Application of machine learning in muon scattering tomography for better image reconstruction

Subhendu Das, Sridhar Tripathy, Priyanka Jagga, Nayana Majumdar, Supratik Mukhopadhyay
Saha Institute of Nuclear Physics, Kolkata, India
Research Foundation, City University of New York, NY, USA

Presented by

Sridhar Tripathy

Research Foundation, CUNY, New York

tripathysridhar@gmail.com

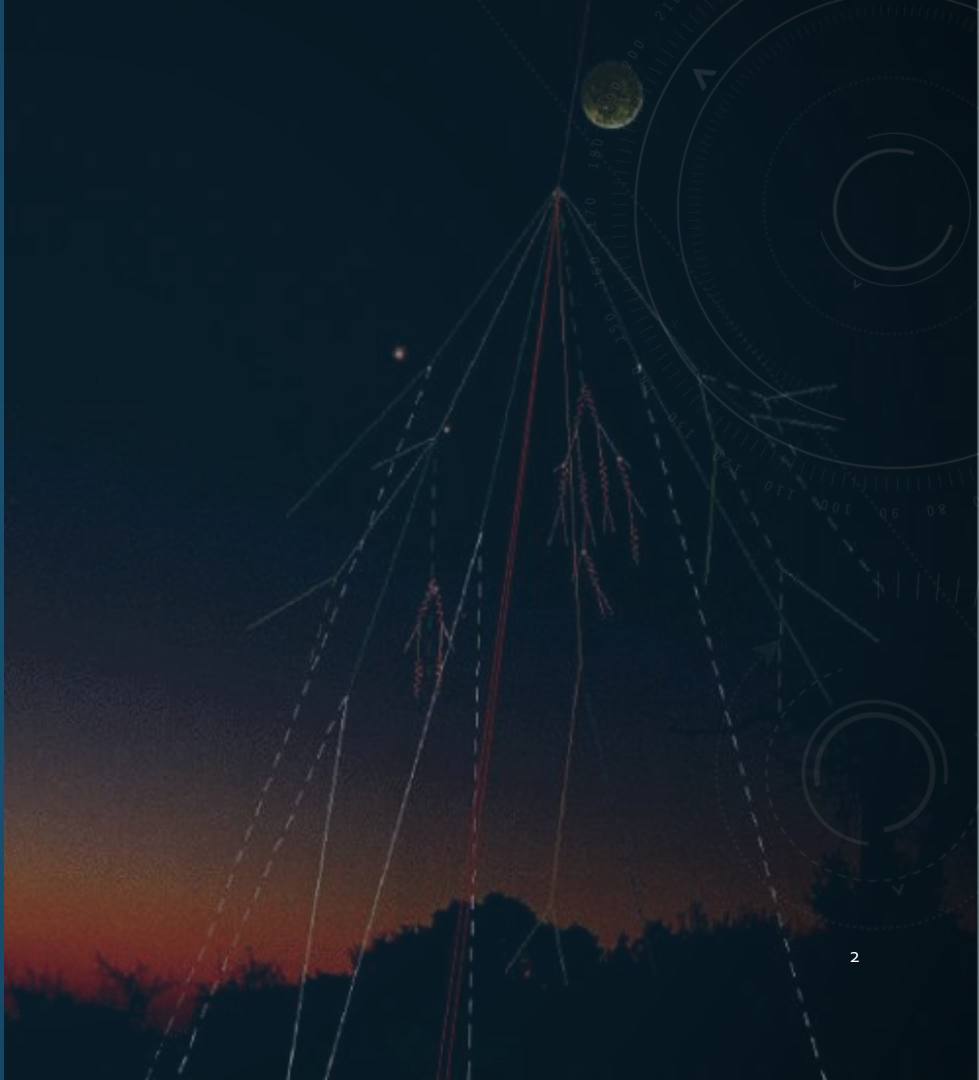
YORK College

RESEARCH FOUNDATION CUNY

Connecting The Dots 2022
7th International CTD Workshop
Princeton University, Princeton, USA
May 31 - June 2, 2022

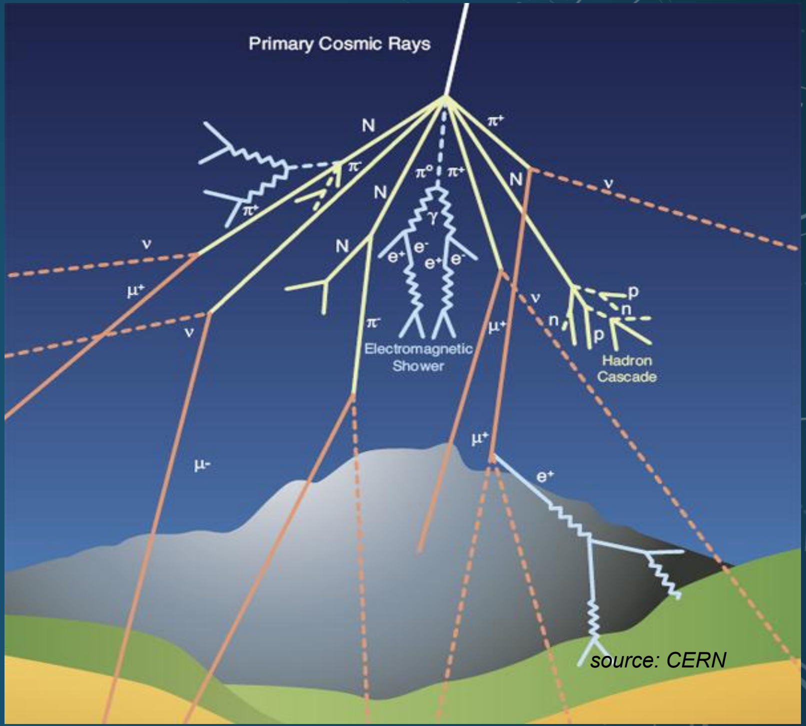
Outline

- ★ Introduction
- ★ Simulation of muon imaging setup for material discrimination
- ★ Use of Clustering Algorithm to find out target locations
- ★ Use of SVM as a classifier
- ★ Results with Concrete Rebar
- ★ Comparison with PRM
- ★ Summary



Introduction: Muon Tomography

- High-energy muons, cosmic source.
- Based on multiple coulomb scattering, and energy deposition of muons.
- These phenomena are measured using tracking detectors: gas detectors/scintillators
- Appropriate for non-invasive imaging
- Applications: examining cargo containers, nuclear waste, monitoring volcano eruptions etc.
- Natural background: Safe for people scanning and object scanned.



Production of cosmic-muons: Primary cosmic rays interact with nuclei/proton at the top of atmosphere produce cascade of particles with many short lived K and $\pi\pm$ (~ 10^{-8} s) which in turn decay to produce many more μ±, e±, ν, γ etc.

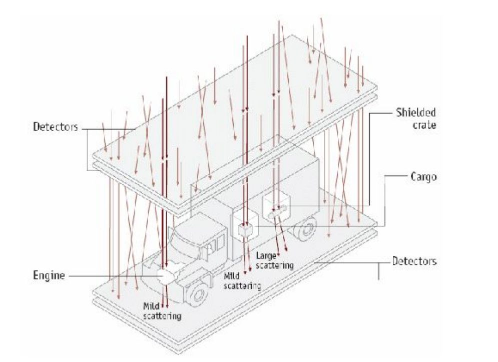
Introduction: Muon Tomography

Scattering Muography (MST):

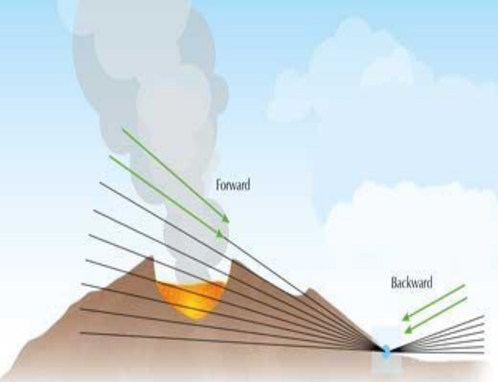
Muons deviate while traversing through the matter due to multiple Coulomb scattering (mcs). Deviation of muon is obtained by placing tracking detectors on either sides of the target region. This method is effective for small targets, such as cargo containers, nuclear dry casks.

Absorption Muography (AM):

After traveling a large distance inside matter muons loose significant energy and eventually get stopped or substantially deviated. Comparing the muon flux obtained from 'free-sky' and target, image map is constructed. This method is effective for humongous targets, like volcanos, pyramids etc.

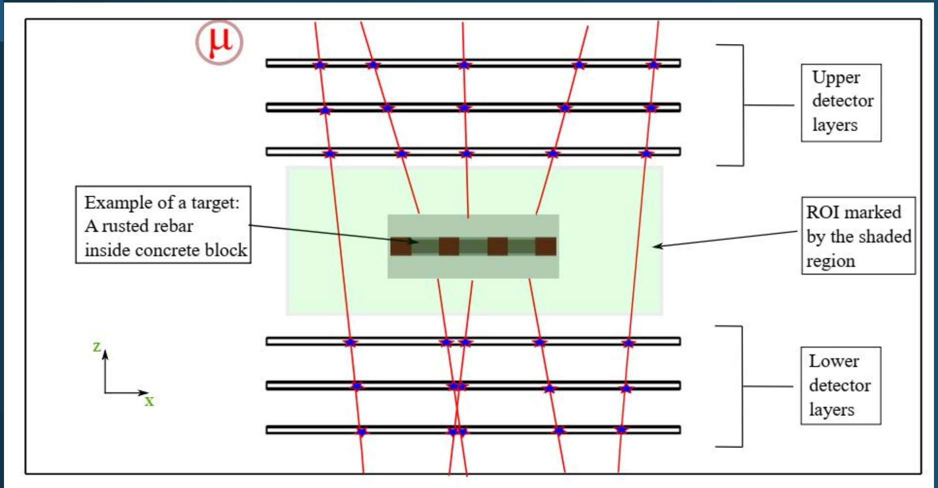


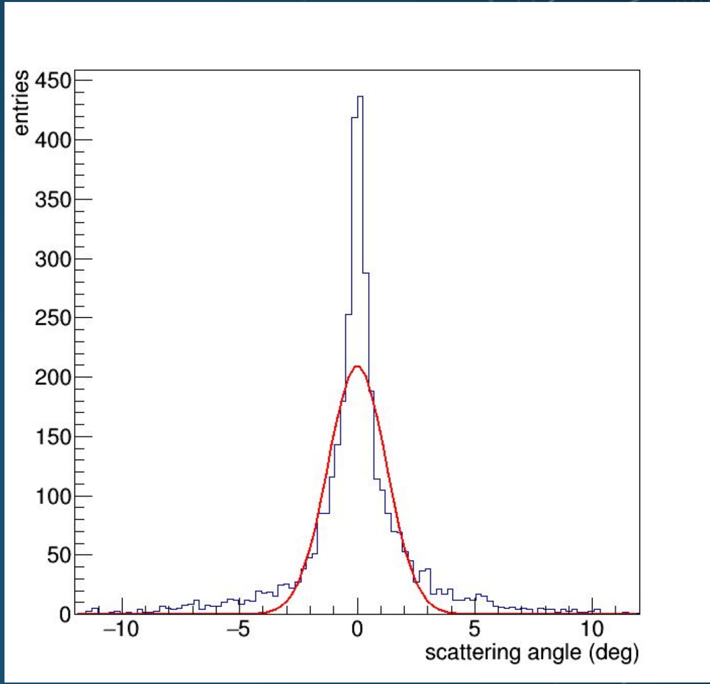
MST for scanning cargo containers



AM for imaging volcanoes.

Strategy of MST





$$\frac{dN}{d\theta_x} = \frac{1}{\sqrt{2\pi}\theta_0} e^{-\theta_x^2/2\theta_0^2}$$

$$X = \frac{716.4\,\mathrm{g/cm^2}\,A}{\rho Z(Z+1)\ln\left(287/\sqrt{Z}\right)}$$

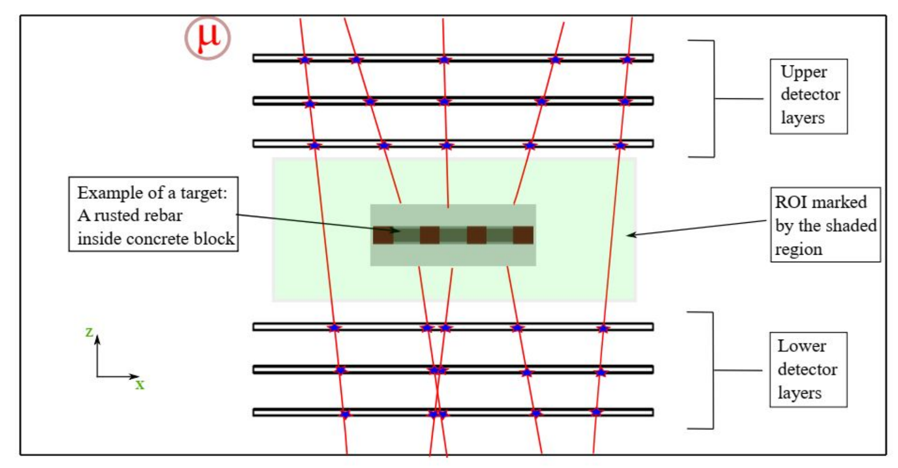
$$\theta_0 = \frac{13.6\,\mathrm{MeV/c}}{p\beta}\sqrt{\frac{L}{X}} \times \left(1 + 0.0038\left(\frac{L}{X}\right)\right)$$

p = Momentum
β = *v/c*
X= Radiation Length
L= Thickness
A= Atomic Mass
Z= Atomic Number
ρ = Density

Scattering angle distribution for 5 cm Pb cube

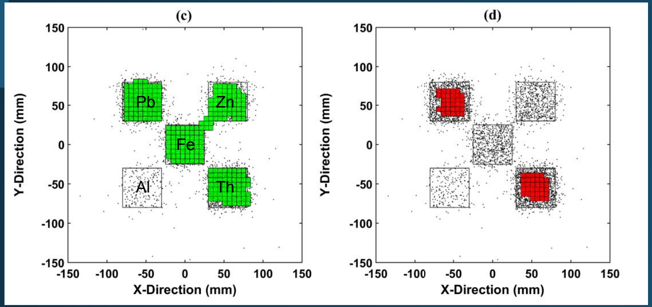
Simulation geometry

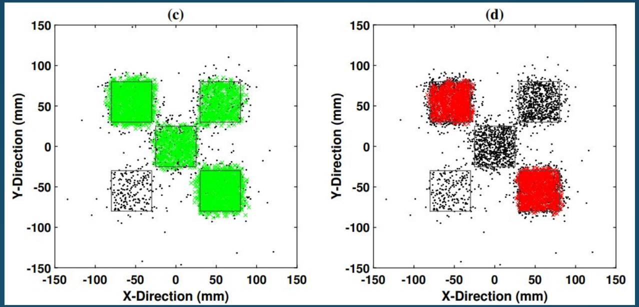
- Simulation carried out in Geant4
- 6 detectors, area calibrated according to the ROI.
- Parallel plate gaseous detectors.
- Detector separation: 7 cm
- CRY generator for muons
- Track reconstruction algorithm: Point of Closest Approach (PoCA)
- 2D image reconstruction
- Analysis based on clustering density(ρ_c) and scattering angle (Θ)
- 30 days equivalent of muon exposure
- Detector spatial resolution: 200 um



Schematic diagram of the simulated geometry.

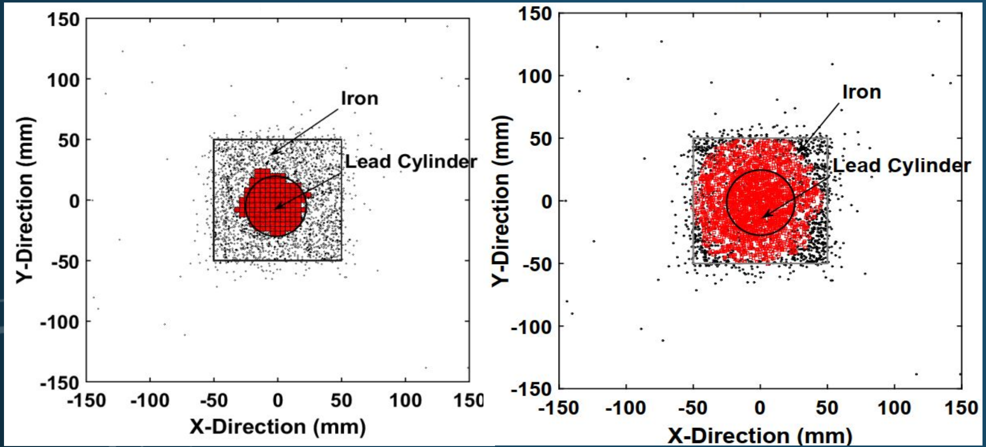
Pattern Recognition Method (PRM)





PRM

DBSCAN

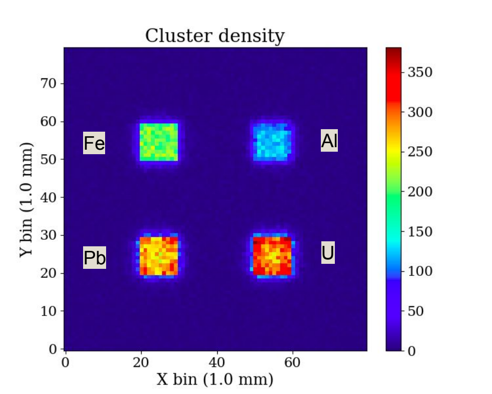


Comparison of DBSCAN and PRM in complex scenarios

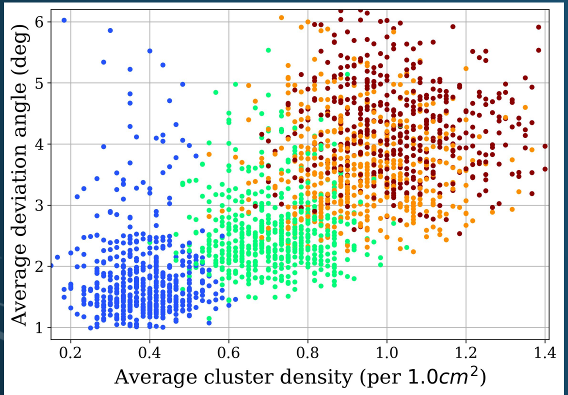
- ➢Clustering algorithms, such as PRM and DBSCAN have been used for detecting objects from noise background; arising from muons not passing through the targets and bad-reconstructed muons
- ➢DBSCAN identifies *minPts* number of scattered vertices in radius **ϵ** after training.
- ➢PRM learns from the bin-contents of S-maps and provides properties of the detected cluster.
- ➢The PRM has been further extended to material discrimination using scattering parameters,
- ➢Several parameters, such as muon exposure, scattering threshold (5 mrad, 10 mrad), pixelation, shape and dimensions of the target can hamper the clustering.
- ➢The clustering algorithm provides the idea regarding the position, shape and dimension of the targets in the ROI and hence the experimenter can focus on these areas in the detection algorithm.

S. Tripathy et. al.,JINST 15 (2020) 06,P06029

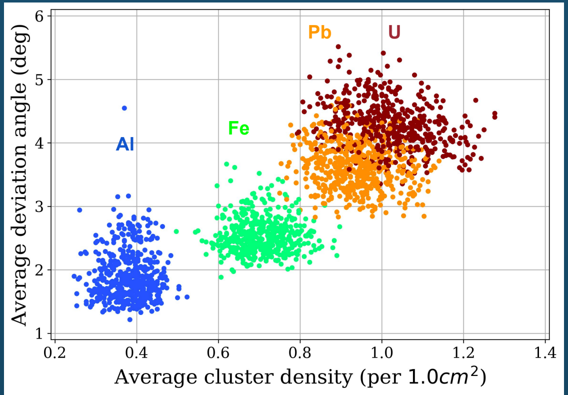
Detection Algorithm: Support Vector Machine (SVM)



Four block target clusters



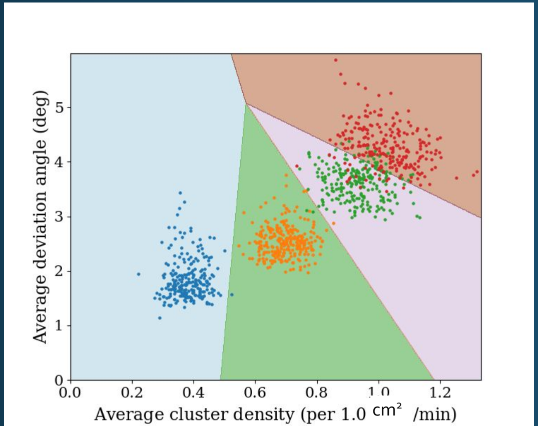
Features for 1 hr data



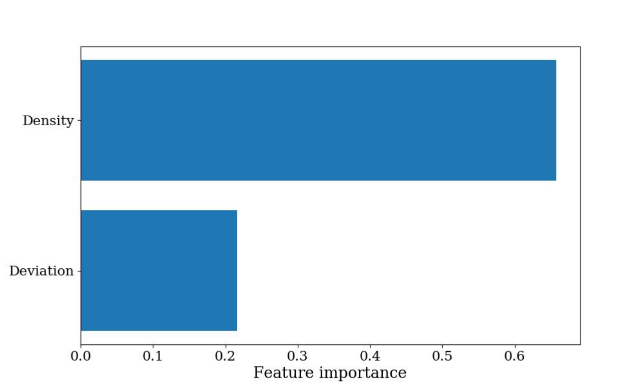
Features for 5 hrs data

- Algorithm used to differentiate high-Z/low-Z materials based on their scattering parameters after then clusters have been identified.
- To achieve this classification, SVM, implemented in Python has been used, which is a supervised learning.
- Above-mentioned: ρ_c, and Θ are used as features of this classification.
- To eliminate extra-noise, a low bound of scattering angle: 10 mrad has been provided. ➽
- The plot of features after background segregation for different time-frames have been given.
- The mean value of the features, from the training pixels and classification has been carried-out by linear kernel function.
- The feature importance has also been given. ➽
- The results have been expressed on the basis of confusion matrix and mean error rate.

Classification based-on features



Decision boundary for several classification

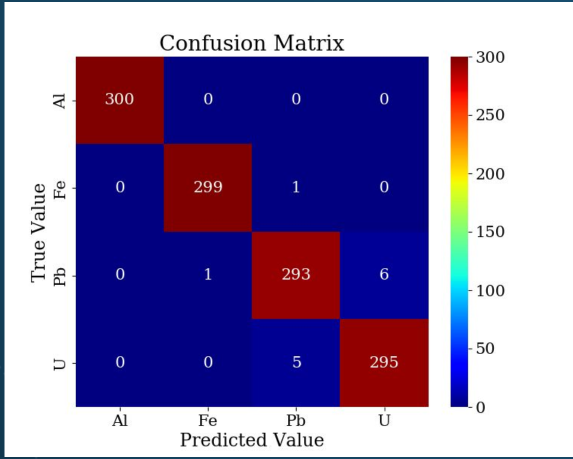


Importance of features wrt classification

- ➢Algorithm used to differentiate high-Z/low-Z materials based on their scattering parameters after then clusters have been identified.
- ➢To achieve this classification, SVM, implemented in Python has been used, which is a supervised learning.
- ➢Above-mentioned: ρ_c, and **Θ** are used as features of this classification.
- ➢ To eliminate extra-noise, a low bound of scattering angle: 10 mrad has been provided.
- ➢The plot of features after background segregation for different time-frames have been given.
- ➢The mean value of the features, from the training pixels and classification has been carried-out by linear kernel function.
- ➢The feature importance has also been given.
- ➢The results have been expressed on the basis of confusion matrix and mean error rate.

Detection results

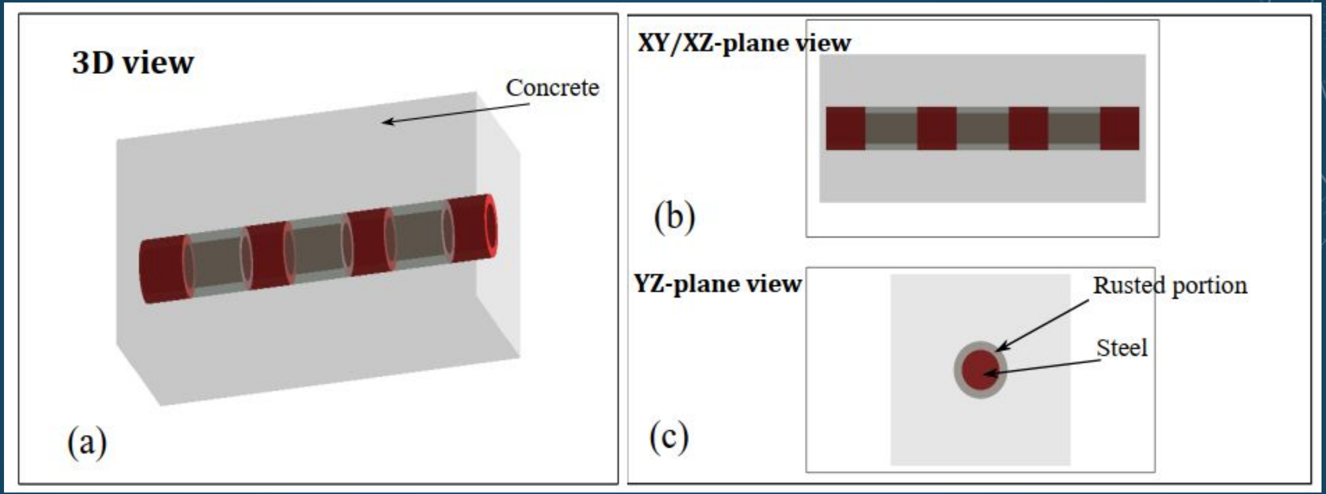
Muon Exposure (hr)	Mean Error Rate (%)							
	Al		Fe		Pb		U	
	Low-Z error	Extreme error	Low-Z error	Extreme error	high-Z error	Extreme error	high-Z error	Extreme error
1	4.5	0	7.3	12.5	25.7	10.8	31.4	1.5
5	0.2	0	0	0.3	14.5	0.8	13.6	0
24	0	0	0	0	2.9	0	1.6	0



low-Z/high-Z error: A low/high Z material wrongly identified in the same category
Extreme error: A low/high Z material wrongly identified in a different category

- ➢ Algorithm used to differentiate high-Z/low-Z materials based on their scattering parameters after then clusters have been identified.
- ➢ To achieve this classification, SVM, implemented in Python has been used, which is a supervised learning.
- ➢ Above-mentioned: ρ_c, and Θ are used as features of this classification.
- ➢ To eliminate extra-noise, a low bound of scattering angle: 10 mrad has been provided.
- ➢ The plot of features after background segregation for different time-frames have been given.
- ➢ The mean value of the features, from the training pixels and classification has been carried-out by linear kernel function.
- ➢ The feature importance has also been given.
- ➢ The results have been expressed on the basis of confusion matrix and mean error rate.

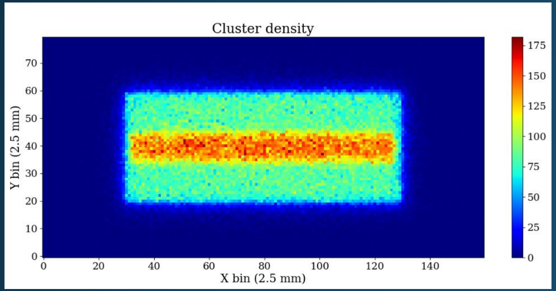
A practical problem: Rusted Rebar

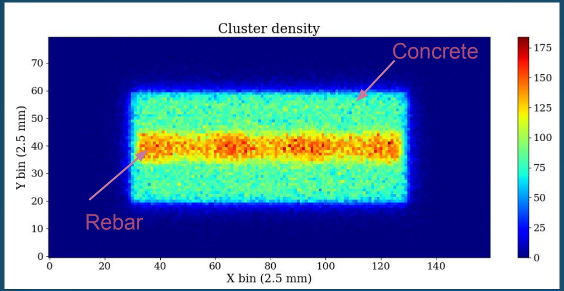


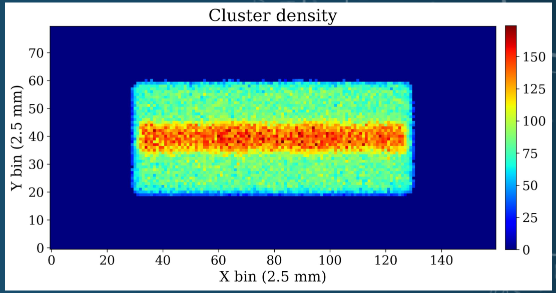
Rusted Rebar

- Steel rebar placed at the center concrete block.
- Partially corroded.
- Outer portion has been corroded with central portion intact.
- Three different materials, concrete, steel, rust.
- Cases, such as 30% and 15% coaxial and circumferential defects have been studied

Rusted Rebar: Reconstructed Images







Rusted Rebar: whole, 30% defect. 15% defect

- The algorithm has been trained with the whole rebar and tested on the defective cases.
- The classification, is between concrete and rod materials.
- Clusters identified using PRM and target identified using SVM.

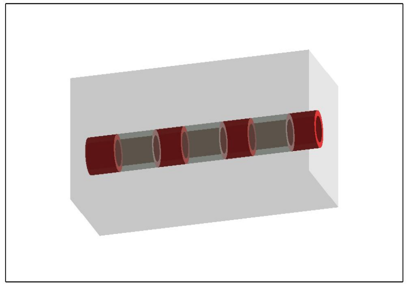
Rusted Rebar: Defect identification

Exposure (days)	Mean Error Rate (%)		
	Whole Rod	15% defect	30% defect
30	1.06	1.4	7.3
15	4.5	10.9	14.7
3	22.3	23.0	27.8

- The classification has been carried out for 3 different muon exposures: 30, 15 & 3 days.
- The 15% defect case, has not been convincingly detected, where are the 30% defect case is identifiable.
- The increase in rust % will increase the mis-identification rates

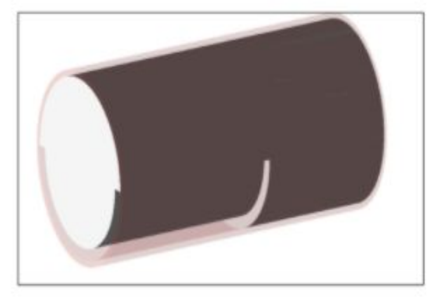
Other Test Cases with PRM and *t*-statistics

Three unique test cases, which are crucial in civil engineering problem have been considered. A variation of defect dimension has been simulated to validate consistency of imaging technique.



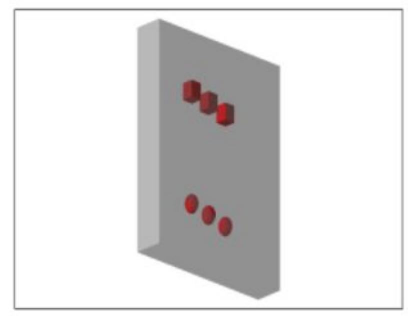
Rusted Rebar

- Steel rebar placed at the center concrete block.
- Partially corroded.
- Outer portion has been corroded with central portion intact.
- Three different materials, concrete, steel, rust.



Voids in CFST

- Steel tubes filled with concrete for better strength.
- Circumferential void.
- Reduces strength of pillars
- Three different materials, concrete, steel, air-void.



Voids in concrete deck

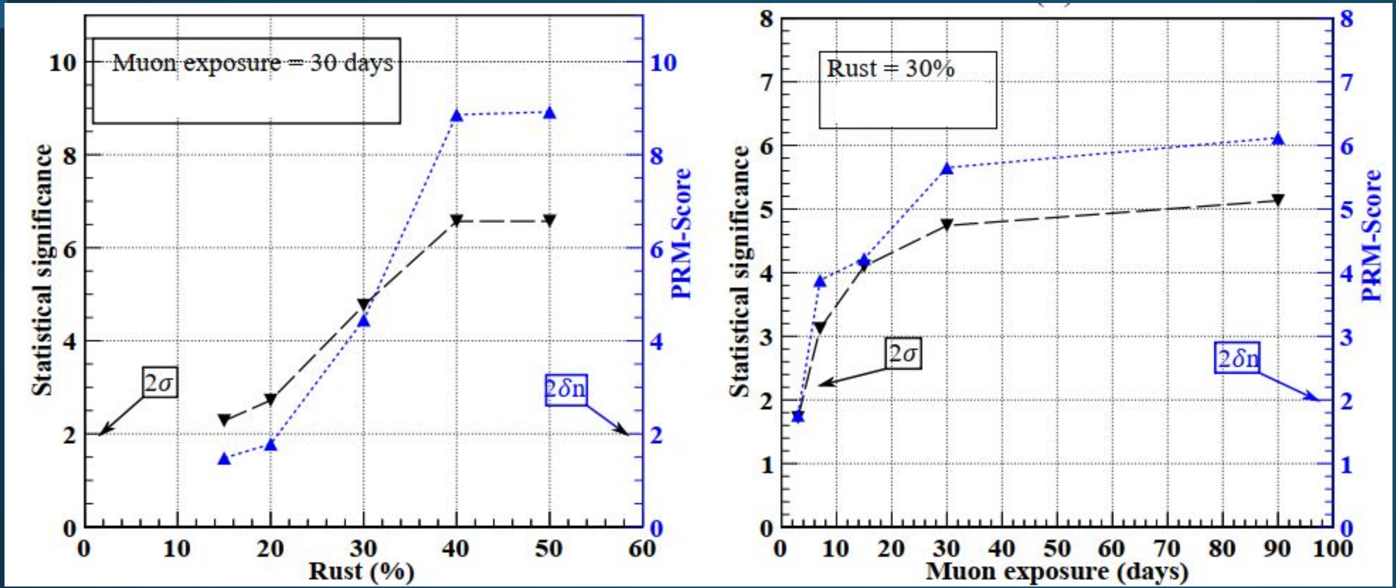
- Voids and delaminations appear in concrete decks.
- Void of two different shapes have been simulated.
- Voids placed at different depth.
- Discrimination between two light materials, concrete , air-voids

Comparison with results from PRM & statistical Analysis

Target type	Defect dimension (mm)	Statistical significance	PRM-score
Rebar in RCC	2.25 4.5	2.29 4.75	1.48 4.45
CFST	7 (side-on) 10 (side-on) 10 (bottom)	5.85 8.10 7.22	5.52 7.94 7.50
Concrete deck	50 60	4.2 4.89	5.62 7.98

S. Tripathy et. al., *The European Physical Journal Plus* volume 136, Article number: 824 (2021)

Comparison with results from PRM & statistical Analysis



(a) Rust thickness variation

(b) Muon exposure variation

Limit of discrimination capablity of MST in concrete structure studied.

- ➔ With increased defect thickness discrimination improves.
- ➔ Identification of defects improves with increasing muon exposure.

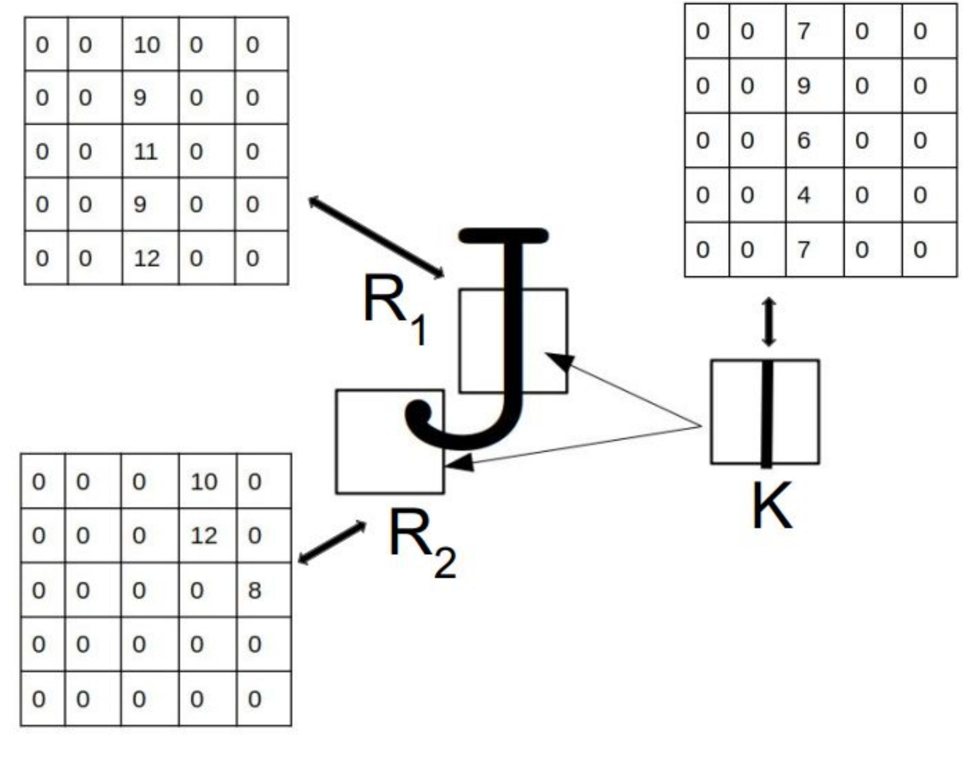
S. Tripathy et. al., The European Physical Journal Plus volume 136, Article number: 824 (2021)

Summary

- ★ Geant4 simulation; whose available physics-lists provide a good platform of monte-carlo simulations, has been used for optimization of MST setup, also used in Imaging Civil structures.
- ★ An PRM technique has been devised, and t-statistics has also been used for material identification.
- ★ This ongoing work demonstrates a cluster-based algorithm to identify target location and supervised learning to classify the targets into pre-trained materials
- ★ The results have also been compared to previously used methods based on examples from civil structures.
- ★ The goal is to achieve faster and quick results for civil structure defects, with similar parameters that can be used in experiments.

Thank you

Pattern Recognition Method (PRM)



A filter ‘K’ searches for a similar pattern ‘R_1’.

- *S*-map represented in terms of a matrix.
- **PRM** searches for similarity with sample in the test image.
- The algorithm learns on the basis of scattering parameter.
- Helps identify position, dimension, shape of target.
- Size of kernel and pixel as per user decision.

$$R_1{*}K : 7*10+9*9+6*11+4*9+7*12=337$$

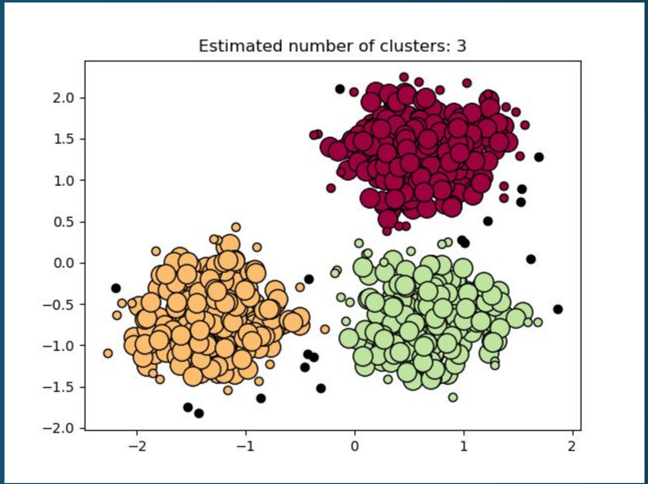
$$R_2{*}K : 7*0+9*0+6*0+4*0+7*0=0$$

S. Tripathy et. al.,Material discrimination in cosmic muon imaging using pattern recognition method *JINST* 15 (2020) 06, P06029

Density Based Spatial Clustering Applications with Noise (DBSCAN)

ε = Neighborhood Distance
minPts = Minimum number of Data in neighbourhood

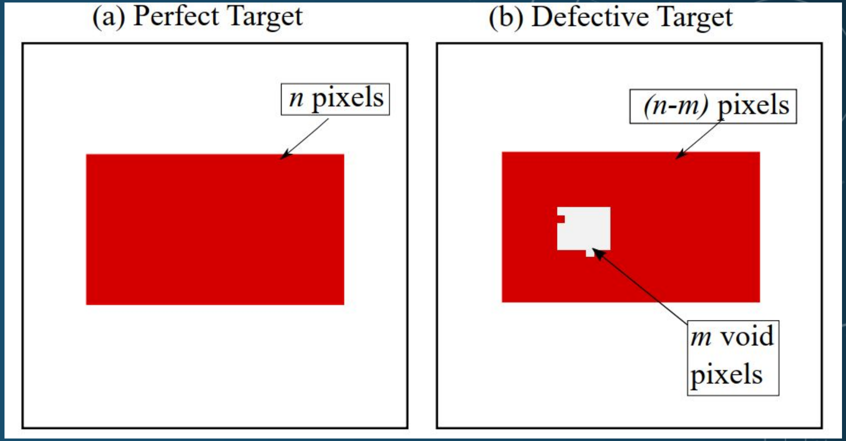
Algorithm:
Input: Several data points
Output: Clusters , noise
Step-1: Begin with arbitrary point: d_i
Step-2: For d_i, check ε and ***minPts***.
Step-3: categorize noise or begin a cluster.
Step-4: mark visited.
Step-5: move to **:** $d_j (i \neq j)$
Step-6: expand cluster or begin new cluster or mark noise.
Step-7: go to *Step-5*.



Source: scikit-learn.org

Analysis based on *PRM-score*

- To numerically quantify the degree of discrimination a metric, namely, *PRM-score* has been introduced.
- The *PRM-score* specifies the difference between images reconstructed after PRM-processing of the test and the reference target measured in the units of *δn*, which is the random error arising out of repeated measurement of PRM on the reference target.
- Higher the PRM-score, the reference and test are more distinguishable.
- PRM-score > 2*δn* considered as threshold for defect identification.



Comparison of images for perfect target (reference), R_p, and defective target (test), R_d, based on PRM-score.

$$\text{PRM-score} = \frac{\text{No. of pixels in '}R_p\text{' }(n)\text{-No. of pixels in '}R_d\text{' }(n-m)}{\text{No. of pixels in '}R_p\text{' }(n)} \times \frac{1}{\delta n}$$

After simpifying,

$$\text{PRM-score} = \frac{m}{n \times \delta n}$$

S. Tripathy et. al., arXiv:2102.08913 (under review)

Results: Rusted Rebar

- Defective rebars with 15% and 30% defects have been shown. The defects have been analyzed using t-test and PRM-score.
- The 30% case has been clearly distinguished with $> 4\sigma$ with t-statistics and $> 4\delta n$ from the without-defect case.
- The 15% case has been identified with $> 4\sigma$ but with $< 2\delta n$ PRM-score.

S. Tripathy et. al., *The European Physical Journal Plus* volume 136, Article number: 824 (2021)