**Naval Research Laboratory**
Washington, DC 20375-5320

U.S. NAVAL RESEARCH LABORATORY

**NRL/6920/MR—2021/2**

# Enhanced Detection Algorithm for Navy Relevant Chemical Sensing

ANTHONY P. MALANOSKI

*Laboratory for Biomaterials and Systems*
*Center for Biomolecular Science and Engineering*

SCOTT N. DEAN

JEROME E. ALVAREZ

BRANDY J. WHITE

DAN ZABETAKIS

*Laboratory for Bio/Nano Science and Technology*
*Center for Biomolecular Science and Engineering*

JEFF S. ERICKSON

*Laboratory for Molecular Interfaces*
*Center for Biomolecular Science and Engineering*

December 15, 2021

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 15-12-2021 | NRL Memorandum Report | 06/01/2021 – 09/30/2021 |

**4. TITLE AND SUBTITLE**

Enhanced Detection Algorithm for Navy Relevant Chemical Sensing

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**
NISE

**6. AUTHOR(S)**

Anthony P. Malanoski, Scott N. Dean, Jerome E. Alvarez, Brandy J. White, Jeff S. Erickson, and Dan Zabetakis

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**
1X51

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

**8. PERFORMING ORGANIZATION REPORT NUMBER**

NRL/6920/MR--2021/2

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

**10. SPONSOR / MONITOR'S ACRONYM(S)**

NRL-NISE

**11. SPONSOR / MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

**DISTRIBUTION STATEMENT A:** Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

This effort describes testing of alternate detection algorithms including evaluating the applicability of deep learning methods such as the application of Multivariate Long Short-Term Memory – Fully Convolutional Network (MLSTM-FCN) approaches. The work is specifically intended to generate additional capabilities for these prototype devices allowing for their utilization in providing enhanced monitoring of chemical threats to the health of personnel in confined spaces both during a critical exposure event and as a result of long duration low level exposures. It was determined that deep learning methods are not appropriate for the type of data collected. The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was found to have equivalent accuracy to the already developed algorithm.

**15. SUBJECT TERMS**

Sensors
Detection algorithm
Threat detection

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| | | | | | Anthony P. Malanoski |
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | U | 16 | **19b. TELEPHONE NUMBER** *(include area code)* |
| U | U | U | | | (202) 404-5432 |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

This page intentionally left blank.

## CONTENTS

## FIGURES



## TABLES

# EXECUTIVE SUMMARY

The Center for Bio/Molecular Science and Engineering at the Naval Research Laboratory (NRL) initiated a program in May 2021 to expand the capabilities of NRL-developed chemical and biological monitoring approaches and prototypes. This specific effort describes testing of alternate detection algorithms including evaluating the applicability of deep learning methods such as the application of Multivariate Long Short-Term Memory – Fully Convolutional Network (MLSTM-FCN) approaches. The work is specifically intended to generate additional capabilities for these prototype devices allowing for their utilization in providing enhanced monitoring of chemical threats to the health of personnel in confined spaces both during a critical exposure event and as a result of long duration low level exposures

This page intentionally left blank.

# ENHANCED DETECTION ALGORITHM FOR NAVY RELEVANT CHEMICAL SENSING

## INTRODUCTION

This effort seeks to develop new detection algorithms intended to accelerate the crossover application of an NRL developed chemical sensing device for use in air quality monitoring within confined spaces, such as those of naval vessels and facilities. The existing prototype device, the ABEAM, was developed for outdoor perimeter monitoring with a focus on changes in the content of chemical or biological compounds found in environmental air samples.[1-17] Targets utilized during development included both chemical and biological threats and those targets identified by the TIC/TIM Task Force (toxic industrial chemicals/ toxic industrial materials). The CONOPS for this original application imposes limitations on size, weight and power (SWAP) that would preclude the use of computationally intensive detection methodologies such as deep learning. In a confined space application, on the other hand, power needs are not limited to battery support for long durations. This opens a broader range of options to the application and provides the potential to expand the utility of the prototype devices. More specifically, the effort is designed keeping in mind shortfalls identified in current approaches to monitoring confined atmospheres. "U.S. Navy submarine atmosphere monitoring has remained largely unchanged since the 1970's" currently using a combination of mass spectrometry and IR spectrophotometry. This approach is costly to maintain and not flexible enough to meet needed future capabilities. Systems that minimize mechanical parts and maintenance while offering the capability for measurement of new targets as they become required are needed. For applications in these types of confined spaces, a compliant system including a display must occupy restricted space and allow for mounting in differing orientations.

Beyond space utilization and adaptability, it is desirable to provide continuous monitoring of the chemical composition of the air in a confined space. This offers the potential for earliest warning and minimization of negative effects due to leaks or long term exposure. Currently utilized systems do not provide this capability; however, the NRL ABEAM prototype devices were designed with precisely this type of application in mind. The devices incorporate an array of colorimetric indicators that are continuously interrogated. The combined response of the individual array elements can be used for indication of a change in chemical air composition, identification of the presence of a known threat, and classification of unknown changes as similar or dissimilar to known threat compounds. In other words, the devices provide the potential for identification of known threats and for anomaly detection. Achieving this potential, however, requires a different approach to that previously taken in development of detection algorithms.

The NRL ABEAM prototype device utilizes a unique type of colorimetric sensing approach to enable capabilities not offered by traditional approaches. Other colorimetric sensing approaches rely on image analysis prior to and following exposure of a coupon using the differences between the images for target identification. This is a one-off type of coupon use analogous to the use of M8 Chemical Detection Paper without the option for visual inspection. Other types of traditional approaches include handheld and similar devices that require user intervention for discrete sampling, devices such as those incorporated into the Central Atmosphere Monitor System (CAMS) and those types of chemical sensing that require collection of samples at the point of interest and return to a central laboratory. Each of these types of approaches involved collection of a sample followed by processing time that depends on the technique utilized.

Manuscript approved December 13, 2021.

The coupons of the ABEAM prototype have been developed for continuous utilization over periods of weeks to months.[12, 16] The devices were developed in response to the introduction of the distributed microsensor paradigm. The idea is to develop inexpensive sensors that are small, low power, and autonomous that would be considered inexpensive enough to be disposable and would be suitable to unattended operation. The concept of use involves a population of sensors distributed across an area of interest. While no single device would match the sensitivity of central laboratory equipment, at least one of the sensors would be expected to be closer to the point of a chemical release than manned personnel or central detection facilities. As such, the exposure concentration would be proportionally larger and would compensate for the loss of sensitivity. Furthermore, with sensors spread out over an area, spatial and temporal information can be obtained regarding the movement of the chemicals throughout the area of interest. A demonstration of this function was completed in October 2019.[3]

**Figure 1.** Images of the ABEAM prototype (left) and the prototype during field trials (right). The device includes fifteen surface-mounted RGB color sensors, eight cool white LEDs, and three printed circuit boards. This prototype currently supports utilization as an autonomous device for post-collection data analysis or through wireless tethering to a laptop for real-time detection.

Current and ongoing efforts have focused on operational usage where limitations to power would be of significant concern. Software and algorithm development to date, as a result, has been focused on methods that can be used with low power devices, reducing the computational load as much as possible. [1, 16] The resulting approaches use three tiers. The first two are incorporated on the sensor devices: an initial data pass that evaluates and discards any electronic noise that could contribute to false responses and initial event detection algorithm that utilizes sliding windows and knowledge of the physics of the detection events to require the minimum amount of onboard memory and number of computations. This approach also allows for minimization of network communications except during event detection. A centralized control node is required for the third tier of the detection algorithm. This tier uses more computational capabilities and would be used to resolve the identification of the change in chemical composition.

This approach overall helped to achieve the lowest per unit cost for the sensor devices. Early prototype versions of the sensor devices have been utilized for collection of more than 30,000

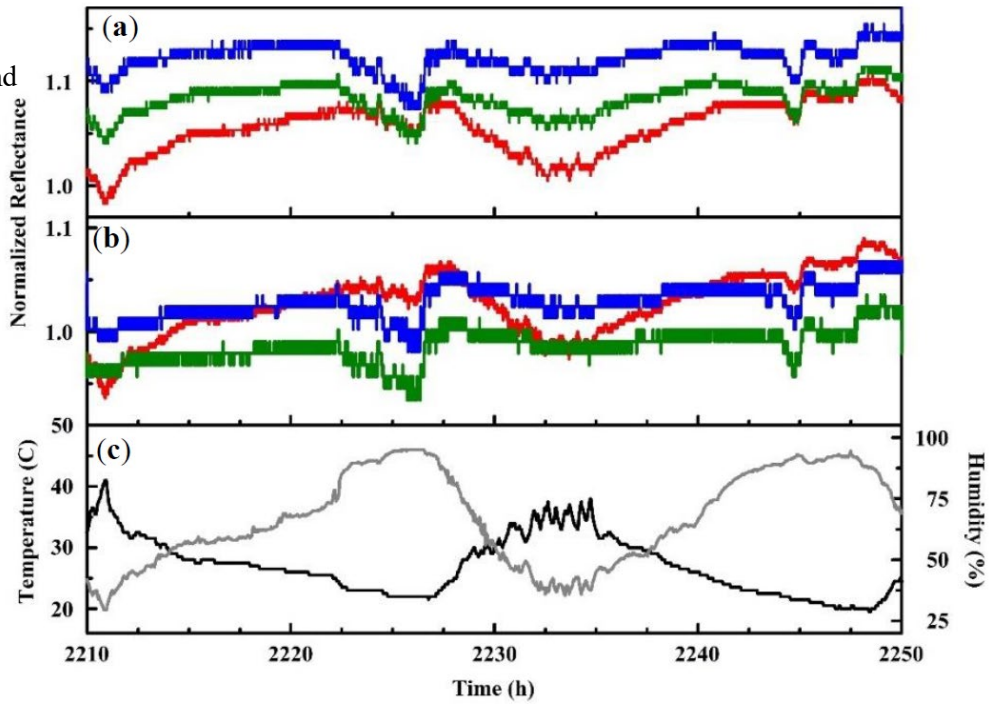


**Figure 2.** Red, green, and blue (RGB) time dependent profiles for data collected in an outdoor environment: indicator Zn N4TPP (a) and Ag DIX (b). The temperature (black) and humidity (gray) as reported by a co-located device are also provided (c).

hours of data in both indoor and outdoor environments on the NRL campus.[12, 16] Figure 2 shows representative data from sensors deployed outdoors. The devices have been independently evaluated by CCDC-CBC against both nerve agents and simulant compounds at a range of concentrations,[4, 5] and they have been utilized in field trials in collaboration with National Center for Spectator Sports Safety and Security (NCS4).[3] This work included exposures to many different chemicals under lab conditions as well as exposures to a limited number of approved for use chemicals in outdoor environments. This represents a large and diverse training set to be used with the selected deep learning technique.

The developed sliding window detection algorithm was selected to meet constrained SWAP requirements and during initial development of the ABEAM platform there was not time to explore other methods. With the SWAP constraints relaxed, a number of other algorithms can be considered for use. Many examples of algorithms that incorporate deep learning and general machine learning methods have been developed for other applications and could potentially be adapted for this application. In measuring a relationship between a sensor measurement and sensing target, one could employ various statistical-based models such as basic regression-based techniques to include linear, multivariate, partial least squares, or random forests. In a classification-based machine learning model, one application could be implemented as identifying and separating sensory data with discrete class labels.[18] In a model training scenario, its development is comprised of two phases: data preprocessing and model training. Application-specific sensor data are usually compiled into unprocessed raw datasets and then processed by standardization, elimination or imputation of null values, and encoding class labels. Additionally, feature extraction is an important pre-processing step before applying the pattern recognition methods since the extracted features could help characterize a system from a user's perspective.[19] At the end-process of model training, the resulting output model's performance is evaluated against a set of validation and test datasets.

## METHODS

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test or peak detection are applied to the set of recorded values on a given window. The steps of the algorithm processing can be represented as five distinct steps as shown in Figure 3 for the use of the KPSS test.

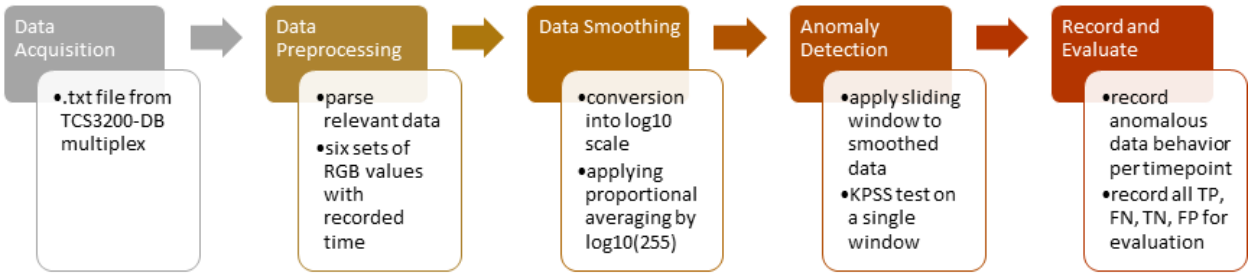


**Figure 3.** Schematic of the KPSS-based detection for ABEAM data.

### *1. Data preprocessing*

To fully utilize the different magnitude of RGB values in the data stream, it is a common practice to scale the features into a manageable range. This process is essential in algorithms that uses Euclidean or Manhattan distances between data points to assure their individual proportionality. Additionally, this step helps to suppress the effects of having extreme low and high magnitudes of a data point. Figure 4 shows the preliminary data for processing after scaling into a common logarithm of base 10.

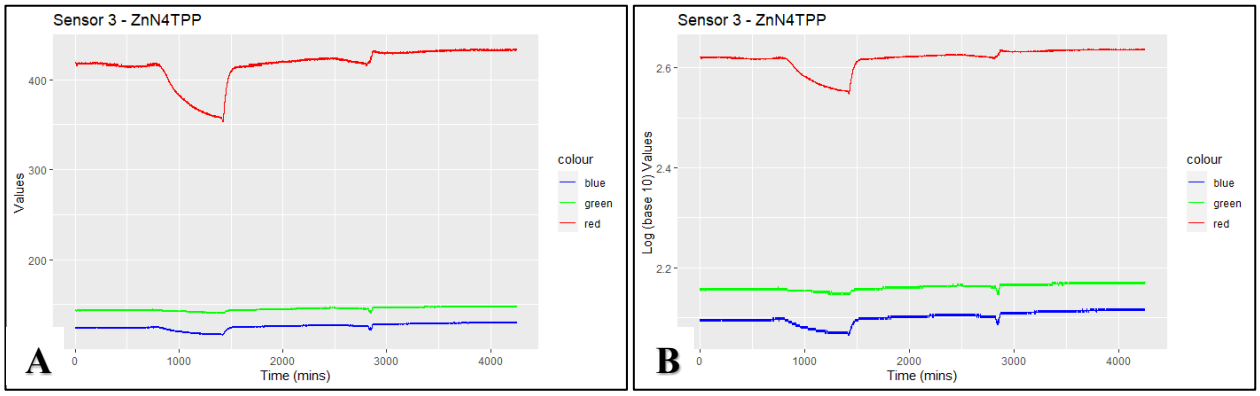


**Figure 4. A)** Sample raw RGB values from four sensor indicators. Values per reading (y-axis) exhibit varying magnitudes depending on sensor indicator used. Original RGB values were recorded every 30-second interval which are condensed into time in minutes. **B)** Transformation of RGB values into $\log_{10}$ scale from four sensor indicators. All values per reading (y-axis) range from 2.0 to 2.6 after processing.

On the other hand, scaled RGB values need to be transformed again into a univariate time-series data to account for the collective contribution of RGB intensities received by the six indicators. Moreover, processing this data into a univariate data stream is a convenient setup of

statistical hypothesis testing for stationarity per sliding window. The chosen scale is the common logarithm of $log_{10}$ and the proportional averaging factor is 255. We apply this as,

$$x_i = \frac{\log_{10} R_i + \log_{10} G_i + \log_{10} B_i}{\log_{10}(255)}$$

Each RGB channel is a bit of information, and a bit is an 8 binary number string. Within a bit, 256 variations can be represented with each value either 0 or 1. The colors red, green, and blue are 8-bit each, hence the $2^8 = 256$ range. Each RGB color has a range of 0-255 with 0 considered as a value and has the same fundamental meaning as light intensity.[20]

*2. KPSS test*

This assessment pertains to the null hypothesis that an observable series is stationary around a deterministic trend wherein the series is expressed as the sum of deterministic trend, random walk, and stationary error. [21] The actual test is the Lagrange-Multiplier diagnostic test of the hypothesis that the random walk has zero variance which stems from a spectral regression framework more often used in econometric work.[22] It is suggested that the Lagrange multiplier statistic may be a particularly useful formulation for testing for model misspecification and its more computational-saving than considering likelihood ratios and Wald tests for analyzing spectral regression residuals.[23, 24]

First, we describe a continuous time-series data stream with elements $x_i$ at times $t_i$, with $i$ representing a specific point in time ($i \in \mathbb{N}$). For a sampling period within the series, we define the window length as,

$$\Delta T = t_{i+k} - t_i$$

where $\Delta T$ represents the length between timepoints $i$ and $i + k$, and with $k$ as the desired length of samples in a window.

In this specific sensor data, various combinations of RGB values are contained in elements $x_i$. Additionally, we can divide the collective red, green, and blue signals into sequential epochs or non-overlapping windows to account for possible non-stationarities in the chemically sensed recordings. Hence, for any time-series data stream with $n$ possible partitions containing elements $x_i$, it can be denoted into a set of multiple, non-overlapping window sets as,

$$W = \{x_p, x_{p+1}, \dots, x_{p+n-1}\}, (n \in \mathbb{N})$$

where $p$ is the set index until the last partition at $n - 1$.

The chemical sensor data consists of red, green, and blue (RGB) readings from six specific TCS3200-DB color sensing breakout board indicators. The text files produced by TCS3200-DB have data structures of six simultaneously recorded RGB readings registered every 30 seconds. Each element inside a sliding window, denoted by $x_i$, contains a set of RGB values, $x_i = \{R_i, G_i, B_i\}, (R, G, B \in \mathbb{R})$ through real-time, but with a varying difference in magnitude.

The overall process starts by data acquisition from the ABEAM box. The text files extracted from the device are preprocessed by parsing relevant data. And then data are normalized into a manageable set of values for denoising. Detection of anomalous readings by KPSS tests is conducted within the set of data on a given sliding window length. Finally, test results are evaluated by recording the anomalous behavior of a dataset with respect to their time of exposure. Figure 3 shows the summary of the overall process.

### *3. Peak detection technique*

Peak detection algorithms have been used in several studies that leverage a certain data's behavior such as epidemiological studies, exposure assessments, or tactile sensors. [21-23] Here, we explore the possibility of utilizing the data's z-scores. Processing it through similar measures, sensor data is converted to a smoothed z-score peak detection setup which uses a moving mean and standard deviation that will select the peaks that are outside a set threshold. The algorithm is based on dispersion: if a new data point is a specific number of standard deviations away from a moving mean, the algorithm classifies this point as a peak.[24]

## RESULTS

Deep learning (DL) technologies have been implemented in recent literature for the collection, analysis, and interpretation of a considerably large amount of information. For example, DL methods using convolutional neural networks have been used in medical image analysis and real-time human action recognition captured by a low-cost RGB camera. [18, 25] With a higher data throughput, the challenge of effective management in systematic labelling and processing of data collection could be achieved by computationally inexpensive object detection. The advantage of machine learning techniques is its adaptability to learn and automate the extraction of patterns and features from a given dataset. Overall, the versatility and robustness of machine learning techniques can easily adapt to various applications that have an existing dataset with properly allocated data structures.

In reviewing the data so far generated for the ABEAM system, it was determined that the data itself currently represents very sparse data with few features and in addition while many hours of negative data have been generated the number of annotated positive events are too few to adequately train deep learning methods. As more data is generated, it may be possible to apply such methods but currently it is not practical, and thus methods with lower data requirements were used in this study.

Two traditional machine learning methods, the KPSS test and peak detection, developed for different applications were evaluated for use with the ABEAM 6 sensor prototype data. The data output from the ABEAM sensor platform also contain metadata of exposure times which are used as a reference point for accuracy measures. Figure 5 depicts an example of these exposure times with three hours of assumed exposure over a 3-day period, and the data transformation by $log_{10}(255)$ proportional averaging with applied data segmentation.

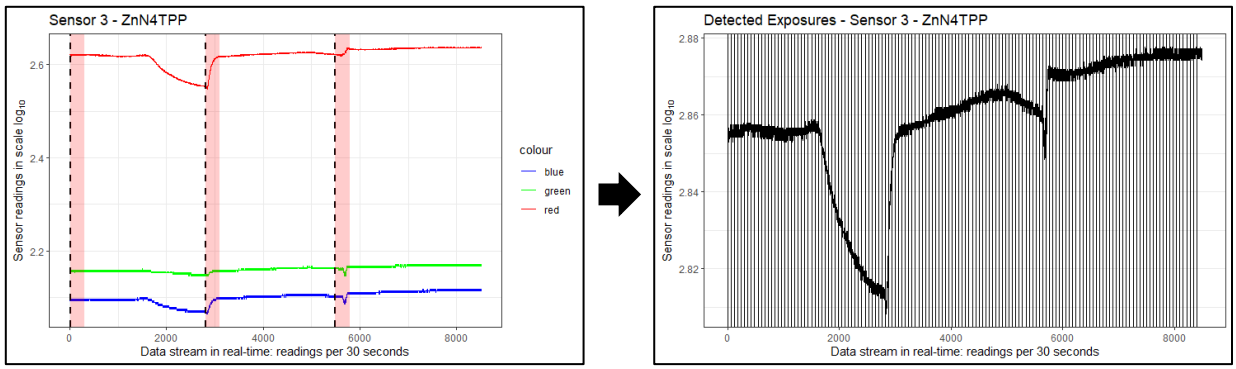


**Figure 5.** Sample exposure times from a specific TCS3200-DB sensor (Sensor 3 – ZnN4TPP). The exposure times are represented by vertical dashed lines with the assumed 3-hour exposure duration to chemical represented by red shade (left). RGB readings are transformed into $log_{10}$ scale which are then averaged by $log_{10}(255)$ to render as a univariate time-series data (right).

In testing either method, all data output from six indicators are processed and tested successively in multiple looping functions as cascaded algorithms. The application of a non-overlapping sliding window is successively applied after the six data streams have been transformed into a univariate time-series data by 60-point increments or 30-minute segments. Within a 30-minute window, KPSS stationarity testing or peak detection method is applied. Figure 6 illustrates this procedure using the KPSS test on one specific sensor.

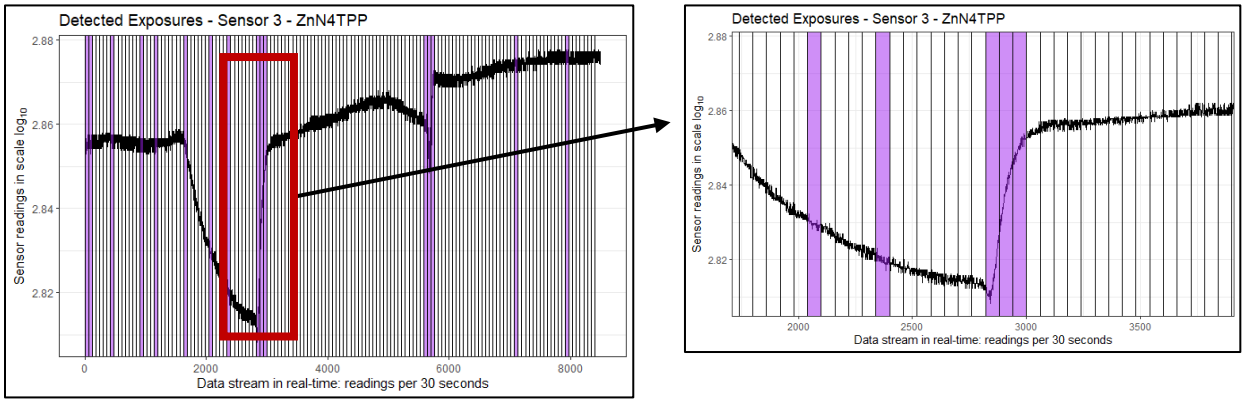


**Figure 6.** Detected exposures by KPSS stationarity testing (sample: Sensor 3 – ZnN4TPP). The transformed data partitioned by 30-minute window increments were subjected to KPSS testing per test window (left). Purple highlights depict the detected exposures as an occurrence of an event. The zoomed window (right) provides a closer look of the detected exposures between 5400 to 6000 seconds of recorded data.

In reviewing the results from all data gathered for the sensor platform, depending on the results of a single sensor seat yields a significant number of false positives. In further evaluation and fine-tuning from the readings from a sensor platform, using the results of more than one type of porphyrin-induced sensor is equally applicable to depending on the results from a single sensor material. The reason for this is that several of the different porphyrin materials may respond to the different exposure chemicals. The data was also analyzed using a 30-minute sliding window with recursive test readings where it was required that more than three sensors detect the same event. Moreover, this is to consider the performance of the six TCS3200-DB breakout boards as a whole. Since there are slight differences in the behavior of the data per different sensor, the results of the

testing are further evaluated to account for the robustness and ability of all the six sensors to detect a sample. Figure 7 shows the summary of all the six sensors combined with a total run time of 2-3 days.

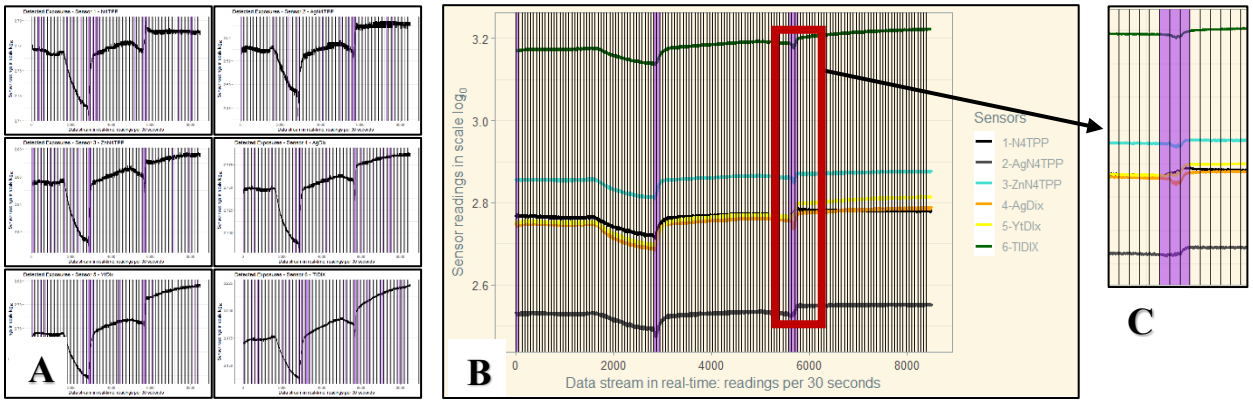


**Figure 7. A)** Individual readings of six TCS3200-DB sensors with purple highlighted windows as detected events. **B)** Aggregate readings of all six TCS3200-DB sensors with >3 sensors detecting an event. **C)** Zoomed-in detected readings between 5400-6000 showing varying fluctuations of recorded exposures.

## 4. *Effects of sliding window sizes in testing accuracy*

Several techniques for data segmentation that divide sensor signals into manageable smaller parts are widely implemented for research purposes. [26-28] This includes the sliding window technique and selecting the window length directly affects the amount of information that could be extracted from a series of data points. There also exists variations of this technique using 50% overlap between adjacent windows to avoid information loss.[29]

In exploring the effects of various sliding window sizes, Table 1 shows the accuracy measurements with multiple sliding window sizes applied on all six TCS3200-DB sensor indicators.

**Table 1.** Accuracy measurements from varied sliding window sizes.

| | **Individual** | | **Aggregate** | |
|---|---|---|---|---|
| **Window Size** | **Sensitivity** | **Specificity** | **Sensitivity** | **Specificity** |
| **5 minutes** | 100% | 86.94% | 66.67% | 99.74% |
| **10 minutes** | 83.78% | 95.92% | 50% | 100% |
| **20 minutes** | 95.45% | 94.86% | 100% | 98.95% |
| **30 minutes** | 100% | 92.14% | 100% | 100% |
| **60 minutes** | 94.59% | 86.39% | 83.33% | 95.08% |

Using a relatively small window could increase recognition rate, but this also could negatively impact recognition performance due to a decreased amount of information extracted. Increased false positives results with a 5-minute window size could mean that the behavior of the univariate data stream was not fully understood due to the scarcity of information. As such, a compensation between performance and dormancy must be carefully considered during the algorithm design.

User-defined requirements should also be considered in the design that depend on either excellent classification performance regardless of speed or aimed to create speed-critical applications such as a continuous real-time data stream. Thus, understanding the effect of window length can help select the suitable window size during the algorithm design.

### *5. Peak detection technique*

There are three parameters that have to be set for a working algorithm: the 1) lag, which dictates how many data points will be used to recalculate the moving mean and standard deviation – which could be also inferred as the size of the sliding window; 2) threshold, which is how many standard deviations away from the mean a new data point has to be in order to be classified as a peak; and 3) influence determines the future behavior of signals on the algorithm's detection threshold.[24] The future signals have no influence based on a threshold that is calculated with a mean and standard deviation if the influence is set as 0 (i.e., assuming data stationarity). If this is the case, it is expected that the time series data has to return to the same average over the set lag parameter, depending to which signals can effectively influence the trend of the data. However, the overall complications of tuning the aforementioned parameters have exhibited varying results on the sensors' signals which are shown on Figure 8.

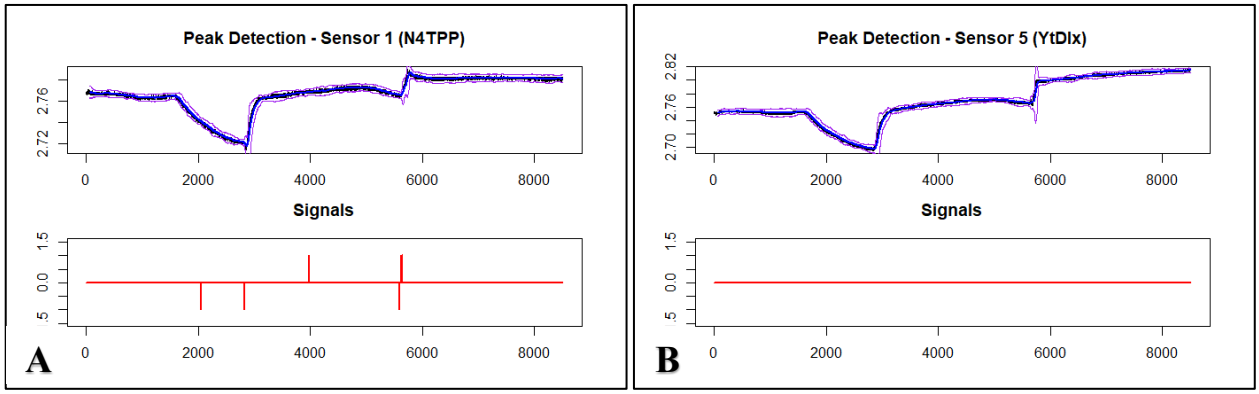


**Figure 8.** Sample peak detections – blue line as sensor data, purple lines as z-scores, and red line as signals. Parameters are set as: 1) lag – similar to the KPSS test sliding window of 30 minutes; 2) threshold – 3.5 standard deviations, and 3) influence – set on 0.5. **A)** Peak detection for Sensor 1 (N4TPP) – signal lines indicate readings close to assumed exposures. **B)** Peak detection for Sensor 5 (YtDIx) – failure detection of exposure. Due to varying results of peak detection techniques, further fine-tuning is required.

## CONCLUSIONS

Having explored several different machine learning methods, it has been confirmed that the initial sliding window slope method developed to meet SWAP considerations is as accurate as other statistical methods, such as the KPSS test. The KPSS test method has a reduced number of parameters relative to the original slope method resulting in a simple process to configure the algorithm for new conditions. Peak detection was not as successful in screening out false positives. The work on developing the new methods has even provided improvements in the methodology used to better tune the slope based method parameters. Deep learning remains an attractive possibility but requires more curated data before it can be used in this application.

**REFERENCES**


1. Malanoski, A.P., et al., *Method for Analysis of Data Related to Use of Reflectance Based Color Changes in Real Time Sensing Applications*. 2017.
2. Johnson, B.J., A.P. Malanoski, and J.S. Erickson, *Development of a Colorimetric Sensor for Autonomous, Networked, Real-Time Application.* Sensors, 2020. **20**(20).
3. Erickson, J.S., B.J. Johnson, and A.P. Malanoski, *Field Demonstration of a Distributed Microsensor Network for Chemical Detection.* Sensors (Basel), 2020. **20**(18).
4. White, B.J., et al., *Reflectance-Based Sensing: Aerosol versus Vapor Targets*. 2019: US Naval Research Laboratory, Washington, DC.
5. White, B.J., et al., *Reflectance-Based Sensing: Post-Evaluation Analysis of Sensor Responses*. 2019: US Naval Research Laboratory, Washington, DC.
6. Malanoski, A.P., et al., *Multiplexed, Optical Reflectance Data in Chemical Detection.* 2019: p. 1-4.
7. White, B.J., et al., *Colorimetric Environmental Sensor: Aqueous Indicator Screening (Part 1)*. 2018: US Naval Research Laboratory, Washington, DC.
8. White, B.J., et al., *Colorimetric Environmental Sensor: Aqueous Indicator Screening (Part 2)*. 2018: US Naval Research Laboratory, Washington, DC.
9. White, B.J., et al., *Colorimetric Environmental Sensor: Aqueous Indicator Screening (Part 3)*. 2018: US Naval Research Laboratory, Washington, DC.
10. White, B.J., et al., *Colorimetric Biosensor: Crosslinker Variations*. 2018: US Naval Research Laboratory, Washington, DC.
11. White, B.J., et al., *Colorimetric Biosensor: Porphyrin Variations*. 2018: US Naval Research Laboratory, Washington, DC.
12. Johnson, B.J., et al., *Reflectance-based detection for long term environmental monitoring.* Heliyon, 2017. **3**: p. e00312.
13. Erickson, J.S., et al., *Practical Implementation of Detection Algorithm for Reflectance-Based, Real-Time Sensing*. 2018: US Naval Research Laboratory, Washington, DC.
14. Johnson, B.J., et al., *Reflectance-based detection of oxidizers in ambient air.* Sensors and Actuators B-Chemical, 2016. **227**: p. 399-402.
15. Johnson, B.J., et al., *Porphyrin-modified antimicrobial peptide indicators for detection of bacteria.* Sensing and Bio-Sensing Research, 2016. **8**: p. 1-7.
16. Malanoski, A.P., et al., *Development of a Detection Algorithm for Use with Reflectance-Based, Real-Time Chemical Sensing.* Sensors, 2016. **16**(11).
17. Johnson, B.J., et al., *Miniaturized reflectance devices for chemical sensing.* Measurement Science and Technology, 2014. **25**(9): p. 095101.
18. Ha, N., et al., *Machine Learning-Enabled Smart Sensor Systems.* Advanced Intelligent Systems, 2020. **2**(9): p. 2000063.
19. Reid Turner, C., et al., *A conceptual basis for feature engineering.* Journal of Systems and Software, 1999. **49**(1): p. 3-15.
20. Soda, Y. and E. Bakker, *Quantification of Colorimetric Data for Paper-Based Analytical Devices.* ACS Sensors, 2019. **4**(12): p. 3093-3101.
21. Committee, *Exposure Assessment in Epidemiologic Carcinogenicity Studies*. Board on Environmental Studies and Toxicology; Division on Earth and Life Sciences; National Research Council. 2014: National Academies Press (US).

22. Checkoway, H., et al., *Peak Exposures in Epidemiologic Studies and Cancer Risks: Considerations for Regulatory Risk Assessment.* Risk Analysis, 2019. **39**(7): p. 1441-1464.
23. Lima, B.M.R., et al., *Heart Rate Detection Using a Multimodal Tactile Sensor and a Z-score Based Peak Detection Algorithm.* CMBES Proceedings, 2019. **42**.
24. Brakel, J.-P., *Peak signal detection in realtime timeseries data - Smoothed Z-score Algorithm (peak detection with robust threshold).* Stack Overflow.
25. Lee, J. and B. Ahn, *Real-Time Human Action Recognition with a Low-Cost RGB Camera and Mobile Robot Platform.* Sensors (Basel, Switzerland), 2020. **20**(10): p. 2886.
26. Wang, G., et al., *Impact of Sliding Window Length in Indoor Human Motion Modes and Pose Pattern Recognition Based on Smartphone Sensors.* Sensors (Basel, Switzerland), 2018. **18**(6): p. 1965.
27. Gjoreski, M., et al., *How Accurately Can Your Wrist Device Recognize Daily Activities and Detect Falls?* Sensors, 2016. **16**(6).
28. Atallah, L., et al., *Sensor Positioning for Activity Recognition Using Wearable Accelerometers.* IEEE Transactions on Biomedical Circuits and Systems, 2011. **5**(4): p. 320-329.
29. Jiang, M., et al., *A method to deal with installation errors of wearable accelerometers for human activity recognition.* Physiological Measurement, 2011. **32**(3): p. 347-358.