Isotope Ratio Features for Classification of Dissolution Events Using Effluents Measurements

Nageswara S. V. Rao[†], Christopher Greulich[†], Michael P. Dion[†], Jason Hite[†], Ken J. Dayman[†], Andrew D. Nicholson[†], Daniel E. Archer[†], Michael J. Willis[†], Irakli Garishvili[†], Riley Hunley[†],Jared Johnson[†], Andrew J. Rowe[‡], Ian R. Stewart[$], James M. Ghawaly[†]
† Oak Ridge National Laboratory, ‡ Cadre5, $ University of Tennessee

Abstract

Gamma-ray measurements of effluents at the off-gas stack of a radiochemical processing facility are used to train a set of classifiers to identify events associated with radioactive material dissolution and processing. Datasets from two ^{238}Pu subsequent campaigns involving the dissolution of irradiated ^{237}Np targets of possibly different source levels are utilized. Gamma-ray count rates estimated through spectral analysis for isotopes of Iodine, Krypton, and Xenon, have been utilized as classifier features in the past. The cumulative distribution functions (cdf) of select isotopic count ratios during the dissolution of material and other periods show separability, indicating their potential use as classifier features. The isotopic ratios, when properly selected, have some inherent invariance and could be more stable indicators of dissolution source activity than (solely) count rate estimates of individual isotopes. Hence classifiers using them as features are expected to be less sensitive to count rates which may vary across dissolution campaigns especially given the difference in irradiation time and cooling time. Multiple classifiers are trained to detect the dissolution events using seventeen isotopic count rates. Additionally, classifiers are trained using six isotope ratios as features and the results of various classifiers are compared. Under 5-fold cross validation, the classification performance with isotope ratios closely matches the performance with count rates, for example, within 0.53% of the classification error for the best classifier and 0.03% for the classifier-fuser.

Primary Area: Nuclear Security & Physical Protection Technologies: Nuclear and Radiological Detection Technologies

1 Introduction

Inferring the occurrences of events related to irradiated material dissolution operations at a radiochemical processing facility testbed were investigated using features derived from a radiometric monitoring sensor system at the off-gas stack. A simplified version of on/off classification associated with the dissolution process of an irradiated ^{237}Np target used for ^{238}Pu isotope production has been studied in [1, 2], using gamma spectral measurements of effluents at the facility's off-gas stack. Count rate estimates within spectral regions of several volatile isotopes are used as the features to train the classifiers in this research. The production processes of the fission products are

governed by nuclear fission and radioactive decay. Therefore, in assuming the elemental transport through the facility into the off-gas stack is constant, their isotopic ratios are inherently more stable, in the context of learning theory, indicators of the source than just count rate measurements, particularly when the detectors are prone to calibration drifts and calibration error. Hence, classifiers that utilize isotopic ratios as features are expected to be less prone to over-fitting to source levels reflected in count rates. In this paper, we present a study utilizing volatile isotope ratios as features of these classifiers using datasets collected at the Radiochemical Engineering Development Center (REDC), Oak Ridge National Laboratory (ORNL) for two ^{238}Pu dissolution campaigns.

The dissolution process of irradiated material at the REDC consists of multiple events, including Al cladding removal by physical and chemical means, chemical dissolution of actinide material, and follow on activities that involve mixing, boil down and re-work, among other operations. Radioactive isotopes are released and transported to the off-gas stack during these events, which together constitute a complex time series of radioactive measurements. Their combined gamma spectrum is measured using a High Purity Germanium (HPGe) sensor located at the off-gas stack. The classification task is to use the radiometric measurements to infer the occurrence of dissolution events, which is complex due to multiple factors including: multiple primary and daughter isotopes with different half-life profiles; transport from the facility hot cells to the off-gas stack; sensor and software limits in generating count rate estimates in the gamma spectral regions; differences in events as the dissolution process is refined across campaigns; and other processes occurring at the facility.

Eight different classifiers, namely, Classification Trees (CT), Discriminant Analysis Classifier (DAC), Error Correcting Output Codes (ECOC), Ensemble of Trees (EOT), Kernel Method (KM), k Nearest Neighbors (kNN), Naive Bayes (NB), and Support Vector Machine (SVM), are trained first, and the top 3, namely, CT, EOT, and kNN, are fused by Ensemble of Trees Fuser (ETF), as in previous works [2]. As a net result of feature and classifier fusion, combined with the incorporation of decay chain and isotope half-life information, this approach achieves 99% detection with under 1% false alarm rate on the training set for kNN. This performance is mainly a result of the incorporation of physics information, namely, decay chains and half-lives, to fine-tune the ML solution by selecting all seventeen isotope count rates and 2-day measurement window. Count rates estimates of radionuclides, including isotopes of Iodine, Krypton, and Xenon, are utilized as features for these classifiers, and their outputs as used as features for the classifier-fuser ETF. However, the dissolution target amounts, irradiation time, and processing may differ among the campaigns resulting in varying amount of effluent releases and thus varying associated count rate estimates. The potential for these classifiers to over-fit to count rates motivated the study of classifier features that are less prone to such variations.

The cumulative distribution functions (cdf) of select isotopic count ratios during the dissolution and other periods show separability, indicating their potential to be effective as classifier features. Thus, classifiers using them as features are expected to be less sensitive to variations in count rates. The above set of classifiers are trained using six isotope ratios calculated from the count rates as features, and the results of various classifiers are compared. Overall, for 5-fold cross validation, the classification performance with isotope ratios closely matches the performance of count rate estimates. For example, based on 2-day time window for the measurements spanning two ^{238}Pu campaigns, the classification error for the best classifiers is within 0.53% (namely, 0.1% of kNN using count rates and 0.63% of EOT using ratios), and within 0.03% for ETF in both cases.

The organization of this paper is as follows. A radiochemical facility and the dissolution process are briefly summarized in Section 2, and the HPGe sensor system and data processing are

briefly summarized in Section 3 (more details can be found in [2]). The measurements and estimates are described in Section 4; the count rate estimates and their ratios are discussed in Sections 4.1 and 4.2, respectively. The classifiers and fusers are described in Section 5; the classification methods are briefly outlined in Section 5.1, and their performance is compared in Section 5.2. A summary of our contributions and directions for future work are described in Section 6.

2 Radiochemical Dissolutions and Facility

A multi-modal sensor testbed has been established at ORNL under the Multi-Informatics for Nuclear Operations Scenarios (MINOS) project, by the Department of Energy's (DOE) National Nuclear Security Administration (NNSA). Its purpose includes the collection of sensor, parametric, and ground truth data sets to support the development of data analytics for nuclear nonproliferation scenarios. The testbed is deployed at the conjoined site of the 85 MW_{th} High Flux Isotope Reactor (HFIR) and REDC shown in Figure 1.The testbed consists of a variety of sensors that collect measurements of multiple modalities, including radiation, thermal imagery, seismic, infrasound, electromagnetic, and biological. For the specific task of inferring occurrences of events related to ^{238}Pu campaign operations at REDC addressed in this paper, the gamma spectra of radioactive volatile effluents are measured at the REDC off-gas stack using a HPGe sensor system.

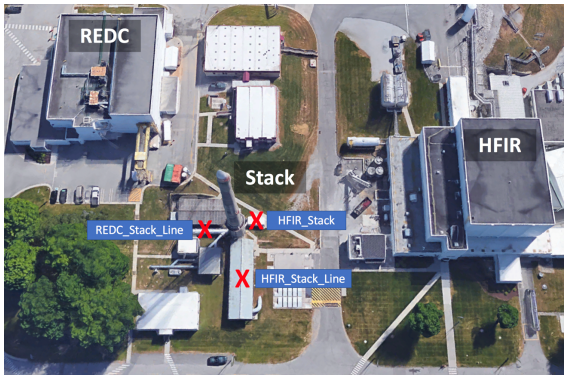

Figure 1: MINOS testbed deployed at co-location site of HFIR and REDC.

For the production of ^{238}Pu using neutron irradiation, ^{237}Np targets are prepared and utilized. The targets are purified from feed stock material and converted to the desired chemical form, and then fabricated into aluminum-clad rods at the REDC. Targets are then transferred to HFIR (co-located with REDC), where they are loaded into HFIR target locations and irradiated. The irradiated targets are transferred back to REDC for processing, where chemical separations are conducted in hot cells and/or glove boxes to separate and purify the product material. Specifically, ^{238}Pu is produced through neutron capture on targets of neptunium oxide in an aluminum matrix, $^{237}NpO_2$/Al, clad in an Al alloy through irradiation in HFIR's Beryllium reflector. The multistep capture and decay processes are shown in Equation (1).

$$^{237}\mathrm{Np} + {}^{1}_{0}\mathrm{n} \longrightarrow {}^{238}\mathrm{Np} \xrightarrow[2.1d]{\beta^-} {}^{238}\mathrm{Pu} \tag{1}$$

The irradiation of targets in the HFIR lead to a complex composition of radioactive isotopes upon discharge. Because fission is induced, the composition continues to change during cool down, transport, and processing in REDC due to radioactive decay.

3 Effluents and Measurements

Through the dissolution of irradiated material at REDC, fission products are released at different times and potentially transported to the off-gas stack in various ways. The gamma spectra measured at the off-gas stack reflect the dynamic transport of the radioactive fission product isotopes.

These spectra are measured using an ORTEC GEM series HPGe detector located at the REDC off-gas stack line.The detector is shielded with lead bricks to reduce the gamma background. Gamma-ray spectra are collected continuously and are processed using standard domain techniques. Further details of this measurement and processing system are provided in [3]. Seventeen individual gamma-ray peaks (corresponding to certain radionuclides) are targeted in the gamma-ray energy spectral analysis. From this analysis, the associated count rate of those gamma-rays are found. A peak fitting methodology, commonly applied in this domain space, is used on the time aggregated gamma-ray spectra and the resulting background subtracted count rates are used as input features for developing and testing classifiers and their fusers described in Section 5.

The target dissolutions give rise to a complex time series of source term $S(t), t \in [0, T]$, where T represents the time where the source term has decayed to zero, or more practically, below the minimum detectable threshold of the detector. The source term $S(t)$ can be represented by a series of coupled equations for each isotope produced with the individual evaluation in the form of Bateman Equation described in [1]. For each isotope, the source window $[0, T]$ is based on its half-life and those of its parent nuclides. For fission reactions that occur during irradiation at HFIR, the abundance of some nuclides reach a maximum, and some (progeny) isotope's concentrations continue to increase and reach a maximum after irradiation ceases. In addition, in the case of spontaneous fission, radionuclides are continuously produced and even short-lived isotopes can be present as long as there is sufficient activity of the isotope decaying by spontaneous fission (e.g., ^{252}Cf). The fission product isotopes are transported and measured in gas effluents at the off-gas stack during $[0, T + T_M]$. Thus, these measurements reflect a combination of operations involving a series of processes and systems in addition to the physics of decay chain kinematics. Additionally, the dissolution events affect the amount of material entering the off-gas system and result in complex profiles of time series $S(t)$. Examples of discrete events include mechanical penetration of targets and sparging and mixing of dissolution tanks during and post primary dissolution, which may result in varying source terms.

4 Classification Features

The gamma spectral measurements collected under MINOS project's continuous data collection campaign during May-June and Sept-Oct 2020 periods are considered here. They include two ^{238}Pu campaigns (i.e., P5 and P6), and also "buffer" periods before and after the campaigns, as shown in Figure 2 for ^{138}Xe count rates.

4.1 Gamma-ray Spectral Features

Seventeen gamma-ray peaks (corresponding to certain radionuclides) are targeted in the gamma-ray energy spectral analysis acquired by the HPGe system. These radionuclides were identified from subject matter expert analysis of peaks present following ^{252}Cf dissolution campaigns. Because of the spontaneous fission decay path, several short lived isotopes are present and easily detected because of their high specific activity. The corresponding isotopes that correlate with the gamma-ray peaks fit during the analysis are I-131, I-132, I-133, I-134 and I-135; five krypton isotopes, Kr-85, Kr-87, Kr-88, Kr-89 and Kr-90; five xenon isotopes, Xe-135, Xe-135m, Xe-137, Xe-138 and Xe-139; Ba-138; and Cs-139. Their count rates are found through analysis of the integrated gamma spectrum every hour by peak fitting or by a regions of interest method [3]. The Xe-138 count rates shown in Figure 2 exhibit high variations both during and outside ^{238}Pu dissolution periods. It should be noted that the largest count rates (smallest uncertainty) are ac-

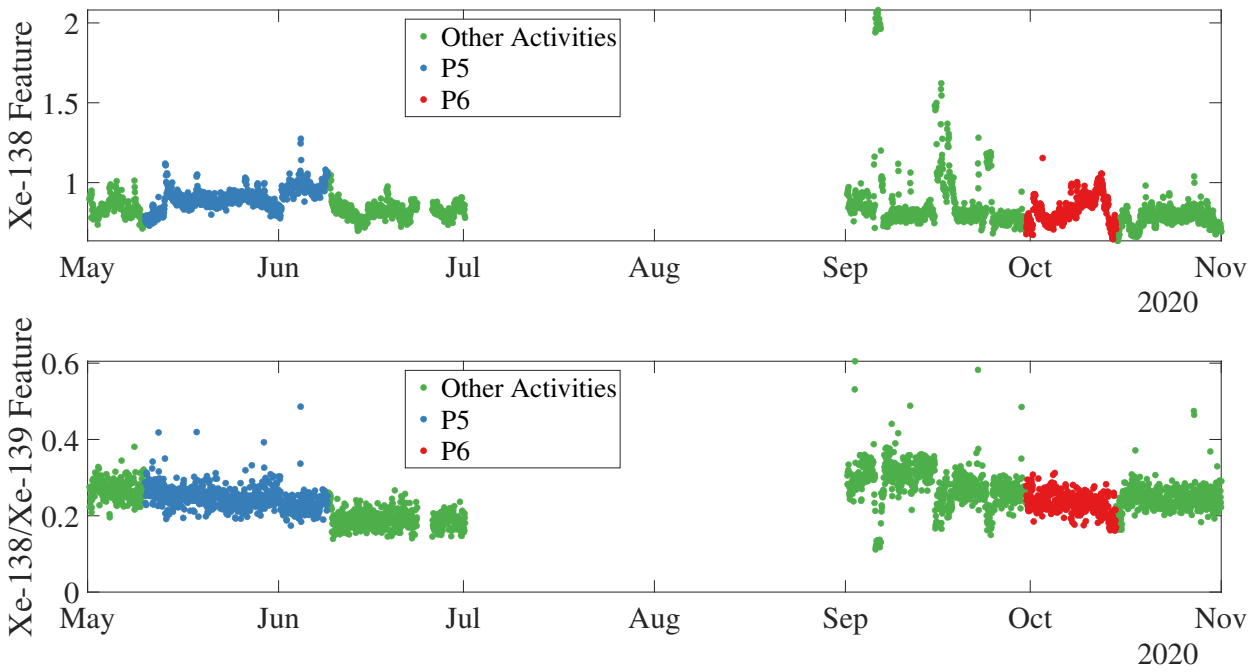

Figure 2: Gamma spectra are collected by a HPGe sensor located on the REDC stack line (a) Xe-138 count rates [cps], and (b) ratio of Xe-138 and Xe-139 count rates. Periods of ground truth ^{238}Pu dissolution events at REDC are indicated by the red regions.

tually observed outside of the P6 campaign, an indication that other activity in REDC could be responsible for these signals.

4.2 Isotope Ratios and Ratio-Ratio Regions

The cumulative distribution functions (cdf) of isotope count rates during the dissolution and non-dissolution periods indicate certain separability, as shown in Figure 3(a) for a select Xe isotope and one of its ratios. It is an indicator of their potential use as classifier features, although classifiers may not explicitly exploit this separability concept. For the ratios of count rates, the corresponding cdf plots, one of which is shown in Figure 3(b), also show a separability, which is a first indication of their potential use as classifier features.

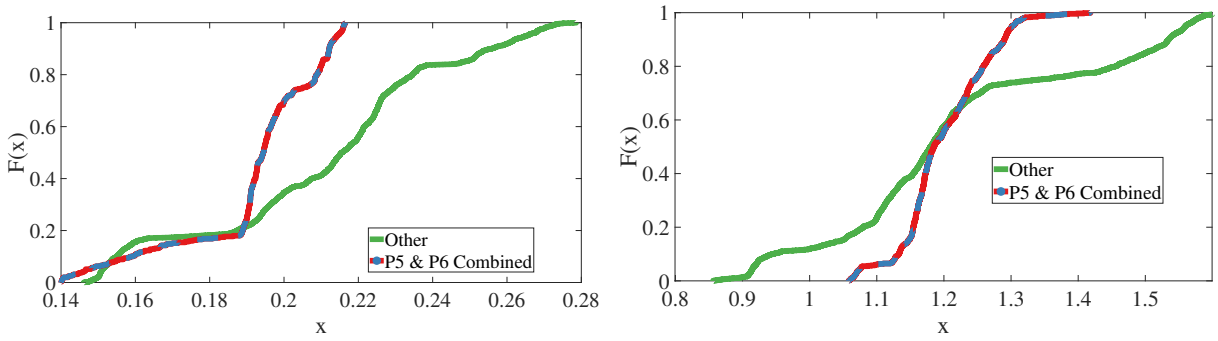

Figure 3: Cumulative distribution function of count rates and ratios (a) Xe-135m and (b) Xe-135m/Xe-135.

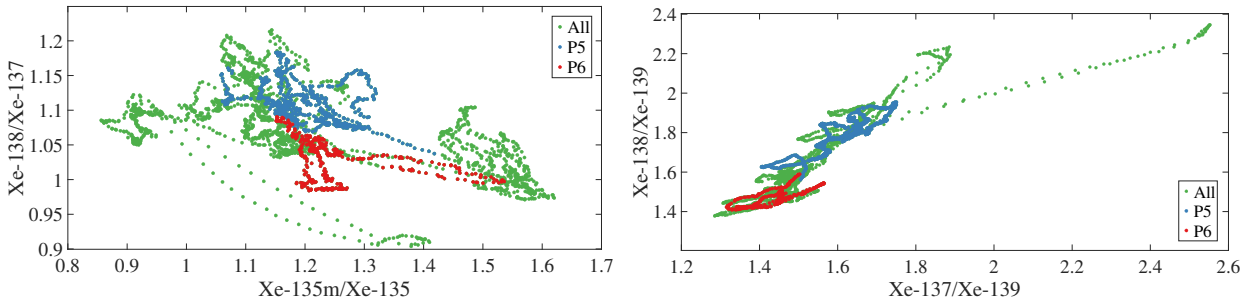

Figure 4: Ratio-ratio regions for two pairs of isotope ratios.

The count rates for Xe-138 during two ^{238}Pu campaigns, and the corresponding ratio of Xe-138 and Xe-139 are shown in Figure 2. The pair-wise ratio plots, called ratio-ratio regions, are shown for three cases in Figure 4. These regions for P5 and P6 overlap in the middle plot but not as much in the other two plots. Also, the regions for non-dissolution periods overlap these regions and also extend beyond them. These ratio-ratio regions are instructive in relating to the decision regions of the classifiers, as will be shown in the next section.

Some of the subtle variation between the two campaigns can be explained by the ongoing research and development of the dissolution process. The P5 campaign decayed the target for roughly 90 days after irradiation before processing while the P6 campaign waited about 130 days after irradiation before its material was processed. Additionally, the material processing workflow was slightly altered. P5 featured an additional caustic dissolution following the cladding dissolutions to ensure that aluminum within the pellet matrix was dissolved, but this step was removed from the P6 workflow as it was discovered the cladding dissolutions were sufficient to dissolve the aluminum within the individual pellets. Furthermore, P6 streamlined the process during the actinide dissolution by maintaining the solution at a high temperatures for an extended period instead of allowing it to cool every 24 hours, as shown in Figure 5.

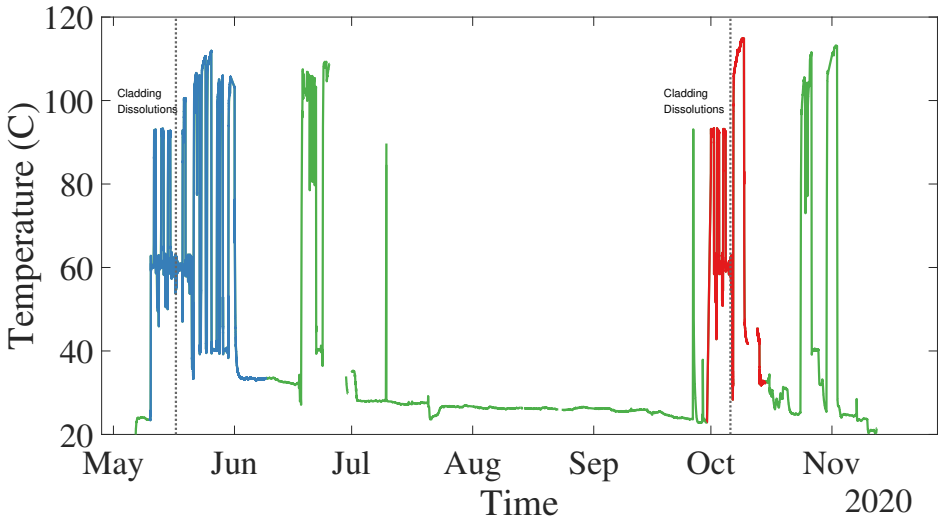

Figure 5: The temperature of the main dissolver tank during the two campaigns.

5 Classification Methods and Performance

Eight disparate classifiers are tested using two types of input features (i) the count rates of all gamma-ray peaks, and (ii) six count rate ratios Xe-135m/Xe-135, Xe-137/Xe-139, Xe-138/Xe-135, Xe-138/Xe-137, Xe-138/Xe-139, and Kr-88/Kr-89. These ratios are chosen to limit the number of classifier features by an order of magnitude below the total number of all pairs. This choice is based on previous results obtained when classifier features are limited to count rates of individual elements. It is intended to mitigate systematic effects that might have been introduced by mechanical or chemical means during the processing, filtration, and ventilation. Xenon isotopes' features achieved the lowest classification error followed by Kr isotopes' features but by a significant margin. The Iodine isotopes' features resulted in much higher classification error and hence not selected for ratio features; a detailed description of these results is provided in [2]. The choice of five Xe ratios and one Kr ratio reflects the differences in their classification errors. In both cases, the classifier-fuser ETF uses the classifiers' outputs as its input features both for training and testing.

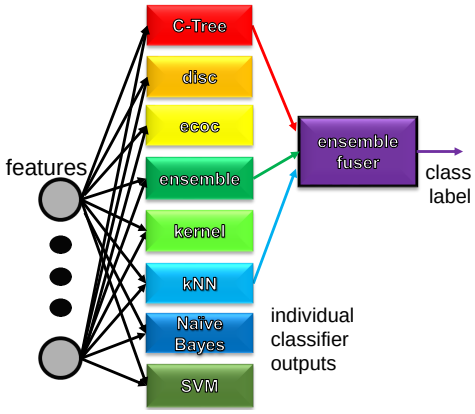

Figure 6: Eight disparate classifiers are trained using count rate or ratio estimates as features, and outputs of top tree are combines using a classifier-fuser.

5.1 Classifiers and Fusers

We use eight classifiers that represent diverse designs, and fuse the outputs of three classifiers as in [2] (illustrated in Figure 5). The eight classifiers provided by the Matlab ML toolkit are:

- **Classification Trees**: CT method employs a decision tree based whose internal nodes represent binary decisions based on values of an input feature and the leaves correspond to classification labels [4]. The underlying classifier function is non-smooth with discrete jumps corresponding to decision values of input features.
- **Discriminant Analysis Classifier**: DAC method is based on quadratic discriminant analysis that minimizes the expected classification cost estimated using both means and covariances of the classes [5]. It is based on a statistical approach and the underlying cost estimation function is typically smooth.
- **Error Correcting Output Codes**: ECOC method is based on employing a collection of binary SVMs under a coding design framework, and the underlying SVM functions are smooth [6].
- **Ensemble of Trees**: EOT method is based on boosting of a collection of classification trees that are customized to fit the training data using the AdaBoost method. It leads to highly non-smooth classification function, and its function class consists of a large collections of decision trees.
- **Kernel Method**: KM method employs a linear combination of Gaussian kernels as a classification function which is smooth.
- **k Nearest Neighbors**: Based on a structural property in feature space, kNN method utilizes the proximity information to classify a feature vector. It leads to a non-smooth classification function.

- **Naive Bayes**: NB method is based on an implementation of Bayes principle, and it uses the training sample to estimate the underlying probabilities needed for implementing an optimal Bayes fuser.
- **Support Vector Machine**: SVM method is based on non-linearly transforming the feature space, X, so that it gets "suitably separated" into classes in the transformed space. It leads to a smooth classification function which is compared to a threshold. Its function class consists of smooth functions as a result of radial basis function kernels used in our implementation.

There eight classifiers represent different types of methods: CT and ET are tree-based with non-smooth classification functions; kNN is based on the nearness concept in the feature space; KM, SVM and ECOC utilize smooth underlying functions; and DAC and NB are based on statistical methods.

5.2 Classification Performance

The performance of a classifier is quantified by its 5-fold training classification error, which is a weighted average of missed detection and false detection errors by the fraction of positive and negative examples. The classification errors of the classifiers and classifier-fuser are shown in Table 1 for the two cases using count rates and ratios as features.

	CT	DAC	ECOC	EOT	KM	KNN	NB	SVM	ETF
counts	0.0085	0.0854	0.2440	0.0031	0.2380	0.0010	0.1227	0.2426	0.0003
ratios	0.0177	0.2493	0.2564	0.0063	0.1096	0.0390	0.2522	0.2561	0.0000

Table 1: 5-fold cross-validation error of eight classifiers and the classifier-fuser for the two cases using count rate and ratio features.

Among the eight classifiers, CT, EOT and kNN achieve the lowest overall errors in both cases, and hence are chosen to be fused. The best individual classifier using count rate feature is kNN with 0.1% error, whereas that for ratio features is EOT with 0.63% error. In both cases, the fuser ETF achieved lower error than the best classifier, given by 0.03% and 0.0% error using counts and ratios features, respectively. These errors, however, must be interpreted within the context of 1758 and 1061 positive and negative examples, respectively. We account for them by explicitly considering the number of missed detections out of 1758 positive examples and the number of false detections (alarms) out of 1061 negative examples. The comparative performance of these classifiers is summarized as follows:

- The same three classifiers, namely CT, EOT and kNN, achieved the lowest errors in both cases, but different ones achieved the lowest error in individual cases. They are all non-smooth methods, and CT and EOT are tree-based and kNN is structure-based (that is, nearness in feature space). The smooth and statistical methods did not achieve comparably low error in either case.
- Best classifiers are different in the two cases, although both are non-smooth. For the count rate features, kNN missed detecting 1 out of 1758 positive examples and has 2 false alarms out of 1061 negative examples, whereas kNN with ratio features made more errors, namely, missed 58 detections and has 52 false alarms. Thus, "neareness" in the count rate space did not carry over to the ratio space, thereby resulting in different errors of kNN method. EOT using count ratio features missed 10 detections and has 8 false alarms, and using count rate features, it missed 4 detections and has 5 false alarms. The tree structure of the decision

surfaces used by EOT seems to be effective for classification in both cases, resulting in the error difference of 0.32%.

- The impact of classifier fusion is more pronounced for ratio features, as the error is reduced from 0.63% to 0.0% (EOT to ETF), which is a result of the elimation of 4 missed detection and 5 false alarms. The reduction is relatively less pronounced for count rate features, namely from 0.10% to 0.03% (kNN to ETF), which is a result of the reduction of missed detections from 1 to 0 and false alarms from 2 to 1. It is interesting to note that the classifier fusion is more effective in reducing the classification error when ratio features are used.
- The difference in ETF errors is very small for the two cases, and is a result of 1 false alarm out of 1061 negative examples in the former case and none in the latter. However, this difference is too small to infer their relative superiority.

In addition to the classification errors, the decision regions of the classifiers provide some insights into their sensitivity to ratio-ratio regions. For kNN based on count rate features with the lowest error, the decision region closely matches the ratio-ratio region as shown in Figure 7(a). For the ratio features, kNN's decision region extends beyond the ratio-ratio region, resulting in a larger (albeit still low) error, as shown in Figure 7(b).

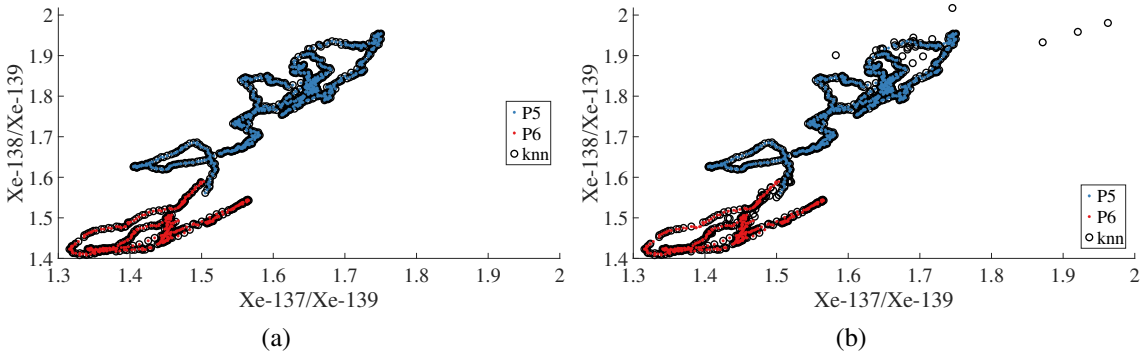

Figure 7: Ratio-ratio regions of training data and kNN classifier.

Overall, the classification performance is comparable in both cases based on the data from two ^{238}Pu campaigns. While these results indicate that the isotope ratios are viable and possibly more justifiable classifier features compare to isotope count rates, these errors must be interpreted with in the context of "small" data set .

6 Conclusions

This paper is a part of continuing efforts towards developing data analytics for inferring the occurrences of dissolutions at the REDC using gamma-ray measurements of the gaseous effluents at the off-gas stack. In previous works, HPGe gamma spectra of the effluents have been used to estimate count rates in spectral regions of radionuclides, which are then used as features to train classifiers [1, 2]. In this paper, we explored the use of the isotope ratios as classifier features since they are considered less sensitive to count rates that could vary within and across the dissolution campaigns. Eight disparate classifiers and their fuser are trained using data from two ^{238}Pu dissolution campaigns to classify dissolution events using the gamma spectra measurements. Two types of classifier feature are utilized, namely, seventeen isotope count rate estimates and six of their ratios. The training classification accuracy results of classifiers are presented, and the ensemble of trees fuser achieved the overall high accuracy in both cases. Overall, the classification

performance is comparable in both cases based on the (limited) data from two ^{238}Pu campaigns at REDC, which indicates that the isotope ratios are viable and possibly more justifiable classifier features compared to isotope count rates.

Several future directions remain to be pursued, including, expanding the study to more campaigns of ^{238}Pu and other isotopes, such as ^{252}Cf, and the performance study of classifiers that are trained with limited data from some campaigns and tested with that from several other campaigns. More generally, it would be of future interest to identify invariants (such as isotopic ratio regions) that are considered stable across multiple dissolutions and also the re-work activities in between them. Another future direction is the physics-based explanations for the classifier performance complemented by the incorporation of details of the radiochemical processes and effluents transport mechanisms of the facility including the abatement and filtering systems.

Acknowledgments

This research used resources at the High Flux Isotope Reactor, a DOE Office of Science User Facility operated by the Oak Ridge National Laboratory. This research was supported by the U.S. Department of Energy, National Nuclear Security Administration, Office of Defense Nuclear Nonproliferation Research and Development (DNN R&D) and was performed at Oak Ridge National Laboratory managed by UT-Battelle LLC for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

References

[1] N. S. V. Rao, C. Greulich, S. Sen, K. Dayman, A. Nicholson, M. R. Chatin, K. M. Buckley, R. D. Hunley, J. Johnson, H. H. Hesse, and R. Hale, "Classifiers for dissolution events in processing facility using effluents measurements," in *Institute of Nuclear Materials Management Annual Meeting*, 2019.

[2] N. S. V. Rao, C. Greulich, S. Sen, J. Hite, K. J. Dayman, A. D. Nicholson, D. E. Archer, M. J. Willis, I. G. R. D. Hunley, J. Johnson, A. J. Rowe, I. R. Stewart, and J. M. Ghawaly, "Classification of dissolution events using fusion of effluents measurements and classifiers," in *Institute of Nuclear Materials Management Annual Meeting*, 2020.

[3] D. E. Archer, A. D. Nicholson, M. J. Willis, and I. Stewart, "Radiation detectors with associated contextual sensors within the multi-informatics for nuclear operations scenarios (minos) program," in *Institute of Nuclear Materials Management 60th Annual Meeting*, pp. 1357–1367, 2019.

[4] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, 1984.

[5] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley and Sons, Inc., 2001. Second Edition.

[6] S. Escalera, O. Pujol, and P. Radeva, "Separability of ternary codes for sparse designs of error-correcting output codes," *Pattern Recognition Letters*, vol. 30, no. 3, pp. 285–297, 2009.