SAND2018-5996C

Unsupervised Machine Learning for Nuclear Safeguards*

Nathan Shoman and Benjamin B. Cipiti

Sandia National Laboratories

P.O. Box 5800, MS 0747, Albuquerque, NM 87185-0747

Abstract

The International Atomic Energy Agency (IAEA) has expressed interest in eliminating the need for on-site analytical laboratories at nuclear facilities. The Department of Safeguards (SG) has set a priority to "explore data analysis methods to strengthen the synthesis and evaluation of information (e.g. optimal random verification schemes, nuclear material flow analysis, material balance evaluation, near real-time accountancy and process monitoring tools)." [1] The goal then is to develop a system of unattended measurements that can reduce the reliance on destructive analysis (DA) measurements, which, are expensive and time consuming. This unattended measurement system may be realized through breakthroughs in measurement technology or through use of advanced data algorithms. In this work the latter approach is considered and a brief introduction to using machine learning is provided with the goal of incorporating process monitoring data into the safeguards evaluation. This approach is particularly valuable for advanced nuclear facilities, such as pyrochemical processing plants, where destructive analysis may be difficult.

Introduction

The inclusion of process monitoring data may increase the confidence of the evaluation while reducing the required number of person hours. Process monitoring systems might include NDA measurements such as gamma and neutron measurements, scales for bulk masses, and tank level measurements. The use of these measurements by the IAEA has thus far been limited due to relatively high measurement uncertainty. The use of gamma spectroscopy, for example, can determine plutonium content, but with a 5-10% measurement uncertainty. Relying on such measurements alone would result in a significant reduction in diversion detection. The purpose of this paper is to introduce machine learning concepts that could be applied to multiple process monitoring measurements to develop a new safeguards approach. A safeguards implementation that does not rely on traditional metrics such as the uncertainty in material unaccounted for (MUF) is explored.

Background

The main goal of the IAEA SG group is to "deter the proliferation of nuclear weapons, by detecting early the misuse of nuclear material or technology, and providing credible assurances that States are honoring their safeguards obligations". The established method of accomplishing this goal in a bulk handling facility is to setup a material balance area around an area of the facility with a nuclear material flow. The inputs, outputs, and change in inventory are measured to calculate the MUF value. Statistical tests are then applied to these measurements and calculations to detect facility

* SAND2018-XXXXC, Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

misuse. These statistical tests use thresholds that are tuned to have a particular false alarm probability (FAP).

While the traditional safeguards approach is sensitive to facility misuse and meets the IAEA timeliness goals it requires a substantial amount of sampling. Transportation of samples from a facility to IAEA's analytical laboratory is problematic due to the large number of samples and significant amount of time required for transit. This has led to the construction of an on-site laboratory at Rokkasho reprocessing plant at a considerable operating cost to the IAEA. The ideal case is to have a facility safeguarded with unattended monitoring supplemented with a much smaller quantity of samples. This paper considers the implementation of machine learning algorithms to better leverage PM data.

Separations and Safeguards Performance Model

The results presented in this work are generated using data from the Separations and Safeguards Performance Model (SSPM) [2] that is developed and maintained at Sandia National Laboratories. The SSPM platform was created in Matlab Simulink that tracks elemental and isotopic material flows through various unit operations. Measurement blocks simulate material accountancy and process monitoring data and are used with traditional statistical tests or external safeguards method development. SSPM has also recently been integrated with the GADRAS (Gamma Detector Response and Analysis Software) [3] code to provide simulated gamma spectra at a variety of locations in the modeled facilities. Several SSPM models exist including several reprocessing models such as PUREX, UREX+, and pyrochemical processing in addition to enrichment and molten salt models. This work uses data from the pyrochemical model to develop and prototype machine learning algorithms for safeguards. Figure 1 shows the flowsheet of a generic pyrochemical reprocessing facility.

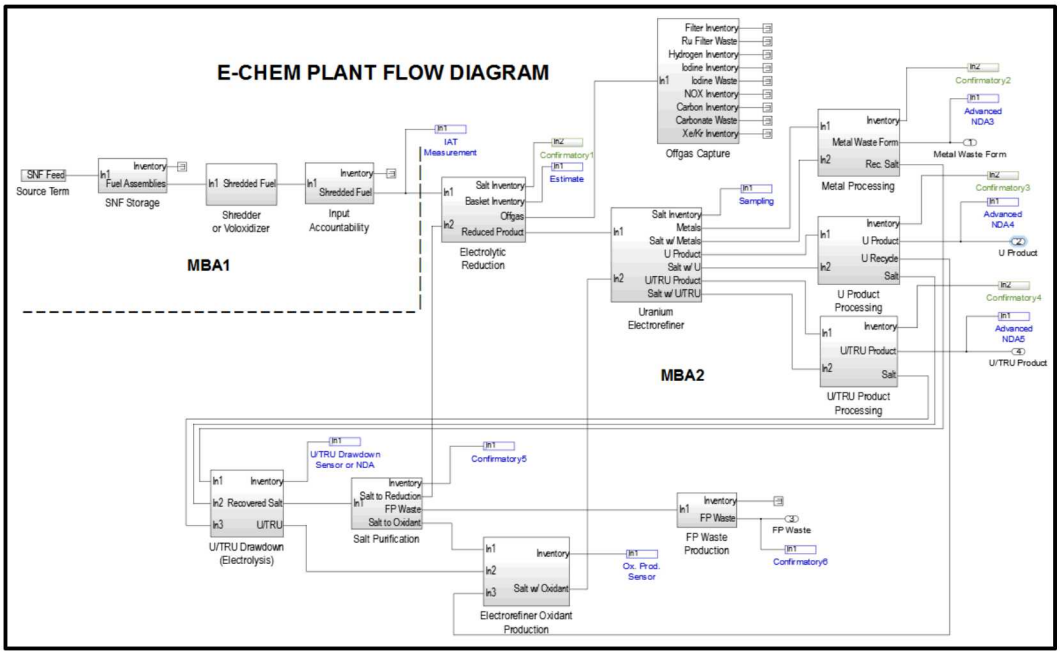

Figure 1: MATLAB Simulink pyrochemical reprocessing model

Machine Learning Introduction

Machine learning algorithms are a class of techniques that "learn" with data without being explicitly programmed. Many of the concepts have been around for several decades, but recent advances in computing capabilities and the rise of big data have led to their successful implementation. Perhaps the most well-known machine learning algorithm, the neural network, consists of layers of nodes that can collectively "learn" to categorize different types of input data. In the following section a brief introduction to machine learning is given to support ideas in this work as well as describe their limitations.

Supervised versus Unsupervised Machine Learning Techniques

The most popular machine learning techniques, the ones that power artificial intelligence systems and self-driving cars, are known as supervised techniques. These techniques require input labels or classification. For example, a convolution neural network (CNN) [4] trained to recognize objects in an image must be trained to recognize stop signs, cars, chairs, people and so on. This requires an input data set containing examples with labels of each possible output category. A Long-Short Term (LSTM) [5] neural network used for machine translation requires a large body of training data of specific labeled training data to map between languages. Most recent state-of-the-art advancements in machine learning have been in supervised techniques.

Unsupervised machine learning techniques, on the other hand, do not require labeled training data. Instead, these methods try to find clusters, groupings, and relationships through the structure of the data itself. This group of techniques includes clustering, manifold learning, mixture models, and outlier detection methods. The unsupervised problem is more difficult because additional human level knowledge is not provided to the algorithm, which is different from supervised methods in which images may be labeled cats or dogs. Figure 2 shows the ways different clustering methods try to find structure in unlabeled data.

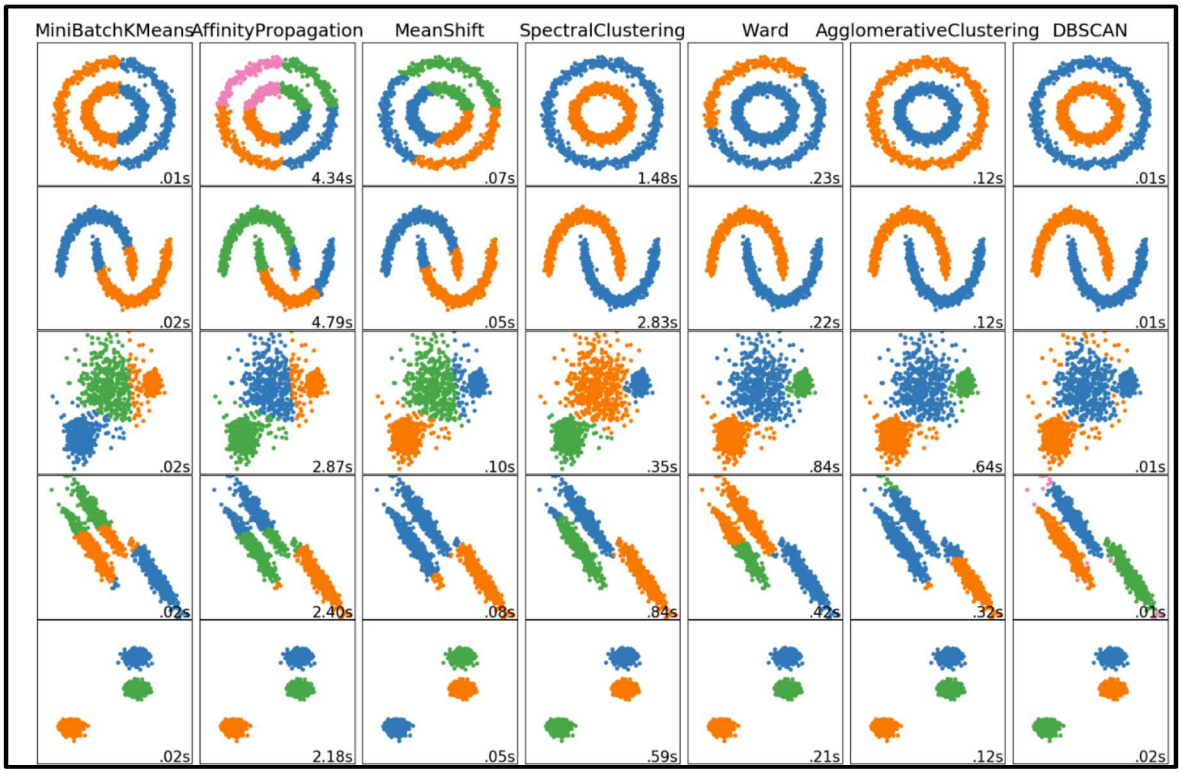

Figure 2: Classification of data from several unsupervised learning methods [6]

For the application of machine learning to facility level safeguards, unsupervised techniques must be used. Examples of all possible misuse cases must be provided during training in order for a supervised method to accurately detect those cases. However, with unsupervised methods only normal data and few, if any, misuse cases are used to train the models. The use of unsupervised methods eliminates the burden of identifying all possible diversion pathways.

One-Class Support Vector Machine

One-class support vector machines (OCSVM) [7] are traditional support vector machines (SVM) [8] that are trained in a one-class sense. To understand the OCSVM and how it enables anomaly detection at nuclear facilities, the traditional SVM must be described. Support vector machines are a type of supervised learning model that is known as a large margin classifier. Put simply, the objective of the SVM is to create a hyperplane that maximizes the margin of separation between two labeled classes. An example of a support vector machine is given in Figure 3 where the two classes only have two features so they can be represented in a two-dimensional space.

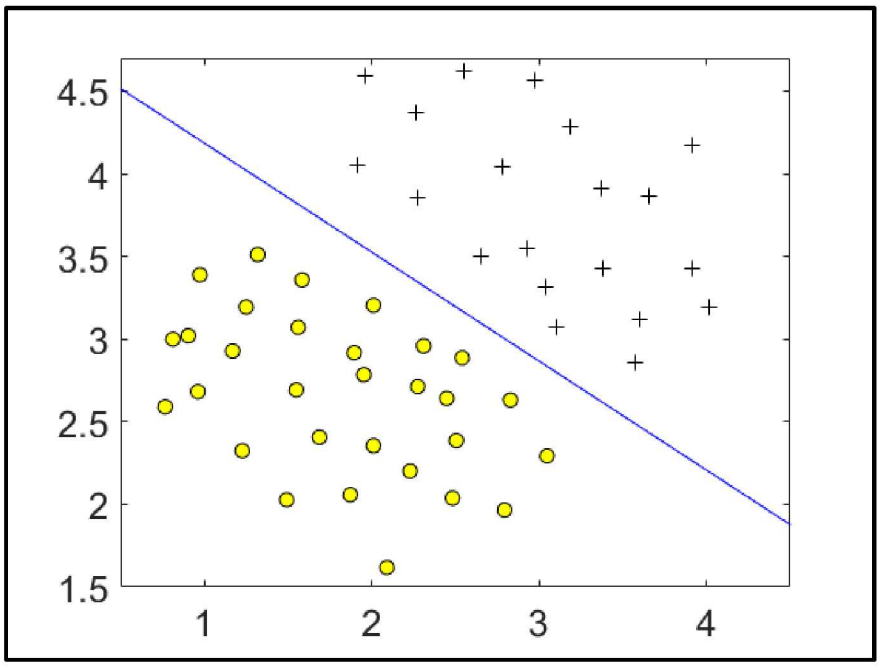

Figure 3: Margin separating two classes of data as determined by a SVM

The case presented in Figure 3 is trivial and could be performed using another method such as linear regression. However, the power of SVMs is derived from the ability to produce a non-linear boundary through a technique called the kernel method. In short, kernels map the input data from the given input space to a higher dimensional space. The motivation is that a set of data that is not linearly separable in the input dimensional space will be linearly separable in a higher dimensional

space. The hyperplane in the higher dimensional space can be represented as a polynomial in a lower dimensional space. A visualization of a support vector machine implemented with the kernel method can be seen in Figure 4.

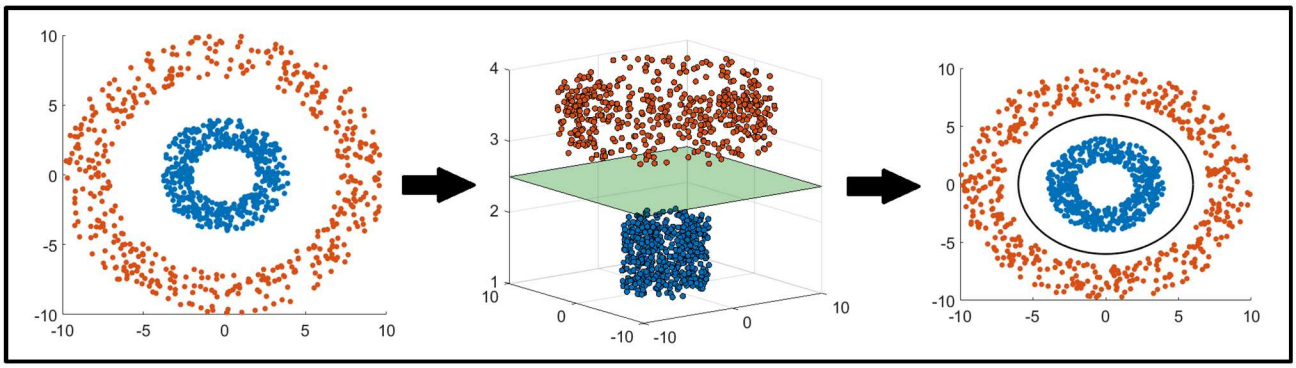

Figure 4: Nonlinear SVM margin determined using the kernel method

As noted previously, the support vector machine is a supervised method that requires knowledge of class membership. However, for this particular safeguards application an unsupervised method is needed. The SVM can be implemented in an unsupervised manner resulting in the OCSVM algorithm. Suppose data is collected with no knowledge of class membership. The SVM can be formulated in a one class sense such that the classifier discriminates between the fractions of the input data set. That is, the margin can be set to discriminate between γ and $(1-\gamma)$ of the input data. Assuming that there are few to no anomalies in the training data, the smaller the parameter γ is the higher the probability data outside the margin is an outlier or anomalous. Conversely, the larger the γ parameter is the higher the probability that the data outside the margin is normal.

New Safeguards Approach

The use of machine learning to classify abnormal operation at nuclear facilities require a different approach to safeguards. As mentioned previously, the OCSVM develops a boundary for classification for data. Once the OCSVM has been trained then new data can be divided into classes depending on the learned margin. The result is that instead of identifying diversion events with material based indicators such as MUF and sigma MUF, events are identified by their classification as shown in Figure 5.

The classifications must then be evaluated to determine if an off-normal event is occurring. In order to determine detection probabilities "windows" are constructed to evaluate the facility operation. For example, consider a window that consists of the classification of 100 observations where γ is set to 0.05. Since the OCSVM margin is set to discriminate between 5% and 95% of the data, it is expected that interval would contain about 5 measurements randomly distributed in the window that are flagged as anomalous. However, in a misuse case several sequential anomalous flags in a short period of time is expected. By changing the size of the window, the false alarm and detection probability can be adjusted. An example of how the OCSVM might be used is provided in Figure 5.

This method does not reveal any information as to the magnitude of a misuse scenario. This paradigm shift results in a safeguards system that does not provide a transparent materials

accountancy and only accounts for facility misuse. This aspect must be addressed in discussions with the IAEA in the future.

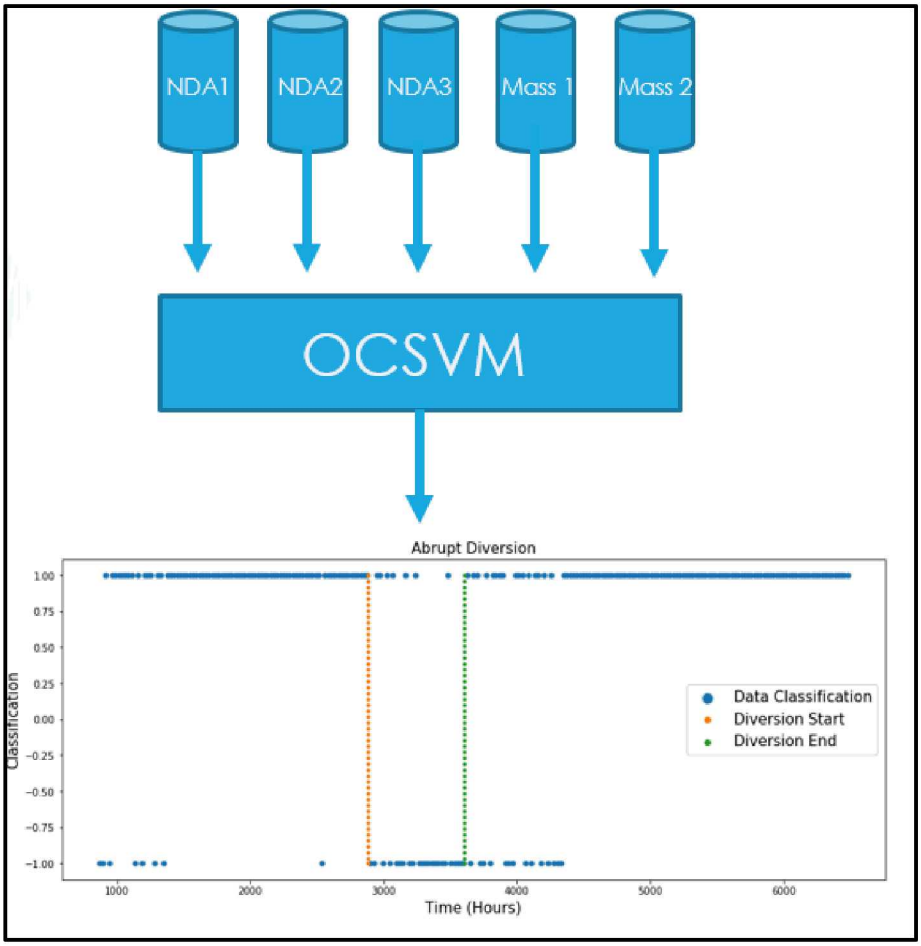

Figure 5: Graphical representation of a safeguards system using the OCSVM

Initial Results

For this preliminary evaluation of the OCSVM, two different measurement types are used; bulk masses and gamma spectroscopy measurements. Bulk mass measurements, which have around 0.1% measurement uncertainty, can be implemented in a cost-effective manner to easily detect direct material loss (either abrupt or protracted). These measurements are already implemented at Rokkahso as the Solution Measurement and Monitoring System (SMMS). The bulk measurements of inputs and outputs of the facility are used as direct inputs to the OCSVM. However, Page's test score on a particular process monitoring balance is used as an input to the model as well. The process monitoring balance is defined as the input minus the output and change in inventory. The resulting process monitoring balance is value that is centered around zero with some noise. The Page's trend test [9] is then applied to this balance. The use of Page's test as an input feature rather than the process monitoring balance itself will likely result in a better detection probability.

While bulk mass measurements can easily detect direct loss, other measurements must be used to detect substitution diversions. A substitution diversion is the removal of material while replacing it with an equal mass of a surrogate in order to fool the mass balance. Surrogate material is unlikely to

match the gamma emissions of the original material without incurring significant effort. NDA measurements, gamma emissions in particular, are often only used as confirmatory measurements due to the relatively high measurement uncertainty. NDA is used to quantify plutonium content though estimates of the transuranics and/or fission product signatures. In the machine learning approach, the plutonium inventory is not estimated with a NDA measurement, instead, the measurement itself is used directly and fed into the OCSVM. Specifically, certain channels relevant to the detection of substitution diversions are used as input features. The gamma spectroscopy measurements are generated using a simulated High Purity Germanium (HPGe) measurement in GADRAS. The code uses geometry and material composition coupled with radiation transport to generate a spectrum. The quantity of material used in the transport calculation was derived from the SSPM simulation with a measurement uncertainty of 5%. Poisson statistics were also simulated to represent the variations that occurs during a measurement.

For this work both protracted direct and substitution diversions are considered. Abrupt diversions are not included here as they are easier to detect than protracted.

Case 1: Normal Operation

This case models the normal operation of the facility. Figure 6 below shows the random misclassifications that occur during normal operation. Note that many of the classifications are 1, which is the label for normal. In this case γ was specified to be 0.10. This means for the training data 90% of the data is classified as normal while 10% is classified as off-normal. The model is trained on a set of 75000 normal operation observations. Then, the OCSVM is tested on a separate set of 5082 normal operation observations that were not used in training. Since it is expected that the measurement errors are random then the data classified as the -1 class will also be random. In the following sections it will be shown that during a diversion scenario there are clusters of off-normal classifications rather than random observations.

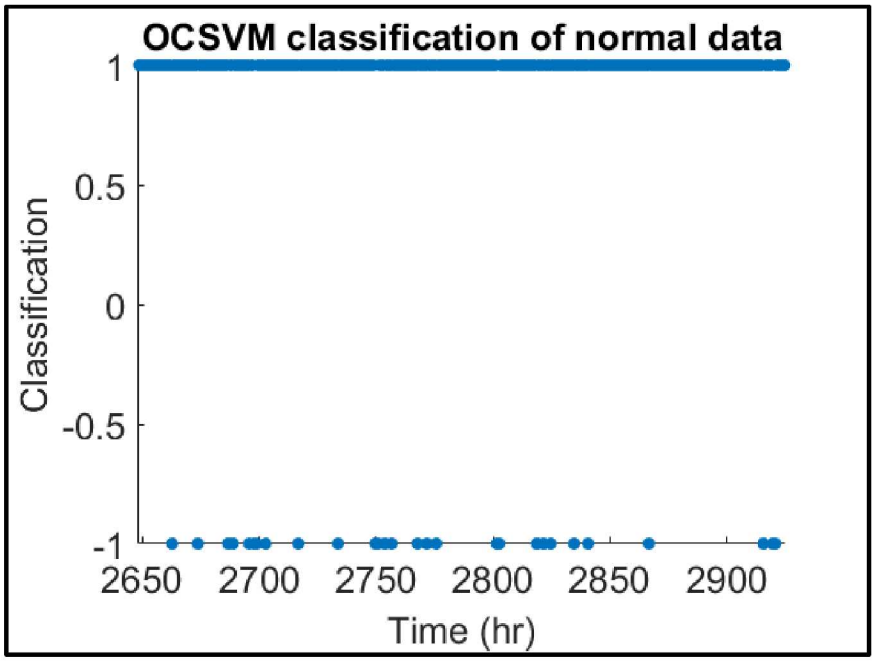

Figure 6: Classification of plant data during normal operation

Case 2: Protracted direct diversion

This case simulates the diversion of material from a unit process over more than one material balance period. This case is relatively simple to detect through the use of statistical trend tests, such as Page's trend test, on the level and mass measurements of unit operations. The advantage of using the Page's score over the tank level measurement directly is that the OCSVM doesn't have to be trained to handle patterns. It would be difficult to the OCSVM to detect diversions where the measurement could be normal. For example, assume a tank level measurement ranged from 10 at the beginning to the process to 7 at the end of the process. If a diversion occurred to reduce the tank level measurement at the beginning of the process from 10 to 8 the OCSVM would likely classify the data as normal as a level measurement of 8 is expected during normal operation. The OCSVM does not incorporate knowledge of cycles in the data such as tanks emptying and filling.

One important feature of using the Page's test in the OCSVM rather than standalone is that since the result of trend tests, such as Page's test, are used directly there is no need to tune multiple threshold criteria for the statistical test. The OCSVM accurately detects the abrupt diversion with the use of the Page's test score on the process monitoring balance. Note that the subsequent misclassifications after the diversion has ended reflects the continued off-normal conditions in the unit operation where the diversion occurred.

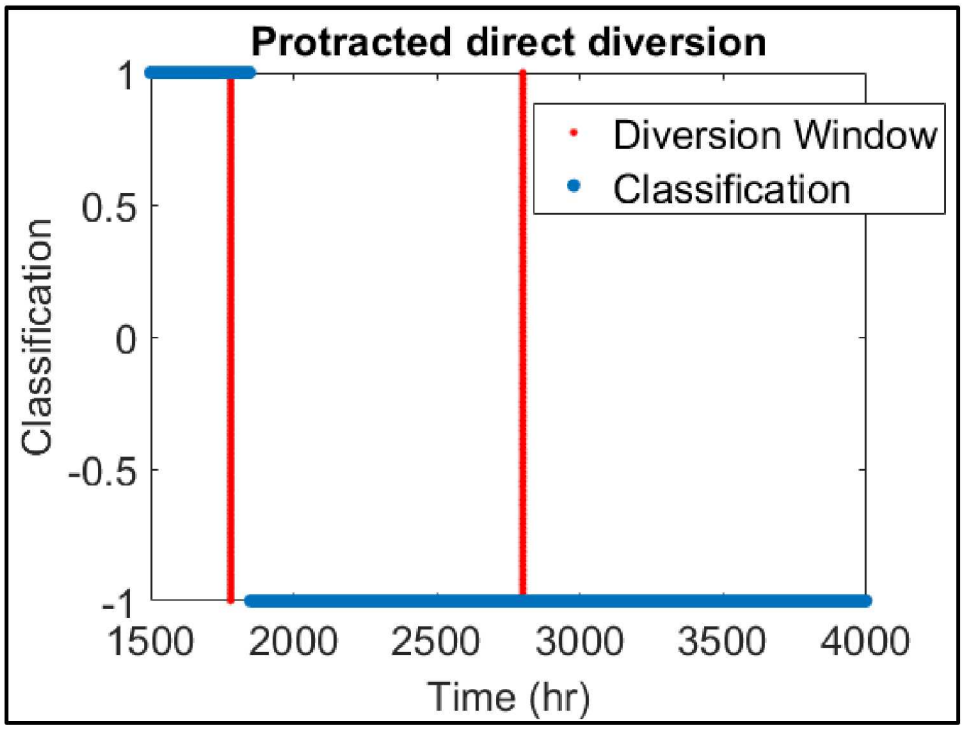

Figure 7: Classification of facility data for protracted direct diversion

Case 3: Protracted substitution diversion

The final case is the most difficult to detect. The mass of material in the unit process at the diversion location will not change as material that is removed is replaced by a surrogate of equal mass. Traditionally this diversion type was only detected through the use of destructive analysis measurements. In this method gamma spectroscopy measurements are used directly to detect these diversions. As mentioned previously, the OCSVM does not account for patterns in the data so care was used to select the right NDA metrics to use as input. In specific, certain gamma channels at specific times during the unit operation cycle were used as input features to the OCSVM.

The OCSVM is able to detect this type of diversion while only using unattended measurement methods. Note that the classifications are not off-normal for the entire diversion due to the characteristics of the unit process in which the diversion occurred. However, the absolute detection probability remains uncertain. The SSPM was run for 100 iterations for each scenario. Each iteration has a different sequence of input fuels to the facility along with different systematic measurement errors. Some data sets produced results that were classified as normal by the OCSVM. Further work is required to further quantify and improve the detection probability.

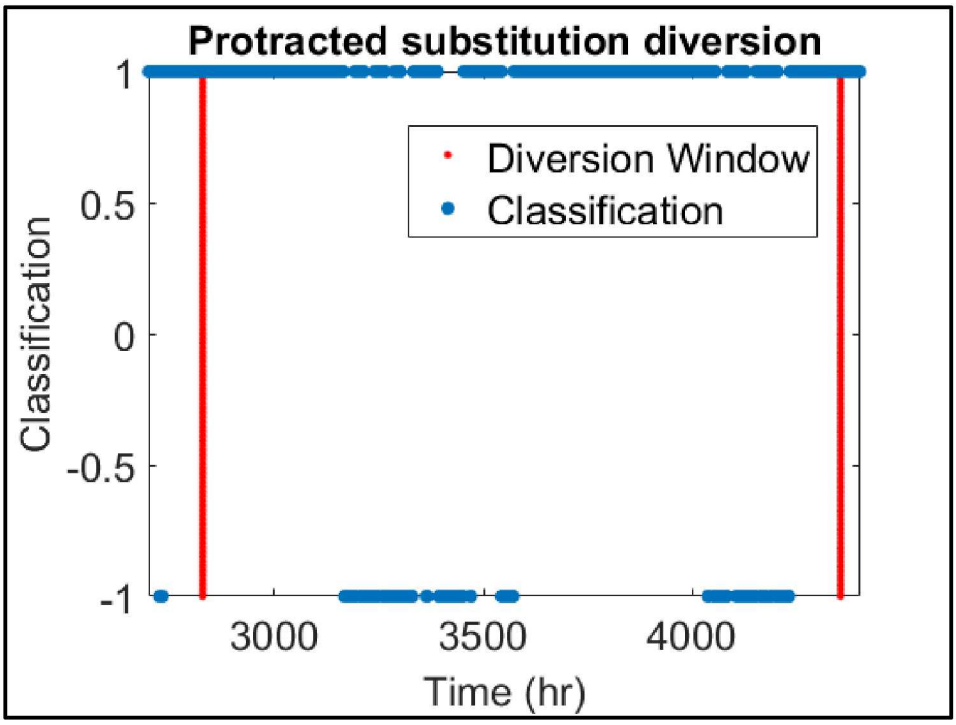

Figure 8: Classification of facility data for protracted substitution diversion

Future Work

While the preliminary results are encouraging there is still work to be done. First, the required amount of training data must be determined. The training in this work used 80,000 observations which is well beyond what could be collected in a typical facility before start-up. If a large quantity of training data is required then perhaps artificial model data could be used to pre-train. A parametric study will be conducted to determine how much training data is required.

Secondly, the "window" used to translate the classification of observations into detection probabilities must be calculated. The window can be adjusted in size to tune the detection and false alarm probabilities. It is possible that two OCSVMs with different margins will be more effective than a single OCSVM.

Finally, the features that are fed to the OCSVM must be evaluated. While results using the bulk masses, process monitoring data, and NDA data are promising perhaps different features could enhance the detection probabilities. As mentioned previously care was used to design inputs that capture features of the data while reducing the cyclical nature. For example, instead of using the level measurement directly, which is cyclical, Page's trend test can be used on the process monitoring balance across a unit operation to determine a score. This score has a range for normal operation and is much higher for abnormal operation. Similarly, the NDA metrics used have to be independent of both the randomized fuel input and the variations due to measurement uncertainty. The proper utilization of signals, perhaps input as ratios, is essential to the effectiveness of the OCSVM.

Acknowledgements

This paper summarizes recent work funded through the U.S. Department of Energy Office of Nuclear Energy and National Nuclear Security Administration.

References

1. International Atomic Energy Agency Office of Safeguards "Research and Development Plan – Enhancing Capabilities for Nuclear Verification" STR-385 (2018)

2. B. Cipiti, "Separations and Safeguards Performance Modeling for Advanced Reprocessing Facility Design", *Journal of Nuclear Materials Management,* **40**/3 pp.6-11 (2012).

3. S. Horne et al., "GADRAS-DRF 18.6 User's Manual" SAND2016-4345. Sandia National Laboratories (May 2016)

4. Y. LeCun et al., "Gradient-Based Learning Applied to Document Recognition" *Proc. of the IEEE* vol.86, no.11 pp. 2278-2324 (1998)

5. S. Hochreiter and J. Schmidhuber "Long Short-Term Memory" *Neural Computation* vol. 9, issue 8, p. 1735-1780 (1997)

6. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python" *Journal of Machine Learning Research* vol. 12, p.2825-2830 (2011)

7. B. Scholkopf et al., "Estimating the Support of a High-Dimensional Data Distribution" *Microsoft Technical Report* MSR-TR-99-87 (1999)

8. C. Cortes and V. Vapnik "Support-vector networks" *Machine Learning* vol. 20, issue 3, p.273-297 (1993)

9. E. Page "Ordered Hypotheses for Multiple Treatments: A significance Test for Linear Ranks" *Journal of the American Statistical Association* vol. 58, issue 301, p. 216-230 (1963)