale development

Joseph C. Wickline

There has been an extraordinary growth in the production of natural gas within the Marcellus Shale region in the past two decades. This sharp increase in production corresponds to an increased risk to surface waters from spill events related to the industry. At present, there are some private organizations and government agencies which monitor water quality in the region using continuous monitoring. However, there are no protocols in place to use data from continuous monitoring stations to quickly detect contamination resulting from spills.

Previous research has found that standard water quality parameters, such as specific conductivity, pH, and dissolved oxygen, measured by “off the shelf” water quality sensors are affected by commonly spilled materials from the natural gas industry. However, it was also found that due to expected low concentrations at monitoring stations, the effects of pollutants on these parameters are not strong enough to use simple threshold value methods to detect contaminants. Thus, there is a need for detection of subtle changes in water quality data which are not congruent with background behavior. The overall objective of this work was to add to the understanding of event detection related to surface spill events resulting from the development of Marcellus Shale.

Municipal water distribution networks commonly deploy event detection systems (EDS) to detect accidental or malicious contamination. The application of EDS in natural systems is less common. In this study, a publicly available EDS, “CANARY”, was applied to water quality data from natural channels to evaluate its ability to detect spill events and its overall performance when applied to a natural system.

Two approaches were used to test event detection capabilities: historical events and a simulated event. First, water quality data from three streams in the Susquehanna River Basin Commission (SRBC) remote water quality monitoring network (RWQMN) which had a spill event (i.e., production water, flowback fluid, and drilling mud) in the watershed were analyzed using CANARY. Then, a contamination event was imposed on water quality data from one uncontaminated stream and analyzed. For spills of flowback and production fluid, events were detected during the week following the spills. However, the algorithms detecting these events produced too many false alarms throughout the data set to confidently attribute these events to the spills. For the drilling mud spill, the only event detected was the result of a sensor malfunction. Therefore, the historical events were not detected. The simulated flowback fluid spill was also not detected. It was therefore concluded that the contaminant concentrations at the sensors were too low for detection. This further demonstrates the need for more extensive water quality monitoring networks in watersheds of concern. Decreasing the distance from spill events to sensors, and optimizing configuration parameters could provide a viable protocol for detecting acute contamination events resulting from natural gas development in the Marcellus Shale.
ACKNOWLEDGEMENTS

This research was supported by the NASA West Virginia Space Grant Consortium (WVSGC) Graduate Research Fellowship. The contents of this project are solely the responsibility of the author and do not necessarily represent the official views of others. I would like to thank the NASA WVSGC for selection of myself and this research project for the Graduate Research Fellowship.

I would like to thank the Susquehanna River Basin Commission (SRBC) for their efforts to provide publicly available continuous water quality monitoring data. Without these data, this research would have been impossible. I would also like to thank SRBC personnel Dawn Hintz and Graham Markowitz for their rapid and helpful responses to inquiries.

Additionally, I would like to thank those involved in the US EPA’s development of the CANARY event detection system. A publicly accessible event detection system was pertinent to this research. I thank Dr. Regan Murray and Jonathan Burkhardt for their personal assistance providing a script to make the many trials involved in this research possible to accomplish.

Next, I would like to thank Dr. Antarpreet Singh Jutla and Dr. Seungho Hong for serving on my advisory and examining committee, and for their helpfulness throughout my graduate education. I extend a special thank you to my advisor and committee chairperson Dr. Leslie Hopkinson. Her guidance throughout the composition of this thesis, and my entire graduate education has been invaluable.

Finally, I would like to thank my family for their continual support and encouragement throughout my academic career. I extend a special thank you to my wife for her enduring love and support throughout my time as graduate student.
Table of Contents

Table of Tables .................................................................................................................. v
Table of Figures .................................................................................................................. vi
List of Abbreviations .......................................................................................................... x
1.0 Introduction .................................................................................................................... 1
  1.1 Justification .................................................................................................................. 1
  1.2 Goals and Objectives ................................................................................................. 2
  1.3 Study Design ............................................................................................................... 2
2.0 Literature Review ......................................................................................................... 4
  2.1 Development of Marcellus Shale .............................................................................. 4
  2.2 Effects of spills on water resources ......................................................................... 5
  2.3 Current monitoring strategies .................................................................................. 5
  2.4 Prior work using threshold values to detect spills .................................................. 7
  2.5 Event Detection Systems ......................................................................................... 8
    2.5.1 CANARY Event Detection System .................................................................. 8
3.0 Methods ...................................................................................................................... 14
  3.1 Data Retrieval ............................................................................................................ 14
  3.2 Case Studies .............................................................................................................. 14
    3.2.1 Meshoppen Creek ......................................................................................... 15
    3.2.2 Hammond Creek .......................................................................................... 16
    3.2.3 East Branch of Wyalusing Creek .................................................................. 18
  3.3 Simulated Event ......................................................................................................... 19
    3.3.1 Site Selection .................................................................................................... 20
    3.3.2 Contamination Event Simulation .................................................................... 20
  3.4 Event Detection with CANARY ................................................................................ 25
4.0 Results and Discussion ............................................................................................... 28
  4.1 Case Studies .............................................................................................................. 28
    4.1.1 Meshoppen Creek ......................................................................................... 28
    4.1.2 Hammond Creek .......................................................................................... 38
    4.1.3 East Branch of Wyalusing Creek .................................................................. 46
  4.2 Simulated Flowback Fluid Event ............................................................................ 55
    4.2.1 Meshoppen Creek 2012: Clean Data Set ....................................................... 56
    4.2.2 Meshoppen Creek 2012: Simulated Data Set ................................................ 59
5.0 Discussion ..................................................................................................................... 64
  5.1 Event Detection in Natural Channels ..................................................................... 64
  5.2 Implications for future event detection ................................................................. 64
6.0 Conclusions ................................................................................................................... 68
7.0 References .................................................................................................................... 70
Appendix A: CANARY Configuration File Used with Script (Hammond Creek) ............ 74
Appendix B: Example Powershell Script Used to Complete Trials (Hammond Creek) ...... 77
Appendix C: CANARY Configuration File Written from Script (Hammond Creek, BED window = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90) ................................. 79
Table of Tables

Table 1. Parameters that are monitored by the SRBC RWQMN, their respective units of measurement, and precision of measurement................................................................. 14
Table 2. Watershed (SRBC 2017b) and event characteristics (Harris et al. 2016b) for the Meshoppen Creek case study.................................................... 16
Table 3. Watershed (SRBC 2018a) and event characteristics (Harris et al. 2016b) for the Hammond Creek case study.................................... 18
Table 4. Watershed (SRBC 2018b) and event characteristics (Harris et al. 2016b) for the East Branch of Wyalusing Creek case study.......................... 19
Table 5. Most probable travel time and concentration estimates for the 2011 Meshoppen creek event used for simulation in 2012........................................... 21
Table 6. The experimental design matrix for each value of BED window tested. There were 24 trials conducted for each BED window. ............................... 26
Table 7. Results for all Meshoppen Creek trials using the LPCF algorithm ......................... 35
Table 8. Results for all Meshoppen Creek trials using the MVNN algorithm .................... 36
Table 9. Results for all Hammond Creek trials using the LPCF algorithm......................... 44
Table 10. The results for all Hammond Creek trials using the MVNN algorithm ............... 45
Table 11. Results for all East Branch of Wyalusing Creek trials using the LPCF algorithm .... 53
Table 12. Results for all East Branch of Wyalusing Creek trials using the MVNN algorithm ... 54
Table 13. Results from all Meshoppen Creek 2012 trials before event simulation using the LPCF algorithm ................................................................. 56
Table 14. Results from all Meshoppen Creek 2012 trials before event simulation using the MVNN algorithm ................................................................. 57
Table 15. Results from all Meshoppen Creek 2012 trials after event simulation using the LPCF algorithm ................................................................. 59
Table 16. Results from all Meshoppen Creek 2012 trials after event simulation using the MVNN algorithm ................................................................. 60
Table 17. Net alarms for the LPCF algorithm with outlier threshold set to 2.5σ following the implemented changes to water quality data as a result of the simulated event. All alarm changes occurred during the time of event simulation. ........................................ 61
Table of Figures

Figure 1. A conceptualization of the LPCF algorithm. The standard deviation from the user specified history window is computed at each time step and used for user the specified outlier threshold. ................................................................. 11

Figure 2. The MVNN algorithm for a single history window. An example history window with 576 data points with normalized specific conductivity, pH, and dissolved oxygen is shown. The red circle represents the current measurement. The residual is the 3-dimensional distance from the current measurement to the next closest point. ...................................................................................... 12

Figure 3. Location of the Meshoppen Creek watershed, spill event, and SRBC monitoring station .......................................................................................................................... 16

Figure 4. Location of the Seeley Creek Watershed, and the contamination event and SRBC continuous monitoring station within the Hammond Creek sub-watershed. .......................... 17

Figure 5. Location of the Wyalusing Creek Watershed, and the contamination event and SRBC continuous monitoring station within the East Branch sub-watershed .................................................. 19

Figure 6. Approximation of concentrations of flowback fluid at the monitoring station using a scalene triangle. The first point \(t = 0\) is the instant the leading edge of the plume reaches the monitoring station. The middle point \(t = 4.2\) hr is point at which the peak concentration arrives. The final point \(t = 12.6\) hr is the point at which the concentration is reduced to 10% of the peak. .......................................................................................................................... 22

Figure 7. Response of specific conductivity to increasing concentrations of production fluid in lab experiments conducted by Harris et al. (2016a, 2016b). .......................................................... 23

Figure 8. Response of pH to increasing concentrations of flowback fluid for concentrations of 0–5 ppm. Regression developed for the average response of all lab experiments conducted by Harris et al. (2016a, 2016b). ...................................................................................... 24

Figure 9. The simulated water quality values for the duration of the simulated event from \(t = 0\) corresponds to the arrival of the leading edge of the plume on July 2, 2012 at time step 21:20:00. ..................................................................................................................... 25

Figure 10. Weekly number of alarms produced by each trial of the MVNN algorithm with outlier threshold set to 1.0σ. Alarms resulting from the reported event were expected in week 2. BED corresponds to the length of the BED window in time steps, and P(event) represents the probability threshold for each trial. .................................................................................. 29

Figure 11. CANARY output plot for Meshoppen Creek during the first quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm. ...................................................................................... 30

Figure 12. CANARY output plot for Meshoppen Creek during the second quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm. ...................................................................................... 31

Figure 13. CANARY output plot for Meshoppen Creek during the third quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue
“plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 14. CANARY output plot for Meshoppen Creek during the fourth quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 15. Decreasing trend of number of alarms versus the outlier threshold for Meshoppen Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 16. Trend in number of alarms versus the outlier threshold for Meshoppen Creek trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity. Unlike the LPCF algorithm, the number of alarms increases from 0.5σ to 1.0σ.

Figure 17. Weekly number of alarms produced by each trial of the MVNN algorithm with outlier threshold set to 0.5σ for Hammond Creek. Alarms resulting from the reported event were expected in week 3. BED corresponds to the length of the BED window in time steps, and P(event) represents the probability threshold for each trial.

Figure 18. CANARY output plot for Hammond Creek during the first quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 19. CANARY output plot for Hammond Creek during the second quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 20. CANARY output plot for Hammond Creek during the third quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 21. CANARY output plot for Hammond Creek during the fourth quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 22. Decreasing trend in number of alarms versus the outlier threshold for the Hammond Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 23. Decreasing trend in number of alarms versus the outlier threshold for the Hammond Creek Trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 24. CANARY output plot for East Branch of Wyalusing Creek during the week of October 29, 2013 to November 5, 2013: BED = 1 hr, outlier threshold = 3.0σ, P(event) = 0.99. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. Markers for both the LPCF and MVNN algorithms indicate an event was detected 9 days after spill was reported, even on the least sensitive algorithms.

Figure 25. CANARY output plot for East Branch of Wyalusing Creek during the first quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 26. CANARY output plot for East Branch of Wyalusing Creek during the second quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 27. CANARY output plot for East Branch of Wyalusing Creek during the third quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 28. CANARY output plot for East Branch of Wyalusing Creek during the fourth quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Figure 29. Decreasing trend in number of alarms versus the outlier threshold for the East Branch of Wy'allusing Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 30. Decreasing trend in number of alarms versus the outlier threshold for the East Branch of Wy'allusing Creek trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.
Figure 31. Decreasing trend in number of alarms versus the outlier threshold for the clean 2012 Meshoppen Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 32. Number of alarms versus the outlier threshold for the clean 2012 Meshoppen Creek trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity. Unlike the LPCF algorithm, the number of alarms increases from 0.5σ to 1.0σ.

Figure 33. CANARY output plot for the clean Meshoppen Creek clean data set during the week of July 1, 2012 to July 8, 2012: BED = 1 hr, outlier threshold = 2.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm. The red box indicates the time frame of event simulation.

Figure 34. CANARY output plot for the simulated event on Meshoppen Creek during the week of July 1, 2012 to July 8, 2012: BED = 1 hr, outlier threshold = 2.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm. The red box indicates the time frame of event simulation.
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWQM</td>
<td>ambient water quality monitoring</td>
</tr>
<tr>
<td>BED</td>
<td>binomial event discriminator</td>
</tr>
<tr>
<td>CIM</td>
<td>continuous instream monitoring</td>
</tr>
<tr>
<td>CSV</td>
<td>comma separated value</td>
</tr>
<tr>
<td>EDS</td>
<td>event detection system</td>
</tr>
<tr>
<td>EWS</td>
<td>early warning system</td>
</tr>
<tr>
<td>HVHF</td>
<td>high volume hydraulic fracturing</td>
</tr>
<tr>
<td>LPCF</td>
<td>linear prediction correction filter</td>
</tr>
<tr>
<td>MVNN</td>
<td>multivariate nearest neighbor</td>
</tr>
<tr>
<td>NY DEC</td>
<td>New York Department of Environmental Conservation</td>
</tr>
<tr>
<td>OH EPA</td>
<td>Ohio Environmental Protection Agency</td>
</tr>
<tr>
<td>PA DEP</td>
<td>Pennsylvania Department of Environmental Protection</td>
</tr>
<tr>
<td>RIBS</td>
<td>Rotating Integrated Basin Studies</td>
</tr>
<tr>
<td>RWQMN</td>
<td>remote water quality monitoring network</td>
</tr>
<tr>
<td>SRBC</td>
<td>Susquehanna River Basin Commission</td>
</tr>
<tr>
<td>TMDL</td>
<td>total maximum daily load</td>
</tr>
<tr>
<td>US EPA</td>
<td>United States Environmental Protection Agency</td>
</tr>
<tr>
<td>UOG</td>
<td>unconventional oil and gas</td>
</tr>
<tr>
<td>WQN</td>
<td>water quality network</td>
</tr>
<tr>
<td>WV DEP</td>
<td>West Virginia Department of Environmental Protection</td>
</tr>
</tbody>
</table>
1.0 Introduction

1.1 Justification

The natural gas industry has rapidly developed in the past two decades with the advent of horizontal drilling and high volume hydraulic fracturing techniques (Rodriguez and Soeder 2015). The Appalachian region overlaying the Marcellus Shale formation has been especially subject to this growth due to quantity of obtainable natural gas (Soeder and Kapell 2009). Hazards to water resources grow in proportion to this industry due to the nature of natural gas extraction techniques (Brantley et al. 2014). According the US EPA (2016), hazards to water resources from the natural gas industry can be lumped into four major categories: 1) water acquisition, 2) chemical mixing, 3) flow back and produced water, and 4) wastewater management and disposal. Flow back and produced waters can impose major threats to water resources when released into the natural environment during accidental spill events. Potential effects of these events are degradation of aquatic habitat, and contamination of ground and surface water drinking resources (Papoulias and Velasco 2013; US EPA 2016). Therefore, awareness and appropriate response to these events should be as timely as possible.

Previous work by Harris et al. (2016) has shown that typical water quality parameters, as measured by “off the shelf” sensors, respond when commonly spilled chemicals are introduced into the sample area. Therefore, it may be possible to detect accidental contamination events utilizing a high frequency continuous water quality monitoring network. However, due to dispersion across the distance traveled by spilled materials, using threshold values to detect contamination are likely ineffective. Harris et al. (2016) instead recommended analyzing data for shifts not concurrent with recent background data. Thus the goal of this research was to develop and evaluate methods of contamination detection aligned with these recommendations.

CANARY event detection software is publicly available software developed by the US EPA in conjunction with Sandia Corporation Laboratories (US EPA 2010). CANARY was developed for implementation into municipal distribution networks to provide drinking water security. While CANARY can be used to detect events based on exceedance thresholds, it also utilizes algorithms that detect events within the normal range of fluctuations for the data set. It does so by comparing the most recently measured event to a specified period of background data.
Based on the recommendations from Harris et al. (2016), these methods could also be useful for detection of contamination events in a natural system.

Several event detection methods have been developed and implemented in event detection systems (Storey et al. 2011; Hou et al. 2013; Che and Liu 2014). Most of these systems have been developed by private entities for municipal water distribution systems to monitor these systems for accidental or malicious contamination. In most of these instances, the event detection software is not available to the public. Fewer cases exist where event detection systems have been implemented in the natural environment. The cases which do, do not implement this tool for the purposes described previously. Therefore, it could be of great interest to implement such a system on a data set from a natural system which has been exposed to contamination.

1.2 Goals and Objectives

Continuous water quality monitoring and spill reports are used to determine the impact of the development of the Marcellus Shale (Betanzo et al. 2016); however, there is little guidance to detect acute spill events (Harris et al., 2016). The overall objective of this work was to add to the understanding of event detection related to surface spill events resulting from the development of Marcellus Shale. I hypothesize that standardized methods utilized in municipal water applications may detect spills related to natural gas development (i.e., produced water, flowback fluid, or drilling mud). The goals of this research were to:

1. Determine if standardized methods utilized in municipal water applications could identify known spill events of produced water, flowback fluid, or drilling mud.
2. Determine if standardized methods utilized in municipal water application could identify a simulated spill event.

1.3 Study Design

First, one year of water quality data was retrieved from the Susquehanna River Basin Commission (SRBC) Remote Water Quality Monitoring Network (RWQMN) for three sites with the following criteria: i) experienced a natural gas related spill event within the watershed large enough to create concentrations at the monitoring station greater than 1 ppm, and ii) existing monitoring data measured at a 5-minute interval. Next, the data sets were analyzed with CANARY, varying the input parameters each time. Finally, the resulting outputs were analyzed
to determine if the documented spill event was detected for any of the parameter configurations, and how the detection software performance changes (number of false alarms triggered) with variations to the parameters.

Because of there is some uncertainty regarding the details of the spill events and whether or not a measurable response should be expected, another analysis was completed. This time, instead of using a dataset with a known event in the watershed, a dataset with no known events was retrieved. Again, an entire year of data was downloaded whose measurements are taken at a 5-minute interval. For this dataset, a contamination event resembling that of what has been previously described was simulated on an arbitrarily chosen date. Once the simulated event was added, the dataset was analyzed with CANARY utilizing the same parameter variations used for the first three datasets. These outputs were analyzed to determine if the simulated event was detected, and further contribute to the analysis of CANARY’s behavior on data from a natural system.
2.0 Literature Review

2.1 Development of Marcellus Shale

The Marcellus Shale is a geologic formation rich in natural gas and underlies much of New York, Pennsylvania, West Virginia, and eastern Ohio of the United States (Soeder and Kapell 2009). Until recently, extraction of the natural gas from the shale was not economically feasible. Advances in horizontal drilling and high volume hydraulic fracturing (HVHF) in the early 2000s made natural gas extraction profitable (Rodriguez and Soeder 2015). In 2015, more than 4.6 trillion cubic feet of natural gas was produced in Pennsylvania, resulting in a 5,769% increase in natural gas production since 2009 (PA DEP 2016).

While the rapid growth of the natural gas industry in this region is economically beneficial, the extent of its environmental impact is unknown, and the impact to water resources remains a concern (Brantley et al. 2014). The effects of HVHF on groundwater resources are still being determined, but the quantity of water used and wastewater and byproducts produced in HVHF are much more certain. Each well uses approximately 2-7 million gallons of source water (Entrekin et al. 2011) and, in the Susquehanna River Basin of the Marcellus Shale region, produces up to 1.3 million gallons of wastewater (US EPA 2016). This wastewater is typically reused in the hydraulic fracturing process, disposed of in wastewater injection wells, or discharged into adjacent rivers or streams either untreated or treated by a wastewater treatment plant (US EPA 2016). Although most discharged wastewater is treated, some studies have shown that it still poses a threat to water quality (Rozell and Reaven 2012). Wastewater which is disposed of via injection wells has also been found to impact surface water quality. Kassotis et al. (2015) found a significant increase in the concentration of endocrine disruptors downstream of a West Virginia wastewater injection well which accepts by wastes from HVHF operations. Other contamination of surface waters occurs when byproducts of HVHF are accidentally spilled or intentionally dumped and flow into surface waters. According to Brantley et al. (2014), wastes from HVHF wells that have potential to contaminate surface waters via spilling in quantities greater than 400 gallons are: flowback water, produced water, fracture fluids, and drilling fluids. These spills typically result from trucking accidents, well blowouts, or leaking tanks, valves, or pipes.
2.2 Effects of spills on water resources

Spills into surface waters can be detrimental to local ecosystems and create problems for drinking water supply. Papoulias and Velasco (2013) reported that the release of HVHF chemicals into Acorn Fork, KY reduced the pH from 7.5 to 5.6, and increased the specific conductivity from 200 µS/cm to 35,000 µS/cm. These changes in water quality were harmful, and in some cases deadly to the fish and invertebrates of the stream, including the threatened black nosed dace. Kentucky water quality standards mandate that pH remain in the range of 6.0 to 9.0 and not change by more than 1.0 in any 24-hour time period. In addition, these standards mandate that specific conductance not increase or decrease to levels that will adversely affect indigenous species (State of Kentucky 2017).

Cozzarelli et al. (2017) documented the effects of an 11.4 ML wastewater spill from an unconventional oil and gas (UOG) pipeline in Blacktail Creek, North Dakota which occurred in January 2015. Chemical contaminants measured in the soil and water were found in higher concentrations downstream of the spill site compared to samples taken upstream of the spill. February water samples showed elevated levels of chloride and bromide downstream of the spill site. Lithium, boron, and strontium were present downstream at levels 5 to 10 times higher than those upstream of the spill. A persistent thermogenic source of methane was detected downstream of the spill by the presence of light hydrocarbons. Persistent sediment contamination was also detected; barium and strontium concentration were both higher downstream of the spill. Radium activity was found to be 15 times higher downstream of the spill, and due to Sr isotope ratios, are believed to be the result of the leaking pipeline. Biological health effects demonstrate the effects of the spill, even months after attempts at remediation. Fish bioassays indicate 89% survival upstream of the spill, compared to 2.5% survival downstream of the spill. Finally, biological impacts resulting from endocrine disruption were detected downstream of the spill. The findings of Cozzarelli et al. (2017) demonstrate that surface spills have the potential to create long term issues for ecosystem health and water quality.

2.3 Current monitoring strategies

Current monitoring strategies for surface spills related to Marcellus Shale development vary depending on the watershed and state monitoring agency. The Pennsylvania Department of Environmental Protection (PA DEP) currently relies on reports from well operators or
anonymous citizens (PA DEP 2017a). The PA DEP also has continuous instream monitoring (CIM) stations which are temporarily deployed into streams of concern. These stations are typically deployed for one year but may be deployed for a shorter period in special cases (PA DEP 2017b). Surface waters are otherwise monitored at fixed stations as part of the Pennsylvania Water Quality Network (WQN). The WQN includes three types of monitoring stations: Standard, Chesapeake Bay Nutrient Loading, and Reference Streams. These stations are sampled monthly or bimonthly, and in 2010 116 stations had chemistry constituents specific to oil and gas signatures added to the list of monitored parameters (PA DEP 2017c).

Similar to the PA DEP, the West Virginia Department of Environmental Protection (WV DEP) has an Ambient Water Quality Monitoring (AWQM) Network which includes 26 fixed (long-term) stations at the mouths of West Virginia’s larger streams and rivers and are sampled bi-monthly (WV DEP 2017a). The WV DEP also has deployable continuous monitoring stations which are used in particular projects where understanding temporal variation is crucial (WV DEP 2017a). As of the time of this writing, the WV DEP has not added any monitored chemical parameters which may be specific signatures of contamination resulting from oil and gas activity.

The New York Department of Environmental Conservation (NY DEC) implements Rotating Integrated Basin Studies (RIBS) to monitor water quality through the state of New York. Under RIBS, each major basin within the state is intensively monitored every five years. These intensive study results are then compared to baseline conditions established by The Routine Trend Monitoring Network, which monitors the basic water quality characteristics within all basins, every year, regardless of where intensive sampling is being conducted (NY DEC 2017a). At the time of this writing, the only continuous monitoring network currently implemented in the state of New York is employed by the Susquehanna River Basin Commission (SRBC) who monitors the sub watersheds of the Susquehanna River Basin with their Remote Water Quality Monitoring Network (RWQMN). Due to the geographic nature of the Susquehanna River Basin, this network extends from New York into much of Pennsylvania. The RWQMN was established to monitor the impact of oil and gas production activities on the basin’s surface waters. (SRBC 2017).

Currently, information regarding water quality monitoring strategies in Ohio is not extensive. According to the Ohio Environmental Protection Agency (OH EPA), five to seven
areas of the state are monitored each year. A total of 400 to 450 sampling sites are examined, of which each site is visited more than once (OH EPA 2017a). The primary goal of this monitoring program is to establish and regulate the Total Maximum Daily Loads (TMDLs) as required by the US EPA. At present, there are no additional monitoring programs implemented by the OH EPA which seek to monitor the impacts of oil and gas development in the state.

Only Pennsylvania has a government monitoring agency which monitors for chemical inputs related to UOG development in the Marcellus Shale play (PA DEP 2017c). However, the PA DEP samples streams bimonthly at the most, a frequency which is likely not great enough to present acute events, such as a spill (Kirchner et al. 2004). The SRBC is the only entity which deploys continuous monitoring year-round throughout a watershed within the region with the intention of monitoring the effects UOG development has on water quality (SRBC 2017). However, the only protocol to date for all entities to detect spill events is by means of self-reporting by UOG development companies, or anonymous reporting by local citizens (PA DEP 2017a, WV DEP 2017b, NY DEC 2017b, OH EPA 2017b). Thus there is a need to have the ability to detect spill events using continuous monitoring data, which in turn creates a greater need for remote continuous water quality monitoring networks.

2.4 Prior work using threshold values to detect spills

Harris et al. (2016a, 2016b) found that detecting surface spills using threshold values of common surface water parameters is largely ineffective. The high variability of water quality requires knowledge and understanding of the background levels of water quality in order to detect contamination with statistical significance (Harris et al. 2016a, 2016b). Harris et al. (2016a, 2016b) found that commonly measured water quality parameters showed promise in detecting spill events: i) specific conductance for produced water, drilling mud, and ethylene glycol, ii) dissolved oxygen for drilling mud and ethylene glycol, and iii) pH for produced water and drilling mud at lower concentrations. The authors recommended these parameters be monitored in conjunction with continuous monitoring as relative change in the aforementioned parameters may be more crucial in detecting spills than the exceedance of threshold values (Harris et al. 2016a, 2016b).
2.5 Event Detection Systems

Event detection systems (EDS) or Early Warning Systems (EWS) have gained significant interest and attention in water quality monitoring in recent years (Storey et al. 2011). In general, an EDS is a computer program which takes incoming data from a monitoring station or a network of monitoring stations and analyzes the data with mathematic and statistical methods to determine if the incoming measurement is somehow different from previous measurements (US EPA 2010, Hou et al. 2013,). The majority of work in the field of water quality EDS is done in municipal distribution networks with the goal of providing water security (Ostfeld and Salomon 2005, US EPA 2010). Few instances currently exist in the literature where an EDS has been implemented in a natural system, such as a lake, stream, or estuary. Glasglow et al. (2004) have shown an EDS can be useful in monitoring dissolved oxygen in estuaries for algal blooms.

2.5.1 CANARY Event Detection System

Currently, most EDS are proprietary and not publicly available (US EPA 2010). However, the US EPA and Sandia Corporation Laboratories recently developed an EDS which is freely available. This EDS is called CANARY and, similar to most other EDS, was developed with the purpose of providing water security in light of the events that occurred on September 11, 2001 (US EPA 2010). However, the event detection methodology utilized by CANARY have potential to be useful for detecting acute pollution events in the natural environment. A major goal in the creation of CANARY was to develop an EDS which could detect contamination events in the subtle fluctuations of water quality time series, where threshold tests are less meaningful (US EPA 2010). CANARY can be used to monitor water quality in real-time, and to process historical data. Processing historical data in CANARY is commonly referred to as “batch processing”; this is the type of analysis which will be discussed for the remainder of this document.

Data processing in CANARY can be broken down into three major phases: read incoming data, utilize algorithms to decide if the most recent measurement is consistent with recent data or is an outlier, and determine whether or not an event has occurred with a binomial event discriminator (BED). For each CANARY analysis, a “YAML Ain’t Markup Language”, or YML, configuration file must be specified. This configuration file specifies real-time or batch processing, the dataset to be analyzed and its corresponding measurement interval, and if each
signal should be analyzed as a regular water quality parameter (used in event detection), an operational parameter (not used in event detection), or a calibration parameter (referenced in event detection to reduce false alarm rate). The configuration file also specifies which algorithms should be used to compute the residual used for event detection, and algorithm parameters which control the sensitivity of event detection. Additionally, the configuration file enables the user to request specialized output files, and, in the case of a water quality monitoring network, which stations should be analyzed. A more extensive overview of the CANARY configuration file is provided by the US EPA (2012).

During the first stage of analysis, data from a specified comma separated value (CSV) file is read by the program at the user specified measurement interval. This interval must be specified so the program can appropriately account for gaps in the dataset resulting from faulty or disabled sensors. While any measurement interval can be entered and analyzed, it is recommended to use data measured from 2 to 15-minute intervals (US EPA 2012).

For the second phase of analysis, CANARY has three built-in algorithms for analysis: set-point, linear prediction-correction filter (LPCF), and multivariate nearest neighbor (MVNN). Additionally, CANARY is capable of utilizing custom algorithms written in JAVA. More information on custom algorithms, including an example is provided by US EPA (2012).

The set-point algorithm allows the user to specify a distance from a set point that incoming measurements should not exceed. Essentially, the set-point algorithm allows the user to define the maximum and minimum threshold values to declare an event by establishing the maximum distance (positive or negative) that measurements may vary from the given set-point. Harris et al. (2016) suggested that threshold values are likely inadequate for event detection in the context of this study. Therefore, the remainder of this discussion will focus on the LPCF and MVNN algorithms provided by CANARY.

Both the LPCF and MVNN algorithms require user specified parameters which define the sensitivity of event detection and timing of event reporting. The required parameters are: history window, outlier threshold, event threshold, event timeout, and event window save. The event timeout parameter specifies how many time steps that may pass before an alarm is silenced, and the event window save parameter specifies the amount of time the program should allow to pass
before reporting an event that has been detected. These two parameters do not affect event detection; rather they allow an operator to adjust how events are reported. More important in the context of this study are those parameters whose values alter the sensitivity of event detection. The history window specifies the number of prior time steps the algorithm uses as historical data. The history window is dynamic, and is used to compute statistics used in event detection, such as standard deviation. For both the LPCF and MVNN algorithms, data in the history window is normalized using a standard Z-transformation function. The outlier threshold specifies the number of standard deviations that the measured value must differ from the expected value in order for the current time step to be declared an outlier. The event threshold specifies the probability that must be exceeded by the BED before a series of outliers are called an event.

In the configuration file, parameters for the BED are specified with each algorithm used. The two required parameters are the BED window, and the outlier probability. The BED window is the number of previous time steps that are included in computing the probability of an event at each time step. The BED window must be less than the history window. Additionally, considerations for setting the BED window should be: the expected duration of the contamination event to be detected, the desired sensitivity of detection, and the desired alarm response time (US EPA 2010, 2014).

Before the BED is used to determine if an event has occurred, a residual value must be produced by the predictive or comparative algorithm. The LPCF algorithm is essentially an autoregressive (AR) model which uses historical data to predict the measurement at the present time step (Figure 1). Then the absolute value of the difference between the predicted and measured value is calculated; this is the residual. Finally, the LPCF algorithm compares the residual against the user specified outlier threshold. If the residual exceeds the outlier threshold, then that particular measurement is declared an outlier. Because the LPCF algorithm computes a residual for each incoming water quality signal, the signal responsible for declaring a time step an outlier can be easily determined (US EPA 2010).
Unlike the LPCF algorithm, the MVNN algorithm does not predict a future value. Instead, the MVNN algorithm combines all analyzed data points into a multi-dimensional space where the measurement for each analyzed parameter is a component of the coordinates (Figure 2). Then, the Euclidean distance from the most recently received measurement and the nearest value from historical data is computed as the residual. Then, similar to the LPCF algorithm, the residual is compared to the user specified outlier threshold; the present time step is declared an outlier if the residual exceeds the outlier threshold. The nature of the MVNN algorithm does not allow for explicit determination of which signal(s) contributed to declaring a time step an outlier. However, CANARY implements a maximum residual classification within the MVNN algorithm which determines the maximum residual across all sensors, and therefore which signals were responsible for the outlier (US EPA 2010).
**Figure 2.** The MVNN algorithm for a single history window. An example history window with 576 data points with normalized specific conductivity, pH, and dissolved oxygen is shown. The red circle represents the current measurement. The residual is the 3-dimensional distance from the current measurement to the next closest point.

The third and final phase of analysis is done for each algorithm. Here, a BED is used to determine if an event has occurred. The probability of an event is computed at each time step using the cumulative binomial distribution function. A binomial experiment requires a fixed number of independent trials whose outcome must be either success or failure. In the context of the CANARY EDS, an outlier is considered a failure. Thus, the probability that the current measurement is the same as the history data is given by the following (US EPA 2010):

\[
P(\text{background}) = b(r; n, p) = \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \tag{1}\]

Here, \( r \) is defined as the number of failures that occurred in the given number of \( n \) trials, and \( p \) is defined as the probability of success on any given trial. For the application in CANARY, the BED window serves as the fixed number of trials, the number of failures is the number time steps within the BED window which are declared outliers, and the probability for success on any given trial is assumed to be 0.5. The complement to this equation is then the probability that an event has occurred (US EPA 2010):
\[ P(\text{event}) = 1 - P(\text{background}) = 1 - b(r; n, p) \]  

Finally, the cumulative binomial distribution function is used to more efficiently detect events by ensuring that the probability of an event increases as the number of failures increases. (US EPA 2010):

\[ P(r \leq z_c) = \sum_{l=1}^{n(r \leq z_c)} b(r; n, p) \] 

Here, \( z_c \) is the probability threshold value defined by the user, and all other variables are as previously defined. Then, if the cumulative binomial probability calculated is greater than the specified probability threshold, CANARY declares an event has occurred at that time step.
3.0 Methods

3.1 Data Retrieval

This study used data from the SRBC RWQMN (SRBC 2017a). These data are publicly available on the SRBC RWQMN website. The SRBC monitors the following water quality parameters: temperature, specific conductivity, pH, turbidity, and dissolved oxygen (Table 1). All data used in this study was measured at a 5-minute interval. Multiple sites with differing years were assessed. For each site, an entire calendar year of data was downloaded and analyzed. Therefore, each data set began on January 1 of the selected year at time step 00:00:00 and ends on December 31 of the same year at time step 23:55:00.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>°C</td>
<td>Hundredths</td>
</tr>
<tr>
<td>Specific Conductivity</td>
<td>mS/cm</td>
<td>Thousandths</td>
</tr>
<tr>
<td>pH</td>
<td>Standard Units</td>
<td>Hundredths</td>
</tr>
<tr>
<td>Turbidity</td>
<td>NTU+</td>
<td>Hundredths</td>
</tr>
<tr>
<td>Dissolved Oxygen</td>
<td>mg/L</td>
<td>Hundredths</td>
</tr>
</tbody>
</table>

Table 1. Parameters that are monitored by the SRBC RWQMN, their respective units of measurement, and precision of measurement.

3.2 Case Studies

The goal of this study was to evaluate the effectiveness of the CANARY EDS in detecting acute contamination events resulting from natural gas development in the Marcellus Shale. To accomplish this goal, a two-fold approach was implemented. First, case studies of historic acute contamination events were evaluated. The chosen case studies must have conformed to the following criteria:

- An acute contamination event sizable enough to produce estimated peak concentrations of at least 1.0 ppm at the sensor must have occurred within the watershed
- Event must have occurred within a watershed which is continuously monitored by the SRBC to ensure the availability of data
- 5-minute interval continuously monitored data must be available during the year in which the event occurred
The criteria for concentrations of at least 1.0 ppm are from the work of Harris et al. (2016a, 2016b), which stated that typical water quality parameters will respond to concentrations greater than or equal to 1.0 ppm. The criteria for the 5-minute interval is relevant to the CANARY EDS for which it is recommended to use 2 to 15-minute interval data (US EPA 2010, 2012).

Case studies were chosen from reports compiled by Harris et al. (2016b), which found events between January 2011 and early 2014, and the US EPA (2015), who found events between January 2006 and April 2012. Additionally, Harris et al. (2016a, 2016b) computed estimated ranges of peak concentrations for each event using methods presented by Jobson (1997). Since the SRBC RWQMN was not implemented until January 2010 (SRBC 2017a), the two sources typically overlapped one another for the specified criteria. Only one event was found in the report by the US EPA (2015) that was not reported by Harris et al. (2016b). However, the sensor measurement interval was 4-hour at the time of the event, so it did not follow the criteria for this study. Finally, given these data and set criteria, three events were selected as case studies as described in the following sections.

3.2.1 Meshoppen Creek

The first selected case study was Meshoppen Creek, located within Susquehanna and Wyoming Counties, PA. Meshoppen creek flows southwest into the North Branch of the Susquehanna River. On January 10, 2011, a 23.8 m³ (150 BBL) spill of flowback fluid was reported by the PA DEP within the Meshoppen Creek watershed (US EPA 2015; Harris et al. 2016a, 2016b). The event occurred 14 km northwest of the monitoring station which is located east of the confluence with the Susquehanna River near Kaiserville, PA (Figure 3). Watershed and spill event data are summarized in Table 2.
Table 2. Watershed (SRBC 2017b) and event characteristics (Harris et al. 2016b) for the Meshoppen Creek case study.

<table>
<thead>
<tr>
<th>Watershed Characteristics</th>
<th>Event Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitored Drainage Area (km²)</td>
<td>134.7</td>
</tr>
<tr>
<td>Average Annual Flow at Monitoring Station (cms)</td>
<td>2.1</td>
</tr>
<tr>
<td>Maximum Flow Time from Headwaters (hr)</td>
<td>32.67</td>
</tr>
<tr>
<td>Natural Gas Drilling Pads</td>
<td>66</td>
</tr>
<tr>
<td>Contaminant</td>
<td>Vol. Spilled (m³)</td>
</tr>
<tr>
<td>Flowback Fluid</td>
<td>23.8</td>
</tr>
<tr>
<td>Aerial Distance to Monitoring Station (km)</td>
<td>14</td>
</tr>
<tr>
<td>Estimated Peak Concentration at Sensor (ppm)</td>
<td>1.1–2.1</td>
</tr>
</tbody>
</table>

3.2.2 Hammond Creek

The next selected case study was Hammond Creek, located largely in Tioga County, PA, with portions of the watershed extending into Bradford County, PA, and Steuben and Chemung Counties, NY. Hammond creek flows northeast into Seeley Creek close to the PA and NY state line. On January 26, 2011, 7.9 m³ (50 BBL) of production fluid was spilled within the Hammond
Creek watershed (US EPA 2015; Harris et al. 2017b). The event occurred 10 km west of the SRBC monitoring station which is located northeast of Millerton, PA, close to the confluence of Hammond Creek with Seeley Creek (Figure 4). Watershed and event characteristics are summarized in Table 3.

Figure 4. Location of the Seeley Creek Watershed, and the contamination event and SRBC continuous monitoring station within the Hammond Creek sub-watershed.
Table 3. Watershed (SRBC 2018a) and event characteristics (Harris et al. 2016b) for the Hammond Creek case study.

<table>
<thead>
<tr>
<th>Watershed Characteristics</th>
<th>Event Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitored Drainage Area (km²)</td>
<td>75.1</td>
</tr>
<tr>
<td>Average Annual Flow at Monitoring Station (cms)</td>
<td>1.0</td>
</tr>
<tr>
<td>Maximum Flow Time from Headwaters (hr)</td>
<td>13.52</td>
</tr>
<tr>
<td>Natural Gas Drilling Pads</td>
<td>12</td>
</tr>
<tr>
<td>Estimated Peak Concentration at Sensor (ppm)</td>
<td>0.85-1.87</td>
</tr>
</tbody>
</table>

3.2.3 East Branch of Wyalusing Creek

The final case study selected is the East Branch of Wyalusing Creek which is located in Susquehanna County, PA. East Branch of Wyalusing Creek flows west toward its confluence with the Middle Branch. On October 22, 2013, 20.7 m³ (104 BBL) of drilling mud, and cement were spilled within the watershed (US EPA 2015, Harris et al. 2016b). The event occurred 18 km northeast of the SRBC monitoring station which is located approximately 0.40 km upstream of the mouth near the town of Lawton, PA (Figure 5). Watershed and event characteristics are summarized in Table 4.
**Figure 5.** Location of the Wyalusing Creek Watershed, and the contamination event and SRBC continuous monitoring station within the East Branch sub-watershed.

**Table 4.** Watershed (SRBC 2018b) and event characteristics (Harris et al. 2016b) for the East Branch of Wyalusing Creek case study.

<table>
<thead>
<tr>
<th>Watershed Characteristics</th>
<th>Event Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitored Drainage Area (km²)</td>
<td>179.5</td>
</tr>
<tr>
<td>Average Annual Flow at Monitoring Station (cms)</td>
<td>2.9</td>
</tr>
<tr>
<td>Maximum Flow Time from Headwaters (hr)</td>
<td>20.23</td>
</tr>
<tr>
<td>Natural Gas Drilling Pads</td>
<td>40</td>
</tr>
<tr>
<td>Contaminant</td>
<td>Drilling Mud, Cement</td>
</tr>
<tr>
<td>Vol. Spilled (m³)</td>
<td>20.7</td>
</tr>
<tr>
<td>Arial Distance to Monitoring Station (km)</td>
<td>18</td>
</tr>
<tr>
<td>Estimated Peak Concentration at Sensor (ppm)</td>
<td>0.53–1.1</td>
</tr>
</tbody>
</table>

### 3.3 Simulated Event

The peak concentration estimates for each of the case studies relies on the assumption that the full quantity of the contaminant was spilled directly into the stream. Because this assumption cannot be confirmed, it is possible that some or all of the contaminant was contained...
before reaching the stream. Thus, it is possible that the peak concentration estimates are overestimated, or the contaminant may never reach the monitoring station at all. To ensure the following analysis was performed on data impacted by an acute contamination event, one was simulated using information from the Meshoppen Creek case study, prior work by Harris et al. (2016a, 2016b), and data from the Meshoppen Creek SRBC monitoring station.

3.3 Site Selection

In order to mimic an acute contamination event, one of the case studies was chosen; then, characteristics of the historical event discussed previously were utilized to simulate an event. Following that, selection of the site for the simulated event was based on the following criteria:

- No reported acute contamination events in the watershed for an entire year
- Watershed must be within the SRBC RWQMN
- 5-minute measurement interval monitoring data must be available for the selected year

Contrary to the case studies, a dataset with no reported contamination events was desired so any events detected could be distinguished as either false alarms or a result of the simulated event. To choose a site, data from Harris et al. (2016b) was analyzed by identifying a year where no significant contamination events were reported. Thus, the chosen site was Meshoppen Creek during the year 2012. There was one reported contamination event during this year; approximately 0.3–0.5 m³ (2–3 BBL) of production fluid was spilled on January 6, 2012. This event occurred 14 km northwest of the monitoring station and results in a peak concentration range of 0.018–0.049 ppm (Harris et al. 2016b). Therefore, this event was considered insignificant, and the selected site was Meshoppen Creek during the year 2012.

3.3.2 Contamination Event Simulation

Working under the assumption that the event described in Table 2 traveled directly into the stream, methods presented by Jobson (1997) were used to simulate the event. When an acute contamination event occurs within a stream, a fixed observer at a downstream point can expect the concentration to rise and fall in the shape of a left-skewed bell curve (Chapra 1997, Jobson 1997). This type of response can be approximated by a scalene triangle (Figure 6) using the
travel time of the contaminant and the estimated peak concentration (Jobson 1997). For this simulated event, the median of the peak concentration estimates from Harris et al. (2016b) was used as the peak concentration, and the most probable travel time was estimated using the methods of Jobson (1997) (Table 5). Using these three points as the basis of the simulated event, concentrations for each time step in between these values are estimated using linear regressions developed in Figure 6.

**Table 5.** Most probable travel time and concentration estimates for the 2011 Meshoppen creek event used for simulation in 2012.

<table>
<thead>
<tr>
<th>Timeline</th>
<th>Spill Event</th>
<th>Leading Edge of Plume Reaches Sensor</th>
<th>Peak Concentration Reaches Sensor</th>
<th>Concentration Reduced to 10% of Peak Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Probable Time (hr)</td>
<td>0</td>
<td>33.4</td>
<td>37.6</td>
<td>46.0</td>
</tr>
<tr>
<td>Concentration at Monitoring Station (ppm)</td>
<td>0</td>
<td>0</td>
<td>1.55</td>
<td>0.155</td>
</tr>
</tbody>
</table>
Figure 6. Approximation of concentrations of flowback fluid at the monitoring station using a scalene triangle. The first point \((t = 0)\) is the instant the leading edge of the plume reaches the monitoring station. The middle point \((t = 4.2 \text{ hr})\) is point at which the peak concentration arrives. The final point \((t = 12.6 \text{ hr})\) is the point at which the concentration is reduced to 10% of the peak.

Harris et al. (2016a, 2016b) found significant correlations between changes in pH and specific conductivity, and increasing concentrations of production fluid. According to the US EPA (2016), flowback and production fluid may be compositionally similar, so the relationships found by Harris et al. (2016a, 2016b) for production fluid were used to simulate water quality response to flowback fluid. Using data from those lab experiments, regression equations were developed to estimate the response of pH and specific conductivity to the simulated concentrations at the monitoring station.

A strong linear relationship was found between flowback fluid concentration and specific conductivity for concentrations from 0–20 ppm (Figure 7). Because the peak concentration of the simulated event at the monitoring station is 1.55 ppm, this linear relationship can be used to estimate the response of specific conductivity. Before directly using this relationship though, the regression equation was adjusted for differences in units. The measurement units used by Harris et al. (2016a, 2016b) were microsiemens per centimeter, while the SRBC were in units of millisiemens per centimeter. The regression equation is reduced then by a factor of 1000 to account for this difference (Equation 4). Therefore, to predict values of specific conductivity resulting from the simulated event concentrations, the slope of the regression was multiplied to

\[
y = 0.369x + 2E-16 \\
\text{R}^2 = 1
\]

\[
y = -0.1661x + 2.2475 \\
\text{R}^2 = 1
\]
the predicted concentration; this value was then added to the existing concentration value (Equation 5).

\[
\text{Predicted Value} = (Slope \times \text{Concentration}) + \text{Existing Measurement} \quad (5)
\]

Predicting the response of pH measurements to the simulated event was done similarly to specific conductivity. Harris et al. (2016a, 2016b) found that pH initially decreased linearly in lab experiments for low concentrations of production fluid, and exponentially reached an asymptote of approximately 6.5 with greater concentrations. Because the peak concentration of the simulated event is low, the pH response is predicted with a linear regression for concentration values of 0–5 ppm (Figure 8). Finally, the simulated response of pH to contaminant concentrations at the monitoring station were calculated using Equation 5.
Figure 8. Response of pH to increasing concentrations of flowback fluid for concentrations of 0–5 ppm. Regression developed for the average response of all lab experiments conducted by Harris et al. (2016a, 2016b).

Finally, having developed predictions for concentrations of flowback fluid at the monitoring station, and the corresponding responses of the measured water quality parameters, the event could then be simulated onto existing water quality data. For this simulation, the center of the 2012 Meshoppen Creek water quality data set was arbitrarily selected. Thus, the simulated event was set to occur on July 1, 2012 at time step 12:00:00. The travel time from the spill to the monitoring station was estimated to be 33.4 hr (Table 5). Therefore, the leading edge of the plume reaches the monitoring station on July 2, 2012 at time step 21:20:00. The event then continues until the predicted contaminant concentration reaches 0 ppm, which occurs on July 3, 2012 at time step 14:35:00. Water quality data from this time period was then used in conjunction with the aforementioned methods to simulate the contaminant passing by the monitoring station (Figure 9). These new water quality values then replaced the measurements from the raw data set for this time period.
Figure 9. The simulated water quality values for the duration of the simulated event from $t = 0$ corresponds to the arrival of the leading edge of the plume on July 2, 2012 at time step 21:20:00.

3.4 Event Detection with CANARY

The CANARY EDS gives the user the ability to adjust the key parameters used in event detection. These parameters affect the sensitivity, type and duration of event to detect, and the timeliness of event detection. The controlling parameters are the history window, outlier threshold, probability threshold, and the BED window. The US EPA (2010) has done extensive analysis determining effective guidelines for the history window. Therefore, the history window was held constant at two days (576 time steps) for all trials. The other three parameters are to be adjusted by the user for their desired level of sensitivity.

For this study, these three parameters were varied to provide several levels of sensitivity to event detection and gain insight into how these parameters influence the number of false alarms produced. First, the outlier threshold was varied from $0.5\sigma$–$3.0\sigma$ (where $\sigma$ = standard deviation of the history window) in increments of $0.5\sigma$. Next, the probability threshold was varied from 0.85–0.99 in increments of 0.05. Lastly, the following BED windows were used: 1 hour (12 time steps), 4 hours (48 time steps), 8 hours (96 time steps), and 12 hours (144 time steps). The BED window was varied at 4-hour intervals to evaluate the performance of event detection at significantly different lengths of the BED window. A maximum BED window of 12
hours was chosen based on: i) travel time estimates for the Meshoppen event (Table 5), and ii) a 12-hour BED causes significant delay between the contaminant reaching the monitoring station and detection of the event. Thus, a 12-hour BED is long enough to capture the entire duration of an event in a smaller watershed, and using a larger BED window is impractical to the goals associated with event detection.

For each analyzed data set (i.e., three case studies and simulated spill), each possible combination of outlier threshold, probability threshold, and BED window was run as a trial. Therefore, 24 trials were conducted for each BED window (Table 6).

**Table 6.** The experimental design matrix for each value of BED window tested. There were 24 trials conducted for each BED window.

<table>
<thead>
<tr>
<th>BED Window</th>
<th>Outlier Threshold ((\sigma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(_{\text{event}})</td>
<td>0.5</td>
</tr>
<tr>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>0.99</td>
<td></td>
</tr>
</tbody>
</table>

Additionally, each set of parameters was applied to both the LPCF and MVNN algorithms. With two algorithms tested on each data set, four BED windows for each algorithm, and 24 trials per BED window, a total of 192 event detection trials were conducted per site. Finally, three case studies were analyzed, and the simulated data set was tested before and after simulation, so a grand total of 960 event detection trials were completed for the study.

CANARY features a user interface that asks the user to select the desired configuration file each time event detection is completed on a data set. Filling out different configuration files and hand selecting them for each individual trial would have been too time consuming to complete the study. US EPA (2014) utilized a script to complete many trials at once. A similar script was implemented for this study in Windows Powershell. With this script, only one configuration file and one script had to be completed to conduct all 192 trials for an individual site. The script was also written to organize the results (alarms) of all 192 trials into a single comma delimited text file.
CANARY analyzes incoming data based on the evaluation type each water quality parameter is given in the configuration file. The three possible evaluation types are water quality (wq), calibration (cal), and operational (op) (US EPA 2014). Water quality parameters which are to be used in event detection are given the water quality evaluation type. Parameters which should be used to calibrate event detection techniques are given the calibration evaluation type. The operational evaluation type is given to water quality signals which are not to be used in event detection. Harris et al. (2016a, 2016b) found that specific conductivity, pH, and dissolved oxygen may be useful for detecting contamination events related to natural gas development. Therefore, these three parameters were designated as water quality evaluation type, while temperature and turbidity were designated as operational signals for all trials. No data was used for calibration.
4.0 Results and Discussion

4.1 Case Studies

4.1.1 Meshoppen Creek

Event detection results varied for the detection of the 23.8 m³ spill of flowback fluid in Meshoppen Creek. Previous research suggests that the spill would impact measurements of specific conductivity and pH (Harris et al., 2016b). In the week following the spill event (week 2), MVNN algorithms with outlier threshold set to 1.0σ were the only algorithms which detected any events; the LPCF algorithm did not detect any events near the known spill date. Across these MVNN algorithms, the range of alarms per week for the second week was 4–10 alarms per trial (Figure 10). While it is possible that these alarms could be in response to the reported contamination event, it is impossible to distinguish this with confidence. Trials in which the outlier threshold is set to 1.0σ produced a significant number of false alarms throughout the entirety of the year where there were no reported events (Figure 10 – Figure 14). Therefore, the event was not successfully detected in this spill of flowback fluid for either algorithm.
Figure 10. Weekly number of alarms produced by each trial of the MVNN algorithm with outlier threshold set to $1.0\sigma$. Alarms resulting from the reported event were expected in week 2. BED corresponds to the length of the BED window in time steps, and $P(\text{event})$ represents the probability threshold for each trial.
Figure 11. CANARY output plot for Meshoppen Creek during the first quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 12. CANARY output plot for Meshoppen Creek during the second quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 13. CANARY output plot for Meshoppen Creek during the third quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 14. CANARY output plot for Meshoppen Creek during the fourth quarter of 2011: BED = 1 hr, outlier threshold = 1.0σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

In application of an EDS, a high false alarm rate is impractical. An EDS manager must be able to examine water quality data at the time of an alarm and quickly distinguish if the alarm was the result of possible contamination, or if it is the result of natural variability (e.g. sudden increase in flow rate) or sensor malfunction. Therefore, it is likely that water quality managers would not implement such sensitive algorithms on natural streams, and can be concluded that the Meshoppen Creek event was not able to be detected by CANARY for the trials implemented in this study.
Although the Meshoppen Creek event could not be detected, the results of all trials provide insight into the behavior of CANARY at different parameter configurations (Table 7 and Table 8). Outlier threshold and BED window had the greatest impact on the sensitivity of CANARY to event detection. Probability threshold does have some impact on the number of alarms produced but is substantially less significant than the other two parameters. For example, the LPCF algorithm produced 159 alarms with outlier threshold = 0.5σ, BED window = 12 time steps (1 hour), and probability threshold = 0.85. When the probability threshold was increased to the maximum of 0.99, the number of alarms is only reduced by 27. However, when the outlier threshold was increased to 1.0σ, the total number of alarms decreased by 110. Similarly, increasing the BED window to 48 time steps (4 hours) decreased the total number of alarms by 65 (Table 7). Furthermore, at higher values of outlier threshold and BED window, the number of alarms does not change with probability threshold (Table 7 and Table 8).
Table 7. Results for all Meshoppen Creek trials using the LPCF algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>159</td>
<td>49</td>
<td>31</td>
<td>22</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>159</td>
<td>49</td>
<td>31</td>
<td>22</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>148</td>
<td>47</td>
<td>27</td>
<td>21</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>132</td>
<td>44</td>
<td>26</td>
<td>19</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>94</td>
<td>31</td>
<td>16</td>
<td>15</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>94</td>
<td>31</td>
<td>16</td>
<td>15</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>94</td>
<td>29</td>
<td>15</td>
<td>14</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>84</td>
<td>27</td>
<td>14</td>
<td>13</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>77</td>
<td>25</td>
<td>14</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>75</td>
<td>25</td>
<td>14</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>76</td>
<td>25</td>
<td>14</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>68</td>
<td>25</td>
<td>14</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>62</td>
<td>23</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>65</td>
<td>23</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>62</td>
<td>23</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>62</td>
<td>23</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>
Table 8. Results for all Meshoppen Creek trials using the MVNN algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td>Total Number of Alarms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>152</td>
<td>167</td>
<td>35</td>
<td>24</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>152</td>
<td>167</td>
<td>35</td>
<td>24</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>143</td>
<td>163</td>
<td>33</td>
<td>21</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>136</td>
<td>158</td>
<td>31</td>
<td>21</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>105</td>
<td>116</td>
<td>21</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>105</td>
<td>116</td>
<td>21</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>103</td>
<td>114</td>
<td>21</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>102</td>
<td>113</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>88</td>
<td>96</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>88</td>
<td>95</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>87</td>
<td>93</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>85</td>
<td>89</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>74</td>
<td>83</td>
<td>18</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>72</td>
<td>82</td>
<td>18</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>71</td>
<td>80</td>
<td>18</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>70</td>
<td>80</td>
<td>17</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

It should be noted that while the length of the BED window clearly has an impact on the sensitivity of event detection, it should generally be set small enough for the duration of an event one hopes to detect, and time of response to an event desired (US EPA 2014). The outlier threshold parameter will then likely be more important to sensitivity adjustments.

As the outlier threshold was increased, the decrease in sensitivity was exponential for the LPCF algorithm (Figure 15). However, this trend was not followed for this data set using the MVNN algorithm (Figure 16). The MVNN algorithm with outlier threshold = 1.0σ was the most sensitive to event detection. Considering the role the outlier threshold fulfills in event detection, this result is unexpected; a decreasing trend as found in the LPCF algorithm was expected.
(Figure 15). Configuration files for the MVNN algorithm trials with outlier threshold $= 1.0\sigma$ were inspected for errors, but none were found. Interestingly, the MVNN algorithms with outlier threshold $= 1.0\sigma$ were also the only algorithms to produce any alarms during the week of expected response from the event.

Figure 15. Decreasing trend of number of alarms versus the outlier threshold for Meshoppen Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.
Figure 16. Trend in number of alarms versus the outlier threshold for Meshoppen Creek trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity. Unlike the LPCF algorithm, the number of alarms increases from 0.5σ to 1.0σ.

4.1.2 Hammond Creek

Analysis of the Hammond Creek data set yielded similar results to what was found in the analysis of the Meshoppen Creek data set. If the EDS were to detect the contamination event, alarms from the CANARY EDS would be expected during the following week (week 3). Previous research suggests that the spill would impact measurements of specific conductivity and pH (Harris et al., 2016b). CANARY only detected events during week 3 with the MVNN algorithm with outlier threshold of 0.5σ, and BED window less than or equal to 4 hours (48 time steps). However, the MVNN algorithm with outlier threshold of 0.5σ produced too many false alarms throughout the year to confidently say that the week 3 alarms are the result of the contamination event (Figure 17 – Figure 21). For practical purposes then, it could be concluded that the reported Hammond Creek contamination event (7.9 m³ of production fluid) was not detected by the CANARY EDS.
Figure 17. Weekly number of alarms produced by each trial of the MVNN algorithm with outlier threshold set to 0.5σ for Hammond Creek. Alarms resulting from the reported event were expected in week 3. BED corresponds to the length of the BED window in time steps, and P(event) represents the probability threshold for each trial.
Figure 18. CANARY output plot for Hammond Creek during the first quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 19. CANARY output plot for Hammond Creek during the second quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 20. CANARY output plot for Hammond Creek during the third quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 21. CANARY output plot for Hammond Creek during the fourth quarter of 2011: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.

Behaviors in event detection sensitivity to the adjustment of the outlier threshold, probability threshold, and BED window for the Hammond Creek trials were similar to those in the Meshoppen Creek trials (Table 9 and Table 10). The sensitivity of the LPCF algorithm with increasing outlier threshold exhibits the same behavior (Figure 22). However, unlike the Meshoppen Creek trials, the sensitivity of CANARY to event detection decreased exponentially for the MVNN algorithm with increasing outlier threshold (Figure 23).
Table 9. Results for all Hammond Creek trials using the LPCF algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>71</td>
<td>25</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>71</td>
<td>25</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>68</td>
<td>23</td>
<td>15</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>67</td>
<td>23</td>
<td>15</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>50</td>
<td>21</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>50</td>
<td>21</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>50</td>
<td>21</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>49</td>
<td>21</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>41</td>
<td>18</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>41</td>
<td>18</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>41</td>
<td>18</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>40</td>
<td>18</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>39</td>
<td>16</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>39</td>
<td>16</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>39</td>
<td>16</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>38</td>
<td>16</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 10. The results for all Hammond Creek trials using the MVNN algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>Total Number of Alarms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td><strong>Probability Threshold</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.85</td>
<td>89</td>
<td>35</td>
</tr>
<tr>
<td>0.9</td>
<td>89</td>
<td>35</td>
</tr>
<tr>
<td>0.95</td>
<td>89</td>
<td>34</td>
</tr>
<tr>
<td>0.99</td>
<td>87</td>
<td>34</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.85</td>
<td>66</td>
<td>29</td>
</tr>
<tr>
<td>0.9</td>
<td>66</td>
<td>29</td>
</tr>
<tr>
<td>0.95</td>
<td>66</td>
<td>28</td>
</tr>
<tr>
<td>0.99</td>
<td>66</td>
<td>28</td>
</tr>
<tr>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.85</td>
<td>59</td>
<td>26</td>
</tr>
<tr>
<td>0.9</td>
<td>58</td>
<td>25</td>
</tr>
<tr>
<td>0.95</td>
<td>57</td>
<td>25</td>
</tr>
<tr>
<td>0.99</td>
<td>55</td>
<td>25</td>
</tr>
<tr>
<td>144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.85</td>
<td>54</td>
<td>22</td>
</tr>
<tr>
<td>0.9</td>
<td>53</td>
<td>22</td>
</tr>
<tr>
<td>0.95</td>
<td>53</td>
<td>21</td>
</tr>
<tr>
<td>0.99</td>
<td>51</td>
<td>21</td>
</tr>
</tbody>
</table>

Figure 22. Decreasing trend in number of alarms versus the outlier threshold for the Hammond Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.
Figure 23. Decreasing trend in number of alarms versus the outlier threshold for the Hammond Creek Trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

4.1.3 East Branch of Wyalusing Creek

Unlike the first two case studies, CANARY produced no alarms during the week following the reported 20.7 m$^3$ spill of drilling mud in the East Branch of Wyalusing Creek. Previous research suggests that the spill would impact measurements of specific conductivity, pH, and dissolved oxygen (Harris et al., 2016b). There was an event detected for several trials on October 31, 2013, 9 days after the reported event. However, upon closer examination of the CANARY output plots at the time of the alarm, these alarms appear to result from a malfunction of the sensor (Figure 24). Temperature, specific conductivity, pH, and dissolved oxygen all exhibit a “spike” behavior at the time of the alarm. Specific conductivity, the parameter for which the event was detected in Figure 24, rapidly and unnaturally drops from approximately 0.15 mS/cm to 0 mS/cm. Therefore, it was concluded that the reported contamination event was not detected by the trials implemented in this study.
Figure 24. CANARY output plot for East Branch of Wyalusing Creek during the week of October 29, 2013 to November 5, 2013: BED = 1 hr, outlier threshold = 3.0σ, P(event) = 0.99. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. Markers for both the LPCF and MVNN algorithms indicate an event was detected 9 days after spill was reported, even on the least sensitive algorithms.
Behavior of event detection sensitivity for the East Branch of Wyalusing Creek trials was similar to that of the Hammond Creek trials. Again, more sensitive algorithms are characterized by high numbers of false alarms (Figure 25 – Figure 28). Sensitivity adjustments are dominated by the outlier threshold for both the LPCF and MVNN algorithms, while the probability threshold affects the number of alarms substantially less (Table 11 and Table 12). Additionally, the relationship between sensitivity and outlier threshold could be characterized as approximately exponentially decreasing for both algorithms (Figure 29 and Figure 30).
Figure 25. CANARY output plot for East Branch of Wyalusing Creek during the first quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 26. CANARY output plot for East Branch of Wyalusing Creek during the second quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 27. CANARY output plot for East Branch of Wyalusing Creek during the third quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Figure 28. CANARY output plot for East Branch of Wyalusing Creek during the fourth quarter of 2013: BED = 1 hr, outlier threshold = 0.5σ, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm.
Table 11. Results for all East Branch of Wyalusing Creek trials using the LPCF algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>157</td>
<td>61</td>
<td>26</td>
<td>14</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>157</td>
<td>61</td>
<td>26</td>
<td>14</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>148</td>
<td>59</td>
<td>26</td>
<td>15</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>144</td>
<td>59</td>
<td>26</td>
<td>15</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>103</td>
<td>39</td>
<td>17</td>
<td>11</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>103</td>
<td>39</td>
<td>17</td>
<td>11</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>99</td>
<td>37</td>
<td>15</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>98</td>
<td>34</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>81</td>
<td>22</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>78</td>
<td>22</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>77</td>
<td>21</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>75</td>
<td>19</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>67</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>67</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>67</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>64</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 12. Results for all East Branch of Wyalusing Creek trials using the MVNN algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>140</td>
<td>87</td>
<td>20</td>
<td>12</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>140</td>
<td>87</td>
<td>20</td>
<td>12</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>132</td>
<td>86</td>
<td>19</td>
<td>12</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>125</td>
<td>85</td>
<td>19</td>
<td>12</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>88</td>
<td>64</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>88</td>
<td>64</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>85</td>
<td>60</td>
<td>12</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>83</td>
<td>58</td>
<td>11</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>74</td>
<td>44</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>75</td>
<td>44</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>73</td>
<td>44</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>70</td>
<td>43</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>63</td>
<td>37</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>62</td>
<td>36</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>63</td>
<td>36</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>63</td>
<td>36</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
Figure 29. Decreasing trend in number of alarms versus the outlier threshold for the East Branch of Wyalusing Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 30. Decreasing trend in number of alarms versus the outlier threshold for the East Branch of Wyalusing Creek trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

4.2 Simulated Flowback Fluid Event

Results for the simulated spill of 23.8 m$^3$ of flowback fluid come from two data sets: the “clean” data which is the raw data from the SRBC with no reported significant contamination
events, and the “simulated” data which features the imposed contamination event onto specific conductivity and pH as previously described. Conclusions regarding event detection sensitivity were drawn from the results of the clean data set. Finally, the clean and simulated data set results were compared at the time of the simulated event to decide if the CANARY EDS detected the simulated event for any of the trials.

4.2.1 Meshoppen Creek 2012: Clean Data Set

The behavior of the EDS with respect to alteration of outlier threshold, probability threshold, and BED window was the same as the Meshoppen Creek case study. The main difference between the two is that there were substantially fewer alarms overall, especially for the more sensitive algorithms with lower outlier thresholds (Table 13 and Table 14). This is likely the result of less natural variability and sensor malfunction throughout the data set.

<table>
<thead>
<tr>
<th>Bed Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>53</td>
<td>22</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>53</td>
<td>22</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>51</td>
<td>20</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>51</td>
<td>20</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>40</td>
<td>16</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>40</td>
<td>16</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>40</td>
<td>16</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>40</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 13. Results from all Meshoppen Creek 2012 trials before event simulation using the LPCF algorithm.
Table 14. Results from all Meshoppen Creek 2012 trials before event simulation using the MVNN algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Probability Threshold</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>70</td>
<td>83</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>70</td>
<td>83</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>68</td>
<td>83</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>66</td>
<td>81</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td><strong>Total Number of Alarms</strong></td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>56</td>
<td>67</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>56</td>
<td>67</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>54</td>
<td>65</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>53</td>
<td>63</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>46</td>
<td>54</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>43</td>
<td>48</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>43</td>
<td>47</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>42</td>
<td>47</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>40</td>
<td>44</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
</tbody>
</table>

The sensitivity to event detection decreased exponentially for the LPCF algorithm (Figure 31). This result is consistent with all of the case studies. For the MVNN algorithm however, the behavior was similar to the Meshoppen Creek case study where the algorithm is most sensitive to event detection when the outlier threshold is set equal to 1.0σ (Figure 16 and Figure 32). Again, the configuration files were inspected for errors on these trials but none were found.
Figure 31. Decreasing trend in number of alarms versus the outlier threshold for the clean 2012 Meshoppen Creek trials using the LPCF algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity.

Figure 32. Number of alarms versus the outlier threshold for the clean 2012 Meshoppen Creek trials using the MVNN algorithm. The average number of alarms was taken across all values of probability threshold since it was found that parameter has little impact on sensitivity. Unlike the LPCF algorithm, the number of alarms increases from $0.5\sigma$ to $1.0\sigma$. 
4.2.2 Meshoppen Creek 2012: Simulated Data Set

Results from the data set featuring the simulated event were almost identical to those from the clean data set. For all trials of the MVNN, the number of detected events from the simulated data set were the same as those from the clean data set (Table 16). The same occurred for the LPCF algorithm with the exception of all trials whose outlier threshold was set to $2.5\sigma$ (Table 15). Each trial using the LPCF algorithm with outlier threshold set to $2.5\sigma$ detected one less event when compared to the clean data set (Table 17). Thus, 16 less events were detected in total.

Table 15. Results from all Meshoppen Creek 2012 trials after event simulation using the LPCF algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Outlier Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>53</td>
<td>22</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>53</td>
<td>22</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>51</td>
<td>20</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>51</td>
<td>20</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>40</td>
<td>16</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>40</td>
<td>16</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>40</td>
<td>16</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>40</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>36</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>34</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 16. Results from all Meshoppen Creek 2012 trials after event simulation using the MVNN algorithm.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Probability Threshold (σ)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.85</td>
<td>70</td>
<td>83</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>70</td>
<td>83</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>68</td>
<td>83</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>66</td>
<td>81</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>56</td>
<td>67</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>56</td>
<td>67</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>54</td>
<td>65</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>53</td>
<td>63</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>48</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>46</td>
<td>54</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>43</td>
<td>48</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>43</td>
<td>47</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>42</td>
<td>47</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>40</td>
<td>44</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 17. Net alarms for the LPCF algorithm with outlier threshold set to 2.5σ following the implemented changes to water quality data as a result of the simulated event. All alarm changes occurred during the time of event simulation.

<table>
<thead>
<tr>
<th>BED Window (Time Steps)</th>
<th>Probability Threshold</th>
<th>Net Alarms (Clean - Simulated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.85</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>48</td>
<td>0.85</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>96</td>
<td>0.85</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>144</td>
<td>0.85</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1</td>
</tr>
</tbody>
</table>

Interestingly, the 16 events which were not detected in the simulated data set occurred during the time period in which the simulated event was imposed upon the raw data. For both the clean and simulated data sets, several trials of both algorithms detected an event on July 3, 2012. Events were detected at this time as the result of decreasing specific conductivity. For both data sets, these events were detected at all variations of BED window and probability threshold. On the clean data set, both the LPCF and MVNN algorithms detected this event for all outlier threshold values up to 2.5σ (Figure 33). On the simulated dataset, results from the MVNN algorithm did not change. However, the LPCF algorithm only detected the event for algorithms with outlier threshold up to 2.0σ on the simulated data set (Figure 34). Since the only difference in these two trials is the imposed changes to water quality data from the simulated event, it was concluded that in this particular case, the simulated contamination event inhibited event
detection by the LPCF algorithm. Because event simulation caused an increase in specific conductivity data, the measurements became more similar to background data and an event could no longer be declared when the outlier threshold was set to \(2.5\sigma\).

**Figure 33.** CANARY output plot for the clean Meshoppen Creek clean data set during the week of July 1, 2012 to July 8, 2012: BED = 1 hr, outlier threshold = \(2.5\sigma\), \(P(\text{event}) = 0.90\). Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm. The red box indicates the time frame of event simulation.
Figure 34. CANARY output plot for the simulated event on Meshoppen Creek during the week of July 1, 2012 to July 8, 2012: BED = 1 hr, outlier threshold = 2.5$\sigma$, P(event) = 0.90. Water quality measurements are shown in the upper five plots. The lower two plots show the probability that an event has occurred. A blue “plus” symbol is added to water quality plots with which an event was detected using the MVNN algorithm. A blue “dot” symbol is added to water quality plots for which an event was detected using the LPCF algorithm. The red box indicates the time frame of event simulation.

Overall, there were no additional alarms from CANARY on the simulated data set resulting from the imposed changes to water quality data. It was concluded then that the CANARY EDS was unable to detect the simulated event. With certainty that water quality measurements of pH and specific conductivity were affected by the simulated event, it can be concluded that these changes were not significant enough for CANARY to detect the event for the trials tested. This indicates that the CANARY EDS may be ineffective on natural channels when contaminant concentrations are low.
5.0 Discussion

5.1 Event Detection in Natural Channels

Most event detection systems, including CANARY, have been developed for use in municipal water supply (pipe flow) systems (US EPA 2010). While these conditions provide water quality variability resulting from changes in demand and distribution network operations, they are generally more predictable and stable than water quality measurements in a natural channel. Water quality measurements in natural channels are prone to variability resulting from ever-changing hydrologic conditions. For example, pH may be relatively stable in a pipe flow system, but in a natural channel exhibits strong diurnal periodicity (Figure 33 and Figure 34). Other factors, such as precipitation and flow rate can have strong effects on water quality parameters (Gerner and Waddel 2003 and Gasim et al. 2006). Small watersheds, like those used in this study, are especially prone to rapid changes as the timescale of hydrologic change is much smaller. For example, the time to peak discharge in a stream is greatly impacted by the size of the watershed. On small watersheds, the time to peak discharge will, in most cases, be much smaller than it would on large watersheds (USGS 2018). Since these hydrologic characteristics are subject to rapid changes, the effects they impose on water quality measurements are as well. Therefore, event detection in natural channels is inherently more challenging than in pipe systems as it is more difficult to distinguish changes due to natural variability from those due to a contamination event.

5.2 Implications for future event detection

Although CANARY was unsuccessful at detecting the historic and simulated events in this study, it may still be a viable option as an early warning system to acute contamination events. Because reported details from the historic events in this study are limited, it cannot be guaranteed that the entire reported contaminant volumes were spilled directly into the streams. Thus, the predicted concentrations may be overestimated. Unless these key details can be guaranteed, future work should focus on the use of simulated events. By using simulated events, the user gains control over the concentrations reported at the monitoring station, providing the study with enhanced flexibility. With the ability to adjust the magnitude of contamination impacts, future work should focus on determining the minimum contaminant concentration which can be detected by CANARY, the maximum watershed monitoring size based on historic
event volumes and concentration estimates, and optimizing CANARY for use in natural channels.

In this study, the simulated event was based on details from one of the case studies. Future work should use similar estimation methods for concentration, but vary the size of the spills, and the distance traveled to the sensor. Varying these two parameters will provide a range of different peak concentrations and contaminant durations at the monitoring station. Then, by imposing a wide range of simulated events, the minimum event size detectable by CANARY could be determined. From these conclusions, further insight could be gained by relating the minimum event size to watershed size. If an approximate peak concentration and event duration are determined, the methods presented by Jobson (1997) could be used to predict the appropriate maximum monitored drainage area. Following determination of maximum monitored drainage area, CANARY should be optimized to achieve high rates of event detection and low false alarm rates.

Optimization of CANARY can be achieved by multiple methods. First, the configuration parameters can be used to adjust sensitivity and reduce the number of false alarms. The outlier threshold had the greatest effect on the number of events reported throughout the trials of this study. Trials using outlier thresholds less than $1.5\sigma$ tend to produce an excessive number of false alarms throughout the year. For most trials, the number of reported alarms became minimized at outlier threshold values of approximately $2\sigma - 2.5\sigma$. Therefore, it is recommended to use outlier thresholds greater than $1.5\sigma$ when working with natural channels. While the BED window has an effect on sensitivity at lower values of outlier threshold, this effect fades away with less sensitive algorithms who produce fewer false alarms. The main effect of BED window at higher values of outlier threshold is the timing of event reporting. This is because it requires more time steps for an event to be declared at higher values of BED window. Additionally, using large BED windows, such as $8 - 12$ hr, could be susceptible to false negatives for events of shorter time frames, such as a spill occurring very close to the monitoring station. Therefore, a shorter BED window of approximately $1$ hr is recommended. The tested values of probability threshold showed little variation in the number of false alarms. US EPA (2014) implemented significantly smaller values of probability threshold on municipal distribution networks than those tested here. While this may be a possible optimization adjustment, higher thresholds, such as those tested in
this study, are likely necessary to aid in reducing the number of false alarms resulting from water quality variability in natural channels.

In addition to using the three discussed configuration parameters for optimization, calibration signals may also be useful. CANARY was designed with the ability to accept data input, such as flow rate or other utility operating parameters, and use them to calibrate the event detection algorithms (US EPA 2010). This was done to help distribution network managers reduce the number of false alarms resulting from expected changes in daily utility operations. Since water quality data in natural channels can vary with hydrologic changes, such hydrologic data may be useful as a calibration signal. For example, high flows have the potential to rapidly change water chemistry and therefore water quality measurements. Therefore, using flow rate data as a calibration signal may prove extremely useful. The difficulty in this will be obtaining these data on the same time scale as the incoming water quality data. The watersheds of interest in this study are all small watersheds with no USGS gaging station available. While extrapolating gaged stream data to these ungagged watersheds has been done (Harris et al. 2016a, 2016b), the data are not measured at a fine enough time scale for use as a calibration signal. Therefore, developing methods to quantify hydrologic changes in a watershed on the same time scale as incoming water quality data may be of future concern.

Finally, other optimization techniques could include creating custom algorithms, and using CANARY’s pattern matching capabilities. CANARY gives the user the ability to write algorithms which perform outlier classification, like the built-in the LPCF and MVNN algorithms. Algorithms must be written in JAVA and specified in the configuration file (US EPA 2012). The LPCF algorithm uses an AR model to predict future concentrations for comparison to measured water quality values. Other model types should be explored by future users in attempt to create better predictions for the natural environment setting. If an algorithm could be found to create better predictions in response to natural variability, it may be possible to decrease the outlier threshold and implement more sensitive algorithms without increasing the false alarm rate. Additionally, pattern matching techniques could be explored to reduce the number of false alarms. CANARY provides the user with the ability to create a library of typical changes in water quality that should not be signified as an event (US EPA 2010). For example, if a water quality sensor malfunction can be characterized by a brief spike in the parameters, this behavior
could be saved into the pattern library for comparison in event detection. Then, when the EDS is implemented with pattern matching enabled, all detected events would be compared to the malfunction in the pattern library. If a detected event is then matched to that pattern, CANARY will not signal an alarm, thereby eliminating false alarms matching this type of sensor malfunction behavior.
6.0 Conclusions

Little is known about the application of EDS in natural systems. In this study, an EDS was applied to three case studies with reported contamination events (i.e., production fluid, flowback fluid, and drilling mud), and one simulated event. In the Meshoppen and Hammond Creek case studies, CANARY detected events in the week following the reported spills. However, the only events during this time were reported for the most sensitive of algorithms. These algorithms were found to report numerous false alarms throughout the year, making them impractical for their designed purpose. Therefore, it was concluded that the EDS was unable to detect the reported contamination events on Meshoppen and Hammond Creek. In the East Branch of Wyalusing Creek case study, an event was detected by several trials, including some of the least sensitive algorithms, nine days following the reported event. However, upon inspection of water quality data at this time (Figure 24), it is evident that these alarms occur as a result of sensor malfunction. Having detected no other events during the week following the spill, it was concluded that the CANARY EDS was unable to detect the reported contamination event on East Branch of Wyalusing Creek. Therefore, none of the tested historical events were able to be detected for the trials of the CANARY EDS implemented.

Following analysis of the three case studies, event characteristics of the Meshoppen Creek contamination event were used to simulate the effects of a contamination event on a data set where no significant events are known to be present. The goal of the simulated event was to provide event detection analysis on a dataset where changes to water quality data resulting from a contamination event are guaranteed. After analyzing the data with CANARY before and after the event was imposed on the data set, it was concluded that the simulated event was not detected. In fact, the effects of the simulated event inhibited event detection of a false alarm by the LPCF algorithm with outlier threshold set to $2.5\sigma$ (Figure 33 and Figure 34). Because no additional events were detected during the simulated event, it can be concluded that the change imposed on water quality data as a result of the simulated contaminant concentration was not great enough for the CANARY EDS to declare the change an event for any of the trials implemented. Therefore, greater contaminant concentrations are most likely necessary for these events to be detected by CANARY.
Ensuring greater concentrations of contaminants at sensors requires either the sensor to be placed closer to the spill event, or a much larger spill event. The three dominating factors of pollutant concentration downstream of a spill are the distance traveled by the pollutant, the quantity spilled, and the flow rate of the stream at the time of the spill. Higher flow rates provide greater velocities and decreased travel times, but also more dilution. Lower flow rates ensure greater concentrations initially as dilution is much more limited, but slower velocities increase contaminant travel times and make them more prone to losses from decay and/or settling. Increased spill volume may provide enhanced event detection further downstream but would be ecologically problematic. Therefore, the best solution to achieve greater contaminant concentrations at monitoring stations is to increase the monitoring presence on watersheds where spills are of concern.

The CANARY EDS was incapable of detecting any of the events presented in this study. However, using simulated events in conjunction with some of CANARY’s more advanced event detection capabilities may provide contamination event detection solutions for natural channels. The application of CANARY in this study has been narrowly focused on contamination resulting from natural gas development in the Marcellus Shale play, but future work with this software need not be limited to this industry. The basic principle of the CANARY EDS is to detect changes in water quality data which are not congruent with background data. Thus, it may be possible to apply these event detection techniques to contamination events from a multitude of sources, so long as there is a measurable response by water quality parameters to the contaminant and high frequency water quality measurements are available.
7.0 References


Appendix A: CANARY Configuration File Used with Script (Hammond Creek)

--- # CANARY Config File
canary:
    run mode: BATCH
    control type: INTERNAL
    control messenger: null
    driver files: null

# Enter the time step options below according to format of data
timing options:
    dynamic start-stop: off
    date-time format: mm/dd/yyyy HH:MM:SS
    date-time start: 01/01/2011 00:00:00
    date-time stop: 12/31/2011 23:55:00
    data interval: 00:05:00
    message interval: 00:00:01

# Enter the list of data sources below
data sources:
    - id: hammond_in
      type: csv
      location: C:\Users\user\Documents\MyCANARY\research\data\hammond\hammond.csv
      enabled: yes
      timestep options:
        field: "TIME_STEP"

# Enter the list of signals and their defined evaluation type below
# evaluation type wq parameters are used for detection. op are not used for detection. cal signals are used for calibration and not detection
signals:
    - id: TEST_TEMP
      SCADA tag: TEMP_VAL
      evaluation type: op
      parameter type: TEMP
      ignore changes: none
      data options:
        precision: 0.01
        units: 'deg_C'
        valid range: [-inf, inf]
        set points: [-inf, inf]
- id: TEST_TURB
  SCADA tag: TURB_VAL
  evaluation type: op
  parameter type: TURB
  ignore changes: none
  data options:
    - precision: 0.1
    - units: 'NTU'
    - valid range: [.-inf, .inf]
    - set points: [.-inf, .inf]

- id: TEST_COND
  SCADA tag: COND_VAL
  evaluation type: wq
  parameter type: SP_COND
  ignore changes: none
  data options:
    - precision: 0.001
    - units: '\{\mu\}S/cm'
    - valid range: [.-inf, .inf]
    - set points: [.-inf, .inf]

- id: TEST_PH
  SCADA tag: PH_VAL
  evaluation type: wq
  parameter type: PH
  ignore changes: none
  data options:
    - precision: 0.01
    - units: 'pH'
    - valid range: [.-inf, .inf]
    - set points: [.-inf, .inf]

- id: TEST_DO
  SCADA tag: DO_VAL
  evaluation type: wq
  parameter type: DO
  ignore changes: none
  data options:
    - precision: 0.01
    - units: 'mg/L'
    - valid range: [.-inf, .inf]
    - set points: [.-inf, .inf]

# Enter the list of event detection algorithms below
algorithms:
- id: test_LPCF
- id: test_LPCF
  type: LPCF
  history window: 576
  outlier threshold: OTT
  event threshold: ETT
  event timeout: 12
  event window save: 1
  BED:
    window: BBED
    outlier probability: 0.5

- id: test_MVNN
  type: MVNN
  history window: 576
  outlier threshold: OTT
  event threshold: ETT
  event timeout: 12
  event window save: 1
  BED:
    window: BBED
    outlier probability: 0.5

# Enter the list of monitoring stations below
monitoring stations:
- id: hammond_SRBC
  station id number:
  station tag name: hammond
  location id number: -1
  enabled: yes
  inputs:
    - id: hammond_in
  outputs:
  signals:
    - id: TEST_TEMP
    - id: TEST_TURB
    - id: TEST_COND
    - id: TEST_PH
    - id: TEST_DO
      cluster: no
  algorithms:
    - id: test_LPCF
    - id: test_MVNN
Appendix B: Example Powershell Script Used to Complete Trials (Hammond Creek)

"Start date and time is: $((Get-Date).ToString())"
$ot="0.5 1.0 1.5 2.0 2.5 3.0"
$bed="12 48 96 144"
$et="0.85 0.9 0.95 0.99"
foreach ($l1 in $ot.Split(" ")) {
    foreach ($l2 in $bed.Split(" ")) {
        foreach ($l3 in $et.Split(" ")) {
            $foldername1 =
            ("Hammond_"+$l1+"_"+$l2+"_"+$l3)
            mkdir $foldername1
            cd $foldername1
            echo $foldername1
            cp C: \Users\user\Documents\MyCANARY\research\config\Hammond\Hammond_script.yml
            cat Hammond_script.yml | %{ $_ -replace 'BBED', ($l2)} > temp1.yml
            cat temp1.yml | %{ $_ -replace 'OTT', ($l1)} >temp2.yml
            cat temp2.yml | %{ $_ -replace 'ETT', ($l3)} >temp1.yml
            cmd /c "type temp1.yml > Hammond_script.yml"
            rm temp*
            cd ..
        }
    }
}$folders= ls Hamm_* -Name
$j=1
foreach ($i in $folders) {
    cd $i
    Start-Process 'C:\Program Files (x86)\CANARY\bin\canary.exe' Hammond_script.yml
    cd ..
    $j++
    while ($j -gt 3) {
        Start-Sleep -s 60
        $j=(ps | findstr "canary").length
        if ($j -gt 30) {$j = 1}
    }
}
Start-Sleep -s 60
$j=$(ps | findstr "canary").length
}
foreach ($i in $folders) {
    cd $i
    findstr "test_LPCF" .\Hammond_script.hammond_SRBC.summary.txt |
    %{ $a+=$_.Split('||')[6]; $b+=$_.Split('||')[22];
        $c+=$_.Split('||')[24]; $d+=$_.Split('||')[26]; if ($a)
    {"$i,$a,$b,$c,$d"};$a=$b=$c=$d=""}>> .\summary.txt
    findstr "test_MVNN" .\Hammond_script.hammond_SRBC.summary.txt |
    %{ $a+=$_.Split('||')[6]; $b+=$_.Split('||')[22];
        $c+=$_.Split('||')[24]; $d+=$_.Split('||')[26]; if ($a)
    {"$i,$a,$b,$c,$d"};$a=$b=$c=$d=""}>> .\summary2.txt
    cd ..
    echo "Hamm, LPCF, MVNN" > results.txt
    foreach ($i in $folders) {
        cd $i
        $lpcf=(cat .\summary.txt).length
        $mvnn=(cat .\summary2.txt).length
        echo "$i, $lpcf, $mvnn" >> ..\results.txt
        Get-Content .\summary.txt | Add-Content ..\LPCF_alarms.txt
        Get-Content .\summary2.txt | Add-Content ..\MVNN_alarms.txt
        cd..
    }
    "Stop date and time is: $((Get-Date).ToString())"
Appendix C: CANARY Configuration File Written from Script (Hammond Creek, BED window = 1 hr, outlier threshold = 0.5\(\sigma\), \(P(\text{event}) = 0.90\))

---

# CANARY Config File

canary:
  run mode: BATCH
  control type: INTERNAL
  control messenger: null
  driver files: null

# Enter the time step options below according to format of data

timing options:
  dynamic start-stop: off
  date-time format: mm/dd/yyyy HH:MM:SS
  date-time start: 01/01/2011 00:00:00
  date-time stop: 12/31/2011 23:55:00
  data interval: 00:05:00
  message interval: 00:00:01

# Enter the list of data sources below

data sources:
- id: hammond_in
  type: csv
  location: C:\Users\user\Documents\MyCANARY\research\data\hammond/hammond.csv
  enabled: yes
  timestep options:
    field: "TIME_STEP"

# Enter the list of signals and their defined evaluation type below

# evaluation type wq parameters are used for detection. op are not used for detection. cal signals are used for calibration and not detection

signals:
- id: TEST_TEMP
  SCADA tag: TEMP_VAL
  evaluation type: op
  parameter type: TEMP
  ignore changes: none
  data options:
    precision: 0.01
    units: 'deg_C'
    valid range: [\(-\inf, \inf\)]
    set points: [\(-\inf, \inf\)]
- id: TEST_TURB
  SCADA tag: TURB_VAL
  evaluation type: op
  parameter type: TURB
  ignore changes: none
  data options:
    precision: 0.1
    units: 'NTU'
    valid range: [-.inf, .inf]
    set points: [-.inf, .inf]

- id: TEST_COND
  SCADA tag: COND_VAL
  evaluation type: wq
  parameter type: SP_COND
  ignore changes: none
  data options:
    precision: 0.001
    units: '{\mu S/cm}'
    valid range: [-.inf, .inf]
    set points: [-.inf, .inf]

- id: TEST_PH
  SCADA tag: PH_VAL
  evaluation type: wq
  parameter type: PH
  ignore changes: none
  data options:
    precision: 0.01
    units: 'pH'
    valid range: [-.inf, .inf]
    set points: [-.inf, .inf]

- id: TEST_DO
  SCADA tag: DO_VAL
  evaluation type: wq
  parameter type: DO
  ignore changes: none
  data options:
    precision: 0.01
    units: 'mg/L'
    valid range: [-.inf, .inf]
    set points: [-.inf, .inf]

# Enter the list of event detection algorithms below
algorithms:
- id: test_LPCF
type: LPCF
history window: 576
outlier threshold: 0.5
event threshold: 0.90
event timeout: 12
event window save: 1
BED:
  window: 12
  outlier probability: 0.5

- id: test_MVNN
type: MVNN
history window: 576
outlier threshold: 0.5
event threshold: 0.90
event timeout: 12
event window save: 1
BED:
  window: 12
  outlier probability: 0.5

# Enter the list of monitoring stations below
monitoring stations:
- id: hammond_SRBC
  station id number:
  station tag name: hammond
  location id number: -1
  enabled: yes
  inputs:
    - id: hammond_in
  outputs:
  signals:
    - id: TEST_TEMP
    - id: TEST_TURB
    - id: TEST_COND
    - id: TEST_PH
    - id: TEST_DO
      cluster: no
  algorithms:
    - id: test_LPCF
    - id: test_MVNN