DATA-DRIVEN EVENT SELECTION IN PIXELATED CADMIUM ZINC TELLURIDE (CZT) DETECTORS FOR IMPROVED GAMMA-RAY SPECTROMETRY

G. Aversano, H. S. Parrilla, M. S. Bandstra, M. Folsom, D. Hellfeld, J. R. Vavrek

Lawrence Berkeley National Laboratory
Berkeley, California, USA

ABSTRACT

Improved detector resolution can aid in the identification and non-destructive assay (NDA) of radioisotopes, which is crucial for nuclear safeguards applications. This improvement would be particularly useful in separating closely-spaced characteristic photopeaks within spectra used for enrichment measurements. A data-driven approach was developed using unsupervised machine learning to cluster segments of an H3D M400 pixelated Cadmium Zinc Telluride (CZT) detector. The candidate clusters were ranked by their resolvability, defined as the square root efficiency divided by the Full Width at Half Max (FWHM), to optimally trade off a modest amount of detector efficiency for large improvements in peak resolution. The unsupervised model was fitted using data collected from long-dwell (64 hour) measurements of a 100 μCi Eu-154 source placed 30 cm away from the front face of the H3D M400 CZT detector. The resulting model can then be applied to spectra outside the training dataset. In one example, a model was applied to spectra obtained from uranium standard measurements from various enrichments at Lawrence Berkeley National Laboratory, demonstrating that the model can generalize to newly seen data from different radioactive sources. Ongoing work will continue to quantify spectral improvement for safeguards-relevant measurement scenarios. This data-driven approach offers a real-time algorithmic solution to improve gamma-ray spectrometry in pixelated CZT detectors. In the future, the model will be accessible to external stakeholders, such as the IAEA, via a python software package, to allow inspectors to select the desired resolvability improvement while conducting measurements in the field.

INTRODUCTION

The International Atomic Energy Agency (IAEA) performs independent verification of nuclear material in states with IAEA safeguards agreements. The accurate and reliable identification and assay of radiological sources in operational environments are therefore critical measurements for these safeguards efforts. For these applications, a compact gamma detector system with high resolution is desirable.

Advancements in crystal manufacturing, Application Specific Integrated Circuit (ASIC) design, and signal processing techniques have led to the development of modern, large-volume, high-resolution, 3D position-sensitive pixelated CZT gamma-ray detector systems that are owned by >75% of US nuclear power generating stations and are commercially available for purchase through companies such as Michigan-based H3D Inc. In addition to excellent energy resolution (nominally <1% FWHM at 662 keV), CZT-based detectors can be operated at room temperature, are compact in size, and can maintain good performance in high-flux environments [1]. These performance and operational advantages have led to an increased interest in the adoption of such detectors into the nuclear safeguards technology suite, as a replacement for scintillator detectors such as NaI(Tl) [2].

Many years of effort have progressively improved the energy resolution of pixelated CZT, making these systems ideal alternatives to High Purity Germanium (HPGe) detectors for nondestructive

nuclear material assay and gamma-ray spectrometry. For example, the M400 system from H3D Inc. [3] now offers a high-resolution option with <0.65% FWHM at 662 keV for individual virtual detector voxels. However, to increase counting statistics and overall detection efficiency, virtual voxels are often combined which, due to spatially varying performance across the voxels, ultimately results in a degradation in energy resolution (<0.8% FWHM at 662 keV). While care has been taken to properly normalize and align individual voxel spectra before summation [4], there remains a capability gap to intelligently and automatically configure, combine, or cluster virtual detector voxels to optimize the detector system performance. Poorly-performing voxels can be rejected to improve the spectral resolution at the cost of lower detector efficiency (and therefore longer measurement times).

Unsupervised machine learning techniques have become increasingly popular for pattern recognition and clustering tasks. Specifically, Non-negative Matrix Factorization (NMF) [5] has been used to model gamma-ray spectra due to its ability to enforce non-negativity while remaining consistent with Poisson statistics. The NMF components, learned directly from measured gamma-ray data, are additive and physically interpretable. For example, NMF has proven successful in discovering physically relevant spectral structure from known sources such as terrestrial KUT and cosmic in gamma-ray background data [6]. As a dimensionality reduction method, NMF provides a compact description of data in a lower-dimension latent space. Enforcing a normalization constraint on the weights results in a parts-based spectral decomposition of spectral basis vectors, where the features describe the importance of the spectral basis to reconstruct the original data. Like other methods of dimensionality reduction, NMF can be used as a basis for clustering data. This clustering interpretation of NMF has been used extensively in the field of document classification [7, 8]. Although NMF can be used for clustering, several other clustering algorithms are explored to cluster the voxels based on the features found during NMF.

This work aims to implement data-driven, unsupervised machine learning methods to enable the automatic and optimal selection of the virtual detector voxels to improve spectral performance using the M400 CZT detector system. The approach is designed to be detector-agnostic and applicable to other pixelated detector systems such as segmented HPGe. The product of this effort will be a python software package and analysis procedure that enables end users, such as IAEA inspectors, to span the efficiency-resolution tradeoff in real-time in the field, ultimately improving the effectiveness of routine nuclear safeguards techniques.

METHODS

Data Collection & Pre-Processing

To train the machine learning models, long dwell measurements of a 100 μCi Eu-154 source were performed using an H3D M400 pixelated CZT detector. Eu-154 was selected because its spectrum contains several characteristic photopeaks across a wide energy range applicable to safeguards applications (in particular, 123, 248, 723, 873, and 1274 keV). Measurements were taken head on[1] at a 30 cm standoff on a laboratory workbench over 64 hours to ensure sufficient count statistics at the voxel level.

[1] The decision to take measurements head on only was made based on feedback from the IAEA as inspectors typically only take head on measurements during non-destructive assay inspections in the field. The detector was placed upside down to reduce chances of overheating. This concern is expected to be addressed in a software patch to be provided by H3D Inc.

The spectrum was collected in listmode, including 3D event position, and converted into binmode. Only single-site events were included in this analysis (estimated to cover >90% of all events). The M400 CZT detector contains four discrete crystals each with an 11x11 grid of pixels for a total of 484 pixels. The XY binning (1.9 mm) was determined by analyzing position differences in distinct horizontal and vertical lines of the collected data. Each pixel was also discretized into 50 virtual depth bins (0.2 mm) between the anode and cathode for a total of 24,200 (484x50) voxels. Small depth bins were used since bins could be combined later, if necessary. The origin is set back into the enclosure such that events at the anode face will have Z=0 mm and events at the cathode face will have Z=10 mm (Figure 1).

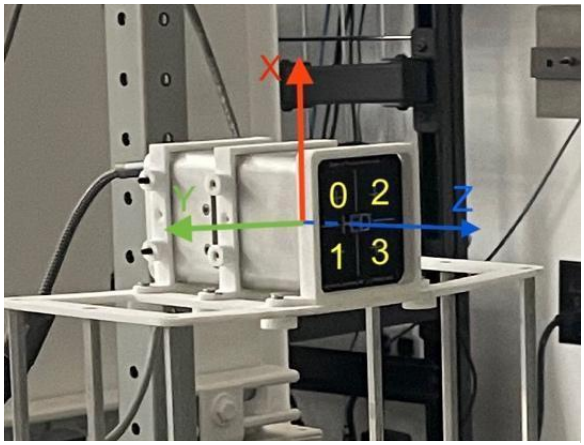

Figure 1. Coordinate system of the M400 (red, green, blue) with individual detector IDs labeled (yellow). The origin is set back into the enclosure.

Regions of interest were extracted around five photopeaks (123, 248, 723, 873, and 1274 keV) to be included in the training data. Pixel-level spectra were calibrated[2] by adjusting the energy 'gain' between the actual and expected photopeak (gain = E_e/E_a). The gain was calculated for each peak of interest and linearly interpolated between the energies. The actual peak energy was determined by fitting the counts to a Doniach [9] lineshape parameterized by:

$$D(E) = \frac{A\cos\left[\frac{\pi\gamma}{2}+(1-\gamma)\left(\frac{E-\mu+\nu}{\sigma}\right)\right]}{(\sigma^2+(E-\mu+\nu))^{\frac{1-\gamma}{2}}}, \quad (1)$$

$$\nu = \frac{2.3548\sigma}{2tan\left(\frac{\pi}{2-\gamma}\right)}, \quad (2)$$

where A is the amplitude, μ is the centroid energy (keV), γ is the asymmetry (i.e., tailing) parameter, σ is the characteristic width parameter (keV) and E is the energy (keV). A linear background model was also included, resulting in a 6-parameter fit to the data. The centroid energy represents the actual photopeak energy observed by the detector.

Initial performance characterization of the detector shows significant variation at the virtual voxel level (Figure 2) suggesting that the poor performing detector regions can be identified and removed to improve the overall spectral performance, at the cost of lower detector efficiency (and longer measurement dwell times), thus motivating the objective of this work.

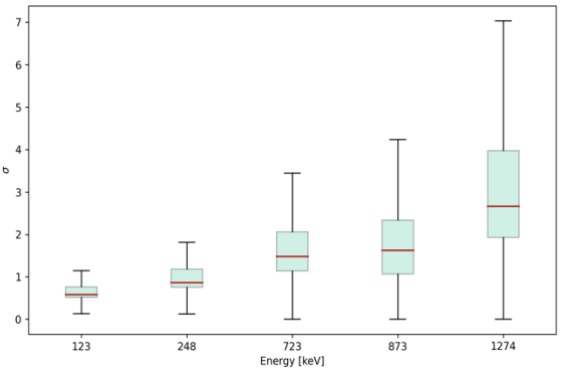

Figure 2. Doniach sigma fit parameter for five photopeaks of interest showing voxel performance variation.

[2] In practice, the detector is typically calibrated prior to conducting measurements. While done in post-processing here, we note that H3D provides an online recalibration tool in its provided software.

Detector Voxel Clustering

NMF was applied to the voxel spectra for each individual photopeak (i.e., a single energy model):

$$X \approx WH \tag{3}$$

where X is an NxM matrix representing the $N = 24\,200$ voxels and M energy bins around each photopeak of interest, W is an NxK matrix with K latent features, and H is a KxM matrix. The matrix H was normalized row-wise after each multiplicative update such that the spectral basis vectors represent spectral basis *shapes* of equal height and the features of W can be interpreted as the importance of each basis on the reconstruction of the original data. The matrix W was corrected by the weighting factor such that the reconstruction was preserved. Regularization (L_1) was enforced to promote sparsity on the matrix W. The number of latent features, K, was varied from $2 \leq K \leq 6$ along with the regularization parameter, α, as either 0 (no regularization), 0.01, or 0.1.

Next, clustering was performed on the feature matrix W (after standardization) using several different clustering algorithms within the scikit-learn python package [10] including Birch [11], Agglomerative Clustering [12], K-Means [13], DBSCAN [14], OPTICS [15], and Gaussian Mixture [16]. The number of clusters, n, was varied from $2 \leq n \leq 6$. Clustering was performed for all combinations of the hyperparameters (i.e., number of NMF features, K, NMF regularization, α, and number of clusters, n) and represents cluster assignments against each detector voxel.

Choosing the Optimal Clusters

For each cluster within each model, the voxels were aggregated and ranked based on the new cluster estimated 'resolvability' which is defined as:

$$r = \frac{\sqrt{eff}}{FWHM} \approx \frac{\sqrt{A}}{2\sqrt{2\,log\,(2)}\,\sigma} \approx \frac{\sqrt{A}}{2.355\sigma} \tag{4}$$

where the FWHM is calculated under a gaussian approximation of the Doniach width parameter, σ, and A is the Doniach amplitude indicative of the efficiency.

For each model, the worst clusters were removed in order of lowest resolvability until only the best cluster was remaining, or the new detector efficiency fell below 10% of the bulk efficiency. Finally, for each model with successive clusters removed, the resolvability was recalculated to find the best overall detector segmentation with low quality voxels removed.

RESULTS & DISCUSSIONS

NMF Spectral Representation

NMF was used to learn a low-dimension, latent representation of the measured spectra. The matrix H shows a decomposition of the spectra into parts-of-a-whole or contributions from spectral basis vectors. The basis vectors are interpretable and can be understood as contributing to the photopeak (orange, green, red, purple), background (purple), and tailing (red) of the spectrum (as seen in the left image in Figure 3). NMF was unable to learn a single linear background basis vector resulting in some unexpected poor fits on noisy data without a discernible photopeak.

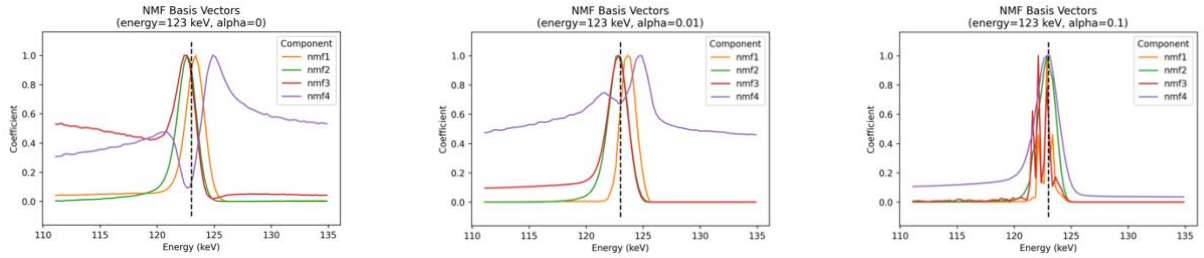

Figure 3. Impact of regularization on NMF basis vectors ($K = 4$).

When including an L_1 regularization penalty, the basis vectors tend to overlap and lose some interpretability where without regularization, the basis vectors tend to stand adjacent (Figure 3). Regularization tends to decrease the accuracy of the reconstruction (measured using the Akaike Information Criterion (AIC) [17] and Bayesian Information Criterion (BIC) [18]) but promotes a sparse feature solution (Figure 4). Although a gain correction was applied to the collected spectra, some voxels still exhibit a centroid bias. With regularization, the basis vectors may not be sufficient to reconstruct the spectra from these voxels.

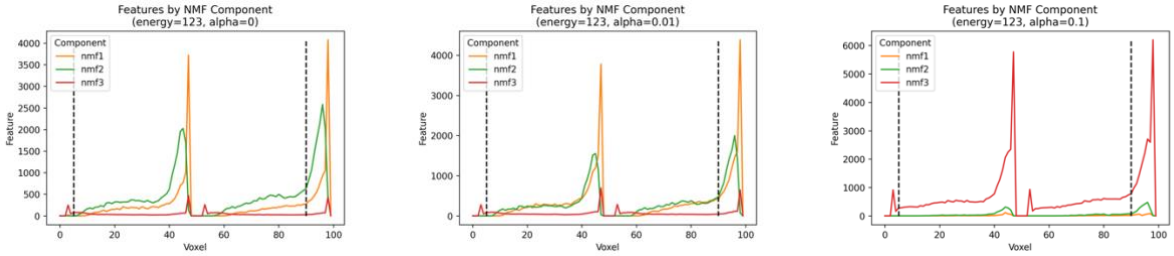

Figure 4. Impact of regularization on NMF features for the first 100 voxels ($K = 3$).

Detector Voxel Clustering

The FWHM heatmap in Figure 5 shows the estimated FWHM (with gaussian Doniach sigma assumption) for each voxel. The regions closest to the anode and cathode have very low counts resulting in poor fits. There are strongly performing voxels largely clustered near the anode-end (but also near the cathode-end for higher energy peaks) with sparse voxels within the body of the detector. In very few cases, poor performing pixels are observed as indicated by vertical bright lines. These plots can be used as a baseline for the clustering algorithms whose objective is to segment the detector into regions based on performance.

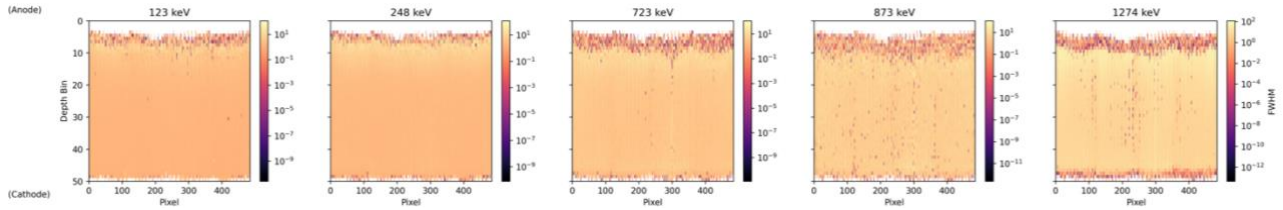

Figure 5. Voxel FWHM for each photopeak of interest assuming a gaussian distribution of the Doniach width parameter (white pixels indicate that a lineshape could not be fit to the data).

Another baseline approach could involve using a 'greedy' algorithm which accumulates voxels in order of highest resolvability. The maximum resolvability improvements compared to the bulk detector are 1.07, 1.06, 1.00, 1.00, and 1.01 respectively for the five photopeaks of interest. The highest improvements tend to occur around 70-90% efficiency. These improvements will be compared to the unsupervised approach. Exploring every voxel combination (on the order of 2^10,000) is intractable thus requiring an intelligent clustering approach.

Various clustering algorithms were used to cluster the NMF features, W as shown in Figure 6. The clusters tend to form depth-wise. In general, models with more clusters were able to better separate different regions of the detector including the regions closest to the anode/cathode and pixel level trends.

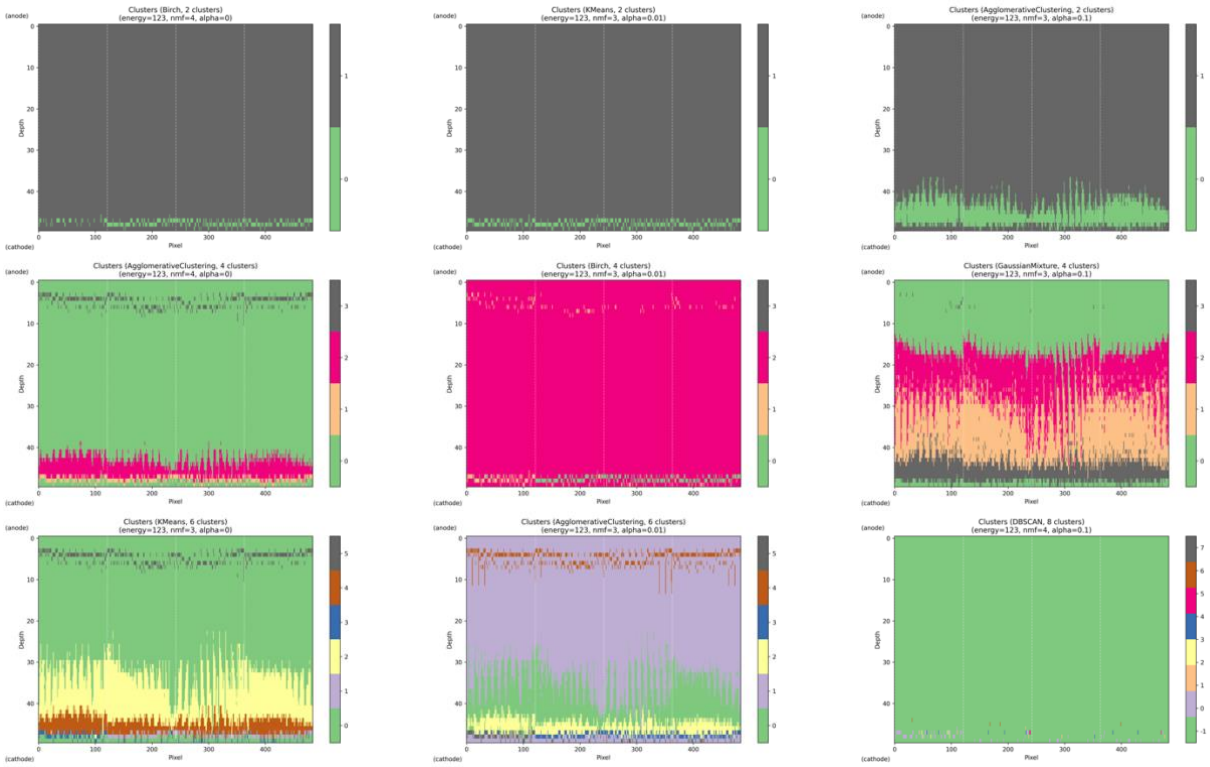

Figure 6. Cluster assignments using various clustering algorithms and NMF models. The number of clusters increases from top to bottom. The regularization increases from left to right.

Choosing the Optimal Clusters

For each NMF + clustering model (>300 scenarios), the worst clusters were successively removed from the bulk spectra until less than 10% of the gross counts remained. The resolvability was calculated again to rank the remaining spectra to find an optimal model and optimal number of removed clusters. Over a thousand scenarios were considered for each photopeak. The optimal parameters, based on the final resolvability, for each photopeak of interest are shown in Table 1.

The best models tend to have more NMF components potentially indicating that several basis vectors are required to reconstruct the data and represent the data in a lower-dimension latent space. There

were several instances where the reconstructions were poor on noisy voxel data resulting in unrepresentative features and ultimately worse clusters. Initializing the NMF basis vectors to include a linear background component may improve the reconstruction.

Table 1. Optimal hyperparameters for the NMF + clustering model.

Energy (keV)	**NMF *K***	**NMF *α***	**Clustering Algorithm**	**# Clusters Removed / Total Clusters**	**Resolvability (best/bulk)**	**Detector Efficiency (best/bulk)**
123	5	0.10	Gaussian Mixture	2 / 6	1.07	0.92
248	4	0.00	Agglomerative	3 / 5	1.06	0.74
723	6	0.01	Gaussian Mixture	2 / 3	1.01	0.74
873	5	0.01	Gaussian Mixture	2 / 4	1.03	0.85
1274	4	0.01	Agglomerative	2 / 5	1.10	0.73

Less regularization improved the model's ability to segment the detector. With regularization, the basis vectors have more overlap and appear to reconstruct the original data by building up the height of the peak. Perhaps the model is unable to find a sparse feature solution with these basis vectors. In the future, the sparsity should be measured to confirm regularization is acting as intended.

Most clustering algorithms were able to learn a detector segmentation resulting in *some* improvement to the peak resolvability. In general, a larger number of clusters were required to allow the model to distinguish between the separate detector regions. The optimal configurations rejected roughly half of those clusters. However, the cluster sizes are uneven and the counts per voxel are not constant. The reduction of detector efficiency ranged from 8 – 27% indicating that most of the voxels are still used in the optimal detector segmentation. Removing the remaining clusters results in a large decrease in counts reducing the resolvability. Lastly, there is opportunity to improve the intra-cluster spectral variation within each cluster (see Figure 7). Expanding the grid search to include an even larger number of clusters or more fine-grained hyperparameter values may be beneficial in finding the optimal solution.

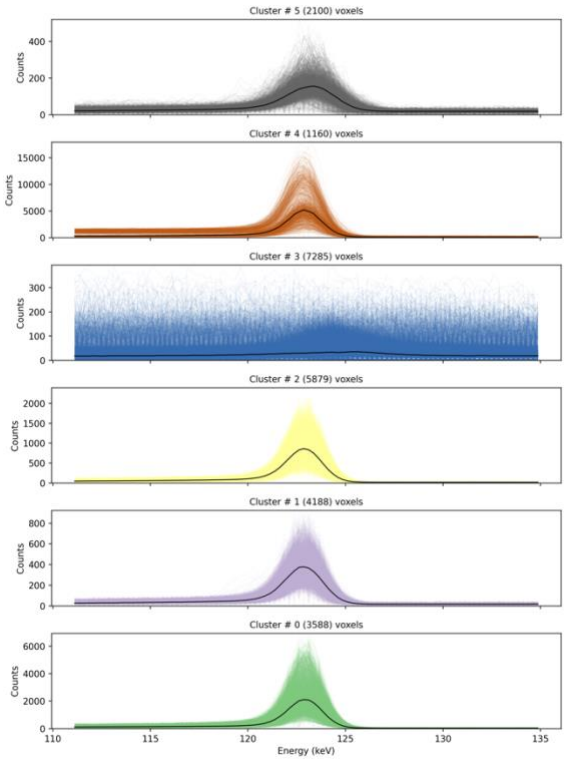

Figure 7. Spectra for each cluster (median in black) showing high variation within each cluster.

The voxel clusters for the optimal models along with the optimal detector segments are shown in Figure 8. Except for the 1274 keV model, all models reject the regions closest to the anode/cathode with low count rates. Pixel level segmentation is also observed in most models.

The improved spectra around the photopeak of interest utilizing the optimal detector segments are shown in Figure 9. The photopeaks have finer resolution and less background counts which will improve resolving the photopeak in the presence of other nearby photopeaks. As a trade-off, the detector efficiency is reduced as a result of rejecting voxels.

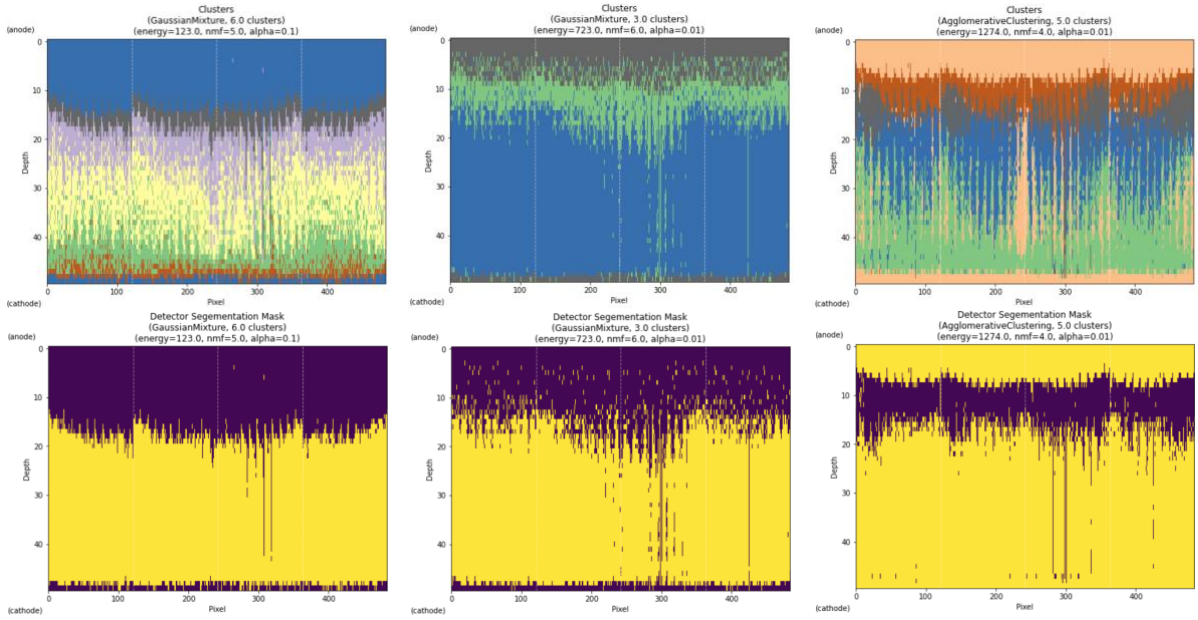

Figure 8. Voxel clusters for the optimal models (top) and corresponding detector segmentations (bottom, yellow=on) for three of five photopeak of interest: 123 keV (left), 723 (middle) and 1274 keV (right).

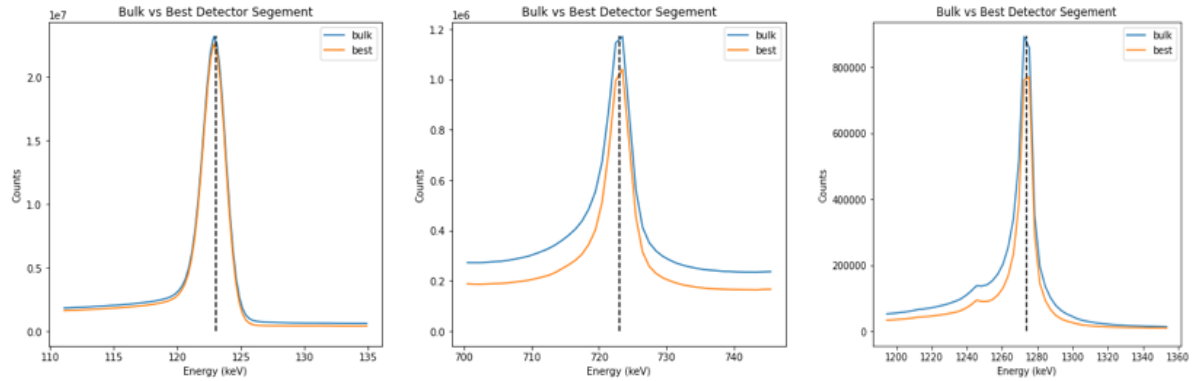

Figure 9. Spectra for the bulk and best detector segments for three of five photopeak of interest: 123 keV (left), 723 (middle) and 1274 keV (right).

Application to Uranium Measurements

The voxel- and/or pixel-level cluster removal results trained on Eu-154 data can then be applied to never-before-seen spectra outside the training dataset, such as uranium standards measurements, to test how the model could perform in the field. Measurements were taken at approximately 20 cm from each source, which was contained in a canister, on a laboratory workbench for 20 minute dwell times. The uranium sources included natural, depleted, and enriched (2-93%) samples. Figure 10 shows the

result of a 4 NMF component, 4 Gaussian Mixture cluster model trained on the 123 keV peak of Eu-154 applied to the low-energy region of a 93%-enriched uranium standard. Clear improvements in resolution can be seen near 80, 190, and 220 keV, though there remains ongoing work to quantify these improvements in terms of the resolvability metric.

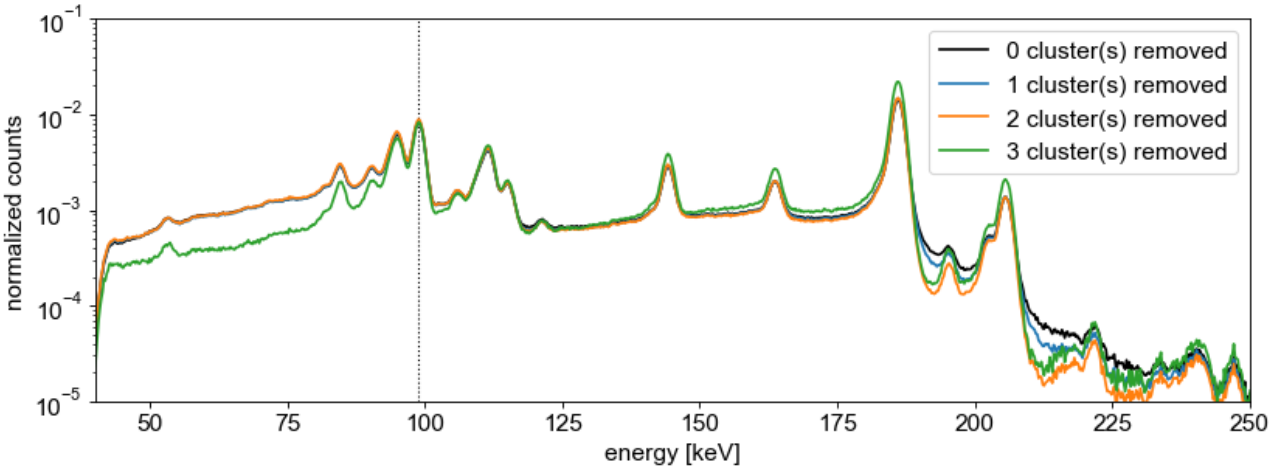

Figure 10. Voxel cluster removal results with a 93% enriched uranium standard. Spectra are normalized to the peak near 95 keV to better show the changes in shape.

CONCLUSIONS

A data-driven approach using NMF and clustering was used to learn an optimal detector segmentation within a pixelated CZT detector to improve the spectral performance (as quantified by the resolvability metric FWHM/sqrt(efficiency)) by rejecting poorly-performing regions of the detector in order to trade off efficiency for resolution. The unsupervised learning approach performs as-well-as or better than the greedy algorithm demonstrating its ability to automatically segment the detector into strong and poor performing regions. This model was applied to never-before-seen uranium standards spectra, demonstrating that it can generalize to sources applicable for nuclear safeguards applications. Improved resolvability will improve nuclear safeguards efforts, such as nuclear material non-destructive assay, particularly as more CZT detectors are used in favor of NaI detectors. This approach should be expanded to include a single model based on data from multiple photopeaks at different energies and offer additional voxel clustering to further traverse the resolvability-efficiency trade-off curve.

ACKNOWLEDGEMENTS

The work presented in this paper was funded by the National Nuclear Security Administration of the Department of Energy, Office of International Nuclear Safeguards. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Berkeley National Laboratory (LBNL) under Contract DE-AC02-05CH11231. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

REFERENCES

1. W. Kaye et al., “In-Situ Sample Analysis with Portable Gamma Spectrometers”, IRPA 2015.
2. “Development and Implementation Support Programme for Nuclear Verification 2022-2023”, IAEA Safeguards Report STR-393, 2022.
3. M400, Custom Integrable Detector Module, H3D Inc. Available online: https://h3dgamma.com/M400Specs.pdf.
4. W. Li et al., “Spatial Variation of Energy Resolution in 3-D Position Sensitive CZT Gamma-ray Spectrometers”, IEEE Transaction on Nuclear Science, vol. 46, no. 3, pp. 628-633, 1999. https://doi.org/10.1109/NSSMIC.1998.775218.
5. D. D. Lee and H. S. Seung, “Learning the Parts of Objects by Non-negative Matrix Factorization”, Nature, vol. 401, pp. 788–791, 1999. https://doi.org/10.1038/44565.
6. M. S. Bandstra et al., “Modeling Aerial Gamma-Ray Backgrounds Using Non-negative Matrix Factorization”, IEEE Transaction on Nuclear Science, vol. 67, no. 5, pp. 777-790, May 2020. https://doi.org/10.1109/TNS.2020.2978798.
7. F. Shahnaz et al., “Document Clustering using Nonnegative Matrix Factorization”, Information Processing and Management, vol. 42, no. 2, pp. 373-386, 2006. https://doi.org/10.1016/j.ipm.2004.11.005.
8. W. Xu et al., “Document Clustering based on Non-negative Matrix Factorization”, Proceedings AXM SIGIR Conf. on Research and Development in Information Retrieval, pp. 267-273, 2003. https://doi.org/10.1145/860435.860485.
9. S. Doniach and M. Sunjic, “Many-electron Singularity in X-ray Photoemission and X-ray Line Spectra from Metals,” Journal of Physics C: Solid State Physics, vol. 3, no. 2, pp. 285-291, 1970. https://doi.org/10.1088/0022-3719/3/2/010.
10. Pedregosa et al., “Scikit-learn: Machine Learning in Python”, Journal of Machine Learning Reaching, vol. 12, pp. 2825-2830, 2011. https://dl.acm.org/doi/10.5555/1953048.2078195.
11. T. Zhang, R. Ramakrishnan, and M. Livny, “BIRCH: An Efficient Data Clustering Method for Very Large Databases,” in Proc. ACM SIGMOID, vol. 25, no. 2, pp. 103-114, 1996. https://doi.org/10.1145/235968.233324.
12. A. S. Vijaya and R. Bateja, “A Review on Hierarchical Clustering Algorithms”, Journal of Engineering and Applied Sciences, vol. 12, no. 24, pp. 7501-7507, 2017. doi:10.36478/jeasci.2017.7501.7507.
13. J. Hartigan and M. Wong. “Algorithm AS136: A k-means clustering algorithm”, Journal of Royal Statistical Society (Applied Statistics), vol. 28, no. 1, pp. 100-108, 1979. https://doi.org/10.2307/2346830.
14. M. Ester, et al., “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”, Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, AAAI Press, pp. 226–231, 1996. doi:10.5555/3001460.3001507.
15. M. Ankerst, et al., “OPTICS: ordering points to identify the clustering structure”, ACM Sigmod Record, vol. 28, no. 2, pp. 49-60. 1999. https://doi.org/10.1145/304181.304187.
16. C. Bouveyron, S. Girard, and C. Schmid, “High-dimensional data clustering”, Computational Statistics & Data Analysis, vol. 52, no. 1, pp. 502-519, 2007. https://doi.org/10.1016/j.csda.2007.02.009.
17. H. Akaike, “A New Look at the Statistical Model Identification”, IEEE Transaction on Automatic Control, vol. 19, no. 6, pp. 716-723, 1974. https://doi.org/10.1109/TAC.1974.1100705.
18. G. Schwarz, "Estimating the Dimension of a Model", The Annals of Statistics, vol. 6, no. 2, pp. 461-464 1978. https://doi.org/10.1214/aos/1176344136.