CHANCE project

(Contract Number: 755371)

Performance of the CHANCE muon system in the reconstruction of known blocks of material (D4.2)

Work Package 4: Muon imaging for innovative tomography of large volume and heterogeneous cemented waste packages

Author(s): **A.F. Alrheli, D. Barker, C. De Sio, D. Kikoła, M. Mhaidra, J.P. Stowell, L.F. Thompson, J.J. Velthuis, M.J. Weekes**

Reporting period 3: 01/06/2020 – 31/03/2022

Date of issue of this report: **May 2022**

Start date of project: **01/06/2017** Duration: 58 Months

This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 755371;		
Dissemination Level		
PU	Public	X
CO	Confidential, only for partners of the CHANCE project and EC	

History chart			
Status	**Type of revision**	**Partner**	**Date**
Draft	Initial version	WUT	May 2022

Reviewed by

Approved by

The Executive Board

Table of Contents

1. Introduction

The CHANCE project aims to address the specific issue of the characterization of conditioned radioactive waste. The characterization of fully or partly conditioned radioactive waste is a specific issue because unlike for raw waste, its characterization is more complex and therefore requires more advanced non-destructive techniques and methodologies.

The objective of CHANCE is to further develop, test and validate techniques already identified that will improve the characterization of conditioned radioactive waste, namely those that cannot easily be dealt with using conventional methods. Specifically, the work on conditioned radioactive waste characterization technology will focus on:

- Calorimetry as an innovative non-destructive technique to reduce uncertainties on the inventory of radionuclides;
- Muon Tomography to address the specific issue of non-destructive control of the content of large volume nuclear waste;
- Cavity Ring-Down Spectroscopy (CRDS) as an innovative technique to characterize outgassing of radioactive waste.

The present report focuses on activities from Work Package 4 related to the development of the Muon Scattering Tomography (MST) for the characterization of large and dense objects, for instance, nuclear waste packages. MST uses cosmic ray muons that are highly penetrating particles; thus, it can allow for three-dimensional imaging of sealed waste packages which use dense matrix material (e.g. concrete or bitumen) to immobilize radioactive waste.

During the CHANCE project, we developed several methods for the identification of objects in the radioactive waste packages. These methods allow for the recognition of the blocks of material, both high-density and low-density ones, and finding their size (or volume) and location.

We first tested these methods using realistic Monte Carlo simulations of the CHANCE muon scattering system, and then we analyzed the experimental data collected with the CHANCE muon detector.

This report is based on results presented at meetings and conferences (1), (2), and published in (3).

2. Muon Scattering Tomography

Muon scattering tomography is a non-invasive method which allows producing 3D images of closed and high-density objects, thus is it well suited non-destructive assay for the characterization of nuclear waste packages. MST uses cosmic rays as probes. Cosmic rays are high energy, charged particles which come to the Earth's atmosphere from outer space. In the atmosphere, cascades of new particles are produced. The main type of particles that reach sea level are muons. Muons are identical to electrons, but 200 times heavier. Muons can go through large amounts of material as they do not scatter very much

due to their high mass. When traversing material, Coulomb interactions take place between the muons and the nuclei of the material. As a result, muons exit the material under an angle. The angular distribution of scattering of muons can be described by a Gaussian distribution with a mean of zero and a standard deviation σ_θ described by (4):

$$\sigma_\theta = \frac{13.6\ MeV}{pc\beta}\sqrt{\frac{T}{X_0}}\left[1 + 0.038ln\left(\frac{T}{X_0}\right)\right] \quad (1.1)$$

$$X_0 \approx \frac{716.4A}{Z(Z+1)ln\left(\frac{287}{\sqrt{Z}}\right)}\ [g \cdot cm^{-2}] \quad (1.2)$$

Where p is the muon's momentum, β the muon's speed divided by the speed of light c, T is the thickness of the material and X_0 its radiation length. A is the atomic weight of the medium in g·mol^{-1}. The standard deviation depends on the atomic number, Z, of the traversed material. Under the assumption that scattering occurs in single locations and by reconstructing the incoming and outgoing trajectories of the muons, the scattering angle distribution can be reconstructed and thus information about the traversed material can be extracted.

Muon tomography requires both the incoming and outgoing muon trajectory to be measured. Hence, the object under inspection needs to be covered on both sides. As muon tomography relies on reconstruction of the scattering angle, the key parameter for the detector system is the angular resolution of the upper and lower detector system.

2.1 Imaging objects with Muon Scattering Tomography

Point of Closest Approach algorithm

Imaging with muon tomography is based on registration the incoming and outgoing muon trajectory. In the simplest approach, called the Point of Closest Approach (PoCA) algorithm, multiple scatterings of a muon are modelled as a single scattering at a single point ('scattering vertex'), see Figure 1. The scattering vertex is found by extrapolating the incoming and outgoing tracks and searching for a point at which the distance between them is minimal.

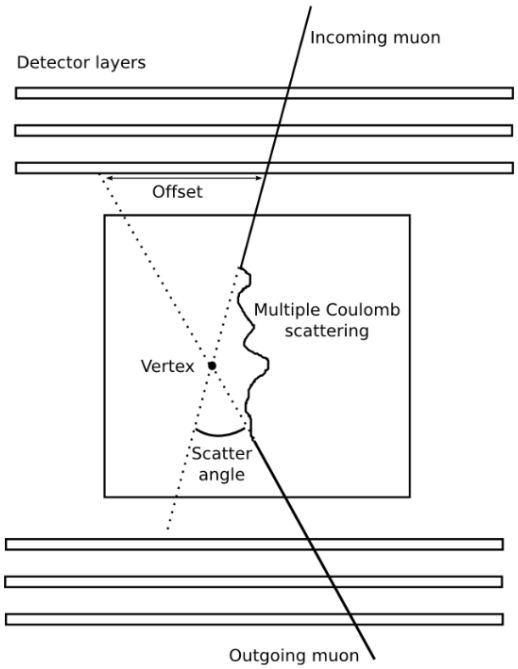

Figure 1: Illustration of scattering vertex reconstruction of a muon (5)

To obtain a 3D image, the scanned volume is divided into cubic voxels. The number of scattering vertices and the scattering angle reconstructed in each voxel depends on the radiation length of the material in that voxel. Thus, analysis of the density of scattering vertices, their distribution and distribution of a scattering angle provide means to discriminate between material with different atomic number Z.

In the PoCA algorithm, the image is obtained as a **3D density map of the scattering vertices**, sometimes weighted by a value of scattering angle. Such an approach is simple, but it suffers from intrinsic noise due to single-scattering-point approximation. In our studies, we developed more advanced methods, based on Binned Clustering (BC) algorithm (6). It builds on PoCA and exploits the spatial density of scattering vertices to improve image resolution and quality.

Binned Clustering algorithm

The Binned Clustering algorithm (6) employs spatial density of the scattering vertices to discriminate between materials of different densities. The principle of the BC method is as follows:

1. The volume is divided into voxels (for instance cubes of side length 1 cm) and location of muon scattering vertices is calculated within each voxel.
2. Within each voxel, scattering vertices are sorted into descending order by the scattering angle. The first n entries in the list are kept and the rest discarded. Voxels with less than predefined value of n scattering vertices are discarded.
3. For each pair of vertices i, j in each voxel, a metric value $m_{i,j}$ is calculated as

$$m_{i,j} = \frac{|\mathbf{v}_i - \mathbf{v}_j|}{(\theta_i p_i)(\theta_j p_j)}$$

where $\mathbf{V}_i$, θ_i and p_i are, respectively, the scattering vertex position, scattering angle, and momentum of muon *i*. Then, $|\mathbf{V}_i - \mathbf{V}_j|$ is a matric distance between vertex *i* and *j*.

For high-Z materials the density of the scattering vertices is higher (this the distance between voxels is shorter) and scattering angles are larger thus the metric value is lower.

The original BC method uses median of the $\ln(m_{i,j})$ distribution in a voxel as a material-discriminating variable. In our studies, we adopted slightly modified definitions of the discriminator to maximize performance for low-Z and high-Z object identification. We provide details in Sec. 5.1 and 6.1.

3. CHANCE Muon Scattering Tomography system

The CHANCE muon tomography system was build using Resistive Plane Chamber (RPC) and Drift Chamber detectors. For the details of the system configuration and its performance, please see **D4.1: Report on Performance of the Muon Detector System**. Here we provide only a short description and information relevant for material identification studies presented in this report.

The CHANCE muon tomography system consists of 30 Resistive Plate Chambers (RPCs), 18 Drift Chambers, and plastic scintillator trigger panels. The panels are located in two perpendicular orientations, namely X and Y: each orientation detects hits in the (X, Z) and (Y, Z) planes, respectively, together forming a 3D track. The detector was operated in a non-laboratory environment, at the Fenswood Farm, 5 miles south-west of Bristol, UK. Each RPC plane is 200 × 200 cm.

During the projects, there were 2 different experimental configurations used. **Configuration A** corresponds to trigger panels, drift chambers and 4 layers of RPCs, and **Configuration B** consists of trigger panels, drift chambers and 5 layers of RPCs. Figure 2 (right) shows the Configuration A of the system, while Configuration B is presented in Figure 2 (left). Figure 2 (left) also shows a mock-up drum during the experimental program. The drum has a diameter of approximately 66 cm and a length of 88 cm, and it was positioned in the center of CHANCE muon tomograph system.

To quantify the system's performance in a size and position reconstruction and quality of the material identification, we placed objects of known material inside the measurement area during the data-taking campaign. We used blocks of lead, tungsten, steel, and aluminum; each of them had a size of approximately 5x5x5 cm, and there were located alongside the mock-up drum.

Figure 2: The CHANCE muon detector: the configuration A (right) and configuration B (left). See text for details.

4. Configuration of the muon tomography system used in simulation studies

We did quantitative studies of the performance of muon tomography algorithms in the reconstruction of known blocks of materials using realistic simulations of the CHANCE muon detector using GEANT4 (7; 8) model. GEANT4 is a tool for Monte Carlo simulation of detector performance used widely in particle and nuclear physics, including medical physics. For example, such simulations are used for the calculation of detector efficiencies and performance by experimental collaborations at the Large Hadron at CERN and Relativistic Heavy Ion Collider in Brookhaven National Laboratory. Thus, the Monte Carlo studies are robust (in fact, GEANT4 is an industry standard), and allow us for reliable performance studies of the algorithms that use muon tomography for the identification of the shape, size, and material content of the bodies in conditioned waste packages.

The muon tomography system in our studies closely reassembles the original configuration (Configuration A) of the CHANCE muon detector (see CHANCE **D4.1: Report on Performance of the Muon Detector System**) and it is presented in Figure 3. It consists of layers of Resistive Plane Chambers (RPCs), drift chambers, and plastic scintillator detectors used as trigger devices. The spatial resolution is 350 μm for RPCs and 2 mm for the drift chambers. The detector modules are approximately 2 m x 2 m.

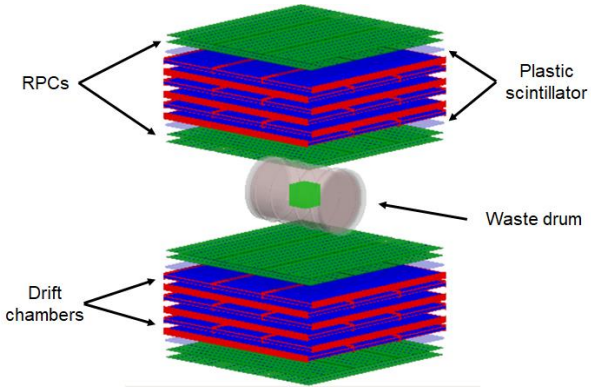

Figure 3: Muon scattering tomography detector system simulated with GEANT4 within the CRESTA framework.

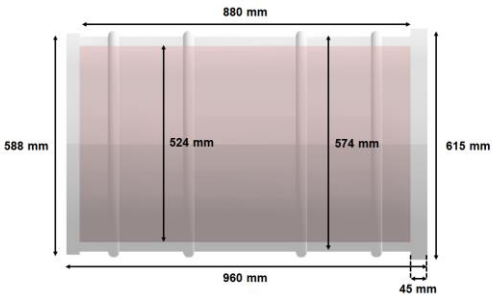

Figure 4: The simulated steel nuclear waste drum

The waste container is a drum made of 25 mm thick steel with a diameter of approximately 60 cm and a length of about 100 cm (see Figure 4). We simulated different cases where the matrix material in the drum was either concrete (density of 2.3 g cm^{-3}) or bitumen (with density of 1.35 – 1.41 g cm^{-3}). The study was performed using cosmic-ray simulation framework CRESTA (9), which is built on the Geant4 particle physics toolkit and the CRY cosmic ray library.

5. High-Z material identification in nuclear waste drums

We developed a method for high-Z material identification that uses multivariate analysis techniques to locate and identify materials in nuclear waste drums (3). We investigate its performance for iron, lead, and uranium samples within the concrete matrix.

5.1 Method for material identification using Muon Scattering Tomography and Multivariate Analysis

The method employs Binned Clustering (BC) algorithm (6), which uses spatial density of the scattering vertices to discriminate between materials of different densities. We use the BC algorithm to build discriminator variable as follows:

1. The volume is divided into voxels of side length 1 cm and location of muon scattering vertices is calculated within each voxel.
2. Within each voxel, scattering vertices are sorted into descending order by the scattering angle. The first $n = 5$ entries in the list are kept and the rest discarded. Voxels with less than $n = 5$ scattering vertices are discarded.
3. For each pair of vertices i, j in each voxel, a metric value $m_{i,j}$ is calculated as

$$m_{i,j} = \frac{|\mathbf{v}_i - \mathbf{v}_j|}{(\theta_i p_i)(\theta_j p_j)}$$

   Where $\mathbf{V}_i$, θ_i and p_i are, respectively, the scattering vertex position, scattering angle, and momentum of muon i. For high-Z materials the density of the scattering vertices is higher (this the distance between voxels is shorter) and scattering angles are larger thus the metric value is lower.
4. The distribution of $\log(m_{i,j})$ values within each voxel is then used as input for Multivariate Analysis using the TMVA machine learning package.

The initial analysis of the performance of available machine learning methods showed that gradient-boosted decision trees offer the best performance; thus, we used this approach in our studies. After the initial decision regarding the voxel content and discrimination between the concrete matrix and high-Z material of interest, the voxels were clustered to reassemble bodies stored in the waste drums. We used *k*-means++ clustering algorithm for this purpose. Finally, we applied multivariate analysis to these clusters to identify which material they consist of. Please refer to (3) for details of the clustering and the MVA analysis.

5.2 Performance of size and position reconstruction and quality of the of high-Z material identification

To evaluate the performance of the method for material identification, we simulated objects made of uranium, lead, and iron of different shapes and sizes dispersed throughout the waste drum. For each test case, we simulated muon track data corresponding to a 10-day exposure.

The first round of validation consists of simulations of 15 and 10 cm cubes of uranium, lead, and iron aligned with the voxel grid. Figure 5 shows the result of the test for 10 cm objects, together with assigned material scores, i.e. the probability that a given object is made of U, Pb or Fe, respectively. For each object the method found its location and assigned the correct material.

Figure 6 shows the material identification result for more complicated geometry, where U, Fe and Pb objects have different sizes, locations and rotation. In this case, the method has accurately identified the material for each object.

Figure 7 present a further test, where objects of various shapes, sizes, orientation, and rotation are distributed in the drum. This is a much more challenging case, but the method still performs well.

The reconstructed clusters reassemble true locations of the stored objects, and the method correctly assigned material for four out of five simulated bodies: both uranium objects, the iron sphere, and one of the lead objects (no. 4). The lead tube (object no. 2) has been incorrectly identified as iron, which may indicate a limitation of the method in determination the materials of non-spherical objects.

5.1 Sensitivity of the method and false positive rate

The method sensitivity was evaluated by using a set of 100 simulations of a waste drum with three spheres of radius 6 cm, randomly located throughout the drum but constrained not to intersect each other. Fifty simulations contained one U, one Pb and one Fe object, and the remaining consisted of two lead spheres and one iron sphere. A true positive identification of a uranium object was defined as the case when an object is identified close to the true location of the simulated U sphere, and the method allocated uranium to that object. The false positive corresponds to the case when the algorithm assigns uranium to a body that does not contain uranium. We found the method sensitivity to uranium is $0.90^{+0.07}_{-0.12}$ and a false positive rate of $0.12^{+0.12}_{-0.07}$ (quoted as 95% Clopper-Pearson confidence intervals).

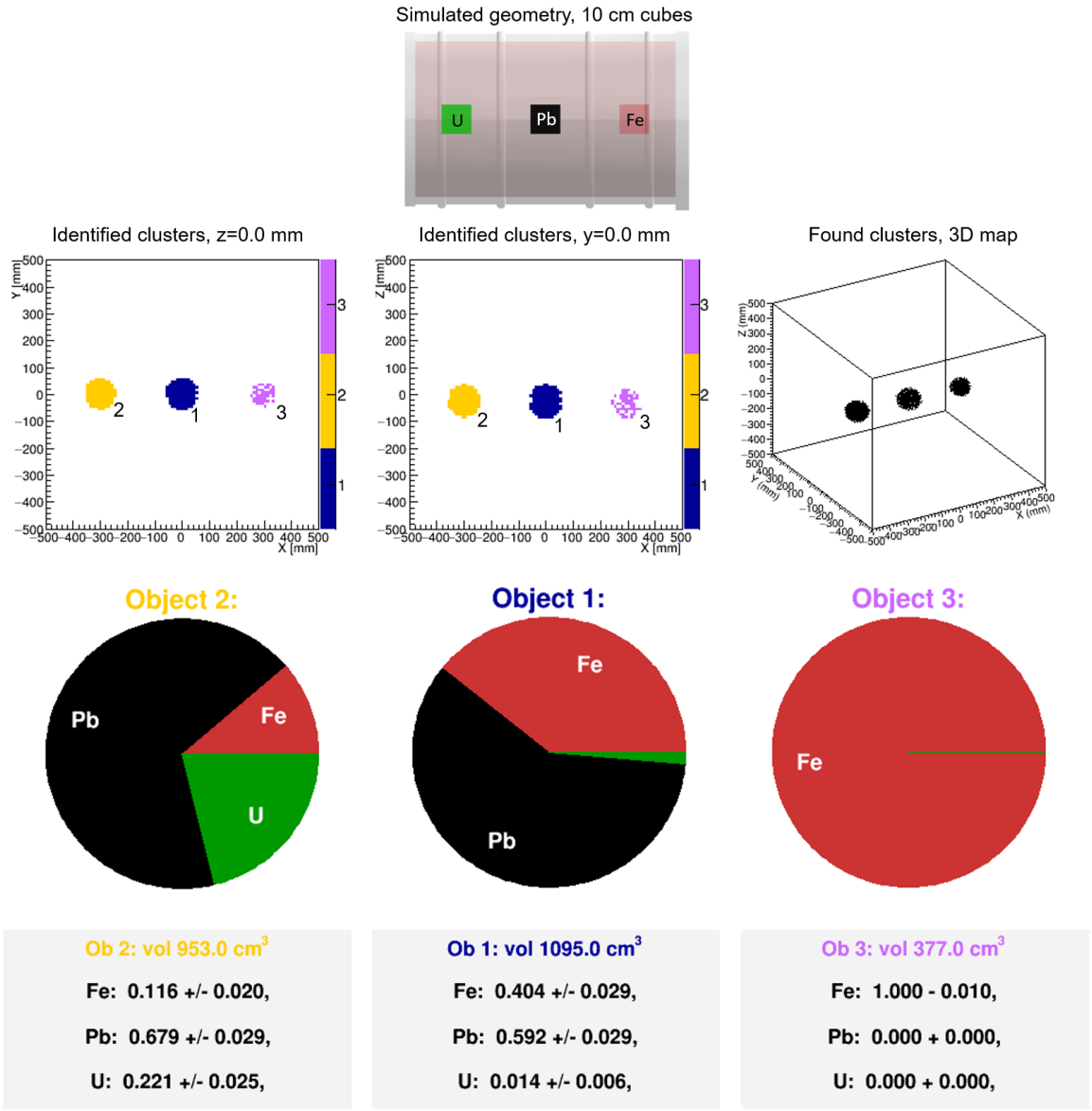

Figure 5: Material identification results for 10 cm cubes, uranium, lead and iron, aligned with the voxel grid. The method has accurately identified the correct material for each object.

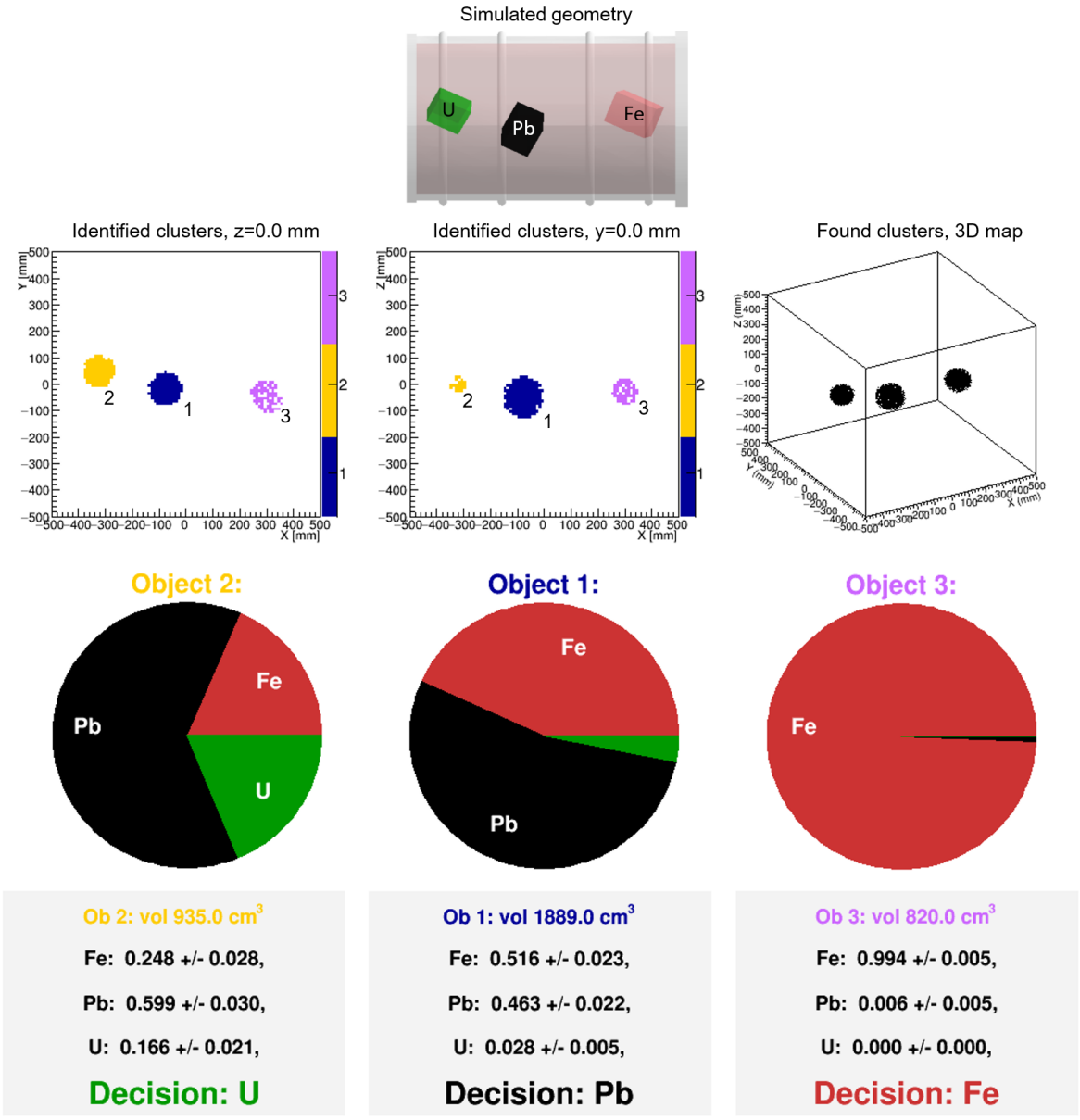

Figure 6: Material scores for three objects made of uranium, lead and iron of different size, location and rotation. The method has accurately identified the correct material for each object.

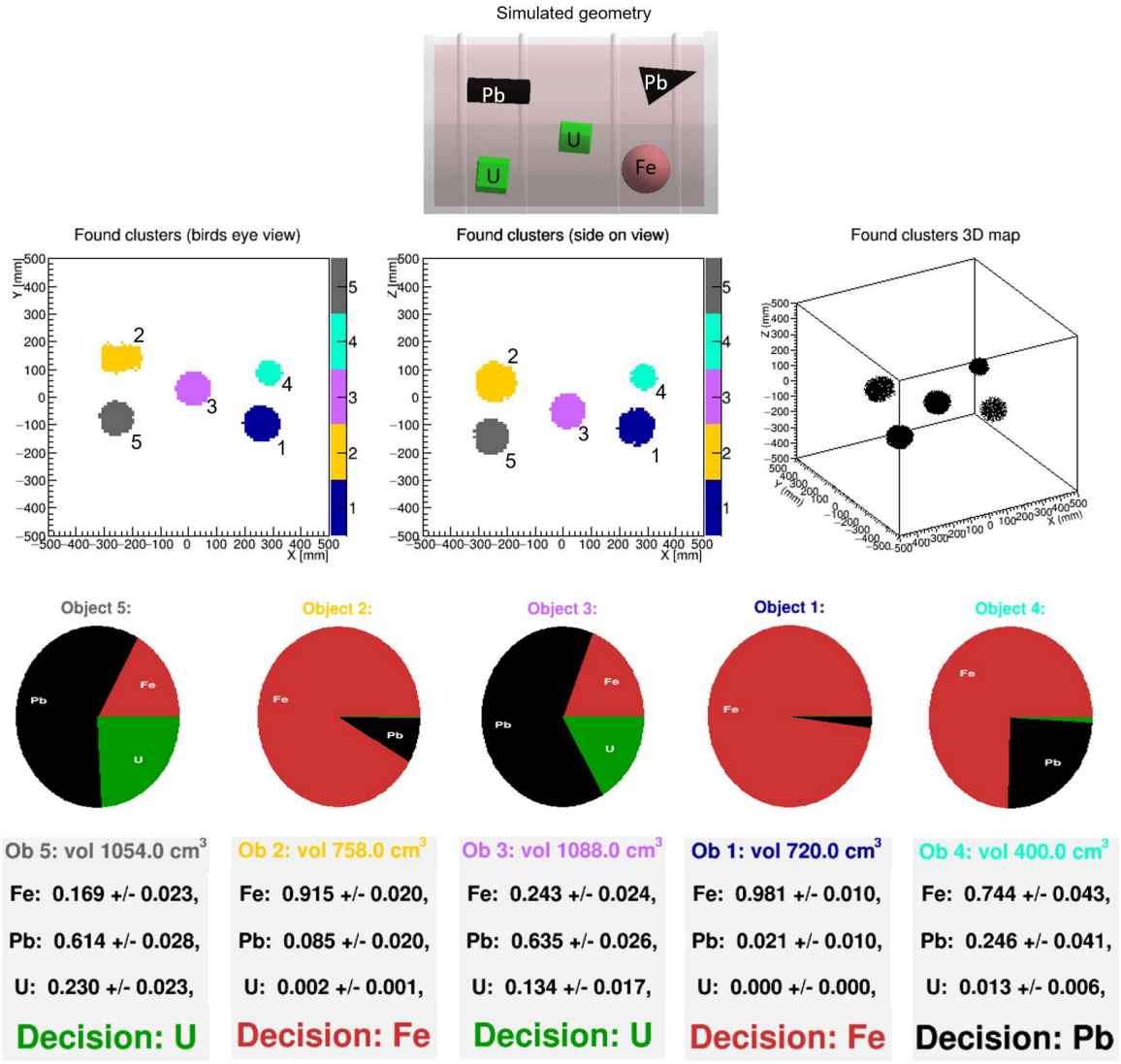

Figure 7: Material scores for three objects made of uranium, lead and iron of different shape, size, and location. The method has accurately identified four objects; one lead object (no. 2) was incorrectly classified as iron.

6. Gas bubble (low-Z material) identification in nuclear waste

During the CHANCE project, a new approach to gas bubble and void detection was developed. It is based on voxel-by-voxel identification of low-density materials and allows for the detection of the location of gas bubbles (or voids) and calculation of their volume.

6.1 Method for gas bubble detection using muon scattering tomography

The principle of the method is similar to the one described in *5.1 Method for material identification using Muon Scattering Tomography and Multivariate Analysis:*

1. The volume is divided into voxels of side length 3 cm. To select a default voxel size, we analyzed Receiver Operating Characteristic (ROC) Curves for signal and background separation for 1, 2, 3, 4, and 5 cm voxels. The 3-cm voxel case provided the best ROC (measured as the area under the curve); thus, we used 3-cm voxels in our studies.
2. The location of muon scattering vertices is calculated within each voxel.
3. Within each voxel, scattering vertices are sorted into descending order by the scattering angle.
4. For each pair of vertices i, j in each voxel, a metric value $m_{i,j}$ is calculated as

$$m_{i,j} = \frac{|\mathbf{v}_i - \mathbf{v}_j|}{(\theta_i)(\theta_j)}$$

   where $\mathbf{V}_i$ and θ_i are the scattering vertex position and scattering angle of muon i.
5. The median of distribution of $m_{i,j}$ values (*Median Metric*) within each voxel is then taken as an observable for material discrimination between gas and other materials.

To establish the decision threshold for gas detection, we analyzed distribution of Metric Median for matrix material (bitumen, density 1.41 g cm^{-3}) and hydrogen. Figure 8 shows these distributions (in black for bitumen-filled voxels and in read for hydrogen voxels), and Figure 9 presents the efficiency and purity of hydrogen voxel identification as a function of a threshold applied to *Median Metric* values. The threshold was selected such purity and efficiency are maximized. Since the value of *Median Metric* may depend on location in the waste container, the drum was divided into 10-cm wide regions (see Figure 10), where such an analysis was conducted separately to calculate the location-dependent decision threshold. The efficiency of hydrogen detection is better than 90% for each of these regions, with false-positive rate lower than 10%.

Such an approach automatically removes any object with a density larger than the matrix since for higher-Z material the median of the metric $m_{i,j}$ distribution will have a lower *Median Metric*. Thus, its takes into account both the bitumen and blocks of higher-Z substances, like iron or uranium.

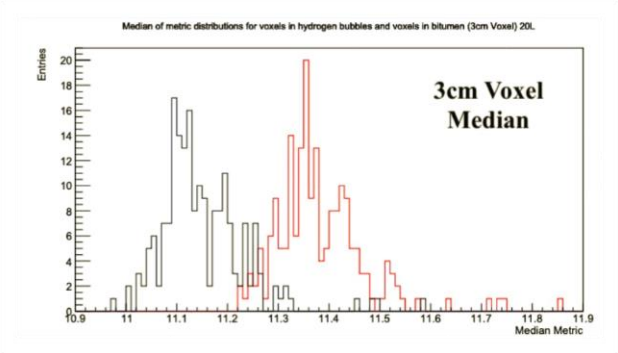

Figure 8: Distribution of Median Metric for matrix material (bitumen, **in black**) and hydrogen (**in red**) for 3 cm voxels.

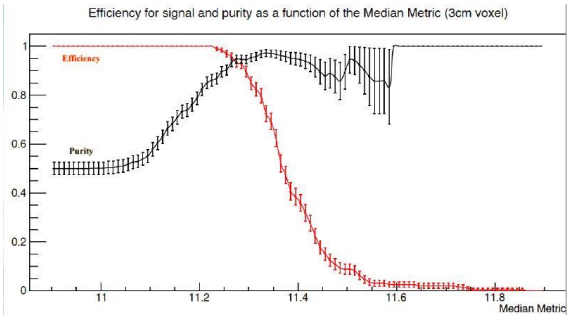

Figure 9: The efficiency and purity of hydrogen detection in the bituminized waste container as a function of decision threshold of Median Metric.

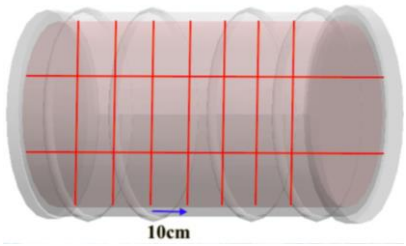

Figure 10: Illustration of the selection of the areas used to establish threshold for separation of hydrogen from waste package matrix and other higher-Z materials.

6.1 Performance of volume and position reconstruction and quality of the of gas bubbles (low-Z material) identification

Figure 11 shows a 3D image of two reconstructed hydrogen bubbles, which were simulated within the bituminized waste drum. Their location matches well the simulated objects.

To evaluate the precision of the method and its detection limits, we simulated different hydrogen volumes within waste container filled with bitumen. Figure 12 shows the obtained reconstructed volume vs the simulated (true) one. The data points show the results calculations, the line represent a linear function fit together with its uncertainties (one- and two-standard-deviation contours).

Our initial investigation indicates that the relative uncertainty on hydrogen volume measurement using muon scattering tomography and voxel-by-voxel approach is below 10% for bubbles larger than 0.85 L. Based on the fit result and its uncertainty, we estimated the detection limit of this method be 0.55 liter at a 95% confidence level.

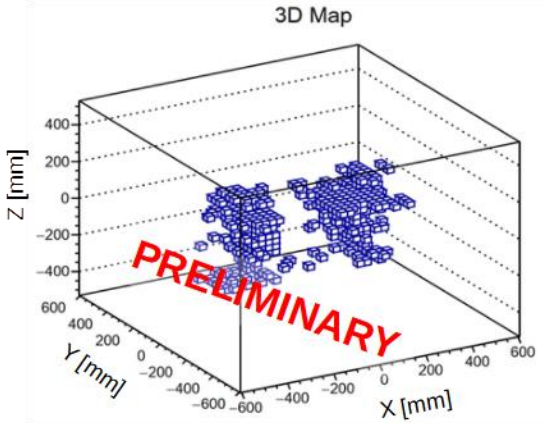

Figure 11: 3D image of two 4-L hydrogen bubbles, extracted for a case where a bitumen-filled drum with two bubbles was simulated.

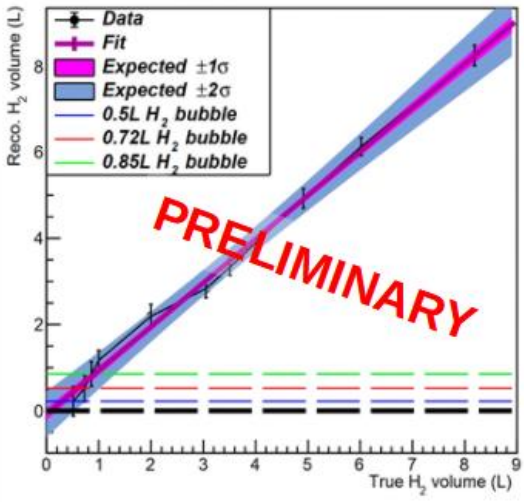

Figure 12: Reconstructed hydrogen volume vs simulated amount of gas for a bituminized waste container.

7. Experimental results

We analyzed experimental data separately for Configuration A and B of the CHANCE muon systems, as described in Section 3. We started with a simple PoCA method, where the scattering points were reconstructed independently in XZ and YZ planes.

The scattering vertex is taken as an intersection of two tracks registered in the top and the bottom parts of the system. Each of these tracks is reconstructed independently in XZ or YZ plane. We required at least two hits in the top or bottom detector for each track, respectively. The image is then created as a density map of the PoCA scattering vertices within the CHANCE muon system geometry.

Figure 13 - Figure 24 show the PoCA image (the density of the reconstructed scattering vertices) for XZ and YZ projections of the data collected with the CHANCE detector with two hardware configurations, i.e. Configuration A and B, respectively. The rectangular and circular shapes indicate the position of the mock-up waste drum during the data taking.

We first used small voxels (2 cm x 2 cm) to check if we can obtain a high-resolution image of the drum. Figure 13 - Figure 16 present results for all available tracks, while results in Figure 17 and Figure 18 are for tracks with the reconstructed scattering angle less than 15º. With such small bins, the track sample per voxel is low, and we next used bigger voxels of 20 cm x 20 cm to try to remedy this issue. Figure 19 - Figure 24 show results with large voxels, again for all tracks (Figure 19 - Figure 22) and those with the reconstructed scattering angle less than 15º (Figure 23 - Figure 24).

As explained in the D4.1 report, we were caught by surprise by a Freon ban that came into force at the early phase of CHANCE. This forced us to use CO_2 in the RPC system. CO_2 yields a much lower hit efficiency. In order to reconstruct tracks, hits in all traversed layers are required. Hence, the efficiency

to detect tracks reduces by the product of the efficiency of all layers. This led to a very small track sample. Due to limited statistics, we were not able to perform more differential experimental studies of the performance of methods of material identification we had developed for the CHANCE muon scattering tomography system.

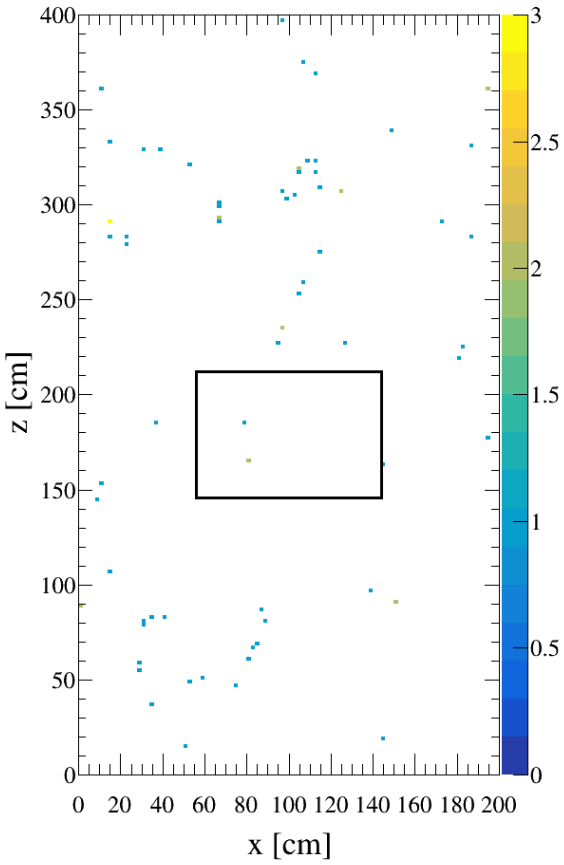

Figure 13: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the XZ plane. Results for Configuration B of the CHANCE muon detector and voxel size of 2 cm x 2 cm. The black rectangle presents the expected location of the mock-up waste drum.

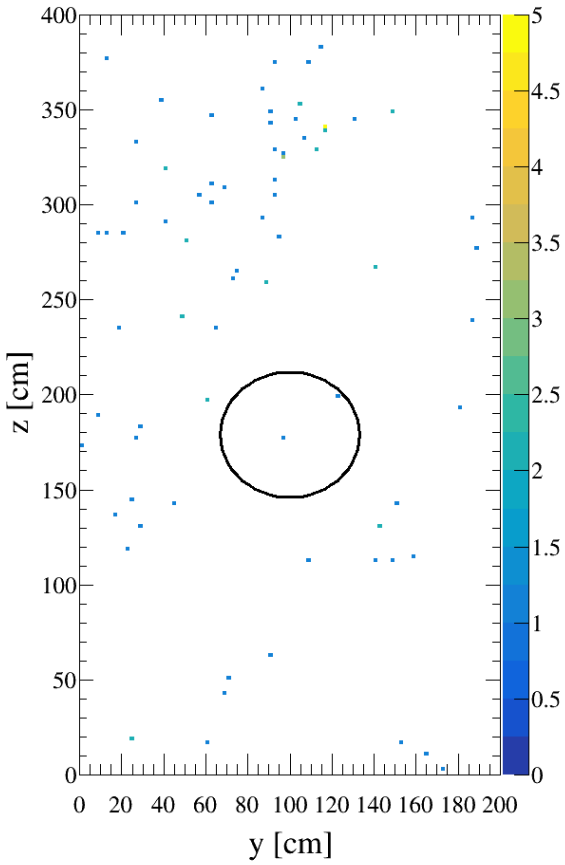

Figure 14: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the YZ plane. Results for Configuration B of the CHANCE muon detector and voxel size of 2 cm x 2 cm. The black circle presents the expected location of the mock-up waste drum.

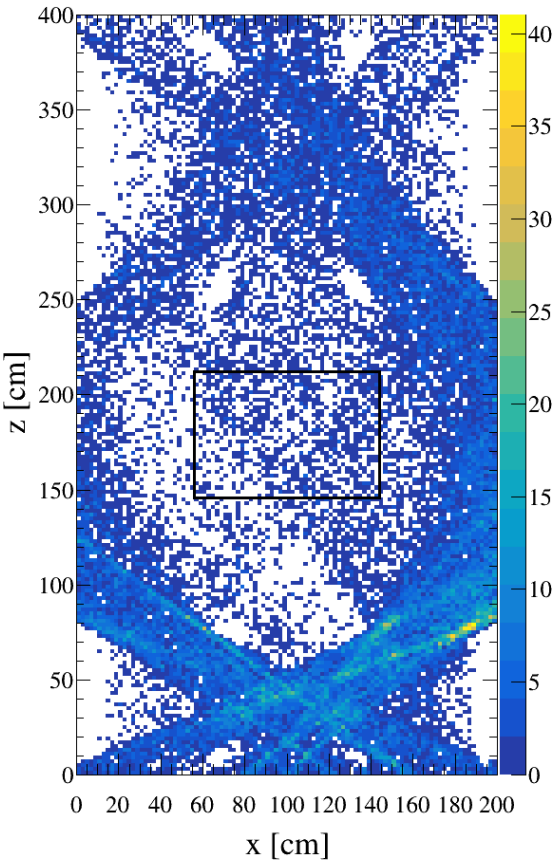

Figure 15: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the XZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 2 cm x 2 cm. The black rectangle presents the expected location of the mock-up waste drum.

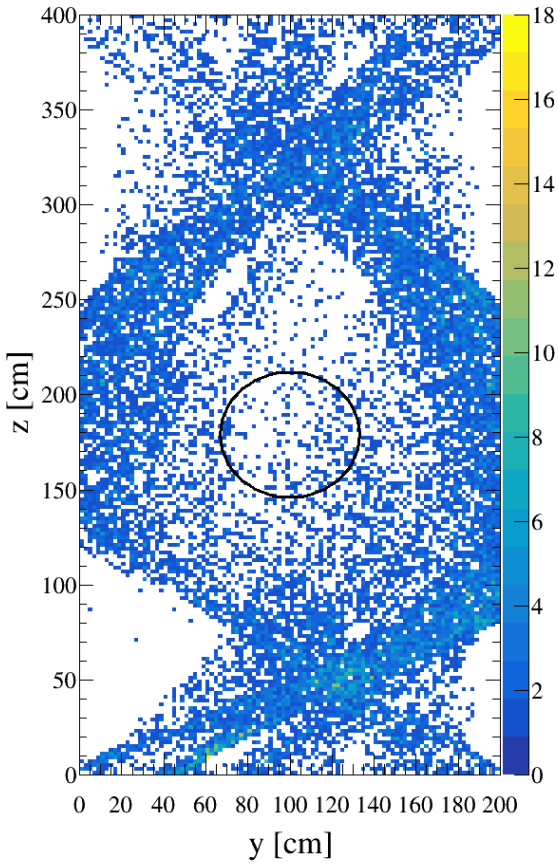

Figure 16: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the YZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 2 cm x 2 cm. The black circle presents the expected location of the mock-up waste drum.

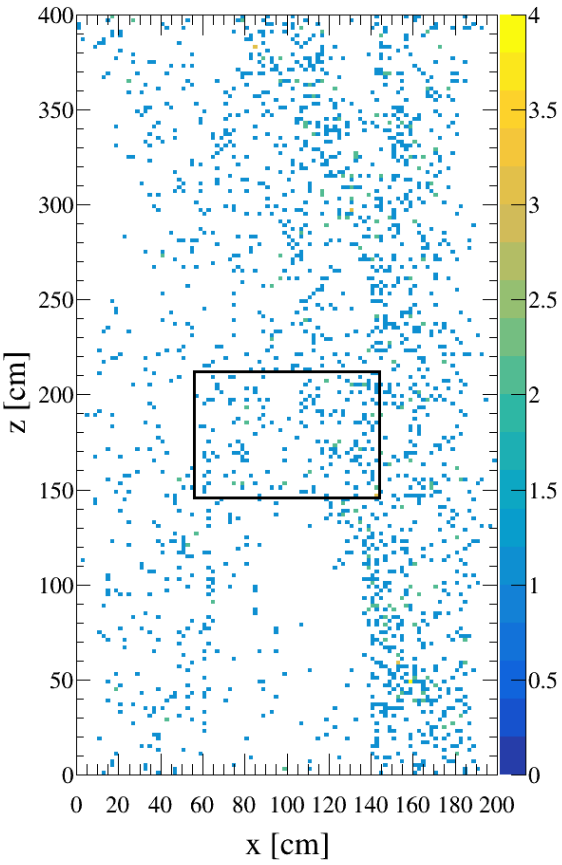

Figure 17: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the XZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 2 cm x 2 cm. Only candidates with scattering angle less than 15º were plotted. The black rectangle presents the expected location of the mock-up waste drum.

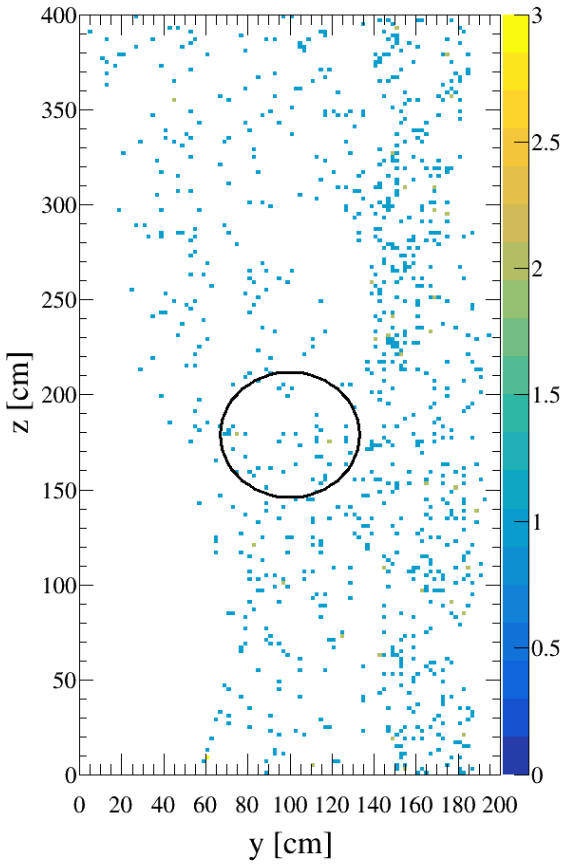

Figure 18: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the YZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 2 cm x 2 cm. Only candidates with scattering angle less than 15º were plotted. The black circle presents the expected location of the mock-up waste drum.

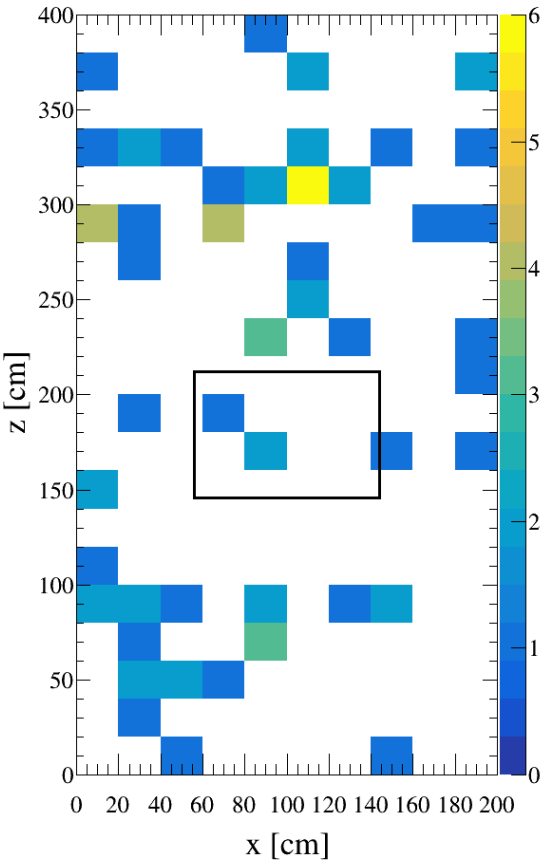

Figure 19: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the XZ plane. Results for Configuration B of the CHANCE muon detector and voxel size of 20 cm x 20 cm. The black rectangle presents the expected location of the mock-up waste drum.

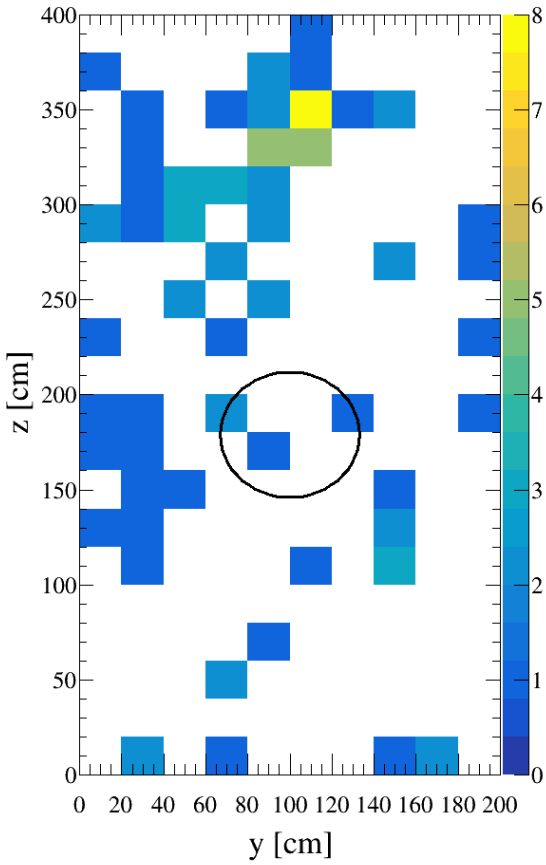

Figure 20: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the YZ plane. Results for Configuration B of the CHANCE muon detector and voxel size of 20 cm x 20 cm. The black circle presents the expected location of the mock-up waste drum.

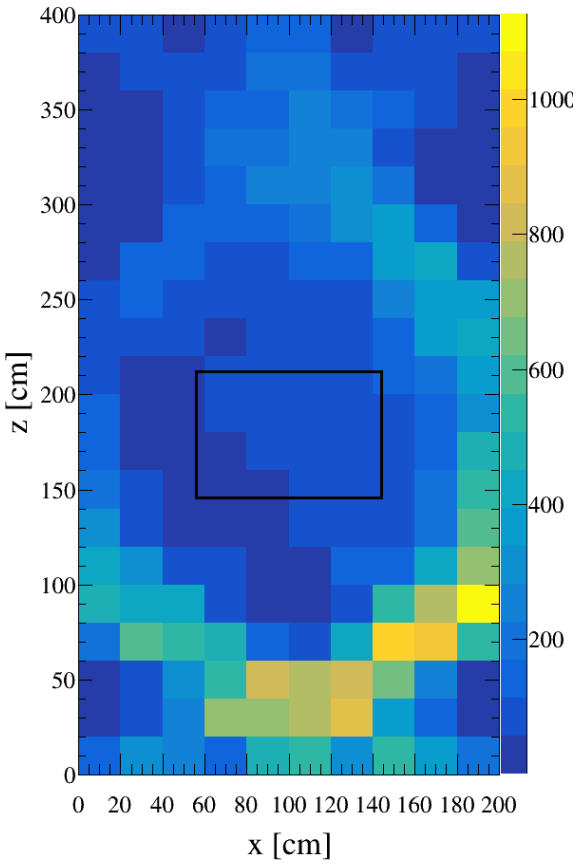

Figure 21: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the XZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 20 cm x 20 cm. The black rectangle presents the expected location of the mock-up waste drum.

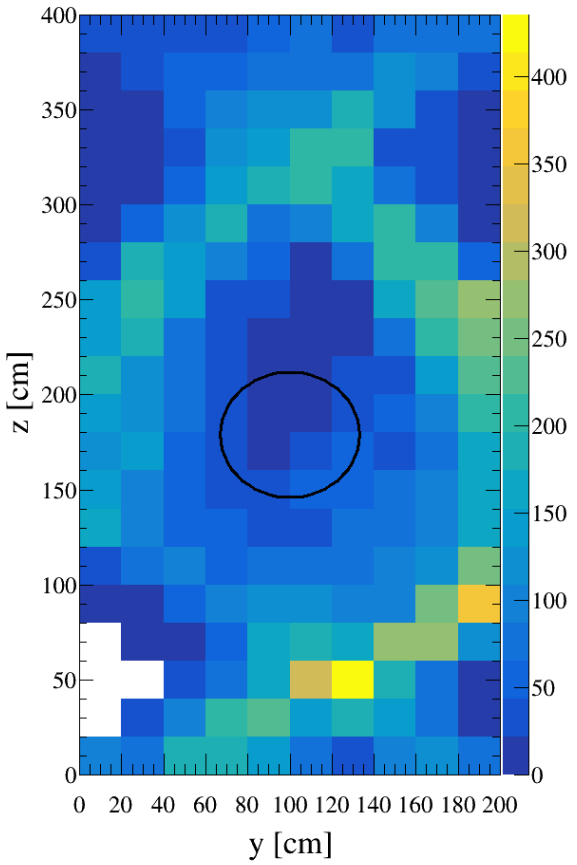

Figure 22: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the YZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 20 cm x 20 cm. The black circle presents the expected location of the mock-up waste drum.

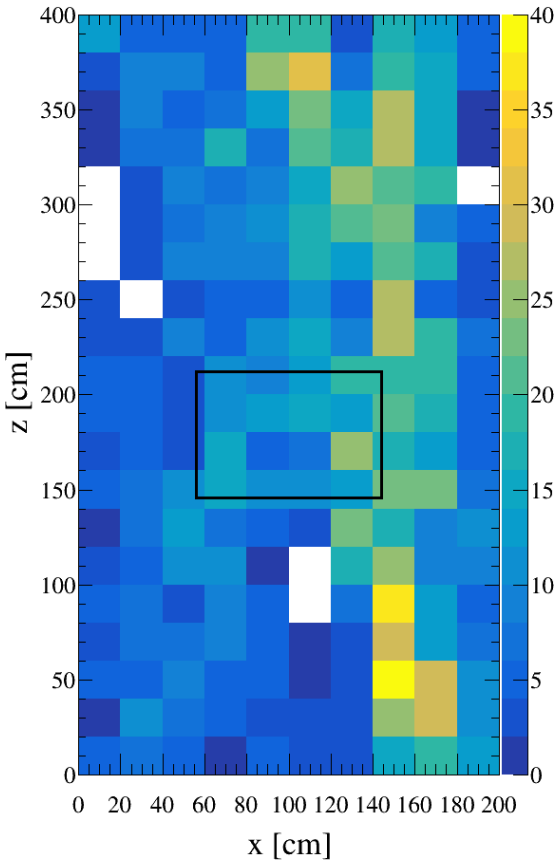

Figure 23: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the XZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 20 cm x 20 cm. Only candidates with scattering angle less than 15º were plotted. The black rectangle presents the expected location of the mock-up waste drum.

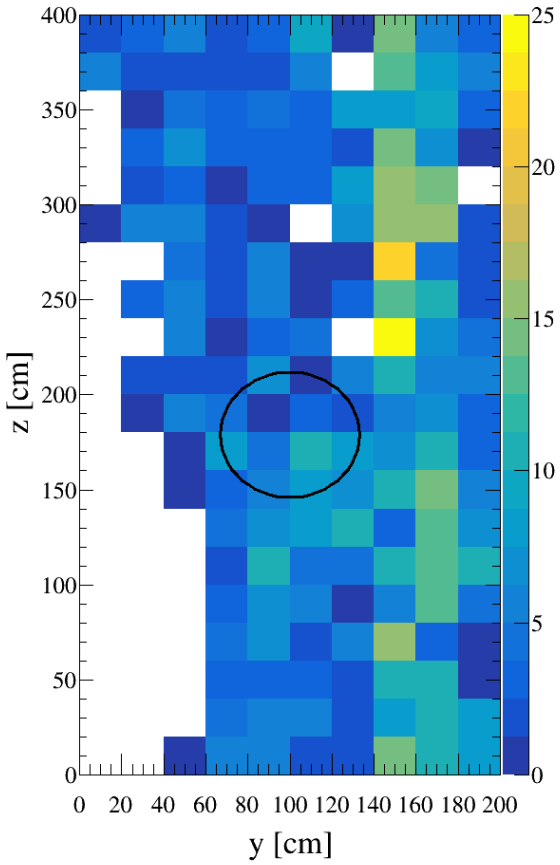

Figure 24: Distribution of scattering vertices reconstructed using the PoCA algorithm with CHANCE muon tomography system in the YZ plane. Results for Configuration A of the CHANCE muon detector and voxel size of 20 cm x 20 cm. Only candidates with scattering angle less than 15º were plotted. The black circle presents the expected location of the mock-up waste drum.

8. Summary

We analyzed performance of the CHANCE muon tomography system in the reconstruction of known blocks of material, with focus on high-Z elements, and gas bubbles (thus low-Z and low-density objects).

We presented the studies done using realistic Monte Carlo simulations of the CHANCE muon detector. They allowed us for reliable performance evaluation of the algorithms that we developed during the project. Our quantitative analysis showed the method for high-Z material identification performs well in the recognition of blocks of uranium, lead and iron in the conditioned nuclear waste packages. We found the method sensitivity to uranium is $0.90^{+0.07}_{-0.12}$ and a false positive rate of $0.12^{+0.12}_{-0.07}$ (quoted as 95% Clopper-Pearson confidence intervals).

The initial results of a novel method of gas detection in bituminized waste containers are very promising, pointing into a gas detection limit of 0.55 L at 95% confidence level.

Unfortunately, due to the limited size of our track sample, we were not able to perform more differential experimental studies of the performance of methods of material identification we had developed for the CHANCE muon scattering tomography system.

Bibliography

1. *Material Identification of Bodies Stored in Nuclear Waste Drums using Muon Scattering Tomography and Multivariate Analysis.* **M. Weekes, A. Alrheli, D. Barker, D. Kikoła, A. Kopp, M. Mhaidra, C. de Sio, P. Stowell, L. Thompson, J. Velthuis.** Phoenix, Arizona, US : s.n., 2021. Waste Management Symposia WM2021. 21102.

2. *Analytical Muon Scattering Tomography Methods for a Nuclear Waste Imagining Detector.* **D. Barker, A. Alrheli,Ch. De Sio, D. Kikoła, A. Kopp,M. Mhaidra, J.P. Stowell, L. Thompson, J. Velthuis, M. Weekes.** 2021. 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference.

3. **M.J. Weekes, A.F. Alrheli, D. Barker, D. Kikoła, A.K. Kopp, M. Mhaidra J.P. Stowell, L.F. Thompson and J.J. Velthuis.** Material identification in nuclear waste drums using muon scattering tomography and multivariate analysis. *Journal of Instrumentation.* 2021, Vol. 16, p. P05007.

4. **Eidelman, Simon et al.** Review of particle physics. *Physics Letters B.* 2004, Vol. 592.1.

5. **L. Frazão, J. Velthuis, C. Thomay and C. Steer.** Discrimination of high-Z materials in concrete-filled containers using muon scattering tomography. *Journal of Instrumentation.* 2016, Vol. 11, p. P07020.

6. **Thomay, C., Velthuis, J.J., Baesso, P., Cussans, D., Morris, P.A.W., Steer, C., Burns, J., Quillin, S. and Stapleton, M.** A binned clustering algorithm to detect high-Z material using cosmic muons. *Journal of Instrumentation.* 2013, Vol. 8, p. P10013.

7. **J. Allison, K. Amako, J. Apostolakis, P. Arce, M. Asai, T. Aso, E. Bagli, A. Bagulya, S. Banerjee, G. Barrand, B.R. Beck, A.G. Bogdanov, D. Brandt, J.M.C. Brown, H. Burkhardt, Ph. Canal, D. Cano-Ott, S. Chauvie, K. Cho, G.A.P. Cirrone, G. Cooperman, M.A.** Recent developments in Geant4. *Nuclear Instruments and Methods in Physics Research A.* 2016, Vol. 835, pp. 186-225.

8. **S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo,.** Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A.* 2003, Vol. 506, 3, pp. 250-303.

9. **C. Steer, P. Stowell and L. Thompson,.** CRESTA: Cosmic rays for engineering, scientific, and technology applications. *https://gitlab.com/cosmicraysim/cresta.* [En ligne]

10. *A robust method to find gas bubbles inside large nuclear waste containers using Muon Tomography.* **M. Mhaidra, A. Kopp, A. Alrheli, C. de Sio, D. Barker, P. Stowell, M. Weekes, M. Dobrowolska, D. Kikoła, L. Thompson, J. Velthuis.** Phoenix, Arizona, US : s.n., 2021. Waste Management Symposia WM2021. 2021151.

11. *Figures of Merit for the Application of Muon Tomography to the Characterization of Nuclear Waste Drums.* **P. Stowell, A. Alrheli, D. Kikola, A. Kopp, H. Tietze-Jaensch, M. Mhaidra, L. Thompson, E. Valcke, J. Velthuis, M. Weekes.** Phoenix, Arizona, USA : s.n., 2019. Waste Management Symposia WM2019. 19253.

9. Appendix

Published by IOP Publishing for Sissa Medialab

Received: *December 4, 2020*
Revised: *February 12, 2021*
Accepted: *February 24, 2021*
Published: *May 6, 2021*

2021 JINST 16 P05007

Material identification in nuclear waste drums using muon scattering tomography and multivariate analysis

M.J. Weekes,[a,*] **A.F. Alrheli,**[a] **D. Barker,**[a] **D. Kikoła,**[b] **A.K. Kopp,**[c] **M. Mhaidra,**[b,c] **J.P. Stowell,**[a] **L.F. Thompson**[a] **and J.J. Velthuis**[c]

[a] *University of Sheffield, Department of Physics and Astronomy, Hounsfield Road, Sheffield, S3 7RH, U.K.*
[b] *Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warsaw, Poland*
[c] *University of Bristol, School of Physics, HH Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, U.K.*

E-mail: mweekes1@sheffield.ac.uk

Abstract: The use of muon scattering tomography for the non-invasive characterisation of nuclear waste is well established. We report here on the application of a combination of feature discriminators and multivariate analysis techniques to locate and identify materials in nuclear waste drums. After successful training and optimisation of the algorithms they are then tested on a range of material configurations to assess the system's performance and limitations. The system is able to correctly identify uranium, iron and lead objects on a few cm scale. The system's sensitivity to small uranium objects is also established as $0.90^{+0.07}_{-0.12}$, with a false positive rate of $0.12^{+0.12}_{-0.07}$.

Keywords: Pattern recognition, cluster finding, calibration and fitting methods; Search for radioactive and fissile materials; Particle tracking detectors

*Corresponding author.

https://doi.org/10.1088/1748-0221/16/05/P05007

Contents

2021 JINST 16 P05007

1 Introduction

It is important to develop non-destructive methods to determine the contents of sealed nuclear waste packages, in order to minimise the risks of environmental contamination and personnel radiation exposure and to allow for more effective safeguarding. Non-Destructive Assay (NDA) techniques in current use include calorimetry and Muon Scattering Tomography (MST).

NDA techniques can analyse drum contents in a variety of ways. For example, calorimetry can be used to measure the mass of nuclear material inside a container by its heat emission [1]. In contrast, MST (with exposure times of several days to weeks) can produce full 3D images of a volume of interest, allowing individual objects inside the drum to be viewed as well as giving information on their atomic number Z and density [2].

Simulation studies are useful tools to assess MST techniques and algorithms; the technique described in this paper was developed and tested via Monte Carlo simulations. It uses MST data in combination with Multi-Variate Analysis (MVA) classifiers and clustering algorithms to approximately identify the locations and shapes of objects stored in a concrete-filled waste drum. Subsequently, additional trained classifiers are applied to each identified object to classify them as 'iron',

'lead', or 'uranium', representing low-threat medium-Z material, low-threat high-Z material, and high-threat high-Z material respectively. The use of these four materials allows three classification problems of interest to be investigated: separation of stored objects from the concrete background, separating medium- and high-Z materials, and distinguishing between two high-Z materials.

Previous applications of machine learning techniques to MST imaging have demonstrated methods for distinguishing between drums containing uranium and lead blocks [3] and for reconstructing the size of uranium blocks [4]. Our system builds on these through the ability to isolate and identify multiple distinct bodies of different materials and sizes in a waste drum. Other previous research into combining machine learning and MST include applications in cargo scanning [5, 6], a related problem for which short exposure times are required.

2 Muon scattering tomography

Cosmic rays interact with the Earth's atmosphere to produce showers of particles, some of which subsequently decay to muons, resulting in a muon flux at sea level of around 1 $\mathrm{cm}^{-2}\,\mathrm{min}^{-1}$ [7]. These cosmic ray muons are highly penetrating due to their large mass and lack of strong interactions. They have an angular distribution that varies approximately as $\cos^2\theta$, where θ is the zenith angle. Muons are also highly sensitive to the atomic number Z of the material they are passing through, making them suitable candidates for tomographic imaging of nuclear waste drums.

Muons undergo multiple elastic Coulomb scatterings in matter, with the projected scattering angles following an approximately Gaussian distribution with width σ given by

$$\sigma \approx \frac{13.6\,\mathrm{MeV}}{\beta c p}\sqrt{X/X_0} \tag{2.1}$$

where β is the muon speed divided by the speed of light in a vacuum, c; p is the muon momentum, X is the thickness of the material and X_0 is the radiation length of the material [8]. The latter is given by

$$X_0 = \frac{716.4A}{Z(Z+1)\ln(287/\sqrt{Z})}\left[\mathrm{g}\cdot\mathrm{cm}^{-2}\right] \tag{2.2}$$

where ρ is the material density and A is atomic mass [9].

A general MST experiment consists of two sets of particle detectors, one above and one below some volume of interest such as a waste drum (see figure 1). Multiple layers of detector are necessary in order to construct a three dimensional trajectory for each muon from the detector hits. This allows the incoming and outgoing trajectories of each muon to be measured and hence the muon scattering angles to be calculated.

Several algorithms have been developed to enable imaging of a volume of interest from MST data. The simplest is the Point of Closest Approach (PoCA) algorithm [10], which models a muon's multiple scatterings as a single scattering at a single point ('scattering vertex'), found by extrapolating the incoming and outgoing tracks into the volume and finding the point which minimises the distance to each. This assumption allows for fast computation at the expense of image quality. A more advanced MST algorithm has been used in this study (see section 2.1) which builds on PoCA by exploiting the spatial density of scattering vertices; a high density of scattering vertices corresponds to the presence of high-Z material as large-angle muon scatterings occur more often in such materials.

2021 JINST 16 P05007

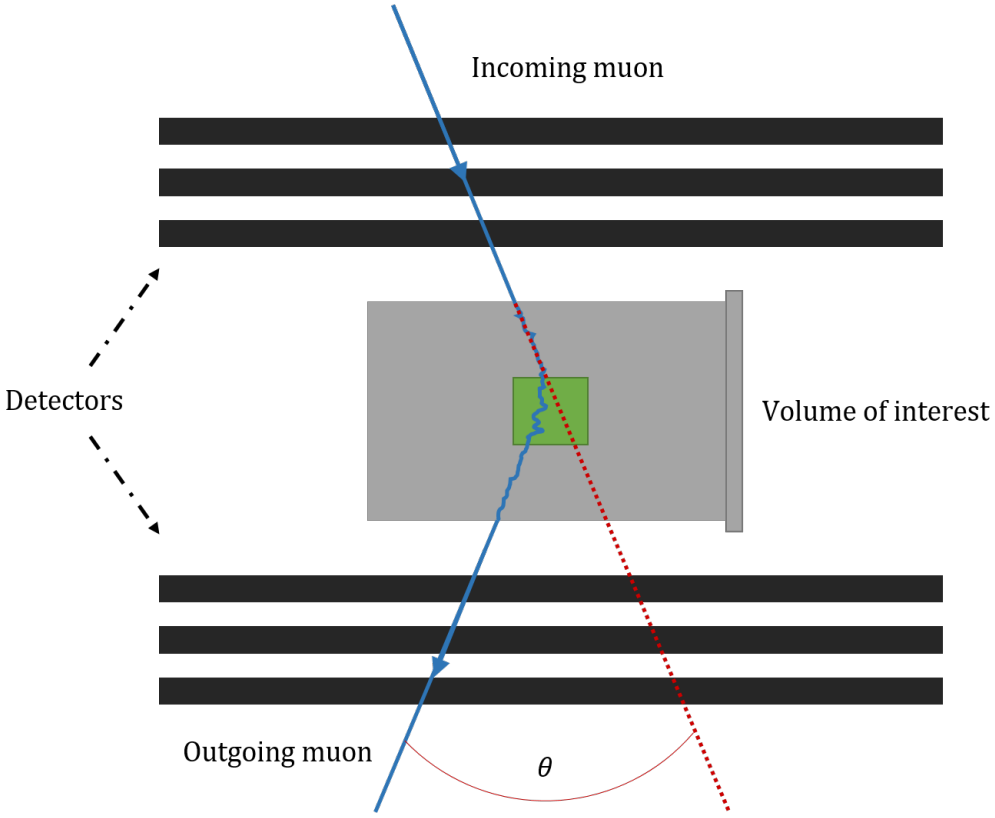

Figure 1. Schematic showing the principle of muon scattering tomography applied to a nuclear waste drum containing a block of high-Z material (in green). Particle detectors measure the trajectories of muons before and after encountering the volume of interest, allowing the scattering angle θ (here exaggerated) to be calculated.

2.1 Binned clustering algorithm

This algorithm, developed in [11], improves on the widely-used Point of Closest Approach (PoCA) muon tomography algorithm [2] by taking into account the degree of spatial clustering of muon scattering vertices. A higher density of vertices corresponds to higher-Z materials (once the muon momentum is accounted for, see below) as strong muon scatterings take place with greater frequency in such materials.

The volume is divided into cubic voxels of side length 1 cm. The incoming and outgoing muon tracks are extrapolated through the volume, and the point at which the distance between the tracks is minimal (the PoCA) is designated as the scattering vertex for the muon. This is repeated for all of the detected muons that encounter the volume of interest. Next, the scattering vertices inside each 1 cm^3 voxel are sorted by the scattering angle of the corresponding muon, and the vertices corresponding to the n largest scattering angles are kept (voxels which contain less than n vertices are discarded). This factor of n is an important tunable parameter of the algorithm. High values of n improve the contrast between high and low-Z materials, as a greater sample of muons are kept, but reduce image 'quality' (i.e. the number of non-empty voxels in the image) as more voxels fall below the cut and are removed from the image.

2021 JINST 16 P05007

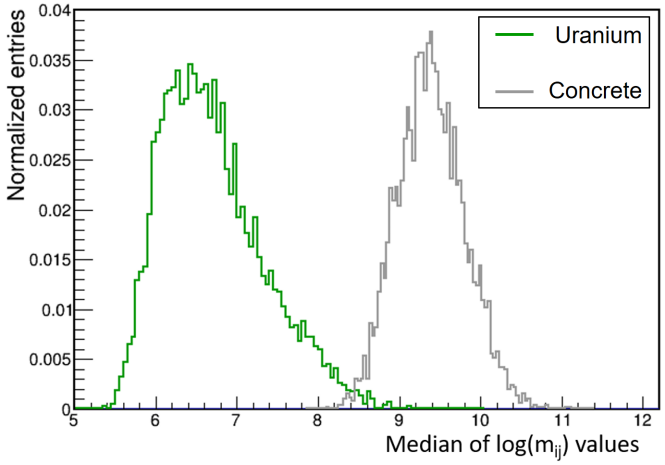

Figure 2. Comparison of distributions of the binned clustering algorithm discriminator, for 20 cm cubes of uranium and concrete. Lower discriminator values correspond to higher Z material.

For each of the $\binom{n}{2}$ pairs of vertices i, j in each voxel, a metric value m_{ij} is calculated according to

$$m_{ij} = \frac{|\mathbf{V}_i - \mathbf{V}_j|}{(\theta_i \tilde{p_i}) \cdot (\theta_j \tilde{p_j})} \tag{2.3}$$

where $\mathbf{V}_i$, θ_i and $\tilde{p_i}$ are respectively the position, scattering angle and normalised (by a factor of 3 GeV/c) momentum of muon i. Weighting by muon momentum is necessary as large scattering angles could indicate low-momentum muons being scattered in low-Z materials instead of strong scattering in high-Z materials. In an experimental system, the muon momentum can be estimated using the muon scatterings between the detector planes, as the planes are of known material and thickness. Following the method of [12], for our simulations the momentum was obtained by adding a smearing factor to the Monte Carlo truth momentum. The smearing factor was drawn from a Gaussian with width 50% of the Monte Carlo truth momentum.

Finally, the median of the distribution of $\log(m_{ij})$ in a voxel is determined; this is the algorithm's discriminator value for that voxel. Comparing the distributions of this discriminator for high- and low-Z materials shows that the discriminator is sensitive to Z (see figure 2).

For imaging purposes, each voxel is filled with its discriminator value as described above, creating a tomogram of the volume of interest. Viewing slices of discriminator values through the image (see figure 3) allows regions of high-Z material to be identified visually. This gives a degree of information about the locations and morphologies of objects stored in the drum. However, it is vulnerable to a vertical smearing effect inherent in the PoCA reconstruction, and without an object of known material for comparison, it is difficult to determine the specific materials of objects 'by eye'. Additionally, without any way to automatically remove background materials such as the steel drum and concrete matrix, the 3D image must be viewed in slices to determine the locations of stored objects.

By default, the binned clustering algorithm only takes into account the median of the $\log(m_{ij})$ distribution in each voxel. To test the possibility that additional material information is encoded in the shape of the $\log(m_{ij})$ value distribution, variables capturing the shape were used to train MVA classifiers. These classifiers are then used to separate the regions of the image corresponding to

2021 JINST 16 P05007

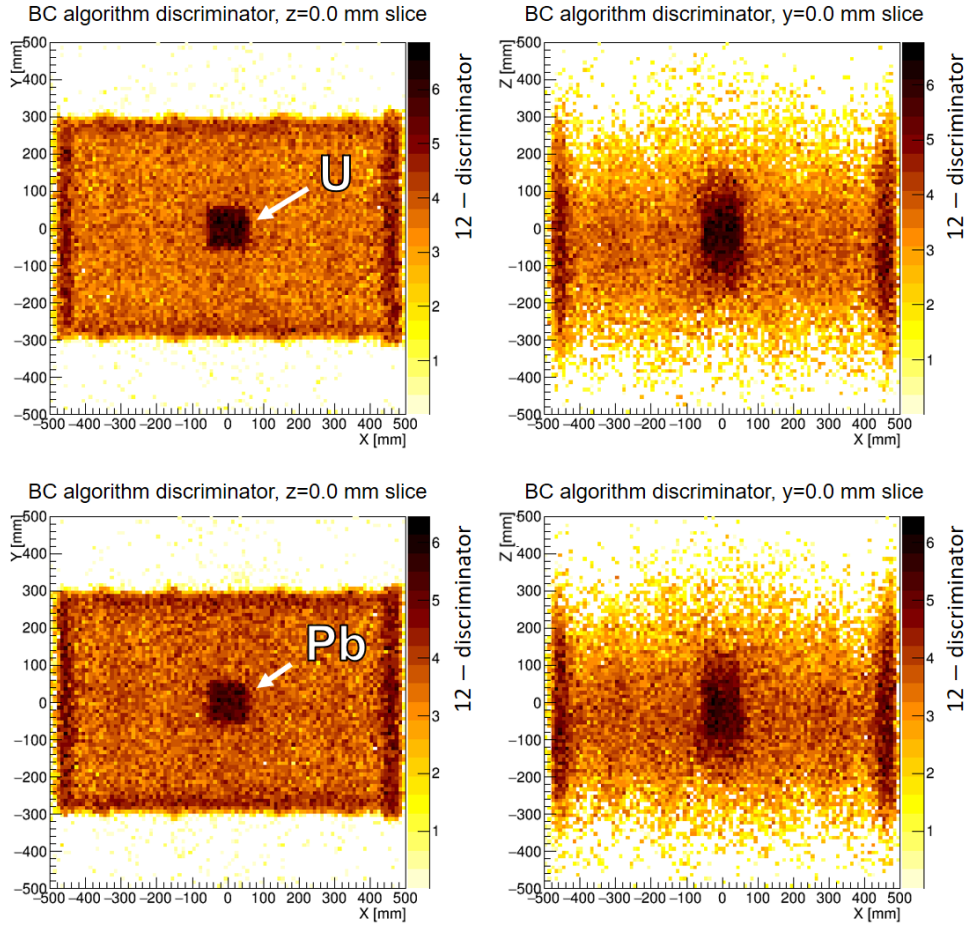

Figure 3. xy (left) and xz (right) slices from binned clustering output images of waste drums containing 10 cm side length cubes of uranium (top) and lead (bottom). Exposure time = 10 days, $n = 5$. The smearing effect along the z axis is due to uncertainty in the scattering vertex z coordinate for tracks with small scattering angles. Note that the plotted discriminator values have been subtracted from 12 for visual clarity.

objects stored in the drum from the concrete matrix. Subsequently the classifiers are used to assign a material to each identified object.

2.2 System configuration

All simulations were performed using CRESTA [13], a cosmic ray simulation platform built on the Geant4 [14] particle physics toolkit and the CRY [15] cosmic ray library. Within CRESTA a MST detector system comprising two particle detector modules above and below a waste drum was simulated (see figure 4). This represents a 'generic' MST detector system, designed for imaging a 1 m waste drum. The detector modules are 2 m by 2 m and each consists of two layers of resistive plate chambers (RPCs), polystyrene scintillator triggers and three layers of drift chambers. The RPCs and drift chambers have spatial resolutions of $\sim$ 350 µm and $\sim$ 2 mm respectively. The detectors are arranged in alternating x and y layers, allowing 3D muon hits to be recorded and the incoming and outgoing tracks reconstructed.

2021 JINST 16 P05007

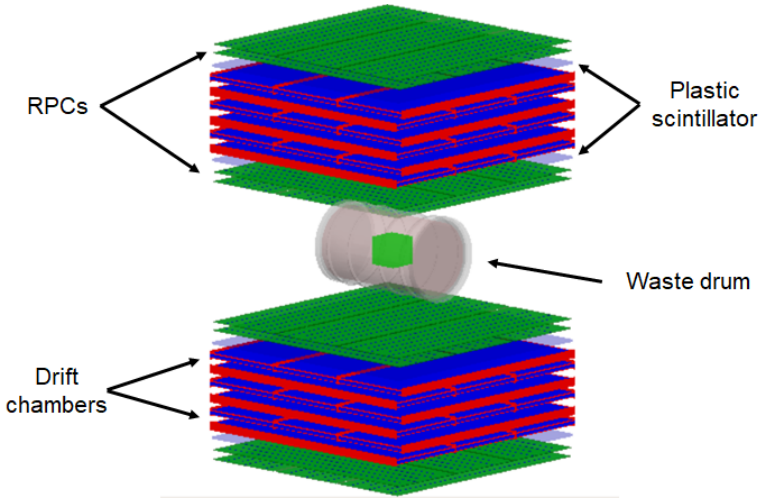

Figure 4. MST detector system simulated in CRESTA: detector modules above and below a waste drum, in which objects can be placed. The detector modules are approximately 2 × 2 m.

The waste drum is made of steel (approx. 91% iron, 9% carbon; element isotopes in natural abundances). It is approximately 100 cm in length and 30 cm in radius (see figure 5 for precise dimensions), and is filled with homogeneous concrete of density 2.3 g cm^{-3}.

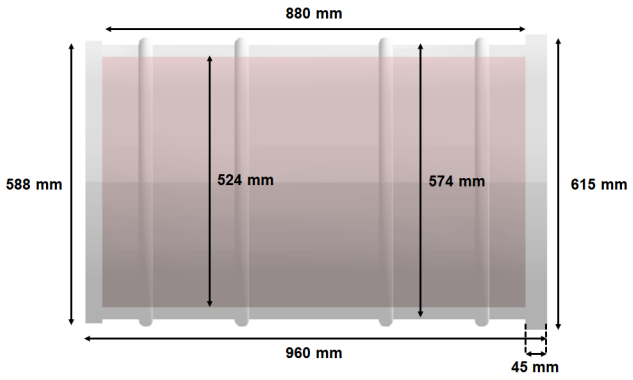

Figure 5. The simulated concrete-filled steel nuclear waste drum used in CRESTA, with its dimensions.

3 Multivariate analysis

3.1 MVAs and muon tomography

Frazão et al. [3] used MVA classifiers trained on simulated MST data to discriminate between waste drums containing lead and uranium blocks. This method can be thought of in a 'global' sense, distinguishing between two categories of waste drum but not analysing the specific drum contents in terms of bodies encased in the concrete. Our approach by contrast is 'local', as we are able to produce localised material information down to the scale of single 1 cm voxels. This approach requires longer exposure times (of the order of several days rather than hours) but gives more

2021 JINST 16 P05007

detailed material information. This allows for the possibility of combining these two techniques. A user could use the former method and a short exposure to identify drums likely to contain high-Z material, then subsequently apply our method and a longer exposure to the flagged drums to identify the stored objects and their materials.

Our MVA classifiers were built, trained and analysed using TMVA [16], a ROOT [17]-integrated machine learning platform. Our set of variables used as input to the MVA classifiers are obtained via the binned clustering algorithm (see section 2.1). The algorithm calculates a set of metric values for each voxel, with each value corresponding to a pair of muon scattering vertices. By default, the algorithm uses the median of the distribution of $\log(m_{ij})$ values only as the discriminator for each voxel. We build on this by first binning the $\log(m_{ij})$ values into 28 bins (see figure 6), calculating the normalised bin counts, and passing the counts to the MVA classifiers as the input variables (figure 7) for that voxel. This approach allows more of the shape of the distribution of metric values to be captured, enhancing the information available to the classifiers.

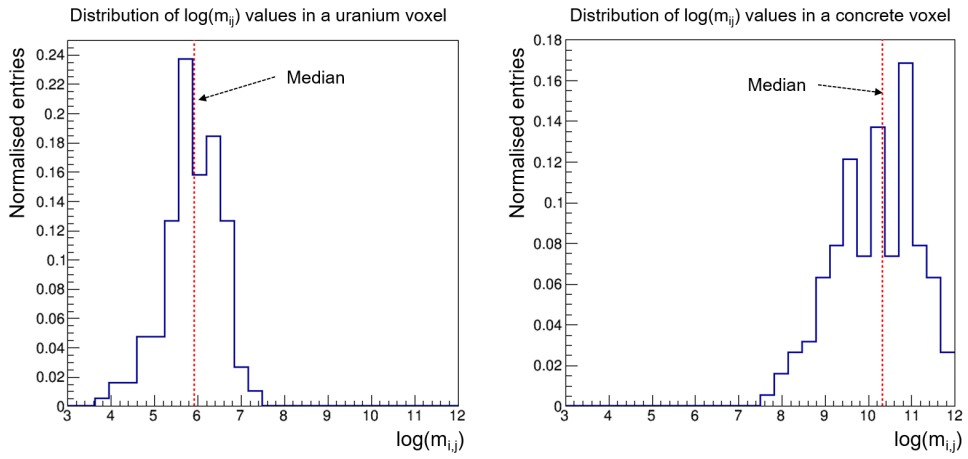

Figure 6. Comparison of distributions of log(metric) values for a voxel corresponding to uranium (left) and concrete (right). The median of each distribution is used as the discriminator in binned clustering algorithm images such as figure 3. The normalised bin counts are used as the MVA input variables.

TMVA allows multiple MVA methods to be trained simultaneously and their efficacy compared. The performance of a binary MVA classifier can be quantified through a Receiver Operating Characteristic (ROC) curve: a plot of the true positive rate (also called sensitivity) against the false positive rate for different cuts on the classifier response for the test sample. The Area Under the Curve (AUC) is a standard measure of the classifier's discriminating power. AUC = 1 indicates a perfect classifier whereas an AUC = 0.5 would indicate the classifier performs no better than random classification.

Applying the classifier to the training datasets and comparing the resulting AUC for a range of MVA methods (figure 8) shows that the Gradient-Boosted Decision Tree (BDTG) method is the most suitable, with AUC = 0.811. The Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM) methods, which overlap in figure 8, have AUCs of 0.808 and 0.804 respectively. For this reason the BDTG method is used hereafter as the default MVA method.

2021 JINST 16 P05007

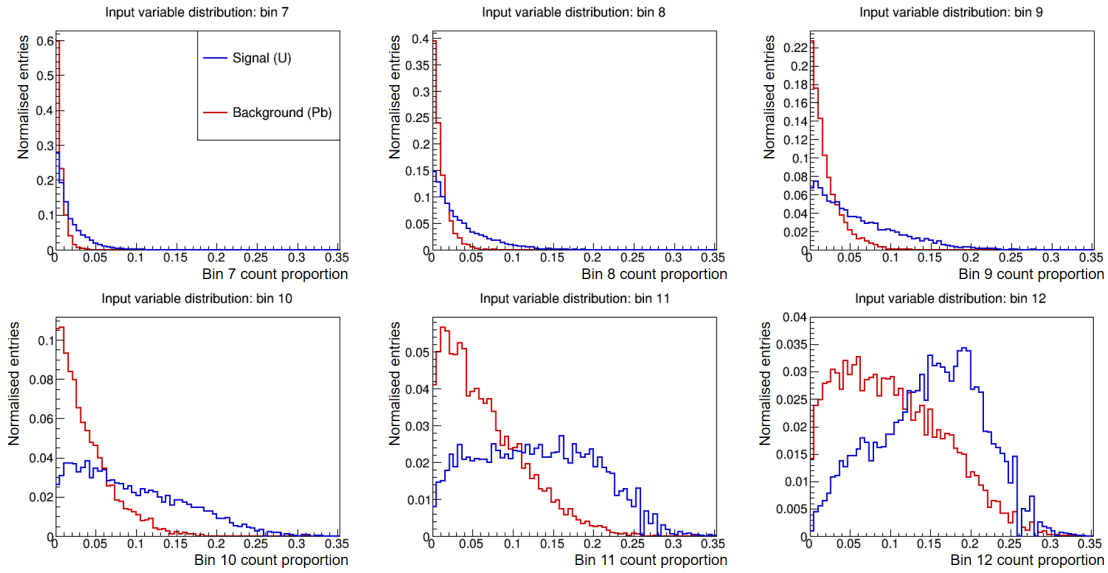

Figure 7. Example distributions of some of the input variables used to train the MVA classifiers, here specifically a binary uranium-lead classifier. The variables are the normalised bin counts (see figure 6) of the $\log(m_{ij})$ values calculated by the binned clustering MST algorithm. The signal set (blue) are voxels in a 20 cm cube of uranium, and the background set (red) an equivalent cube of lead.

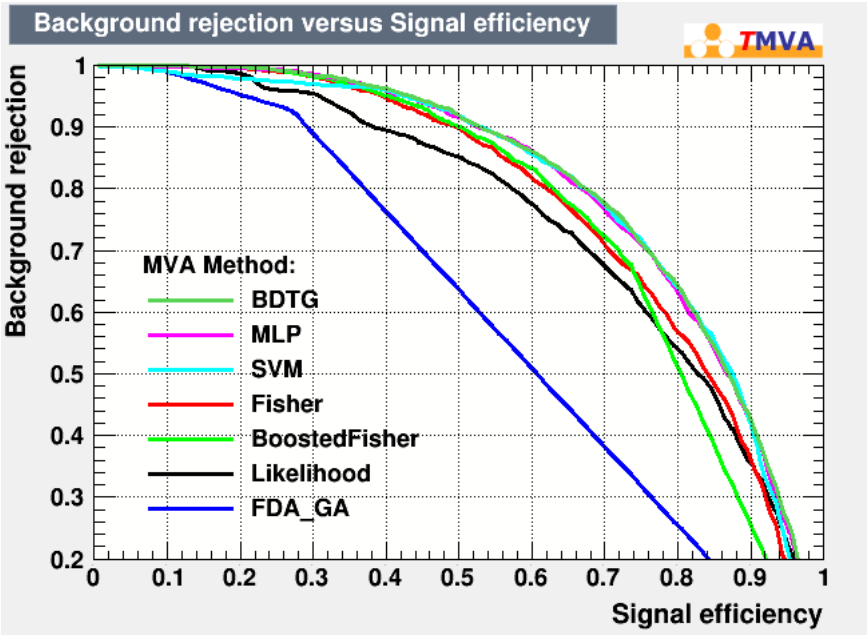

Figure 8. ROC curves showing discriminating power for several TMVA methods when applied to the described binned clustering algorithm variables, for distinguishing voxels in 20 cm cubes of uranium and lead.

2021 JINST 16 P05007

3.2 Training MVA classifiers

The MVAs were trained on a number of simulated MST muon track data corresponding to a 10 day exposure of four different waste drums: an 'empty' drum containing only concrete, and three drums containing 20 cm side length cubes (see figure 9) of iron, lead and uranium, in the centre of the drum and aligned with its central axis. Only the voxels in the cube (or the equivalent volume for the homogeneous empty drum) were passed to the classifier. The binned cluster algorithm's n parameter (see section 2.1) was set to 20. The dataset is split into equally sized 'training' and 'testing' sets; the MVA is trained on the former then applied to the latter as an overtraining check. For a binary classifier, one dataset of voxel variables is designated as 'signal' and the other 'background', whereas a non-binary classifier is passed a single signal dataset and several background datasets. In each case, the classifier attempts to distinguish signal voxels from background(s) voxels, such that when applied to a new voxel it will be classified correctly as often as is possible from the provided variables and the classifier's discriminating power. The non-binary classifiers are trained to distinguish the signal set from all the provided backgrounds (i.e. one-vs-all classification). TMVA calculates an optimum cut value on the classifier response, with a response above the cut being considered 'signal-like' and below 'background-like'. The optimum cut corresponds to the point at which the signal efficiency is equal to the background rejection. On the ROC curve, this corresponds to the point with the maximum Youden index [18], defined as signal efficiency + background rejection − 1; i.e. the length of the vertical line between the ROC curve and the 45° line connecting the curve's ends.

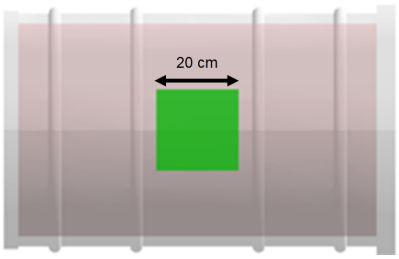

Figure 9. Example simulated geometry used for MVA training: 20 cm side length uranium cube, in the centre of the waste drum.

To check for overtraining, TMVA's standard check was used: the training signal and background datasets of voxels are both randomly split into two equal groups, with one being used to train the classifier and the other reserved for testing. The trained classifier is then applied to the test set. The classifier output distributions for the training and test sets are then directly compared (see figure 10), with a close match between the distributions indicating a low degree of overtraining. A Kolmogorov-Smirnov test is also performed to quantify the similarity of the distributions. In our case, the distributions of the test and training MVA outputs are a close match visually. The Kolmogorov-Smirnov test value is low however, indicating some degree of overtraining has taken place.

3.3 Momentum information

To determine the importance of momentum information for material classification, two alternative approaches to the muon momentum were investigated in addition to the 50% Gaussian smeared truth momentum described in 2.1. These were using the Monte Carlo truth momentum itself,

2021 JINST 16 P05007

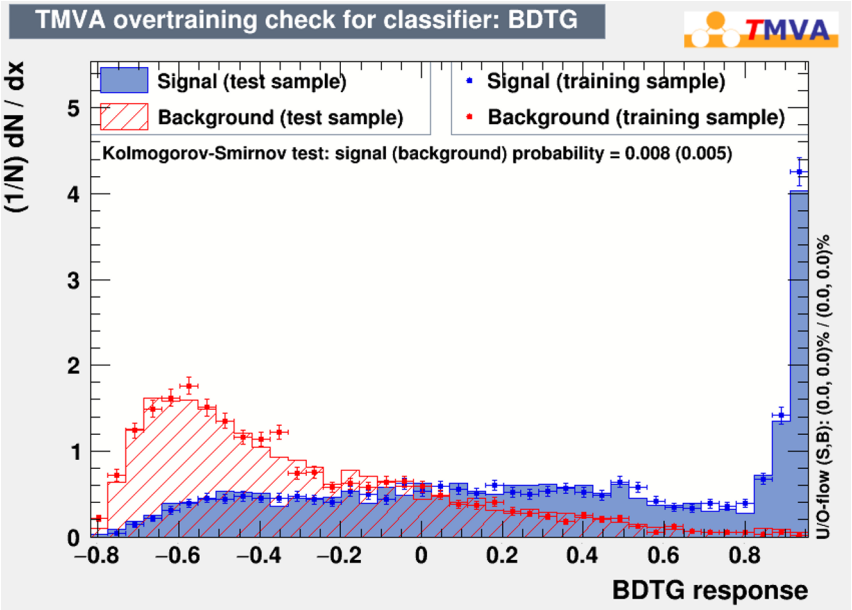

Figure 10. TMVA overtraining check plot for the uranium-lead binary MVA classifier. The MVA output distributions for the signal and background training sets are overlaid with the output distributions for the test sets for comparison and a Kolmogorov-Smirnov test is performed.

with no smearing, and fixing the measured muon momentum at a constant value of 3 GeV/c, i.e. removing momentum information entirely. A comparison of binned clustering algorithm output images of a drum containing 15 cm cubes of uranium, lead and iron for the different approaches is shown in figure 11. Using the Monte Carlo truth momentum results in a slightly sharper image with less variation in the concrete background, whereas using fixed momentum significantly reduces the quality of the image with the iron cube in particular difficult to distinguish from the concrete background.

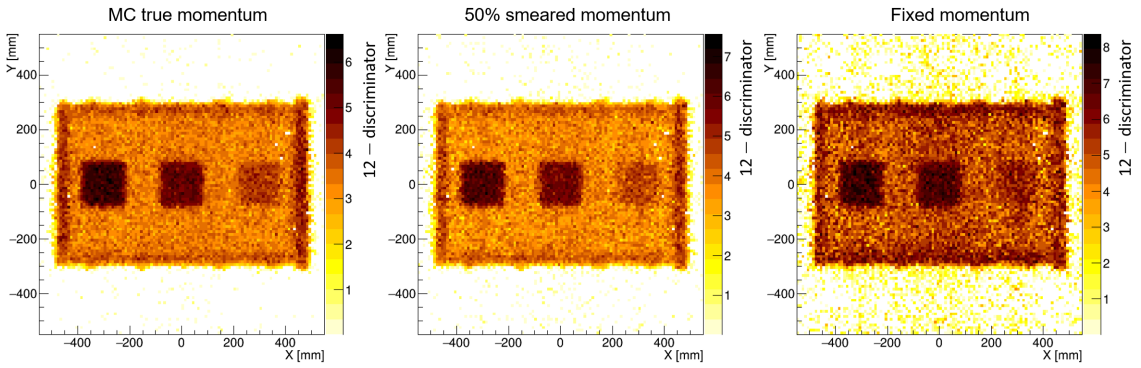

Figure 11. xy slices from binned clustering algorithm output images (with the algorithm's discriminator value subtracted from 12) of a waste drum containing 15 cm side length cubes of uranium, lead and iron, with three different muon momentum approaches: using the Monte Carlo truth momentum (left), applying a 50% Gaussian smear to the truth momentum (centre), and removing momentum information entirely by fixing it at a constant value (right). Exposure time = 10 days, $n = 5$.

2021 JINST 16 P05007

To quantify the effect on material discrimination, binary uranium-lead MVA classifiers trained as described in section 3.2 but with samples obtained using the three different momentum approaches were used to create ROC curves for each scenario (figure 12). Comparing the AUC for each case shows that smearing the momentum slightly reduces the discriminating power of the classifier, with AUC = 0.852 for the truth momentum and AUC = 0.811 for the 50% smeared momentum. The fixed momentum classifier has significantly worse performance with AUC = 0.631. The implication is that momentum information is important for this 'local' i.e. voxel-scale approach to material discrimination, but that a smeared momentum approach gives comparable performance to the idealised Monte Carlo truth.

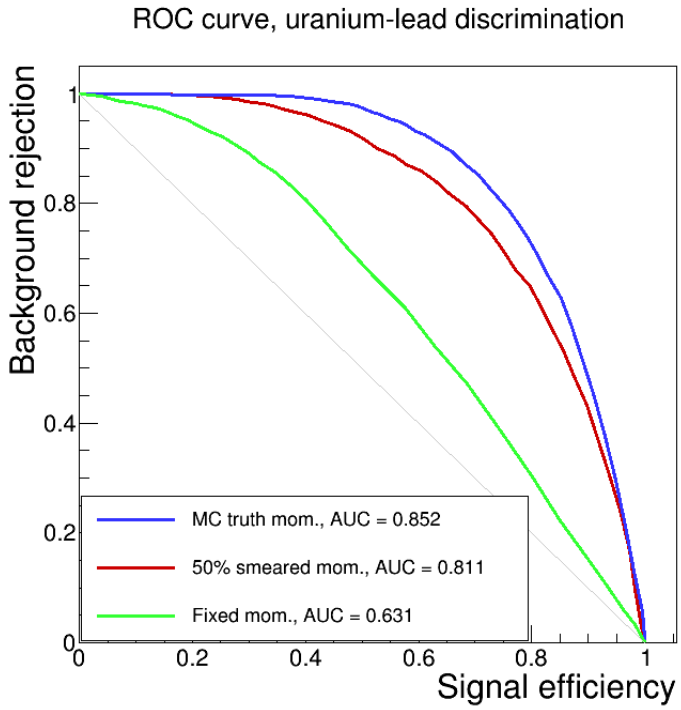

Figure 12. Comparison of ROC curves and their AUCs for the three momentum approaches (Monte Carlo truth momentum, 50% Gaussian smeared truth momentum, and fixed momentum). The MVA classifier trained to discriminate uranium and lead voxels from samples taken from drums containing 20 cm cubes, with exposure time 10 days. Smearing the momentum reduces the discriminating power by a small degree, removing momentum information greatly reduces discriminating power.

4 Identifying stored bodies

4.1 Removal of concrete background

It is necessary to attempt to remove the voxels corresponding to the concrete background and steel shell from the binned clustering algorithm output image. The remaining voxels, corresponding to stored objects, can then be sorted into distinct clusters using the algorithm described in section 4.2. The non-binary concrete classifier's training outputs and ROC curves are shown in figure 13.

2021 JINST 16 P05007

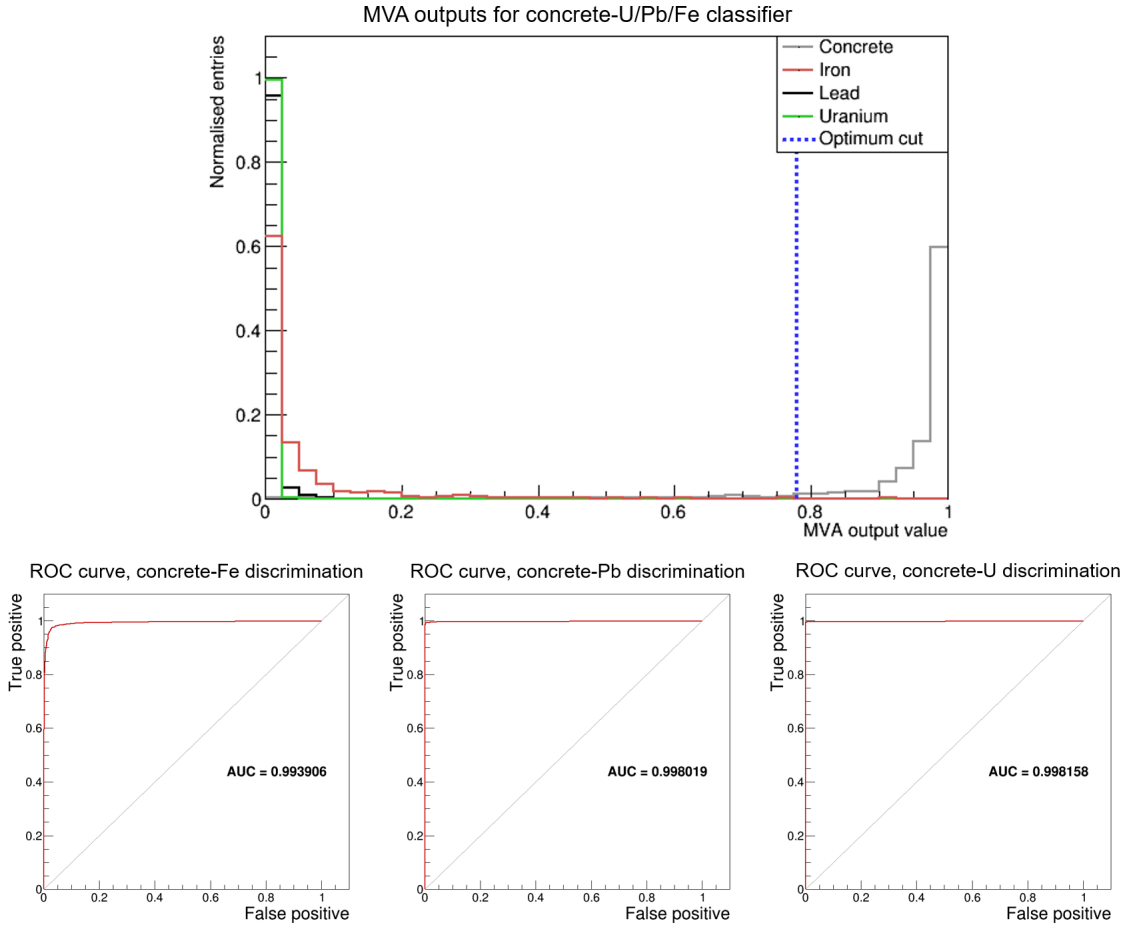

Figure 13. MVA training outputs (top) and ROC curves (bottom) for concrete vs iron/lead/uranium non-binary classifier. The optimum cut (blue) corresponds to the point at which signal efficiency is equal to background rejection.

As the dimensions of the drum are known, the steel outer shell voxels can be removed trivially thorough a cylindrical spatial cut on the image. Subsequently an MVA classifier trained as described in section 3, designating the dataset of concrete voxels as 'signal' and the other materials as 'backgrounds', is applied to the remaining voxels to filter out the concrete voxels. As the classifier is not perfect, some voxels that correspond to concrete in the original simulated geometry remain in the filtered image. The problem is partially mitigated by applying a simple filtering algorithm to remove 'isolated' voxels from the image. Each remaining voxel has its 6 nearest neighbour voxels checked; if they are all empty, the voxel is removed from the image. figure 14 illustrates the result of applying this process to a simulated geometry of three 15 cm cubes. The removed voxels are coloured white in the images; the remaining voxels are black. To test the performance of the nearest neighbour filtering method, the false positive and false negative rates were calculated for this example. Defining a false positive as voxel that does not correspond to concrete being filtered out, and a false negative as a voxel that does correspond to concrete passing the filter, the false positive rate was $0.014^{+0.008}_{-0.005}$ and the false negative rate was 0.497 ± 0.008. The low false positive

2021 JINST 16 P05007

rate indicates that very few non-concrete voxels are being incorrectly filtered out. The high false negative rate however indicates that a large number of concrete voxels remain in the final image; this corresponds to the smearing in the z direction of objects in the drum visible in figure 3.

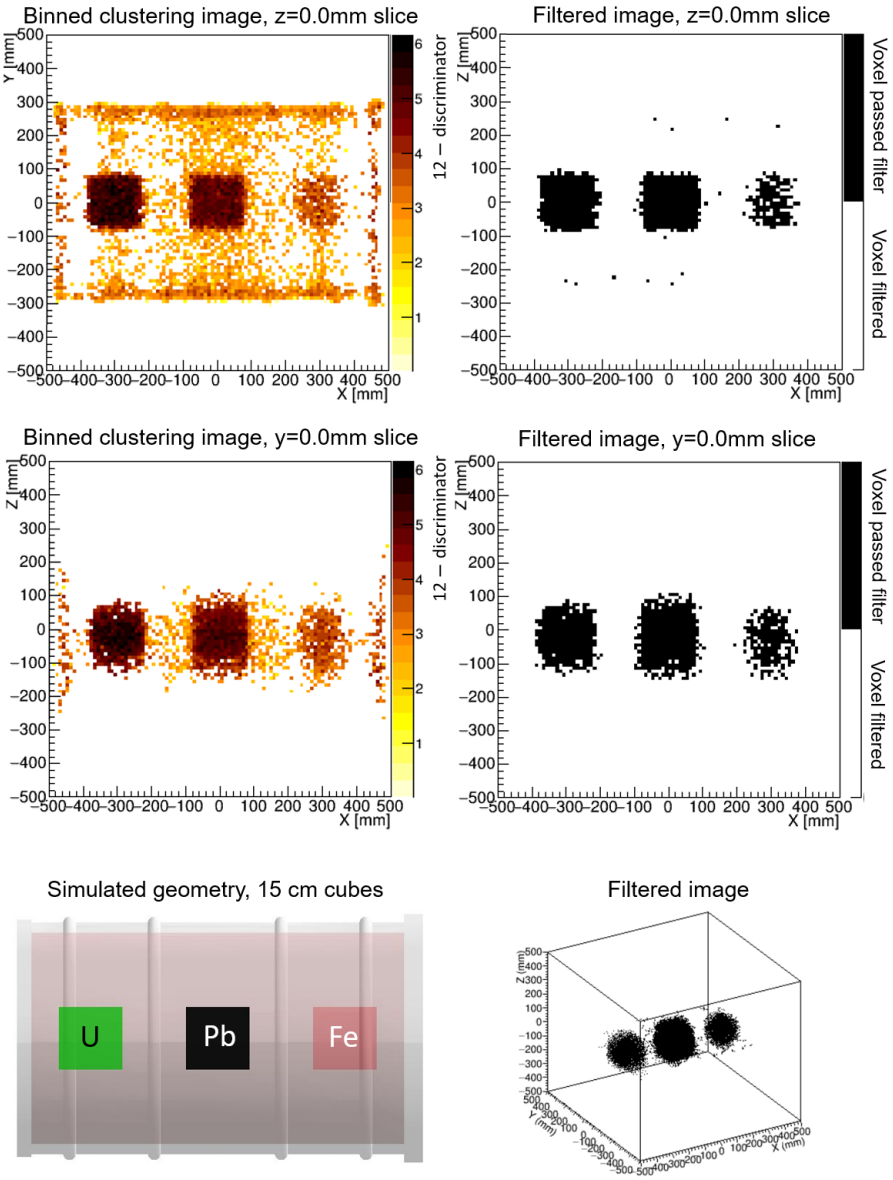

Figure 14. Illustrative example of MVA-filtering algorithm applied to a simulated geometry of a drum containing 15 cm cubes of uranium, lead and iron. Voxels passing the MVA filtering process described above are coloured black.

4.2 Clustering

Subsequently these identified and separated 'object' voxels need to be grouped into individual clusters, each corresponding to a body stored in the drum. This will allow material information to be calculated by applying MVA classifiers to each identified body. The clustering is achieved

2021 JINST 16 P05007

through the widely used k-means clustering algorithm, which in its simplest form operates as follows:

- Choose a value for the number of clusters, k.
- Pick k randomly selected data points to be the initial cluster centroids.
- For each data point, calculate the Euclidean distance (in geometric space) to each of the centroids and assign the point to the cluster with the closest centroid.
- Calculate new centroids as the new centres of the clusters.
- Repeat until the centroid locations converge.

Though this algorithm is fast and easy to implement, it requires the number of clusters k to be known in advance and used as an input. One solution is to run the algorithm multiple times with range of k values as input, and calculate some figure of merit of the clustering output for each. A commonly used figure of merit for clustering algorithms is the Dunn index [19], defined as the ratio between the minimum inter-cluster distance and the maximum intra-cluster distance. A high Dunn index therefore indicates well-separated and compact clusters. The inter- and intra-cluster distances can be defined to suit the problem; in our case the inter-cluster distance metric is the distance between the closest two data points in the two clusters, and the intra-cluster distance metric is the distance between the two furthest-apart points in a cluster. Defined in this way, the k value that corresponds to the maximum Dunn index will represent the most natural choice for k. In most cases this will correspond to the number of bodies stored in the waste drum. In some cases the algorithm can under-estimate k if e.g. two objects are in contact or very close together.

In practise, the simple k-means algorithm often produces poor clustering solutions if the randomly chosen initial centroids are too close together. This problem is avoided by choosing the first centroid only from a uniform distribution and the subsequent $k-1$ centroids from a distribution weighted by the squared distances of the data points from the already chosen centroid(s). This form of the algorithm is often referred to as 'k-means++' [20]. Figure 15 shows the result of applying the k-means++ algorithm to a drum containing 15 cm cubes of iron, lead and uranium.

This algorithm occasionally fails when applied to MVA-filtered binned clustering images such as figure 14, as the 'noise' voxels that do not correspond to a stored object can be treated as a new superfluous cluster. These 'fake' clusters are much more sparse than clusters corresponding to stored objects. This allows the problem to be mitigated by defining a cluster density and removing clusters with densities below some cut. We define cluster density as the ratio of the number of voxels in the cluster to the cube of the mean inter-voxel distance. A density cut of 5×10^{-2} voxel cm^{-3} is effective at removing the sparse clusters.

A small percentage of voxels that correspond to concrete in the drum will be incorrectly passed by the classifier and included in the filtered image. These will be incorporated into one of the clusters, which could cause an incorrect material decision. These voxels will be outliers in the cluster as the majority of the cluster voxels will be close to the cluster centroid; thus they can be filtered out by placing a cut on the distribution of voxel-centroid distances for each cluster. Choosing the cut so as to remove voxels for which the voxel-centroid distance is greater than one standard deviation from the mean of this distribution is effective at removing outlier voxels.

2021 JINST 16 P05007

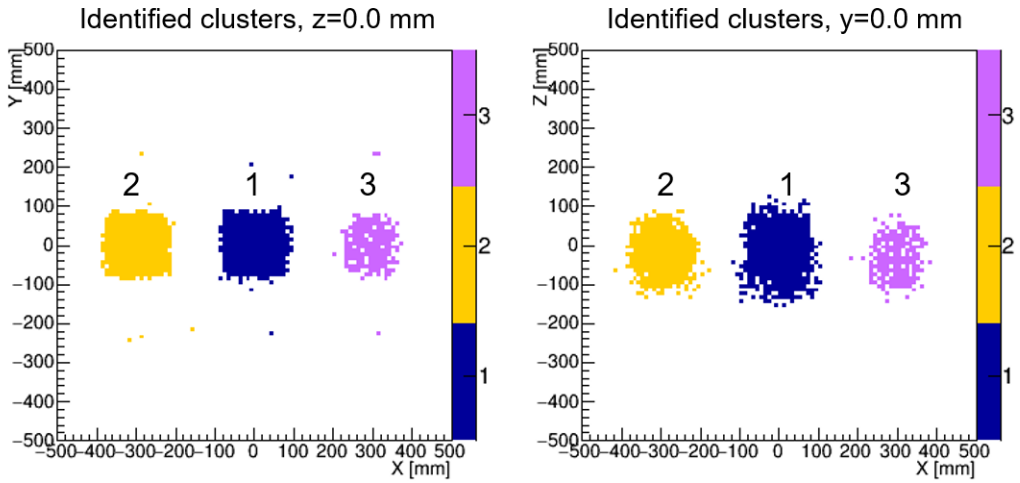

Figure 15. xy (left) and xz (right) slices of the clustering solution for a simulated waste drum containing three 15 cm cubes of different materials. The voxels separated by the method described in section 4.1 have been grouped into three clusters using the k-means++ clustering algorithm.

Finally, a filter is applied to remove approximately the outermost voxel layer (see figure 16) from the surface of each cluster. This is necessary as in general there will be a degree of smearing between a stored body and the concrete background, due to scattering vertices from muons passing close to the object contributing to the algorithm's metric values (see section 2.1) and hence affecting the variables that are passed to the MVA classifiers. The filtering is achieved by calculating the mean of the centroid-voxel distances for each cluster, and removing voxels for which the distance is greater than the mean.

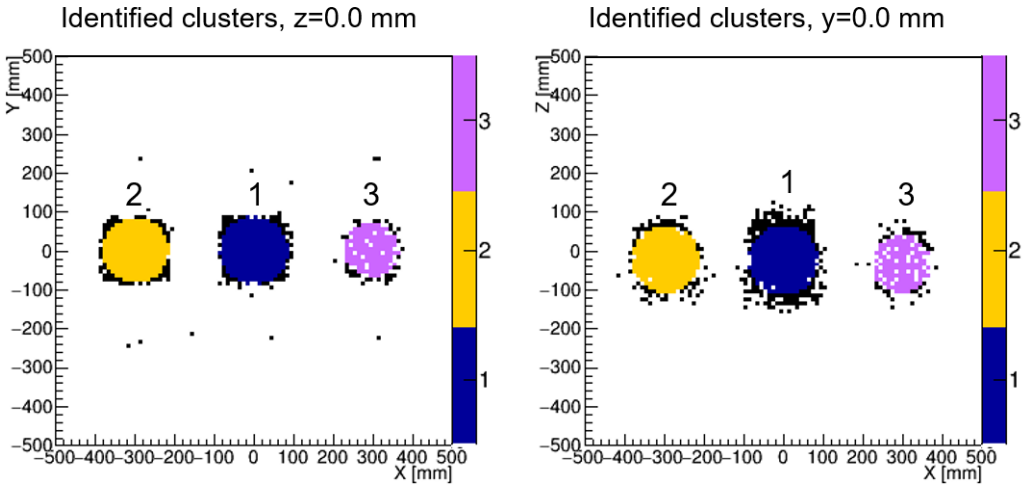

Figure 16. xy (left) and xz (right) slices of the clustering solution of figure 15 after filtering the outermost voxels from each object. Here black indicates voxels removed from the cluster.

2021 JINST 16 P05007

5 Results and analysis

5.1 Applying MVAs to clustered objects

Further MVA classifiers are now applied to the voxels in each identified cluster to obtain material information for the bodies stored in the drum. Two additional MVA classifiers are trained: a non-binary classifier that separates iron signal from lead and uranium backgrounds (see figure 17), and a final binary classifier to discriminate lead and uranium (figure 18). The training ROC AUCs for these classifiers show that the lead and uranium cases are easily distinguished from iron (as the AUC values are close to 1), whereas the lead/uranium classifier does not perform as well, due to the similarity of the materials' Z values.

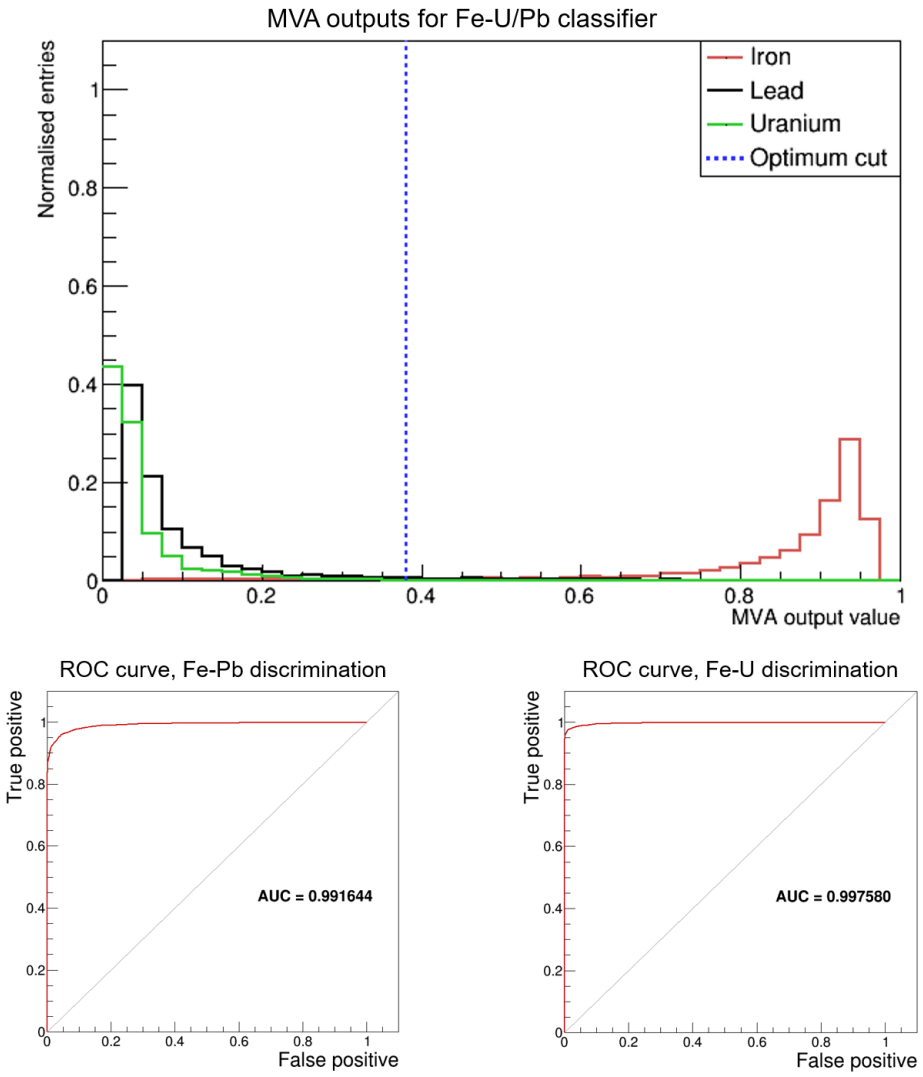

Figure 17. MVA training output and ROC curves for iron/lead/uranium non-binary classifier. The optimum cut corresponds to the point at which signal efficiency is equal to background rejection.

2021 JINST 16 P05007

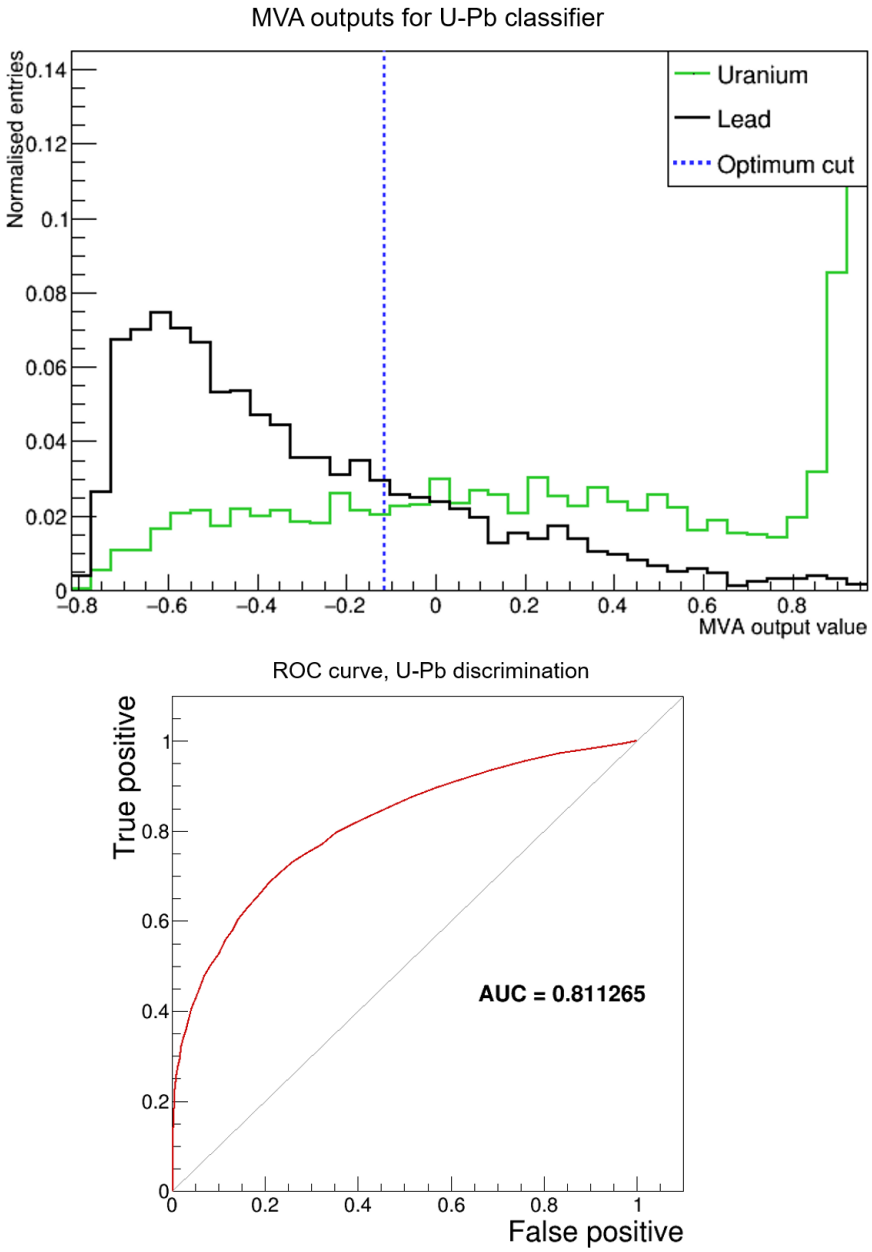

Figure 18. MVA training output and ROC curves for lead/uranium binary classifier. The optimum cut corresponds to the point at which signal efficiency is equal to background rejection.

Each MVA classifier will produce a single response value for each voxel it is applied to. If the value falls above the cut (see section 3.2), the voxel will be considered signal-like, and if it falls below, background-like. Each identified object is a set of voxels; we apply the classifiers to each voxel to obtain the object's distributions of response values, then calculate the proportions of response values that fall above the cuts (i.e. the proportion of the object's voxels that are signal-like) to arrive at a single value from each classifier for each object. Figure 19 shows the MVA classifier response distributions for the three identified objects in the 15 cm cube example simulated geometry.

2021 JINST 16 P05007

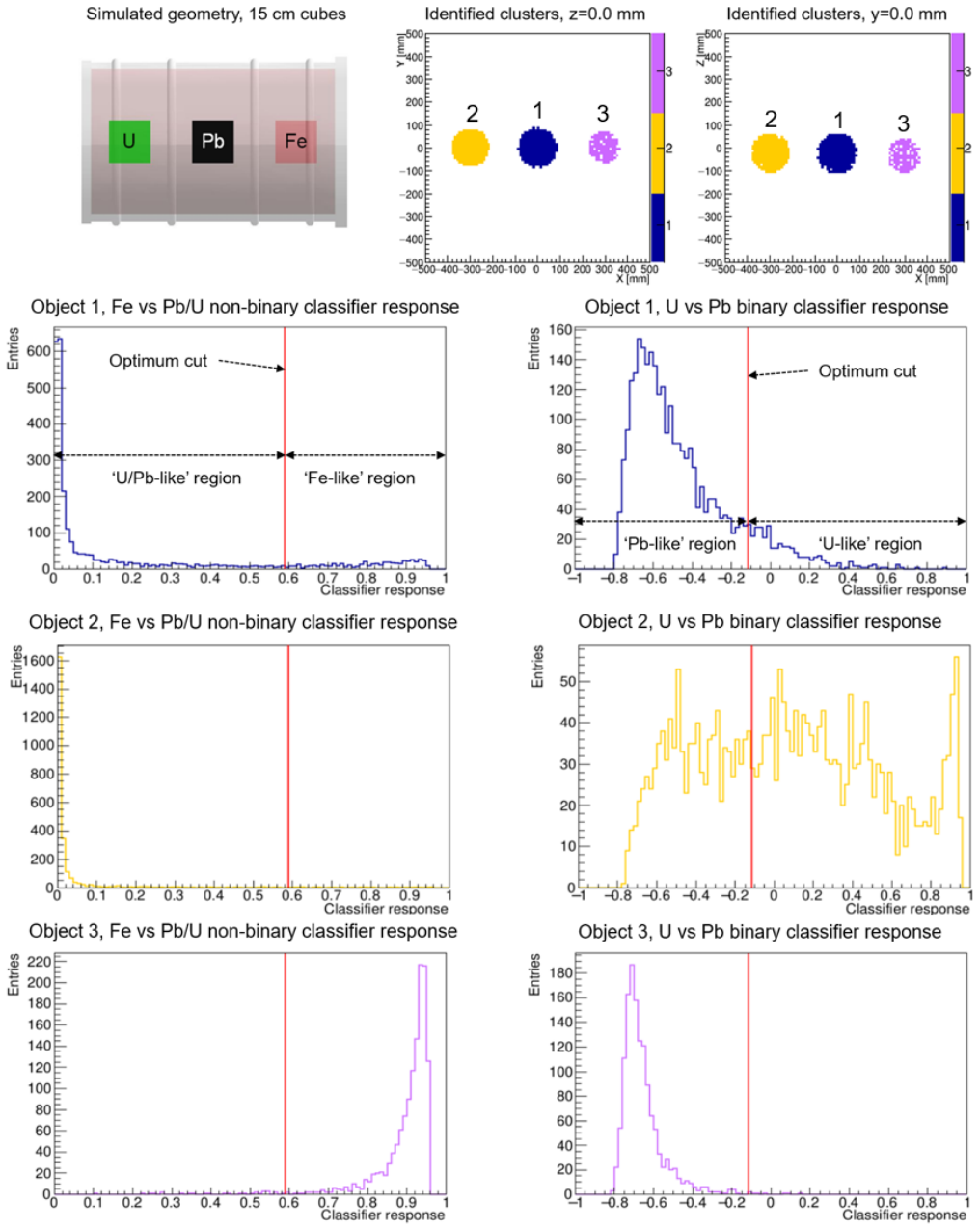

Figure 19. Distributions of responses of MVA classifiers applied to found clusters from a simulated waste drum containing 15 cm cubes of uranium (object 2), lead (object 1) and iron (object 3). The optimum cuts for the classifiers correspond to the points at which the signal efficiency is equal to background rejection.

5.2 Obtaining material decisions

Applying the integral method described above to these distributions results in uranium, lead and iron 'material scores' for each object stored in the drum. The uranium and lead material scores are subsequently multiplied by 1 – iron score, i.e. the 'not-iron' score. These scores are very effective at distinguishing objects of different materials once the sizes of the objects are taken into account. The material scores are intuitively viewed as a pie chart (see figure 20). For the simulated drum contain-

2021 JINST 16 P05007

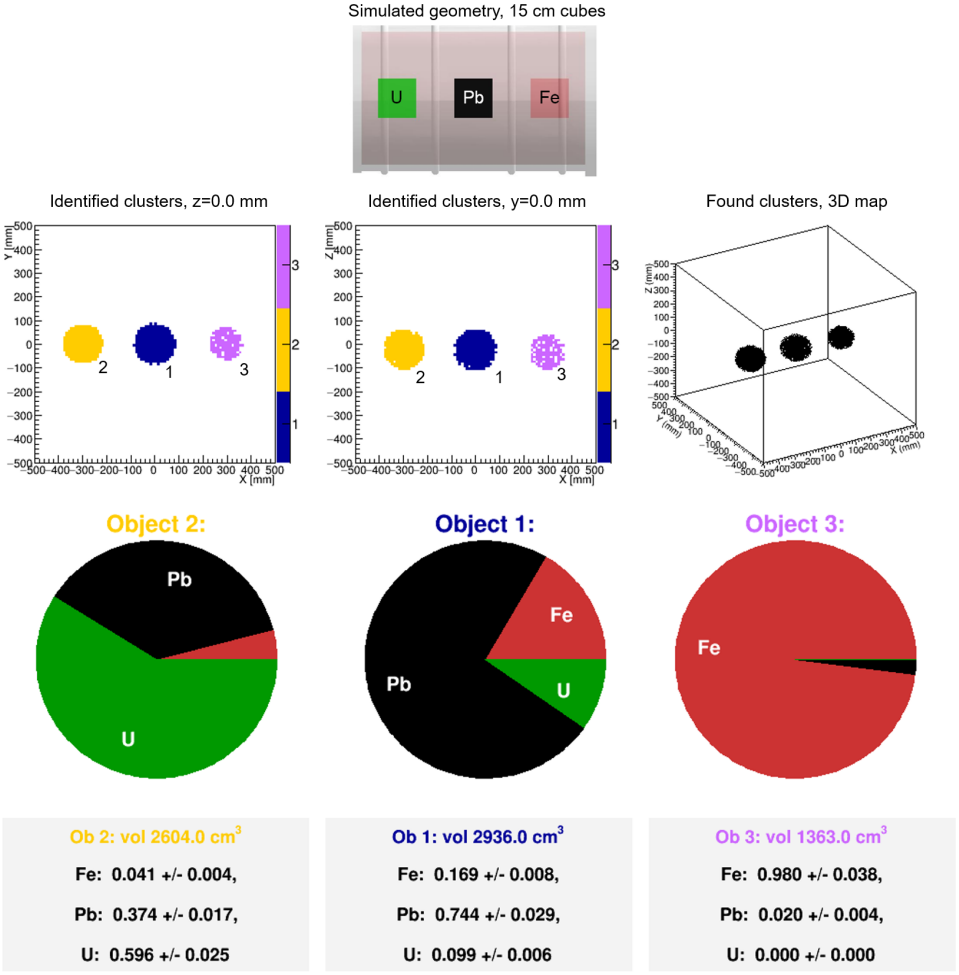

Figure 20. Material scores for simple geometry of three 15 cm cubes, uranium, lead and iron, aligned with voxel grid.

ing three 15 cm side length cubes of uranium, lead and iron, each object has the MVA-calculated material score that corresponds to the true material as the largest score. The scores for the uranium and lead blocks are also clearly distinguished from each other. However, this simulation is an idealised case due to the large size of the objects and their similarity to the 20 cm cube training geometries.

Applying the MVA classifiers to a similar but more challenging geometry of three 10 cm side length cubes (see figure 21), two effects become apparent. Firstly, the classifiers do not perform as well i.e. the score corresponding to the true material is not necessarily the largest. For example, the 'uranium' score has reduced from 0.596 ± 0.025 for the 15 cm cube case to 0.221 ± 0.025. However, the uranium score for the lead cube has also reduced, by a comparable factor. This effect can be explained by considering the repeated scatterings of muons in a large high-Z object: a larger object will lead to larger detected muon scattering angles, and hence a smaller binned clustering metric value (see 2.1). Hence a large lead object can appear more 'uranium-like' than a smaller lead object. The implication is that the size of stored objects must be taken into account to reliably determine their material composition.

2021 JINST 16 P05007

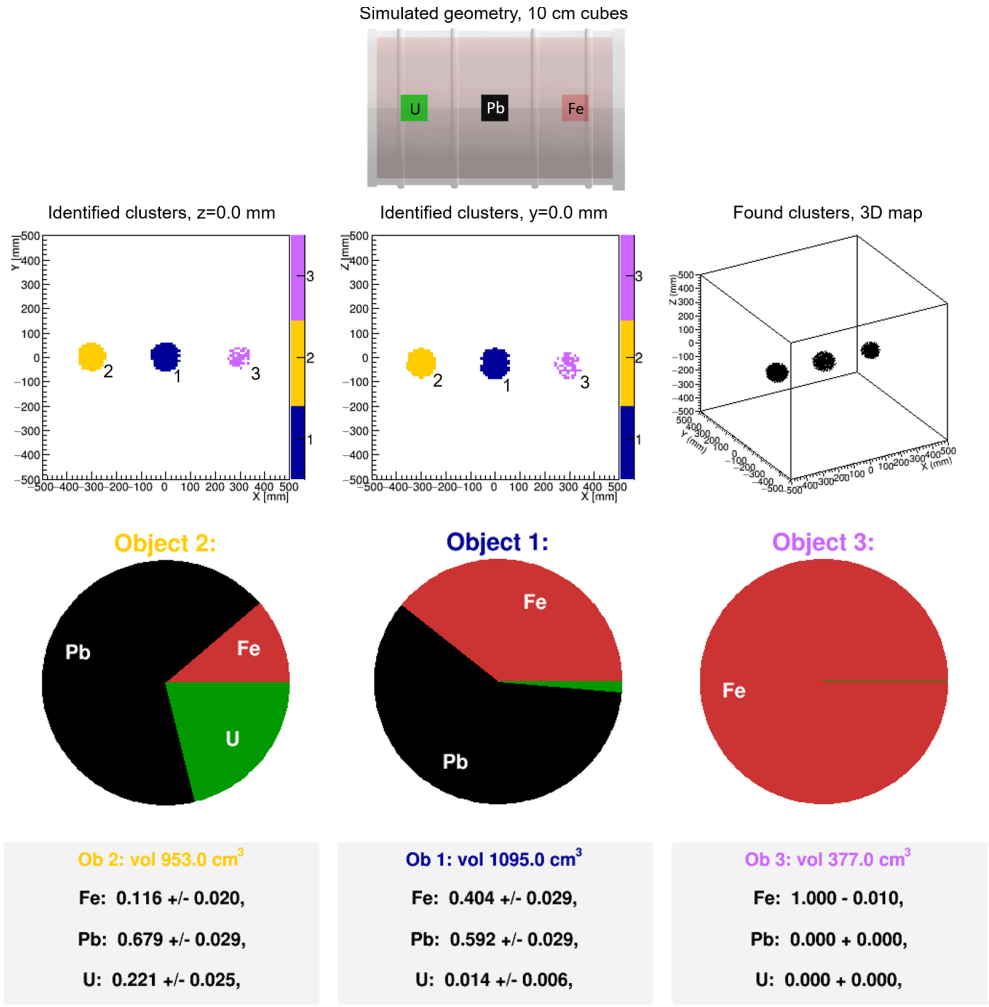

Figure 21. Material estimate results for simple geometry of three 10 cm cubes, uranium, lead and iron, aligned with voxel grid.

To quantify the relations between the object size and the material scores, we applied our system to a series of simulated drums containing spheres of different materials and increasing radii. The results are shown in figure 22. It is apparent that whilst there is no simple relation between the material scores and the object volume, objects of different material are clearly distinguished for a wide range of volumes.

However, these plots can be used empirically to arrive at a single decision material for each identified stored body in the drum. As the volumes of the clusters (equivalent to the number of constituent voxels) are known, the plots in figure 22 give the 'expected' material scores for a cluster of that size if the object was composed of one of the three materials. Finally, a material decision is arrived at by comparing the object's actual material scores with each set of expected values. The material with the best match, i.e. the minimal 3D Euclidean norm between the actual and expected material scores, is selected as the final material decision.

2021 JINST 16 P05007

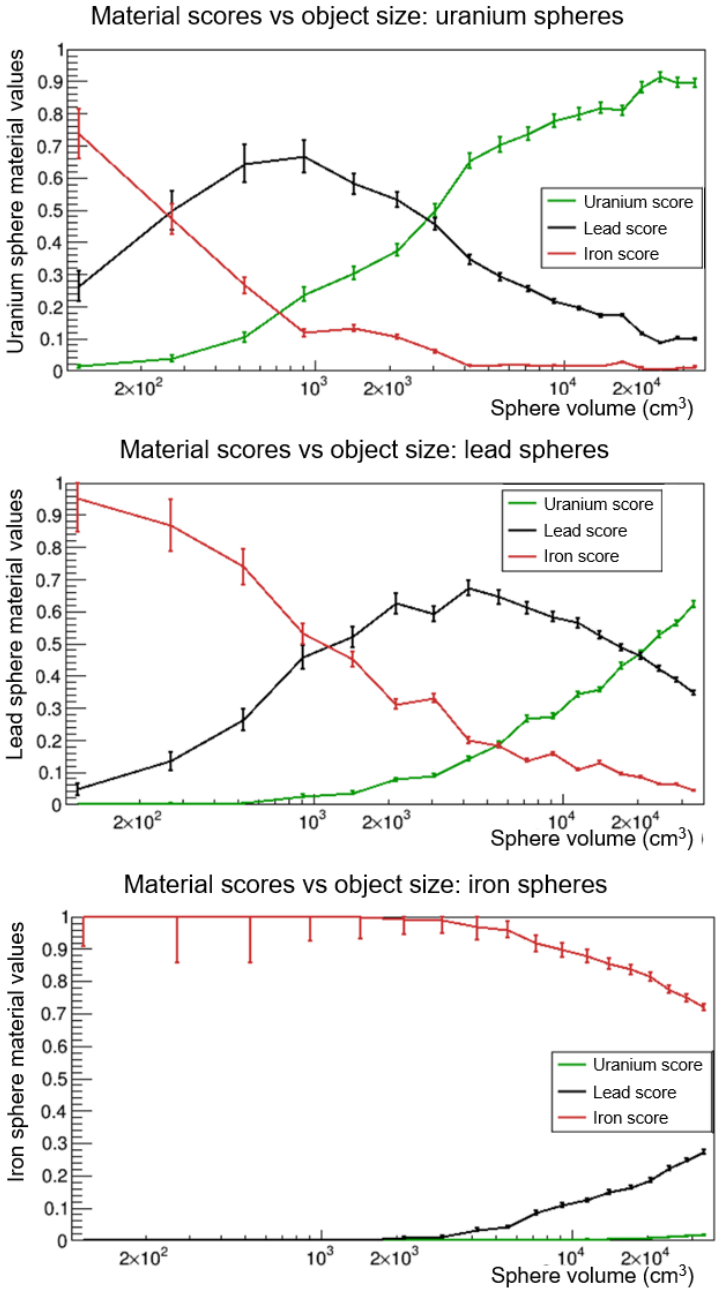

Figure 22. Relationship between the MVA-calculated material scores and the size of the stored object. Each simulated geometry contains a single sphere of increasing radii, composed of uranium (top), lead (middle) or iron (bottom).

This approach was tested on more complex simulated geometries. Figure 23 shows results for a drum similar to the three-cube example of figure 20, but with objects of irregular size, location and rotation. In this case the system has accurately identified the correct material for each object. Despite the uranium block's low uranium score compared to the equivalent 15 cm cube (figure 20), the calibration by volume has correctly identified it as uranium.

2021 JINST 16 P05007

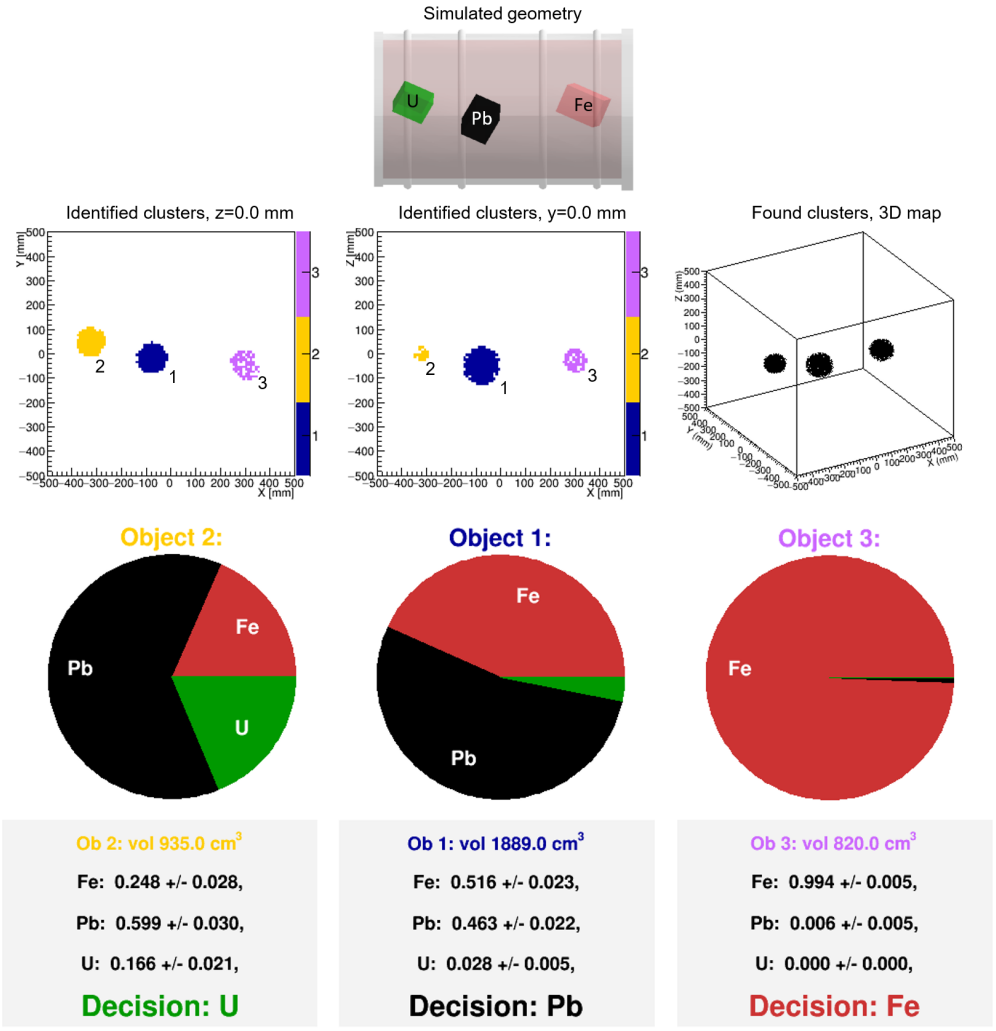

Figure 23. Material estimate results for more complex geometry of three objects, uranium, lead and iron, not aligned with the voxel grid or centred. By calibrating the three material scores against the volume calibration curves, (figure 22), the correct material has been assigned in each case.

A further example with a larger number of objects is shown in figure 24. This drum contains five objects (two uranium, two lead, and one iron) of a wider range of shapes, dispersed more evenly through the drum. However, the system still performs well. The identified clusters are a close match to the true locations of the stored objects. Both uranium objects are correctly assigned, as is the iron sphere and one of the lead objects. One lead object, a tube, has been incorrectly identified as iron. This indicates a limitation of the system when attempting to determine the materials of non-spherical objects.

5.3 Sensitivity

To establish the system's sensitivity and false positive rate, we then applied it to a set of randomly generated waste drum simulations. Each simulation contained three spheres of radius 6 cm, randomly dispersed throughout the drum but constrained to not intersect each other. 100 simulations

2021 JINST 16 P05007

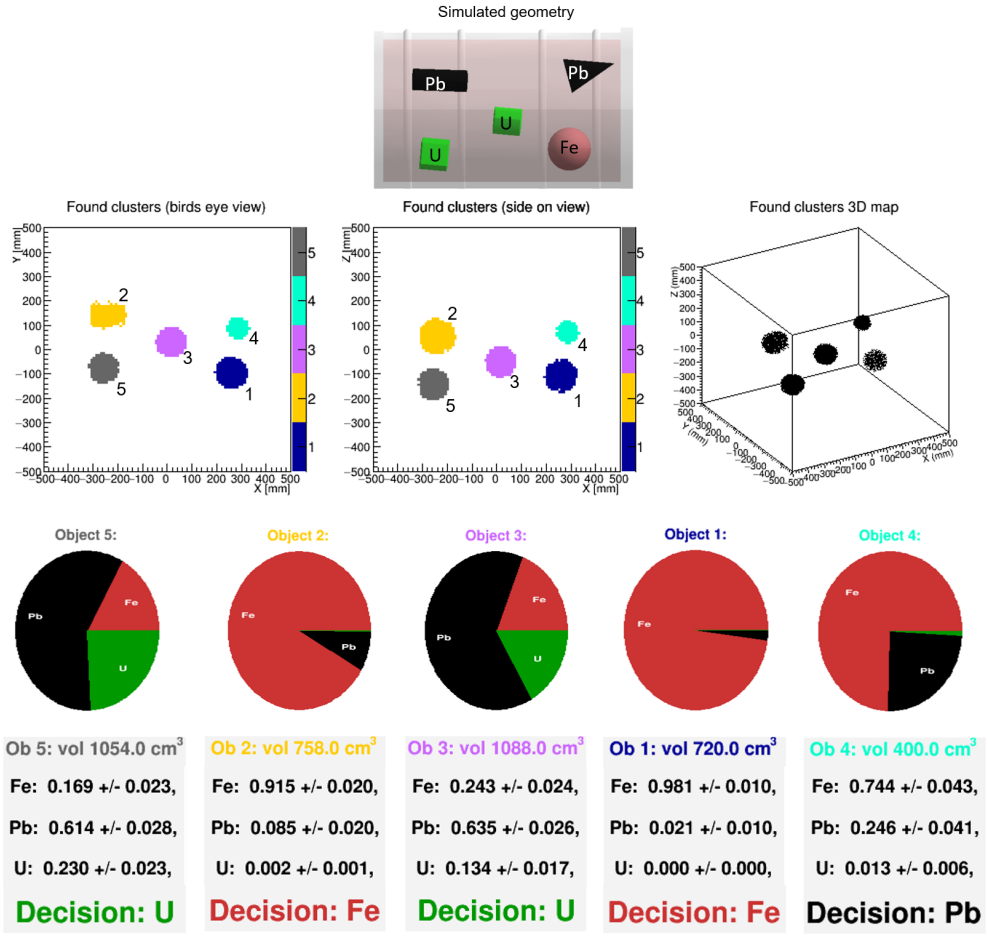

Figure 24. Material estimate results for more complex geometry of five objects of various materials and shapes, dispersed throughout the drum. Note that the 2D cluster plots are viewed as side-on and bird's eye views of the 3D map; this is necessary to view all the clusters as they do not all intersect the central xy and zx planes. Four of the objects have been assigned the correct material; one lead object has been incorrectly classified as iron.

were run in total. 50 simulations contained one uranium, one lead and one iron sphere, and the remaining 50 contained two lead spheres and one iron sphere. A true positive identification of a uranium object was defined to be an object identified close to the true location of a uranium sphere that was designated as uranium by the system. Conversely a false positive comprised any assignment of a uranium decision to an object in a drum not containing uranium. With these criteria, we found a sensitivity of $0.90^{+0.07}_{-0.12}$, and a false positive rate of $0.12^{+0.12}_{-0.07}$ (95% Clopper-Pearson confidence intervals).

6 Conclusions

We have demonstrated that machine learning techniques are a powerful tool for enhancing the information about a waste drum's contents that can be obtained in a muon scattering tomography experiment. MVA classifiers trained on variables obtained from the distribution of binned clustering

2021 JINST 16 P05007

algorithm metric values are effective at discriminating materials in waste drums. The concrete matrix can be distinguished from stored objects of mid- and high-Z material, allowing the voxels corresponding to the matrix to be removed, and the remaining object voxels sorted into clusters.

Additional material information can be obtained with further MVA classifiers, to discriminate first mid-Z (e.g. iron) from high-Z (lead, uranium) objects, and then between materials with similar Z. The effectiveness of the material discrimination is highly dependent on object size. By establishing the empirical relation between object size and the MVA classifiers' material output scores, a final material decision can be made for each identified stored body in the simulated waste drum. This has shown to be accurate for a wide range of object sizes, shapes and drum locations.

When tested against a set of simulated drums containing 6 cm radius spheres of different materials in randomly determined positions, the system performed with a true positive rate of $0.90^{+0.07}_{-0.12}$, and a false positive rate of $0.12^{+0.12}_{-0.07}$, indicating this approach is effective at identifying uranium objects inside waste drums. The main identified vulnerabilities are objects with large differences in Z (e.g. iron and uranium) being very close too each other, and more spatially extended objects being misidentified, although the latter problem could be mitigated by extending the object size-based decision method (see figure 22) to account for a wider range of object shapes.

Acknowledgments

This project has received funding from the Euratom research and training programme 2014–2018 under grant agreement No. 755371.

References

[1] W Kubinski et al., *Calorimetric non-destructive assay of large volume and heterogeneous radioactive waste drums*, *EPJ Web Conf.* **225** (2020) 06003.

[2] L.J. Schultz et al., *Image reconstruction and material Z discrimination via cosmic ray muon radiography*, *Nucl. Instrum. Meth. A* **519** (2004) 687.

[3] L. Frazão, J. Velthuis, C. Thomay and C. Steer, *Discrimination of high-Z materials in concrete-filled containers using muon scattering tomography*, 2016 *JINST* **11** P07020.

[4] L. Frazão, J. Velthuis, S. Maddrell-Mander and C. Thomay, *High-resolution imaging of nuclear waste containers with muon scattering tomography*. 2019 *JINST* **14** P08005.

[5] T. Stocki et al., *Machine learning for the cosmic ray inspection and passive tomography project (CRIPT)*, *IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.* (2012) 91.

[6] X.-Y. Pan, Y.-F. Zheng, Z. Zeng, X.-W. Wang and J.-P. Cheng, *Experimental validation of material discrimination ability of muon scattering tomography at the TUMUTY facility*, *Nucl. Sci. Tech.* **30** (2019) 120.

[7] L.J. Schultz et al., *Statistical reconstruction for cosmic ray muon tomography*, *IEEE Trans. Image Processing* **16** (2007) 1985.

[8] G.R. Lynch and O.I. Dahl, *Approximations to multiple Coulomb scattering*, *Nucl. Instrum. Meth. B* **58** (1991) 6.

[9] Particle Data Group collaboration, *Review of particle physics. Particle Data Group*, *Phys. Lett. B* **592** (2004) 1.

2021 JINST 16 P05007

[10] S. Riggi et al., *Muon tomography imaging algorithms for nuclear threat detection inside large volume containers with the Muon Portal detector*, *Nucl. Instrum. Meth. A* **728** (2013) 59 [arXiv:1307.0714].

[11] C. Thomay et al., *A binned clustering algorithm to detect high-Z material using cosmic muons*, 2013 *JINST* **8** P10013.

[12] C.L. Morris et al., *Tomographic imaging with cosmic ray muons*, *Sci. Global. Secur.* **16** (2008) 37.

[13] C. Steer, P. Stowell and L. Thompson, *CRESTA: Cosmic rays for engineering, scientific, and technology applications*, https://gitlab.com/cosmicraysim/cresta.

[14] GEANT4 collaboration, *GEANT4 — a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[15] C. Hagmann, D. Lange and D. Wright, *Cosmic-ray shower generator (CRY) for Monte Carlo transport codes*, *IEEE Nucl. Sci. Symp. Conf. Rec.* **2** (2007) 1143.

[16] H. Voss, A. Hocker, J. Stelzer and F. Tegenfeldt, *TMVA, the Toolkit for Multivariate Data Analysis with ROOT*, *PoS* **ACAT** (2007) 040.

[17] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.

[18] W.J. Youden, *Index for rating diagnostic tests*, *Cancer* **3** (1950) 32.

[19] J.C. Dunn, *A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters*, *J. Cybern.* **3** (1973) 32.

[20] D. Arthur and S. Vassilvitskii, *k-means++: The advantages of careful seeding*. http://ilpubs.stanford.edu:8090/778/, June 2006.

2021 JINST 16 P05007