UNIONE EUROPEA
Fondo Sociale Europeo

Ministero dell'Università e della Ricerca

PON RICERCA E INNOVAZIONE 2014 - 2020

UNIVERSITÀ DEGLI STUDI FIRENZE
Da un secolo, oltre.

DOTTORATO DI RICERCA IN
Fisica e Astronomia

CICLO XXXVII - PON ex DM1061

Development of AI-based diagnostic systems applied to muon radiography for monitoring applications

Settore Scientifico Disciplinare
Fis/04 (PHYS-01/A)

Dottorando

Dott. Paccagnella Andrea

Supervisori

Dott. Frosin Catalin

Prof. D'Alessandro Raffaello

Coordinatore

Prof. Giovanni Modugno

Contents

1 Cosmic rays and the muon radiography tecnique **5**
- 1.1 Cosmic rays . . . 5
  - 1.1.1 Primary spectra . . . 6
  - 1.1.2 Cosmic rays in the atmosphere . . . 7
  - 1.1.3 Muon flux to the ground . . . 9
  - 1.1.4 Muon flux underground . . . 11
- 1.2 Physical processes used in muon radiography . . . 13
  - 1.2.1 Muon interaction with matter . . . 14
  - 1.2.2 Scattering technique . . . 17
  - 1.2.3 Transmission technique . . . 18
- 1.3 Fast simulation for the muon transmission technique . . . 21
  - 1.3.1 Simulation geometry . . . 21
  - 1.3.2 The muon flux generator . . . 22
  - 1.3.3 Range-based simulation . . . 23
  - 1.3.4 Transmission-opacity conversion . . . 26

2 Detectors for muon radiography **29**
- 2.1 MIMA Detector . . . 32
  - 2.1.1 Concept and design of the MIMA muon telescope . . . 32
  - 2.1.2 The photosensors . . . 35
  - 2.1.3 Front end electronics . . . 36
  - 2.1.4 Track Reconstruction . . . 37
- 2.2 BLEMAB project . . . 40
  - 2.2.1 Concept of detector . . . 42
  - 2.2.2 Construction of the detector . . . 43
  - 2.2.3 Installation . . . 45

3 Monitoring water infiltration inside tailing dams **47**
- 3.1 The study case of the tailing dams . . . 48
- 3.2 Simulation of the infiltration inside the dam . . . 48
  - 3.2.1 First result . . . 52

3.3 Use of adaptive binning . . . 52
3.3.1 Voronoi Tessellation . . . 54
3.4 Use of the Voronoi tessellation in the Tailing dam . . . 59
3.5 Use of the Random Forest classifier in the tailing dams case . . . 62
3.5.1 Introduction to Decision Trees . . . 63
3.5.2 Introduction of Random Forest Classifier . . . 64
3.5.3 Data preprocessing . . . 64
3.5.4 Results of the Random Forest . . . 67
3.5.5 Confusion Matrix . . . 71
3.5.6 Evaluation of test data . . . 73
3.5.7 Visual Test . . . 75
3.6 Tailing Dams Conclusion . . . 77

4 Use of a UNet algorithm for cavity identification and segmentation within mines 79
4.1 The measure at Temperino mine . . . 80
4.2 Generation Dataset . . . 81
4.2.1 Generation of measured transmission maps . . . 81
4.2.2 Preprocessing of data . . . 83
4.3 Neural Networks . . . 85
4.3.1 Introduction to image recognition . . . 86
4.3.2 UNet . . . 89
4.3.3 Metrics and Loss . . . 92
4.4 First UNet Results . . . 94
4.4.1 Performance Visualization . . . 95
4.4.2 Test Networks . . . 96
4.5 Bayesian Optimization for Hyperparameter Optimization . . . 99
4.5.1 Theory of Bayesian Optimization . . . 101
4.5.2 Bayesian Optimization Process . . . 101
4.5.3 Application to the studied case . . . 102
4.6 Results after Hyperparameter Optimization . . . 105
4.6.1 Comparison with Previous Results . . . 105
4.6.2 Test Phase Results . . . 105
4.6.3 Discussion of Results . . . 107
4.7 Model behavior on real data . . . 109
4.8 UNet Conclusions . . . 111

5 Summary and conclusions 115
5.1 Summary of activities . . . 115
5.2 Conclusions and future developments . . . 116

Introduction

In recent years, the application of muon radiography has gained considerable traction across various fields due to its non-invasive and highly effective imaging capabilities. Originating from the study of cosmic rays, muon radiography leverages naturally occurring muons generated in the upper atmosphere to penetrate dense materials. This imaging technique has proven to be particularly advantageous in contexts where conventional methods struggle due to environmental constraints or the need for non-destructive analysis. Building on this premise, this thesis explores and advances the implementation of muon radiography, applying it to both industrial and geological settings.

This doctoral research focuses on two primary objectives. The first is the development and deployment of a scintillating bar detector system for the BLEMAB project. BLEMAB (Blast Furnace Stack Density Estimation through Muon Absorption Measurements) represents a significant advancement in the use of muon radiography for real-time monitoring of density variations in blast furnaces. The thesis delves into the design, construction, and installation of this innovative detector, specially engineered to endure the harsh conditions within an industrial setting. Through muon absorption, the detector enables precise density mapping, a capability that holds the potential to revolutionize blast furnace management by offering insights into internal processes that were previously difficult to observe.

The second objective centers on applying AI-based diagnostic systems to muon radiography applications for monitoring of geological structures. By focusing on tailing dams and underground mining environments, this thesis demonstrates that the use o machine learning techniques, combined with muon radiography, can lead to promising results in high-density environments where traditional methods may fall short. Through the detection of water infiltration in tailing dams and the segmentation of underground cavities in mining applications, this research employs muon transmission data coupled with advanced machine learning algorithms, such as Random Forest and UNet, to enhance imaging accuracy and segmentation. Given the need for a large amount of data to train the aforementioned models and the limited availability of databases with muographic measurements, a fundamental part of this thesis was the development of software capable of replicating

a muographic measurement in a short time. The work addresses critical issues of structural integrity and environmental safety, offering a framework for early detection of potential hazards.

In both industrial and geological settings, the thesis aims to contribute with innovative methodologies and practical solutions, expanding the scope of muon imaging and establishing it as a critical tool for structural and environmental monitoring.

This thesis work is organized as following. Chapter 1 introduces the principles of cosmic rays and muon interactions with matter, providing the foundational knowledge for understanding muon radiography. Chapter 2 discusses the BLEMAB project, outlining the design and capabilities of the scintillating bar detectors developed for blast furnace monitoring. Chapter 3 and 4 addresses the challenges of muon radiography in geological structures, illustrating the novel use of AI and machine learning for water infiltration detection and cavity mapping in mines. Lastly, Chapter 5 summarizes the contributions of this research, presents conclusions, and suggests directions for future exploration in the field of muon radiography.

Chapter 1

Cosmic rays and the muon radiography tecnique

Primary Cosmic rays consist of highly energetic elementary particles and nuclei, which show a sharply declining power spectrum ranging from a few MeV to 10^{20} eV per particle. Direct measurements of primary cosmic rays can be conducted using space or balloon experiments at energy levels where the flux is significant, as show in Figure 1.1. The interactions of primary cosmic rays within the atmosphere generate a flux of secondary elementary particles (secondary cosmic rays), detectable in the atmosphere, on the Earth's surface and underground. Hereafter we will detail the two types and see how the latter can be used in practical applications, like muon radiography. This technique uses secondary cosmic rays, in particular the muons generated by the interactions of primary cosmic rays with the nuclei of the atmosphere, as an imaging tool.

1.1 Cosmic rays

Cosmic rays are a stream of highly energetic particles originating from outer space that constantly bombard the Earth. These particles are primarily composed of protons, but also include heavier atomic nuclei and a small fraction of electrons and positrons. The study of cosmic rays is crucial for understanding a wide range of astrophysical phenomena, from the processes occurring in distant galaxies to the fundamental mechanisms of particle interactions [2].
The discovery of cosmic rays dates back to the early 20th century. In 1912, the Austrian physicist Victor Hess conducted a series of balloon flights to measure ionizing radiation levels in the atmosphere. He found that the radiation intensity increased with altitude, leading to the conclusion that a significant portion of this radiation originated from outer space. Indeed, as explained, primary cosmic

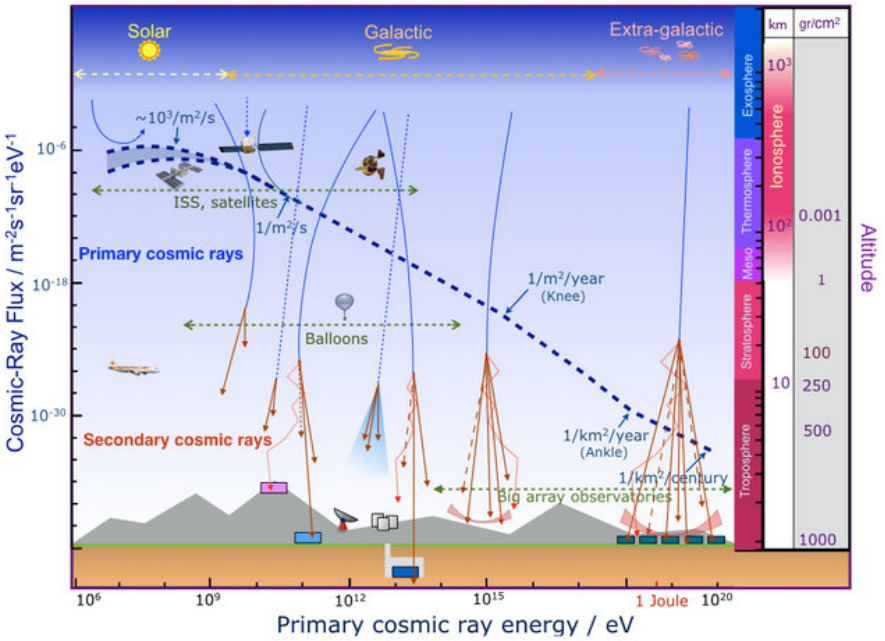


Figure 1.1: Primary cosmic ray spectrum as a function of the energy. The blue dashed line represents the cosmic rays flux using the scale units shown on the left vertical axis. Below energies $\sim 10^{14}$ eV primary cosmic rays are measured directly with detectors placed in satellites and balloons. At higher energies indirect measurements done by ground-based arrays of detectors are used. The right vertical axis shows the atmosphere layers, their heights and the corresponding air pressure [1].

rays originate outside the Earth's atmosphere, while secondary cosmic rays are produced by interactions of primary cosmic rays with the Earth's atmosphere.

1.1.1 Primary spectra

Primary cosmic rays, are composed of approximately 98% nucleons, in the form of protons, helium and other nuclei, stripped of all their orbital electrons. The remaining 2% is composed of leptons. The majority of these particles originate from star-like sources. The components of antiproton, positron and other nuclei (such as lithium, beryllium and boron) are thought to originate from the interaction of the primary cosmic ray particle with interstellar gas. A small fraction of antiproton and positron may be primary, although this is a topic of current interest [2]. A small component of this primary cosmic rays is generated by the sun, but most of these cosmic rays comes from the solar system. Incoming charged particles are

influenced by the solar wind, which decelerates and partially blocks lower energy galactic cosmic rays from entering the inner solar system [2]. There is a notable anticorrelation between solar activity, which follows an alternating eleven-year cycle, and the intensity of cosmic rays with rigidities below approximately 10 GV [2]. Additionally, the geomagnetic field impacts lower-energy cosmic rays, as they must penetrate this field to reach the top of the atmosphere. Consequently, the intensity of any component of cosmic radiation in the GeV range varies depending on both location and time.
The spectra of the components of the cosmic radiation can be describe by particles per unit rigidity. Propagation (and probably also acceleration) through cosmic magnetic fields depends on gyroradius or magnetic rigidity, R, which is the gyroradius r_L multiplied by the magnetic field strength B:

$$R = \frac{pc}{Ze} = r_L B. \tag{1.1}$$

Figure 1.2 shows the major nuclear components for kinetic energies greater than 0.22 GeV/nucleus.

1.1.2 Cosmic rays in the atmosphere

When primary cosmic rays enter the Earth's atmosphere they interact with atoms and molecules, especially oxygen and nitrogen, trough strong interaction and electromagnetic processes (like pair production). The collision produces a cascade of lighter particles (secondary cosmic rays) including X-rays and γ, neutrons, mesons (such as pions and kaons), electrons and muons. If sufficiently energetic, hadrons will themselves initiate new hadronic interactions building up a hadron cascade, which for energetic events constitutes the core of a shower. However, unstable particles such as pions, kaons and other particles are subject to decay.
Due to a very short life ($\tau_{\pi^0} \approx 8.4{\cdot}10^{-17}$ s) neutral pions (π^0) decay almost instantly in to a pair of photons. The latter can produce electron-positron pairs which subsequently undergo bremsstrahlung. Eventually, these repetitive processes build up an electromagnetic shower in the atmosphere.
Charged pions ($\pi^{\pm}$) have a mean life at rest of about $\tau_{\pi^{\pm}} \approx 2.6 \cdot 10^{-8}$ s and decay via the processes:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$

$$\pi^- \rightarrow \mu^- + \overline{\nu}_\mu$$

into muons and neutrinos.
The charged kaons ($K^{\pm}$) are strange mesons and have a mean life of about 12 ns. They also mostly decay (with a branching ratio of 64%) as:

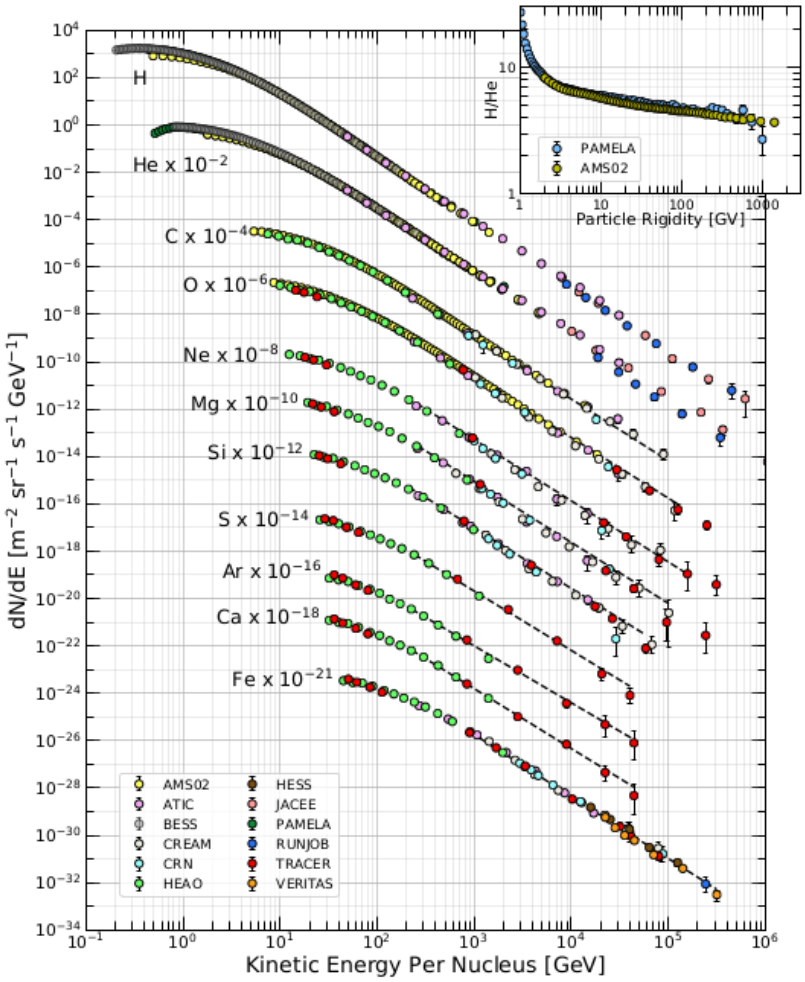


Figure 1.2: The energy spectrum of cosmic rays, spanning from 100 MeV to 10^6 GeV, is presented in the form of a graph, with the x-axis representing the kinetic energy per nucleus in GeV and the y-axis representing the flux of nuclei of a given energy per nucleus. [2]. The inset shows the H/He ratio as a function of rigidity.

$$K^+ \rightarrow \mu^+ + \nu_\mu$$

$$K^- \rightarrow \mu^- + \overline{\nu}_\mu$$

Other short lived particles have decay processes that lead to the formation of pions which in turn decay as described above.

Figure 1.3 shows the vertical fluxes of the major cosmic-ray components in the atmosphere. As we can see from this image, muons are the most abundant particle at ground level.

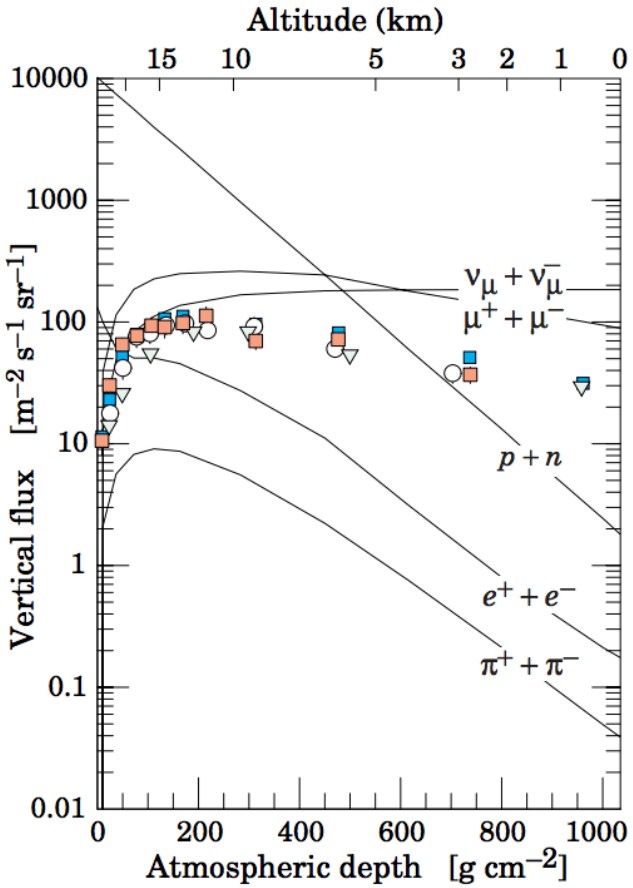


Figure 1.3: Vertical fluxes of cosmic rays in the atmosphere with E>1GeV estimated from the nucleon flux [2]. The points show measurements of negative muons with E_μ >1GeV.

1.1.3 Muon flux to the ground

The muon belongs to the lepton family and has a mass of 106 MeV/c^2 (about 200 times that of the electron). It is an unstable particle and has a mean life of about 2.2 μs, the second longest among unstable particles after the neutron. Being the lightest charged particle after the electron, the muon can almost exclusively decay according to the process:

$$\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu,$$

and its positive antiparticle decays according to

$$\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu,$$

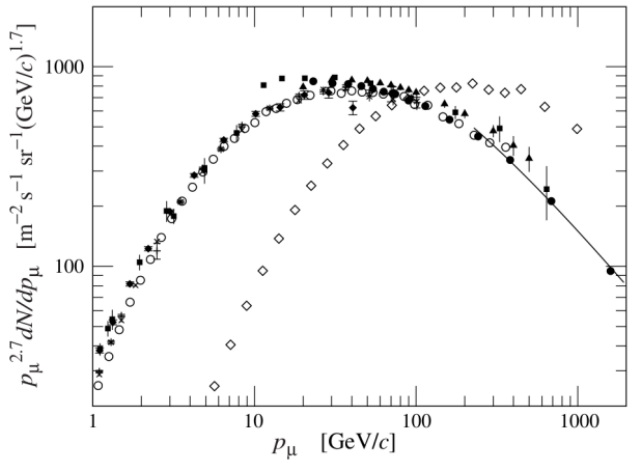


Figure 1.4: The spectrum of muons at sea level for $\theta = 75°$ is indicated by the empty diamonds while the other markers refer to the muon spectrum for $\theta = 0$. The average energy of muons increases as the zenith angle of origin increases. The y-axis is multiplied by $p_\mu^{2.7}$ to compress the spectra [2].

Muons are produced at a height of about 10-15 km in the atmosphere and lose about 2 GeV of energy through ionization before reaching the ground. Their energy and angular distribution at the ground reflects a convolution of the production spectrum, energy loss in the atmosphere, and decay. The energy spectrum is almost flat below 1 GeV, gradually steepening to reflect the primary spectrum in the 10-100 GeV range. The integral intensity of vertical muons above 1 GeV/c at sea level is approximately 70 $m^{-2}s^{-1}sr^{-1}$ [2]. The overall angular distribution of muons at ground as a function of zenith angle θ is proportional $\cos^2\theta$, which is characteristic of muons with $E_\mu \sim 3$ GeV. At lower energy, the angular distribution becomes increasingly steep, while at higher energy, it flattens, approaching a secθ distribution.

Figure 1.4 shows the muon energy spectrum at sea level for two angles. At large angles, low-energy muons decay before reaching the surface and high-energy pions decay before they interact, thus the average muon energy increases. An approximate extrapolation formula, which is valid when muon decay is negligible ($E_\mu > 100/\cos\theta$ GeV) and the curvature of the Earth can be neglected ($\theta < 70°$), is given by [2]:

$$\frac{dN_\mu}{dE_\mu d\Omega} \approx \frac{0.14\ E_\mu^{-2.7}}{cm^2\ s\ sr\ GeV} \times \left\{ \frac{1}{1 + \frac{1.1\ E_\mu\ cos\theta}{115\ GeV}} + \frac{0.054}{1 + \frac{1.1\ E_\mu\ cos\theta}{850\ GeV}} \right\} \tag{1.2}$$

where the two components represent the contributions from pions and charged kaons. Equation 1.2 disregards a minor contribution from charm and heavier fla-

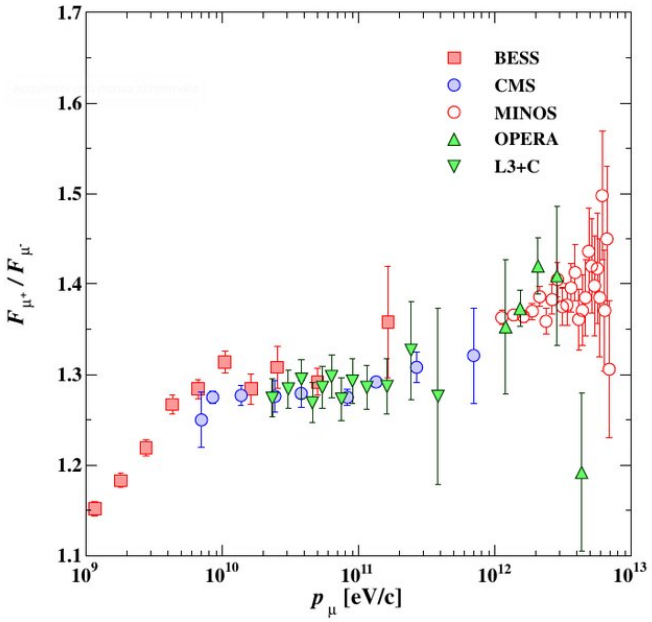


Figure 1.5: Muon charge ratio as a function of the muon momentum [4].

vors, which becomes significant only at extremely high energies [2]. The muon charge ratio reflects the prevalence of π^+ over π^- and K^+ over K^- in the forward fragmentation region of interactions initiated by protons, The higher abundance of protons, both free and bound within nuclei, compared to neutrons in the primary beam [cf. [3]] has significant implications. This imbalance influences the dynamics of nuclear reactions and the production of secondary particles, as the interaction properties of protons can differ substantially from those of neutrons. Such differences are particularly relevant in applications like modeling cosmic-ray bombardment effects or in radiation analyses, where the predominance of protons can largely determine the overall behavior of the system under study. Figure 1.5 illustrates the muon charge ratio as a function of momentum for several experiments. It can be observed that the ratio is consistently greater than one, indicating a prevalence of positive muons relative to negative muons. This finding is consistent with the observation that the majority of primary cosmic rays are protons.

1.1.4 Muon flux underground

Only muons and neutrinos can penetrate significant depths underground. Muons dissipate their energy through ionization and radiative processes. These aspects will be further discussed in subsequent sections. The total energy loss of muons can be expressed as a function of the material traversed:

E_μ	R	a	b_{brems}	b_{pair}	b_{nucl}	$\sum b_i$
[GeV]	[km.w.e.]	[MeV g^{-1} cm^2]	—	10^{-6} g^{-1}	cm^2	—
10	0.05	2.17	0.70	0.70	0.50	1.90
100	0.41	2.44	1.10	1.53	0.41	3.04
1000	2.45	2.68	1.44	2.07	0.41	3.92
10000	6.09	2.93	1.62	2.27	0.46	4.35

Table 1.1: Average muon range R and energy loss parameters calculated for standard rock. Range is given in km-water-equivalent, or 10^5 g cm^{-2} [2].

$$\left\langle \frac{-dE}{dx} \right\rangle = a(E) + b(E)E, \tag{1.3}$$

where E is the total energy, $a(E)$ is the electronic stopping power (ionization loss), and $b(E)$ is due to radiative processes: *bremsstrahlung*, pair production, and photonuclear interaction:

$$b \equiv b_{brems} + b_{pair} + b_{nucl}.$$

This will be clarified in more detail in the next sections. The quantity $\epsilon \equiv a/b$ (approximately 500 GeV in standard rock) defines a critical energy below which continuous ionization loss becomes more significant than radiative losses.

Table 1.1 shows a and b values for standard rock as a function of muon energy. The second column of the table shows the muon range in standard rock (A = 22, Z = 11, ρ = 2.65 g cm^{-3}). These parameters are highly sensitive to the chemical composition of the rock, which needs evaluation for each location. This aspect is crucial to consider during a muographic study. When conducting simulations for the targeted area, an average density representing the entire region is typically selected. Since part of this project involves developing a neural network for anomaly detection within mines, choosing the correct density becomes fundamental for successful training. The underground muon intensity can be estimated based on the muon intensity in the atmosphere and their rate of energy loss.

Figure 1.6 depicts the vertical muon intensity versus rock depth. The flat section of the curve arises from muons locally produced by charged-current interactions of ν_μ. The inset shows the vertical intensity curve for water and ice. It is not as steep as the one for rock because of the lower muon energy loss in water [2]. Therefore, Figure 1.6 provides insights into the feasibility of underground muographic measurements. For example, we can consider the reference detector of the Muon Radiography Group of Florence, MIMA (which will be discussed in detail in the following chapter), shaped like a cube measuring 50 x 50 x 50 cm^3. Aiming to conduct muon radiography at the national laboratories of Gran Sasso

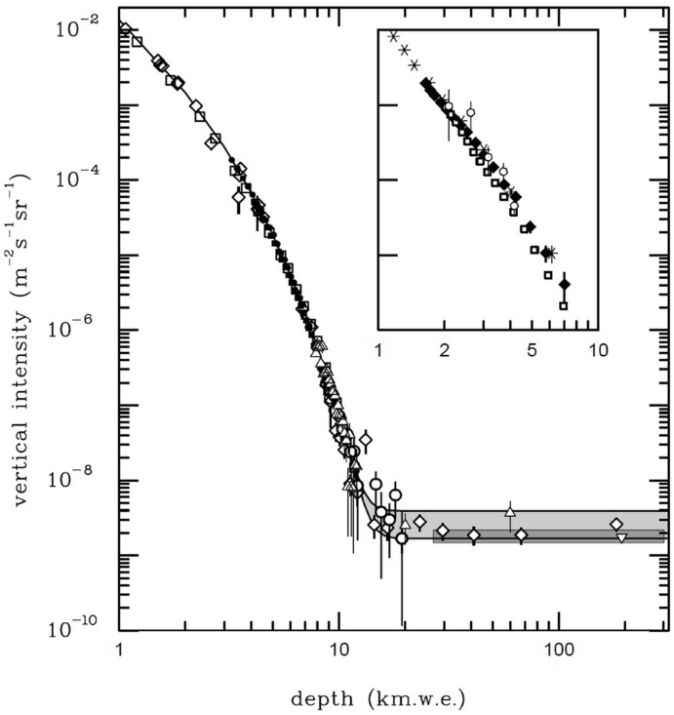


Figure 1.6: Vertical muon intensity vs depth (1 km.w.e.= 10^5 g cm^{-2} of standard rock). The points come from some experiments [2]. The shaded area at large depths represents neutrino-induced muons of energy above 2 GeV. The upper line is for horizontal neutrino-induced muons, the lower one for vertically upward muons. Darker shading shows the muon flux measured by the SuperKamiokande experiment.

(located under 1400 m of rock) using this detector, we would observe a muon only every 3 months. This makes such measurements practically impossible.

1.2 Physical processes used in muon radiography

After the discussion of the origin of the muons used in muon radiography and their flux on and beneath the ground, we will now discuss how these particles interact with matter. In fact, a fundamental part of a muographic measurement involves simulations of the environment to be investigated. Therefore, it is essential to understand how muons, depending on their energy, interact with matter, in order to have highly accurate simulation models.
Muons lose energy through scattering with the nuclei of the material they pass through. This energy loss is described by the quantity called *stopping power*. Due

to their high mass, muons undergo relatively low energy loss per unit distance traveled compared to electrons. This allows muons to penetrate large thicknesses of material, making them particularly useful in applications such as muon radiography. Another effect of muon interaction with matter is multiple Coulomb scattering, which describes the deflection of the muon's trajectory as it passes through a material.
Based on these two types of interactions, two muon imaging techniques can be defined: muon transmission radiography (also known as muon absorption radiography) and muon scattering tomography (often referred to as muon tomography). The first technique exploits the variation in muon flux depending on the type and density of the material traversed, while the second relies on the deflection of muon trajectories due to multiple Coulomb scattering within the material. In the muon radiography group of Florence, the muon transmission radiography technique is employed.

1.2.1 Muon interaction with matter

The mean stopping power for high-energy muons in a material can be described by formula 1.3.

The quantity called critical energy ($E_{\mu c}$) is defined as the energy at which radiative and ionization losses are equal, and can be found by solving $E_{\mu c} = a(E_{\mu c})/b(E_{\mu c})$. Below $E_{\mu c}$ ionization losses dominate, and above $E_{\mu c}$ radiative effects dominate. The dependence of $E_{\mu c}$ on atomic number Z is shown in Figure 1.7. The critical energy is higher for gases than for solids or liquids of the same atomic number due to a reduction in the density effect and the associated reduction in ionization losses.

In muon radiography most of the muons that interact with the environments that we want to investigate, have an energy below of $E_{\mu c}$, for this reason the dominant contribution of the Eq. 1.3 is $a(E)$.

Figure 1.8 shows the stopping power, defined as the energy loss per unit of crossed opacity, for a positive muon in copper. The a parameter of equation 1.3 is given by the Bethe-Bloch formula [2]:

$$\left\langle -\frac{dE}{dx} \right\rangle = 4\pi N_A r_e^2 m_e c^2 \frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ \ln\left(\frac{2m_e c^2\beta^2\gamma^2 W_{max}}{I^2}\right) - \beta^2 - \frac{\delta}{2}\right], \qquad (1.4)$$

where I is the mean excitation energy, W_{max} is the maximum possible energy transfer to an electron in a single collision given by

$$W_{max} = \frac{2m_e c^2\beta^2\gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2},$$

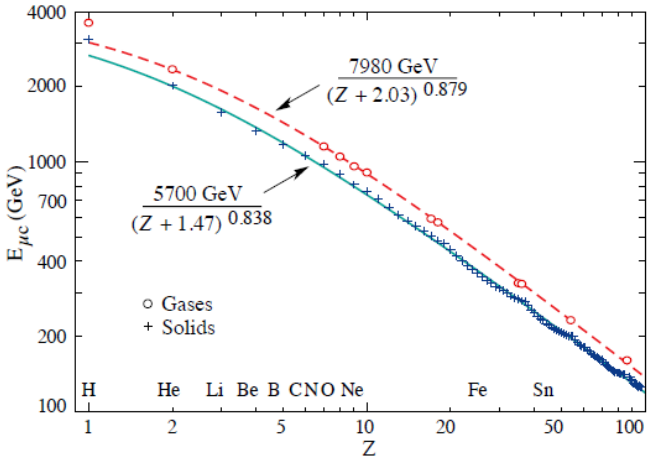


Figure 1.7: Muon critical energy for different chemical elements, defined as the energy at which radiative and ionization energy loss rates are equal. The fits shown are for solids and liquids (solid line) and gases (dashed line). [2].

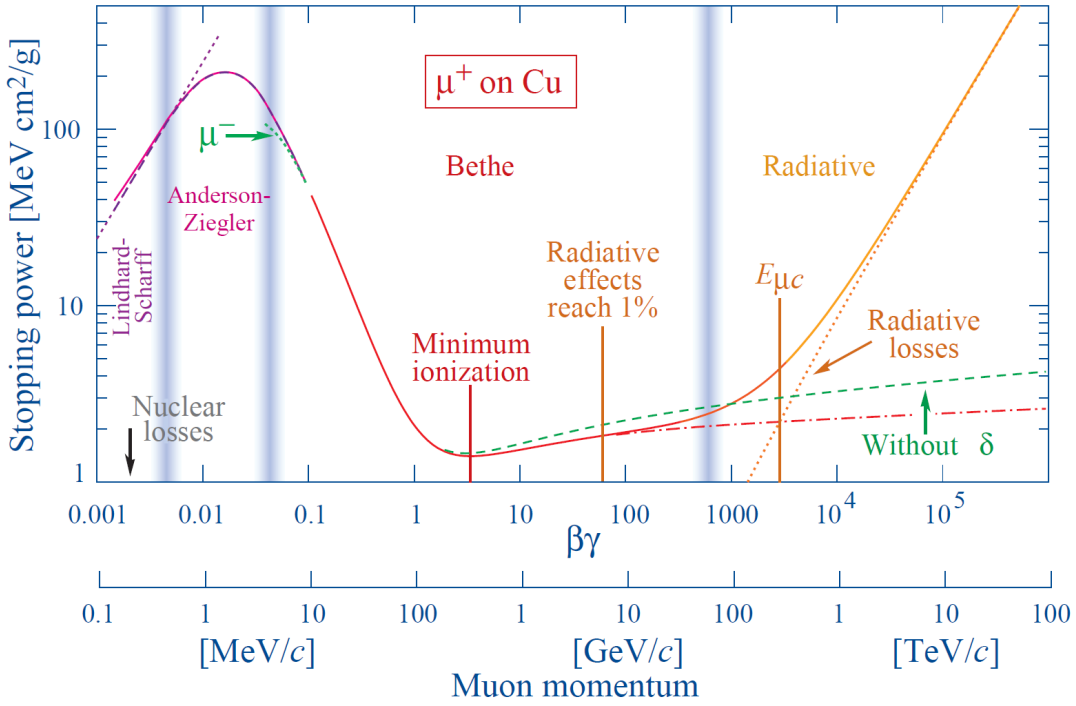


Figure 1.8: Stopping power of a positive muon in copper as a function of its momentum. Solid curve indicate the total stopping power. The vertical bands delimit the regions where different approximations have been used [2].

N_A is the Avogadro number, r_e is the classical electron radius, m_e is the mass of the electron, z, βc and M are the charge, velocity and the mass of incident particle, Z and A are the atomic number and mass number of the absorbed and δ is the correction factor for the density effect. The stopping power function are characterized by a rather large minimum zone whose position ranges from $\beta\gamma = 3.0$ to 3.5 for Z ranging from 7 to 100. Most relativistic cosmic ray particles have stopping power close to the minimum. The continuous-slowing-down-approximation (CSDA) range $X(E)$, of a muon with initial energy E_o, is given by the inverse of equation 1.3

$$X(E) = \int_{E_0}^{E} [a(E') + b(E')E']^{-1} dE'. \tag{1.5}$$

Another process that is important to muon radiography is scattering in the matter. A charged particle passing through some material is deflected by many small-angle scattering. For a muon, most of this deflection is due to Coulomb scattering from nuclei, and hence the effect is called multiple Coulomb scattering. For many small-angle scatters, the net scattering and displacement distributions are Gaussian by the central limit theorem. Defining θ_0 as the characteristic scattering angle, we have

$$\theta_0 = \theta_{plane}^{rms} = \frac{1}{\sqrt{2}} \theta_{space}^{rms}, \tag{1.6}$$

where θ_{plane}^{rms} represents the root-mean-square (rms) scattering angle when the deflection is projected onto a given plane, and θ_{space}^{rms} is the rms scattering angle considering the full three-dimensional (space) distribution. In other words, while θ_{space}^{rms} describes the overall angular deviation in 3D, θ_{plane}^{rms} corresponds to the component of this deviation observable in a two-dimensional projection, as illustrated in Figure 1.9. This relation arises from the assumption of isotropic scattering in the transverse directions.

It is sufficient for many applications to use a Gaussian approximation for the central 98% of the projected angular distribution, with an rms width given by Lynch and Dahl [2]:

$$\theta_0 = \frac{13.6\ MeV}{\beta c p} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038\ \ln \left(\frac{x z^2}{X_0 \beta^2} \right) \right] \tag{1.7}$$

Where p are the momentum of the incident particle, and x/X_0 is the thickness of the scattering medium in radiation lengths. As we can see the scattering angle depends on the material and the momentum of the incoming particle.

In the following subsections the two main techniques used in muography will be explained.

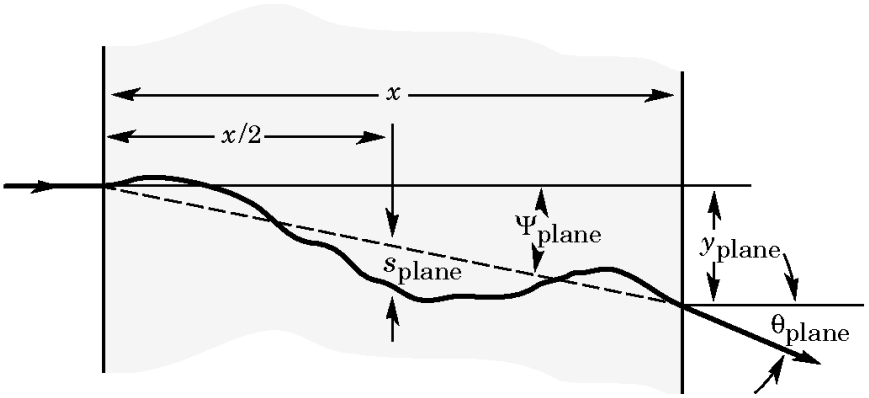


Figure 1.9: Schematic representation of the quantities used to describe multiple Coulomb scattering. The diagram defines the projected (planar) scattering angle, θ^{rms}_{plane}, and the space scattering angle, θ^{rms}_{space}, highlighting the relation $\theta_0 = \theta^{rms}_{plane} = \frac{1}{\sqrt{2}}\theta^{rms}_{space}$, where θ_0 is the characteristic scattering angle. The figure also illustrates the path of the incoming particle and the geometric components associated with the cumulative small-angle deviations due to scattering [2].

1.2.2 Scattering technique

Muon scattering technique uses two detectors, placed upstream and downstream of the object to be investigated. Scattering exploits the deflection of muons in matter due to multiple coulomb scattering (MCS). The scattering angle distribution (Eq. 1.7) of a muon of momentum p projected in a plane can be approximated by a Gaussian with mean 0 and width:

$$\sigma = \frac{13.6\ MeV}{pc}\sqrt{\frac{L}{X_0}}, \tag{1.8}$$

where X_0 is the radiation length (which depends on the atomic number Z of the material) and L is the thickness of the material crossed by the particle. Measuring the deviation angles allows for the reconstruction of information about the radiation length of an unknown material. The deflection of muons is particularly relevant for low-momentum muons. In fact, muons with greater momentum will result in a smaller r.m.s. Therefore, identifying the minimum momentum plays a fundamental role. It must be noted that the momentum of an individual particle is in general unknown, but it can be substituted by a fixed value computed from the average value $\langle 1/p^2 \rangle$ of the $1/p^2$ distribution. The simplest reconstruction method is based on the single scattering approximation (SSA), which assumes that the scattering of any individual muon is concentrated in a single point. This point coincides with the point of closest approach (POCA) of the straight lines (in space) measured by the two detectors. A map of the material linear scattering density

(LSD) can be obtained by assigning to the POCA reconstructed for the i^{th} muon a weight proportional to $\Delta\theta_i^2$, where $\Delta\theta_i$ is the measured projected scattering angle. The method is computationally very fast and works well for cases where an object has a much higher LSD than the rest of the volume. However, it tends to fail in the presence of multiple scattering centers. To improve spatial resolution and provide a more accurate density reconstruction, advanced statistical techniques such as the Maximum Likelihood Expectation Maximisation (MLEM) algorithm can be used. MLEM iteratively refines the estimation of the target's density distribution by maximizing the likelihood function of the measured scattering angles. This method is particularly effective in handling complex geometries and multiple scattering sources, as it enhances the separation of different materials within the scanned volume [5]. Scatter tomography allows the direct reconstruction of the target's density and atomic number by measuring the deflection of muons passing through it. Examples of this technique can be found in [6] and [7].

1.2.3 Transmission technique

To gain insight into the fundamental principles of muography, we may consider the schematic representation illustrated in Figure 1.10. By placing a detector below an environment to be investigated and plotting the number of muons observed as a function of position on a histogram, it becomes evident that a greater number of these will be detected in the presence of a cavity than in the presence of a dense body. This is due to the fact that, as previously described, a portion of the muons stops inside the dense body and decays. The remaining muons are deflected, and only the most energetic components reach the detector.

Muon transmission radiography uses only one detector placed downstream of the object to be investigated (i.e. mines or pyramids), comparing the flux of muons in clear sky (or in jargon called free-sky measurement) configuration with the flux of muons coming from the object of the study. This technique is based on the energy loss (and the eventual decay) of the muons when they interact with matter. The relationship between muon momentum and average crossed opacity can be found in Ref. [8] and, it will be discussed in detail in the next section.
We refer to the reference system shown in Figure 1.11, for all future discussions and the presentation of the results, where the angle θ represents the zenith angle, the angle φ represents the azimuth angle, and e represents the elevation, that is, the complementary of the zenith angle.

Fixed a certain angular direction (θ, φ) and a certain opacity $(X(\theta, \varphi) = \int_0^L \rho(\theta, \varphi) \cdot dl)$, the expected value of the flux transmitted through the target can be estimated as the integral from $p_{\min}$ to infinity of the differential flux in that

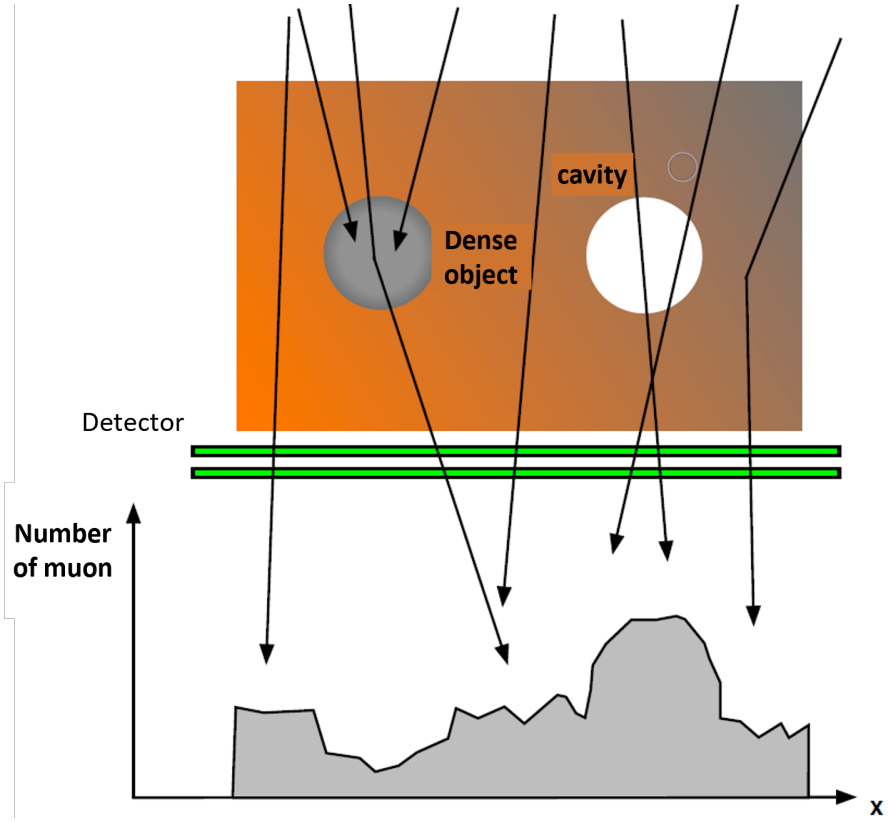


Figure 1.10: The following schematic example illustrates the principle of muon radiography. The histogram illustrates the relationship between the Number of muons detected by a detector positioned in the environment under investigation and their location.

direction [9]:

$$\Phi_{transmitted} = \int_{p_{\min}}^{\infty} d\Phi(\theta, \varphi, p) dp. \tag{1.9}$$

The methodology employed to generate a two-dimensional image of the inner structure of the target consists of three steps. The first step involves observing the number of counts in the presence of the target ($N_{target}(\theta, \varphi)$). The number of muons in each direction depends on the structure of the target, the direction of observation, and the detector's efficiency. The effect of the latter can be reduced by performing the second step, which is the counting measurement of the free-sky ($N_{free-sky}(\theta, \varphi)$). This measurement involves detecting the counts in the free-sky with the same detector orientation as the target measurement. In fact, since the measurements are conducted with the same detector when we take the ratio between the two counts, the common terms relating to the efficiency of the detector

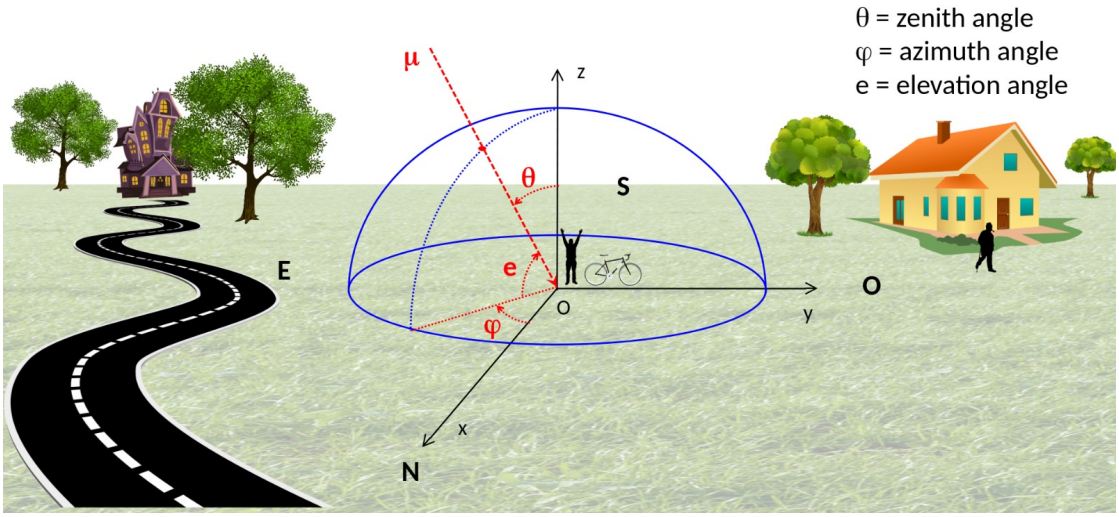


Figure 1.11: A cartoon like example of sistems of reference used in muography. The angle θ is the Zenith angle, the angle φ is the azimuth angle, and e is the elevation angle.

cancel out. Therefore, we obtain the measured transmission:

$$T_{measured}(\theta,\varphi) = \frac{\Phi_{target}}{\Phi_{free-sky}} = \frac{N_{target}(\theta,\varphi)}{N_{free-sky}(\theta,\varphi)} \cdot \frac{\Delta t_{free-sky}}{\Delta t_{target}}, \tag{1.10}$$

where Φ_{target} and $\Phi_{free-sky}$ are the muon flux in the presence of the target and the free sky. Since the detector for the two measurements is the same and therefore the efficiency is the same, the formula can be written as a ratio of counts weighted by the acquisition time. Δt_{target} and $\Delta t_{free-sky}$ are the acquisition times in the target configuration and free-sky configuration, respectively. The times are taken into account as weighting factors. Indeed, a free sky measurement is typically conducted for less time than a measurement in the presence of the object under investigation.
The third step involves comparing the measured transmission with an expected transmission. The expected transmission can be obtained from simulations disposing of an accurate model that reproduces the differential flux of cosmic rays on the Earth's surface and the known geometry of the target [3]. By varying the density of the target ρ in the simulation, a certain number of simulated transmissions $T_{\text{expected}}(\theta,\varphi,\rho)$ are obtained. Then, the relative transmission is obtained by taking the ratio of the measured transmission to the simulated transmission:

$$T_{relative}(\theta,\varphi,\rho) = \frac{T_{measured}(\theta,\varphi)}{T_{expected}(\theta,\varphi,\rho)}. \tag{1.11}$$

Since the transmission is inversely proportional to the density, when the latter is greater than one, it indicates the presence of a region where the density is lower compared to the simulated density. Conversely, when the transmission is less than one, it indicates a denser region.

By further varying the density in the simulation and matching the values of density until the relative transmission is equal to one, a density map of the studied area can be obtained. For more details and examples of typical plots please refer to [10].

1.3 Fast simulation for the muon transmission technique

A fundamental part of muon radiography is the simulation of the environment to be investigated. This simulation is vital when training a predictive model, as in the case of this work. In fact, what we want is to have a fast simulation (so that we can create many images in a short time) that is as faithful as possible to the environment we want to investigate. The simulation must contain: the known structure of the target (geometry), a realistic model of cosmic ray flux on the ground (generator) and a description of the interaction between muons and matter. To reach these goals we purposley designed a C/C++ code integrating all these parts into a single framework.

1.3.1 Simulation geometry

The simulation geometry must accurately replicate the known structure of the target to identify any anomalies relative to this known structure. The required accuracy of the target geometry also depends on the size of the anomalies being searched for. In general, the precision must be higher than the dimension of the anomalies we want to investigate. The target geometry can be generated through:

1. 3D design software (such as AutoCAD [11]);
2. point cloud acquisition (Laser-Scanner, LiDAR, digital photogrammetry (drone), GPS).

For relatively small targets such as blast furnaces, architectural structures, and dams, the first method is generally used since the dimensions of the various components are known by design. For targets such as hills or mountains, a geometry acquired using the techniques described in the second method is employed, as the surfaces can be irregular and discontinuous, making them difficult to realistically

reproduce with 3D design software. In some cases, it is possible to use both methods to obtain a unique and precise geometry of the known exterior and interior of the target. We will address each part in the following paragraphs.

1.3.2 The muon flux generator

In all simulations conducted for this study, the model of the cosmic muon's differential flux at ground level, integrated into a Monte-Carlo generator [12] and [13], is based on data collected using the magnetic spectrometer ADAMO [14] [15], and data from DEIS [16]. ADAMO's data were gathered in Sesto Fiorentino, Florence, Italy, at coordinates Lat. 44°16′ N, Long. 11°26′ W, 71 m above sea level. We opted to utilize these data due to their proximity to our measurement location in central Italy and their inclusion of muons with low momentum, essential for free-sky simulations. ADAMO's magnetic spectrometer studied momentum in the range $(0.1 < p < 130)$ GeV/c and zenith angles $(0° < \theta < 80°)$. Azimuth angle dependence was not examined. For zenith angles $(80° < \theta < 90°)$, DEIS data from Tel Aviv, Israel, at Lat. 32°4′ N, Long. 34°46′ E, 5 m above sea level, were selected, covering a momentum range $(10 < p < 104)$ GeV/c.

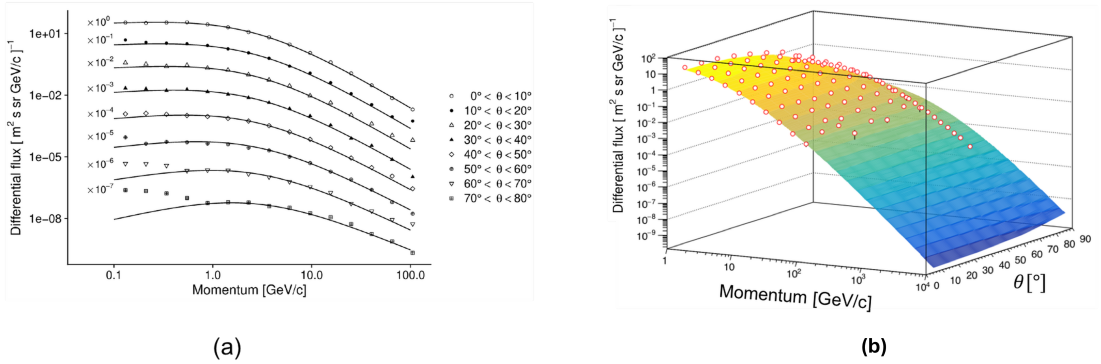
Figure 1.12 (a) shows the differential flux $\Phi(\theta, p)$ derived from ADAMO data, The differential flux of particles on the ground as a function of their momentum is shown, for different ranges of zenith angles. Each of these factors is scaled by gradually decreasing powers of 10. In the last range $70° < \theta < 80°$, distinguishable contributions from low-energy electrons can be seen for momentum values below 1 GeV/c.

Thes low-energy electrons could introduce a background in muographic measurements that is absent in target measurements, potentially causing systematic errors when comparing target and free-sky measurements.

To facilitate simulations, it is necessary to establish a parameterized function describing the differential flux derived from these data as a function of p and θ. A suitable function for parameterizing the data is the function proposed by M.Guan et al. [17], obtained starting from the function proposed by T.K. Gaisser [18] and adjusted for Earth's curvature. The used function is:

$$\frac{d\Phi_\mu}{dE_\mu} = 0.14 \left[\frac{E_\mu}{\text{GeV}} \left(1 + \frac{a\,\text{GeV}}{E_\mu \cos^b \theta^*} \right) \right]^{-2.7} \left[\frac{1}{1 + \frac{1.1 E_\mu \cos\theta^*}{115\,\text{GeV}}} + \frac{0.064}{1 + \frac{1.1 E_\mu \cos\theta^*}{850\,\text{GeV}}} \right] \tag{1.12}$$

where a and b are the free parameters of the fit that assume the following values (after fit): $a = 3.69 \pm 0.021$ and $b = 1.368 \pm 0.003$ and



(a) (b)

Figure 1.12: in (a) the ADAMO data acquired in various ranges of zenithal angles $0° < \theta < 80°$ [14]. In (b) the two-dimensional fit on the data of ADAMO and DEIS [12][16] are shown.

$$\cos\theta^* = \sqrt{\frac{(\cos\theta)^2 + P_1^2 + P_2(\cos\theta)^{P_3} + P_4(\cos\theta)^{P_5}}{1 + P_1^2 + P_2 + P_4}},$$

where P_1, P_2, P_3, P_4 and P_5 are defined in [17]. Figure 1.12 (b) shows the two-dimensional fitting result as a function of (θ, p). The fitting process focused on momentum values exceeding 1 GeV/c to reduce contributions from low-momentum electrons. Once parameterized in this manner, the integrated muon flux Φ_I in the $\hat{\theta}$ direction can be derived for any specified minimum momentum $\hat{p}$:

$$\Phi_I(\hat{\theta}) = \int_{\hat{p}}^{\infty} \Phi(\hat{\theta}, p)\, dp. \tag{1.13}$$

The ADAMO data can be fitted using the relation 1.12, incorporating a term to account for the electron contribution $\Phi(p) = p^{-\gamma_e}$, where γ_e represents the spectral index of electrons.

1.3.3 Range-based simulation

As mentioned above, it is essential to develop high-speed simulation software because training a neural network to detect anomalies requires generating a significant volume of images (approximately 20k), and using simulation tools such as Geant4 [19] would require an impractical amount of time to generate these images within a reasonable time frame for the size of the object under study, thereby making the process infeasible. This method allows for the development of a constant-density simulation of the Temperino Mine in about ten minutes, whereas using software like Geant4 would take several days, if not an entire week, to generate a

simulation. In Chapter 4, we will present a technique we developed to obtain expected transmission maps with variable density. This allowed us to generate more than 20,000 different images in just three days. Consequently, within the Florence muography team a faster simulation software has been devised to optimize simulation time for certain applications. Leveraging the range-momentum conversion tables available in [20], a simulation can be constructed based on muon energy loss in materials. These tables rely on the CSDA hypothesis outlined in section 1.2.1. This type of simulation excludes the multiple scattering effect, making it less realistic compared to Geant4 simulations. To illustrate the process, consider the example of a standard rock in Figure 1.13 (a), where we observe the variation in muon range in a material with opacity define as: $X = \rho \cdot l$, as a function of momentum p. Understanding the geometry of the target and the distance a muon must travel through matter to reach the detector $l(\theta, \varphi)$, allows us to determine the minimum momentum p_{min} required for a muon to traverse the known thickness of the rock. This relationship generally varies with different materials, as illustrated in Figure 1.13 (b). However, for a target composed of homogeneous material with a given density $\overline{\rho}$, the inversion of the $X(p)$ relationship enables conversion of the target's opacity distribution $X(\theta, \varphi)$ into a distribution of minimum momentum ($p_{min}[X(\theta, \varphi)]$) that each muon must possess to reach the detector for every observation direction.

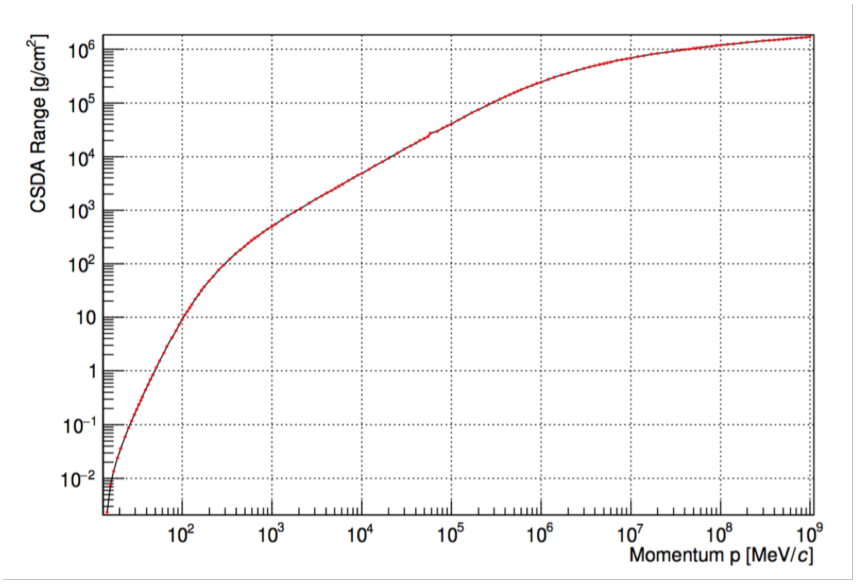
The simulated transmission is computed in a similar manner to actual measurements (Equation 1.10), comparing target and free-sky simulations based on the flux parameterized using ADAMO and DEIS data.

$$\Phi[X(\theta, \varphi)] = \int_{p_{min}}^{\infty} \Phi_{ADAMO}(\theta, \varphi, p)\, dp \tag{1.14}$$

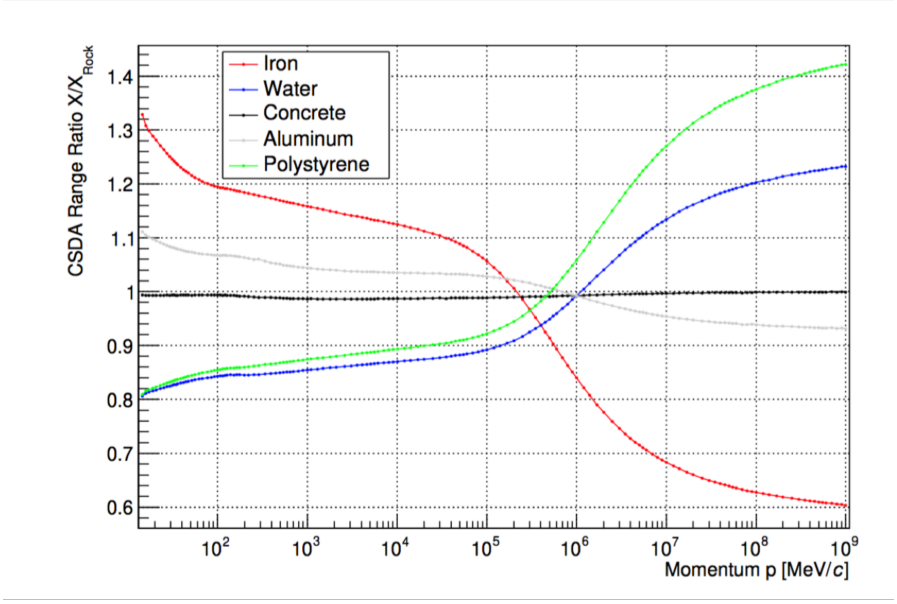
For determining p_{min}, assuming a homogeneous target is adequate in scenarios like hills and mines where the material composition remains consistent. However, in cases such as glaciers, earth dams, blast furnaces, or architectural structures, this approximation becomes unreliable due to significant variations in material density across layers. Here, evaluating the minimum momentum p_{min} requires accounting for distinct material layers, achieved by converting each layer's opacity to its equivalent water opacity X_w. For instance, if there are i layers with opacity X_i, each is transformed into X_{w_i}, and the total minimum momentum would be $p_{min}(X_{w_{\text{tot}}}) = p_{min}\left(\sum_i X_{w_i}\right)$.

In free-sky simulations, the minimum momentum chosen for calculating the integral flux (equation 1.14) is 133 MeV/c, determined in [21].

It would be more correct to add in equation 1.14 the survival probability of the muon $P_S(X, p')$ in the stretch of material at opacity X with initial momentum p' and therefore consider:



(a)



(b)

Figure 1.13: in (a) the average CSDA range for a muon in standard rock as a function of particle momentum, in (b) the average CSDA range for different materials.

$$\Phi[X(\theta,\varphi)] = \int_{p_{min}}^{\infty} P_S(X,p')\Phi_{ADAMO}(\theta,\varphi,p')\,dp'. \tag{1.15}$$

As described in [22] this correction becomes important for large opacities where fluctuations in radiative losses become dominant.

1.3.4 Transmission-opacity conversion

With the range-based simulation it is therefore possible to evaluate the simulated transmission for a given value of opacity X of the target and of elevation angle. From this type of simulation, using equation 1.14, it is possible to create a conversion table $(\theta, X) \rightarrow (\theta, T_{expected})$ as shown in figure 1.14. The conversion table has been evaluated for opacity values corresponding to the range 0.3-3000 m.w.e and for elevation angles from 0° to 90°. With this conversion table it is possible to convert transmission values to opacity values and vice-versa.

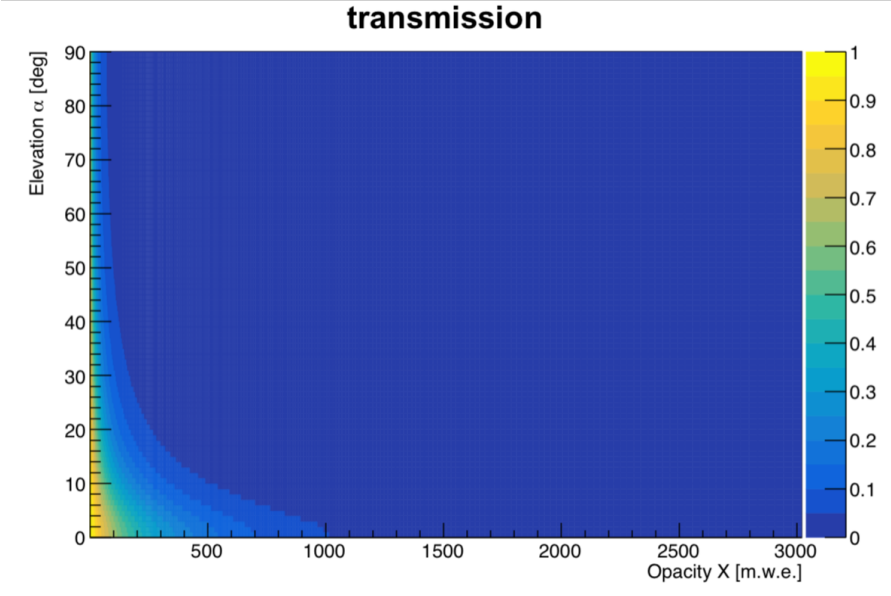


Figure 1.14: Transmission as a function of elevation angle and opacity, evaluated in the range 0.3 - 3000 m.w.e., where 1 m.w.e. correspond to 100 g cm^{-2}.

If a measurement has yielded a two-dimensional angular transmission distribution, this table allows for its conversion into an opacity distribution. With the knowledge of the geometry of the target, a two-dimensional distribution of average density can be generated. Conversely, if the geometry of the observation site and the average density values that constitute it are known, it is possible to switch from an opacity distribution to a simulated transmission distribution.

This describes the functionality of the Monte-Carlo simulation software developed by the Florence muography group. The software permits to undertake of precise preliminary studies with a reduced simulation time (less than an hour), and it provides an expected transmission map for comparison with the actual measurement. From the perspective of neural network training, it is of the utmost importance to have software that is both precise and fast when it comes to general training images. Indeed, the software will subsequently be modified (it will be described in detail in chapter four) to enable the generation of artificially measured transmission maps. Subsequently, these will be employed to train a neural network capable of identifying anomalies within a muon radiography. The modified version of the simulation software allowed us to create the necessary realistic muografic images in less than a week.

Chapter 2

Detectors for muon radiography

Muography represents a cutting-edge technique in non-invasive imaging, leveraging the distinctive property of cosmic muons to examine the interior of natural and artificial structures. As demonstrated in the preceding chapter, muons are especially well-suited for imaging applications in environments of considerable size or density. The efficacy of muography is contingent upon the detectors employed to reveal and quantify muons. These detectors are engineered to detect particles traversing the object, thereby reconstructing their direction of origin. The data obtained is then processed to construct a muographic image of the object. There are diverse types of detectors utilized in muography, each exhibiting distinctive characteristics that render it more suitable for specific applications. The main types of detectors used in muography are:

- **Scintillation Detectors:** Scintillation detectors employ materials that emit light when traversed by charged particles, such as muons. Such scintillating materials are frequently coupled with photomultipliers or photodiodes, which convert the emitted light into electrical signals. These are relatively simple to construct and can be produced in various shapes and sizes, making them suitable for modular configurations that can cover large areas.

- **Gas Detectors:** These devices exploit the ionization of a gas by charged particles (such as muons) to generate an electric charge, which is then collected by electrodes. Several types of gas detectors exist, each with distinct characteristics suited for various applications. Drift chambers, for instance, utilize the drift of ionization electrons in a uniform electric field to accurately determine particle positions. By assembling multiple layers of these chambers, it is possible to reconstruct muon trajectories, making them particularly effective in muon scattering tomography [6, 7, 23]. Another widely used technology is the multi-wire proportional chamber (MWPC), which consists of an array of thin wires acting as anodes. These chambers detect particles

through the avalanche multiplication of electrons near the wires, enabling precise track reconstruction [24]. For applications that require excellent time resolution, resistive plate chambers (RPCs) are a preferred choice. These detectors are composed of two resistive electrode plates separated by a thin gas layer, allowing them to achieve precise timing measurements, making them ideal for experiments that demand fast response times [25]. Gas electron multipliers (GEMs) represent another category of gas-based detectors. They employ thin foils perforated with micro-holes to amplify the charge of electrons produced by ionization. This design provides high spatial resolution and the ability to function at high counting rates, making GEMs particularly useful in environments where tracking a large number of events is necessary [26].

- **Silicon Detectors:** The detectors utilise silicon strips that generate an electrical signal when crossed by muons. These detectors are valued for their high spatial resolution, which makes them suitable for use in experiments requiring precise measurements on the millimetre or sub-millimetre scale. On the other hand, the high cost of these detectors limits their use in this field. An illustrative example of its use can be found in [27].

- **Time Projection Chambers (TPC):** TPCs are capable of tracking particle trajectories over time, enabling reconstruction of the muon transition in three dimensions with high precision. However, the calibration of these detectors is a complex process, but the resulting resolution in terms of position and time is exceptional. An illustrative example of its use can be found in [28].

- **Emulsion Chambers:** Emulsion chambers are devices that employ layers of photographic emulsion to detect the tracks left by muons. Upon traversing the emulsion, a muon ionises the molecules within its path, thereby creating minute traces that can be developed and analyzed in a manner analogous to that of a photograph. These detectors offer an extremely high level of spatial resolution, rendering them particularly useful for applications that require the precise reconstruction of particle trajectories.An illustrative example of its use can be found in [29].

The different detectors are selected according to the specific requirements of the muography application:

- **Volcanoes and Mines:** In order to study the internal structures of volcanoes and predict possible eruptions, or to detect the presence of cavities within mines, scintillation detectors are often the preferred choice due to

their robustness and ability to operate in harsh environments. A further advantage of scintillation detectors is that they are relatively inexpensive to construct and maintain, particularly in comparison to the other aforementioned types of detectors. Some example of the deployment of these detectors for muographic imaging can be found in reference [30, 10].

- **Pyramids and Archaeological Structures:** The utilisation of drift chamber detectors and emulsion chambers is based upon their capacity to facilitate the generation of comprehensive visual representations of the concealed chambers within pyramids, whilst simultaneously ensuring the preservation of the structures themselves. Nevertheless, it is worth noting that scintillation detectors have also been employed in archaeological studies, as evidenced by the reference [31].

- **Nuclear Safety:** Silicon strip detectors are employed in nuclear reactors and storage sites for the monitoring of fissile material distributions due to their high resolution. As an alternative, drift chambers may be a viable solution, as illustrated in reference [23].

- **Blast furnaces:** In the context of blast furnaces, scintillation detectors are employed due to the inherently challenging and hostile working environments. This necessitates the utilisation of a detector that is not only highly robust and autonomous in operation, but also capable of functioning without the need for continuous monitoring of the gas. An illustrative example of this utilisation is provided in [32].

Each type of detector possesses distinctive advantages. For instance, scintillation detectors are relatively inexpensive and straightforward to implement, yet may exhibit a diminished spatial resolution in comparison to silicon strip detectors. Drift chambers offer excellent spatial accuracy, but their use is contingent upon the availability of a controlled environment and the capacity to manage complex gas handling. Emulsion chambers, conversely, offer unparalleled spatial resolution and comprehensive recording of muon tracks; however, they necessitate a sophisticated and laborious photographic development process. In general, the selection of the most appropriate detector for a given measurement is based on a number of factors, including the required resolution, the cost of the detector, the complexity of production, and the difficulty of data analysis. In recent times, there have been notable advancements in the development of novel scintillator materials and the optimisation of data processing techniques, which have further enhanced the imaging capabilities of muography. Moreover, the integration of sophisticated technologies such as artificial intelligence and machine learning is paving the way

for new avenues in the analysis and interpretation of muographic data.

One of currently used by the muon radiography group in Florence is MIMA, which will be explained and introduced in the next section, as related to one of the machine learning application of this thesis. As part of the PhD program, I have also been heavily involved in the constuction, testing and deployment of two new trackers for the BLEMAB project. These are new and upscaled versions of MIMA that can be used as an alternative in some applications, as it will be explained in the last part of this chapter.

2.1 MIMA Detector

The Muon Imaging for Mining and Archaeology (MIMA [33]) hodoscope aims to obtain images, through muon absorption radiography, of very dense structures or voids hidden underground or otherwise surrounded by large volumes of matter. As it is a natural multidisciplinary endeavour, the ultimate objective is to validate this methodology for applications in a variety of fields, including archaeology, geology, mining, civil engineering and civil protection.

2.1.1 Concept and design of the MIMA muon telescope

The telescope's design is based on the extensive experience of the muon radiography group of Florence in volcanological studies [34, 30] and adapted for use in confined environments such as mines and archaeological sites.

The MIMA detector was designed to meet a number of key requirements, including low energy consumption, robustness, compactness, lightweight construction, and the ability to operate across a wide temperature range. The detector is composed of three X-Y tracking modules, each of which is formed by the stacking of two orthogonal planes of plastic scintillator bars. This configuration enables the precise tracking of muon trajectories with an angular resolution of 0.001 steradians [33].

The MIMA telescope is constructed according to a modular design, with dimensions of 50 cm x 50 cm x 50 cm and a total weight of approximately 60 kg, inclusive of both mechanical and electronic components. The three tracking modules are encased within a cubic aluminium frame, which is mounted on an altazimuthal platform. This configuration permits continuous adjustment of the azimuth angle and stepwise adjustment of the zenith angle, thereby ensuring optimal alignment with the target volume under investigation.

The detector's tracking planes are composed of high-quality plastic scintillator bars, each of which is read out by silicon photomultipliers (SiPMs) that are

directly coupled to the scintillators. This direct coupling eliminates the need for intermediate wavelength shifting fibres, thereby simplifying the system and reducing costs. The tracking planes employ triangular cross-section scintillator bars, which optimise performance and robustness.

MIMA muon hodoscope	
Developed/build: 2016/2017	Weight: $\sim$ 60 kg
Number of XY tracker modules: 3	Spatial resolution: $\sim$ 1.5 mm
Technology: bars of plastic scintillator	Angular resolution: 6.5 mrad (0.4°)
Total size: $\sim (50 \times 50 \times 50)$ cm^3	Altazimuth orientation system: yes
Angular acceptance: $\pm 65°$	Power consumption:30 W

Table 2.1: The main characteristics of the MIMA muon hodoscope.

In order for a muon to pass through the detector, it must have a minimum momentum of 133 MeV/c, which is determined by the total opacity of the detector. The detector exhibits an acceptance of approximately $\pm 65°$ with an angular resolution of 0.4° (6.5 mrad). For further details, refer to reference [33].

Table 2.1 presents a summary of MIMA's features and a visual representation of its structural configuration. The following subsections will provide a detailed account of the detector's composition.

The tracking plane

The tracking plane of the MIMA muon telescope represents a critical component, designed with the specific purpose of accurately reconstructing the trajectory of muons. The tracking plane is composed of plastic scintillator bars arranged in two orthogonal layers (X and Y), which provide precise two-dimensional position information (1.5 mm for spatial resolution and 6.5 mrad for the angular resolution).

Figure 2.1 illustrates the scintillating bars utilized in the construction of MIMA. As can be observed, a number of bars are wrapped within the mylar sheets, with the reflective surface oriented inwards. It is of great importance that the light does not escape from the scintillation bar and scatter, as this would compromise the efficacy of the apparatus. Due to the long and thin bar configuration, this is of particular importance, as the number of collected photons may be limited.

The detection plane has dimensions of (40 x 40) cm^2, thereby ensuring a compact yet effective detection area. The triangular cross-section of the scintillator bars optimises both spatial resolution and light collection efficiency.

The utilisation of SiPMs presents a number of advantages, including high sensitivity, a compact size and low power consumption, which are crucial for field applications. Each scintillator bar is aligned and calibrated to ensure high precision in muon tracking. This alignment directly influences the detector's spatial

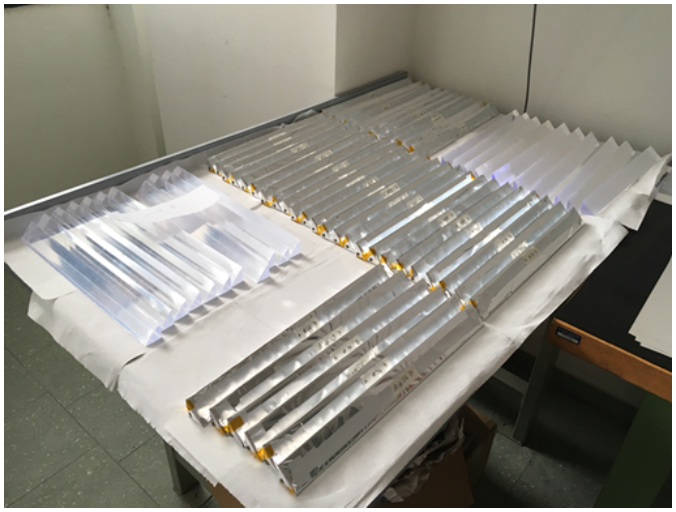


Figure 2.1: Scintillator bars with triangular section used for the MIMA tracking planes. These are shown wrapped inside reflective mylar, while others are in their final constructed state [33].

resolution, which is approximately 1.15 cm per detection plane, derived from the standard deviation of a uniform distribution over the bar width. The configuration of the bars in orthogonal layers, combined with a separation of 80 cm between the two detection planes, allows for the reconstruction of muon tracks with an angular resolution of approximately 14.3 milliradians. This precise calibration and detector geometry are essential for obtaining high-fidelity muographic images, enabling the detailed imaging of subsurface structures and cavities with improved contrast and accuracy.

Mechanical structure

The mechanical structure of the MIMA muon telescope has been designed with the objective of ensuring stability, durability and ease of deployment in a variety of environments. The detector is encased within a robust cubic aluminium frame, with dimensions of 50 cm on each side. The compact and lightweight structure, with an approximate weight of 60 kg including all mechanical and electronic components, ensures the ease of transport and setup.

The aluminium frame provides a stable platform for the three X-Y tracking modules, which are securely mounted within the cube. The frame's design allows for straightforward assembly and disassembly, thereby facilitating rapid deployment and maintenance in the field. Furthermore, the frame's material and construction ensure resistance to environmental conditions, thereby making the detector suitable for use in diverse and challenging settings.

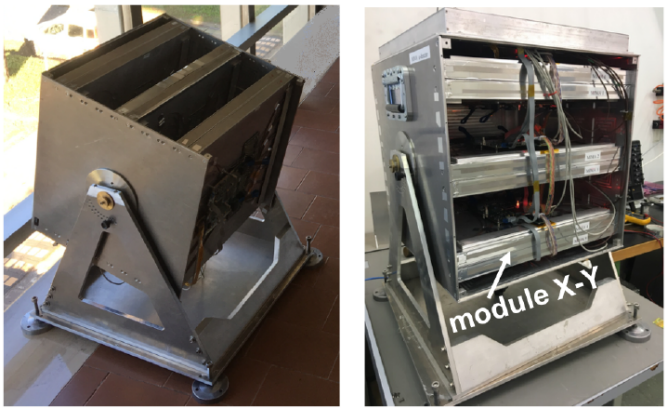


Figure 2.2: The MIMA structure in which the three tracking modules XY and the altazimuthal orientation system are visible.

The MIMA telescope is furnished with an altazimuthal platform, which enables the detector to be oriented with good precision. The platform allows continuous adjustment of the azimuth angle and five-degree step adjustment of the zenith angle.

Figure 2.2 illustrates the mechanical structure of MIMA. One can see the detector that the detector is equipped with an altazimuthal system, which enables it to be tilted and rotated. Furthermore, the interior of the apparatus, in which the six tracer planes are situated, is also depicted.

In conclusion, the mechanical structure of the MIMA muon telescope is characterised by robustness, portability and precise adjustability, which collectively make it an ideal tool for detailed muon imaging in a variety of field applications.

2.1.2 The photosensors

The photosensor system of the MIMA muon telescope has been designed with the objective of achieving high efficiency and reliability. The configuration employs SiPMs directly coupled to plastic scintillator bars, thereby eliminating the need for intermediate wavelength-shifter fibres. The direct coupling simplifies the design, reduces the potential for failure, and enhances the overall robustness of the detector. The direct coupling of SiPMs to the scintillators enhances the signal-to-noise ratio, which is pivotal for precise muon detection and tracking.

SiPMs are selected for their high sensitivity to the scintillation light generated by the plastic scintillators. Their compact size, minimal power consumption, and low light detection performance make them ideal for this application.

Table 2.2 illustrates the model characteristics associated with the SiPM employed in the MIMA assembly process. The reference values are presented at a

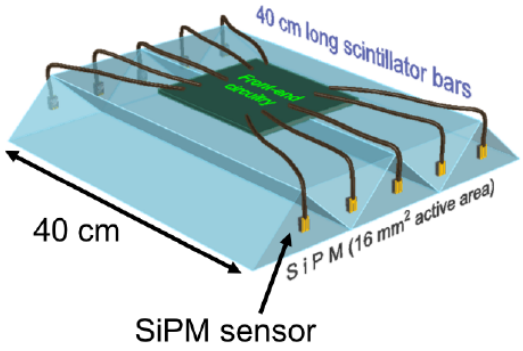


Figure 2.3: The arrangement of the triangular bars to form a tracker plane and the coupling with the SiPMs sensors are schematized.

temperature of 20°.

Detection area (4×4) mm^2	Peak Sensitivity Wavelength 420 nm
Dimension cell (40×40) μm^2	Breakdown voltage 26 V
Number of cells 9340	Gain 3.6×10^6
Quenching resistance 800 kΩ	Breakdown Voltage Temp. Coeff. 26 mV/°C
Fill factor 60%	Dark Count Rate ($\Delta V_{OV} = 2$ V) ≤ 50 kHz/mm^2
PDE 43%	Refractive index of epoxy resin 1.51

Table 2.2: The main characteristics of the ASD-NUV4S-P SiPMs used. The values refer to a temperature of 20°C.

Each scintillator bar is equipped with SiPM detectors at both ends, thereby ensuring efficient light collection and signal detection. Figure 2.3 illustrates the configuration of two SiPMs coupled to each bar. The redundancy permits a more precise signal, as the charge accumulated in each is subsequently summed on the electronic board.

The high granularity of the SiPMs enables the precise localisation of muon interactions within the scintillator bars, thereby contributing to the overall spatial resolution of the detector.

2.1.3 Front end electronics

Figure 2.4 shows the electronics board mounted on the MIMA module, each individual plane is coupled to a sum board and a so-called slave board. The front-end electronics of the MIMA muon telescope have been designed with the objective of achieving optimal performance and efficiency. At the core of the system is the EASIROC (Extended Analogue SiPM ReadOut Chip), a bespoke

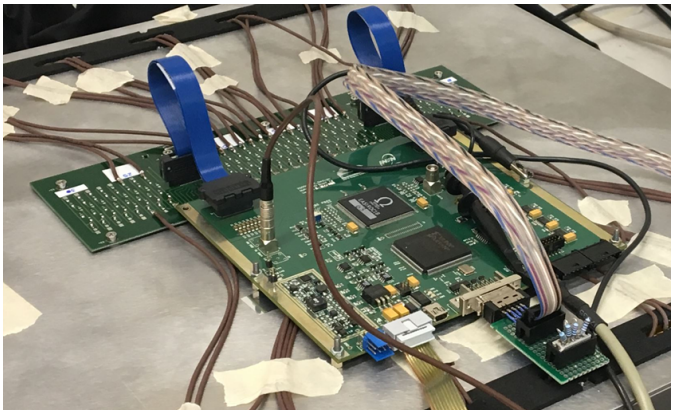


Figure 2.4: Front End electronics mounted on a MIMA's module. On top of the module's frame the summing card (back) and the slave card (front) housing the EASIROC [35] chip and the XILINX Spartan II FPGA [36] are visible.

chip designed for the readout of SiPMs. The chip integrates a number of functions that are necessary for the processing of signals generated by the SiPMs, including amplification, shaping and digitisation. The electronics have been designed to minimise power consumption while ensuring high performance. The entire front-end electronics system has a power consumption of approximately 30 W, making it suitable for use in remote and challenging environments where power availability may be limited.

The signal from each SiPM is initially combined by a custom summing circuit, which aggregates the outputs from multiple SiPMs into a single signal. Subsequently, the summed signal is fed into the EASIROC chip, which processes it and prepares it for further analysis. The EASIROC chip provides both analogue and digital outputs, thus facilitating detailed signal analysis and precise muon tracking. A significant attribute of the front-end electronics is their capacity to detect weak signals generated by muon interactions in scintillator bars, which is contingent upon their low-noise performance. Furthermore, the electronics have been designed to be robust and reliable, capable of operating over a wide range of temperatures and environmental conditions.

2.1.4 Track Reconstruction

The reconstruction of the track is conducted independently for the two views (X and Y), and is comprised of numerous steps that originate from the initial acquisition of raw data. For each event initiated by the trigger logic, the Analog to Digital Converter (ADC) value of each bar is acquired, irrespective of whether a particle signal has been generated or whether the particle has not passed through

it. This methodology allows for the evaluation of the pedestals (i.e., the values corresponding to the average voltage levels at the input of the front-end chip channels) of each bar, thus enabling the definition of a channel's signal as follows:

$$S_{ij} = ADC_{ij} - ped_{ij} \tag{2.1}$$

the index i denotes the variable that is being evaluated on the plane, while the index j denotes the variable that is being evaluated on the channels. The value ped_{ij} is determined by calculating the centroid of a Gaussian fit to the pedestal peak of each bar. Figure 2.5 illustrates the signal distribution of a bar for 5000 events. The standard deviation, σ_{ij}, of the fit is, in contrast, linked to the noise of the bar and is employed for certain cuts in cluster determination, as will be explained in the following. As illustrated in Figure 2.5, both the pedestal peak and particle signals are evident. The signal contribution is observed at ADC values exceeding approximately 40. Subsequently, the process of equalising the gains of all the bars is initiated. In order to facilitate comparison between planes, the first channel in each plane is taken as a reference point, and a process of gain normalisation is carried out for all channels in a given plane.

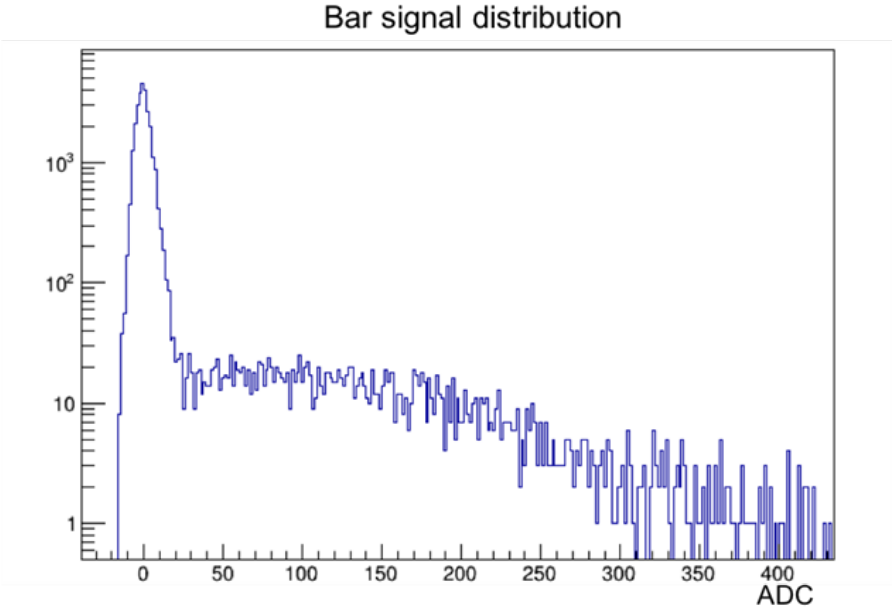


Figure 2.5: the distribution of the variable S_{ij} (equation 2.1) for a scintillator bar of a tracking plane. In addition to the signal distribution (ADC greater 40), the pedestal peak is also observed on which the σ_{ij} was estimated.

Once the gains have been equalised, the cluster search procedure is initiated for each plane and for each event. A cluster is defined as a group of adjacent channels

on which the particle signal has developed. Given the triangular bar geometry, it is possible to identify clusters comprising one (vertical tracks), two, four, or six bars. The following cuts are defined for the cluster search:

$$\begin{cases} S_{ij} > 4\sigma_{ij} \text{ for cluster seed} \\ S_{ijadj} > 3\sigma_{ij} \text{ for bars next to the seed} \\ S_{ij_{next}} > 2\sigma_{ij} \text{ for the bars next to the bars adjacent to the seed.} \end{cases} \tag{2.2}$$

Moreover, a cut on the total signal of a cluster is also defined with the objective of minimising any noise clusters. It is possible that multiple clusters may be identified in a single plane, due to either random coincidences or the presence of noise. For this reason, a best fit process is implemented, whereby the optimal alignment is sought for all possible combinations of clusters. Once the cluster has been identified, the associated coordinate x_{ij} is estimated. In the event that the cluster comprises a single bar, the coordinate of the vertex of the bar is taken. In the case of multiple bars, a centre of gravity algorithm is employed.

$$x_i = \frac{\sum_j^n x_{ij} S_{ij}}{\sum_j^n S_{ij}} \tag{2.3}$$

where j is the index that runs over the number n of bars involved in the cluster and x_{ij} is the coordinate of the vertex of the bars. Finally the process of track reconstruction begins, that can be divided into the following cases:

1. all six planes have at least one cluster;
2. four adjacent planes have one cluster;
3. five planes have one cluster.

In the first scenario, the slope of the straight line that best approximates the cluster coordinates across all planes is determined using a linear fitting method. When multiple clusters are present on a plane, the fit that yields the best result is chosen. In contrast, in the second scenario, due to the limited number of available points, no fitting procedure is carried out. Instead, the inclination of the trajectory on the XZ and YZ planes is evaluated according to the following formulae: $(x_{up} - x_{down})/(z_{up} - z_{down})$ and $(y_{up} - y_{down})/(z_{up} - z_{down})$, where z_{up} and z_{down} are the z coordinates at which the tracker planes involved in the track reconstruction process are located inside the detector. The third case is useful for evaluating the inefficiency of the planes when five planes were involved in the reconstruction of the track. Once the track has been reconstructed, all information (i.e. impact points coordinates, track inclination) are available.

2.2 BLEMAB project

Part of the PhD work was dedicated to building two detectors for the European BLEMAB project. The European project BLEMAB (BLast furnace stack density estimation through online Muon ABsorption measurements) is a further development of the previous European project Mu-Blast to improve muon radiography techniques for imaging the interior of blast furnaces. A blast furnace is a large industrial structure used for smelting iron ore into molten iron, which serves as the primary raw material for steel production. These furnaces are typically 30 to 40 meters in height and have an internal diameter ranging from 8 to 15 meters, depending on their production capacity. They operate continuously, processing thousands of tons of raw materials per day.

Taking the figure 2.6 as a schematic example, a blast furnace is charged at the top with alternating layers of coke and burden, which includes pellets and lump iron ore. In the lower part of the furnace, hot air is injected from hot stoves through tuyeres. This hot air reacts with the coke in front of each tuyere to produce carbon monoxide. The carbon monoxide rises through the furnace, reducing the iron oxides in the iron ore. Molten metal is then collected at the bottom of the furnace. The cohesive zone is the area where charged materials, such as iron ore and coke, begin to melt and form a pasty or semi-liquid material. This zone is typically above the hearth, where the temperature is high enough (this varies between 1,200 and 1,400 degrees Celsius) to cause the ores to partially melt, but not high enough to make them completely liquid. The cohesion zone is critical to the reduction process as it determines the flow of gases and molten material within the furnace.

Knowing the evolution of this zone allows a significant improvement in the blast furnace production process, bringing several advantages, including:

- **Optimisation of product quality:** By monitoring the cohesion zone, ideal conditions for iron ore reduction are maintained, improving the quality of the iron produced.
- **Energy efficiency:** Precise control of the cohesive zone allows fuel to be used more efficiently, reducing overall energy consumption.
- **Extended furnace life:** Proper management of the cohesion zone can prevent damage to the furnace structure, extending its life and reducing the need for costly maintenance.
- **Increased productivity:** Understanding and controlling the cohesive zone can help maintain a continuous and stable flow of material through the furnace, increasing overall productivity.

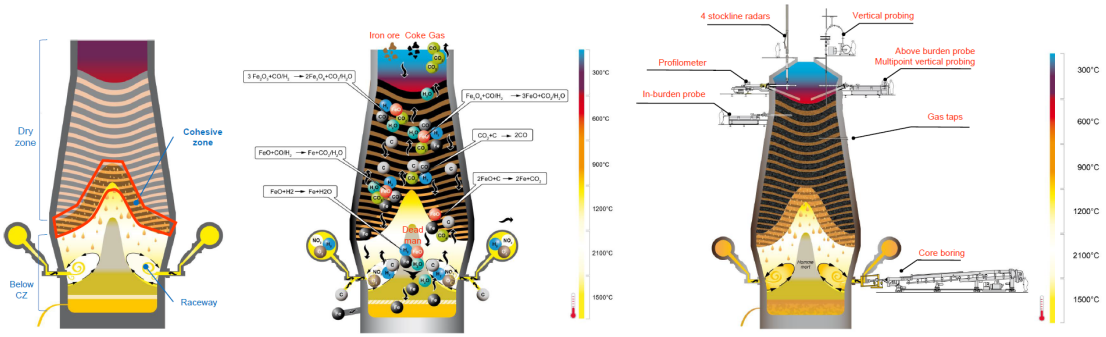


Figure 2.6: Simplified schematic of a blast furnace, divided into distinct operational zones: dry zone, cohesive zone, and below the cohesive zone (CZ). The central diagram illustrates the predominant chemical reactions occurring inside the furnace, including the reduction of iron oxides (FeO, Fe_3O_4) by CO and H_2, and the formation of CO_2 and H_2O. On the right, measurement techniques such as temperature probes, stockline radars, profilometers, and core boring are highlighted. The color scale indicates the vertical thermal profile, with temperatures rising to over 1500°C at the base.

- **Reduced emissions:** By optimising the melting and reduction process, emissions of greenhouse gases and other pollutants can be reduced, contributing to a greener process.

Obtaining information on the reduction process during operation is challenging due to the extreme conditions inside the furnace, such as high temperatures and pressures. Typically, data on material distribution or reaction parameters can only be obtained during stoppages by core drilling or by dismantling the furnace at the end of its life [37].

The BLEMAB project explores the use of muography to address these challenges, employing detectors based on scintillating technology to perform transmission radiography. The main goal is to improve real-time control of blast furnace operations through accurate detection of the cohesive zone [38]. This technique proves to be an extremely advantageous choice for tackling this issue. As previously explained, it allows continuous monitoring of the target over time in a non-invasive and passive manner.

It was therefore necessary to develop a detector that is accurate, robust and easily transportable in an environment such as a steel factory. Based on the long experience of the Florence group in detector construction, such as MIMA [33] and MURAVES [30], two detectors based on the scintillating rod technology. The following sections describe the concept of the detector, its construction and its

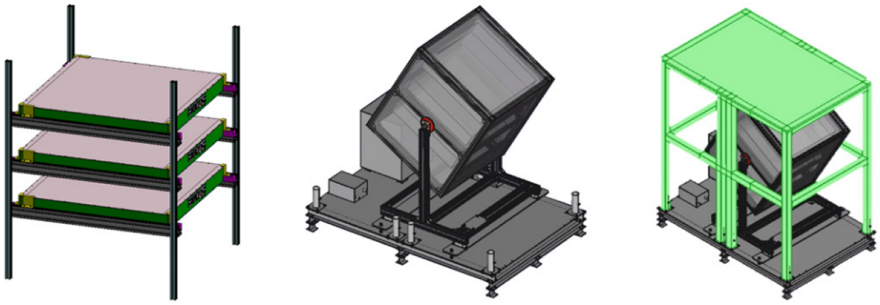


Figure 2.7: Left: single BLEMAB tracker, composed of three XY tracking modules. Centre: the BLEMAB tracker installed on its mounting and placed on the base plate. Right: the skeleton of the protection frame with a steel cover plate at its top (in green) [32].

installation in the vicinity of a furnace. This was a unique opportunity within this thesis, as it allowed to follow all the phases of an experiment, from its construction to its installation in the field, and to fully understand the difficulties that can be found in these phases of an experiment.

2.2.1 Concept of detector

Muon transmission radiography requires the installation of at least one muon tracking system downstream of the target area. Within the BLEMAB project, two separate muon trackers have been developed. This allows flexibility in deployment: both trackers can be installed in the same blast furnace to provide a stereoscopic perspective, or one can be placed in front of a blast furnace while the other is positioned in an open space to provide an free-sky reference measurement. Alternatively, the detectors can be placed in two different blast furnaces for a comparative analysis of their internal structures.

Each tracker consists of three independent tracking modules, each measuring 80 cm x 80 cm, which record the muon impact position along two orthogonal axes (XY). These modules use the same technology as in the MIMA project. An XY tracking module is built up of orthogonal tracking planes, each containing 63 scintillator rods with triangular cross sections and a length of 80 cm. Each scintillator bar is equipped with two silicon photomultipliers, each with an area of $4x4mm^2$.

The tracking module includes an acquisition system consisting of four custom DAQ slave boards, each containing an EASIROC1B 32-channel front-end chip and an ASIC for controlling the front-end circuitry and data transmission to a central DAQ master board. This master board handles the trigger logic and collects data

from up to 16 slave boards. A Raspberry Pi computer connects to the master DAQ board to configure the entire electronic chain, receive data packets and store them on physical media. Network access is facilitated by a SIM-operated commercial modem, allowing online control and data synchronisation with a remote server.

Each muon tracking detector (consisting of three XY modules) is housed in a lightweight aluminium box and mounted on a custom support that allows remote adjustment of the detector's orientation. The entire measurement system is enclosed in a large metal frame (as shown on the right side of Figure 2.7), which is shielded by metal panels to protect against shocks, liquid splashes and corrosive vapours. Figure 2.7 also shows the design of both the detector and the housing frame. Due to the high temperatures in the blast furnace environment, a cooling system was installed within the protective enclosure to maintain stable operating temperatures for the optical sensors.

2.2.2 Construction of the detector

This section will provide a detailed account of the construction of the two detectors, designated as BLEMAB01 and BLEMAB02. The assembly process of the muon tracking detector involves several detailed steps to ensure accuracy and functionality. Here is a complete overview of how the detector is constructed. The reference system is that depicted in Figure 2.8. The module is mounted in accordance with the specifications set forth in this figure.

The assembly of a detector module is divided into two parts: the assembly of the "X" and "Y" planes. The "X" plane assembly is described below, while the "Y" plane is constructed in the same manner but positioned perpendicularly to the "X" plane. To assemble the module, the right partition is first removed, and a Mylar sheet is placed at the base. Then, the first 32 triangular bars are glued into position, forming what is called the lower comb. The remaining 31 bars are inserted, interlocking with those already glued. Once all 63 bars have been positioned, the assembly of the "Y" plane is carried out, placed perpendicularly to the "X" plane.

Once both planes are complete, the plate on which the electronics will be installed is attached. At this point SiPMs, which are glued to the ends of each scintillator rod, are inserted. SiPMs are essential components in the muon tracking detectors used in the BLEMAB project. The SiPMs are responsible for detecting these photons. Each scintillator rod in the tracking module is read by two SiPMs, which are glued to the bars, providing a higher probability of scintillation light detection and redundancy. The signal from each of the two SiPMs is then summed. The SiPMs amplify the weak light signals produced by the scintillator. This is critical because the initial light signals can be quite weak. SiPMs consist of a series of microcells, each of which acts as a Geiger-mode avalanche photodiode

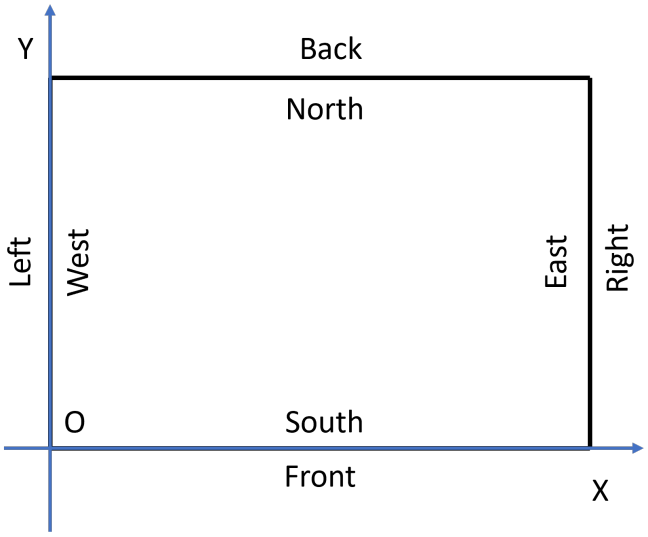


Figure 2.8: The reference system of a detector plane is illustrated below. The X-plane, the inferior plane, extends from left to right in accordance with the origin, while the Y-plane, the superior plane, extends from front to back in a direction perpendicular to the X-plane.

(APD). When a photon strikes a microcell, it triggers an avalanche of charge carriers, producing a measurable electrical signal. The SiPMs that we had use, have an effective area of (4x4) mm^2. For each scintillator, a pair of similar gain SiPMs was chosen. The SiPM were also diveded into groups of 63 with similar gain voltage in order to have one common operating plane voltage.

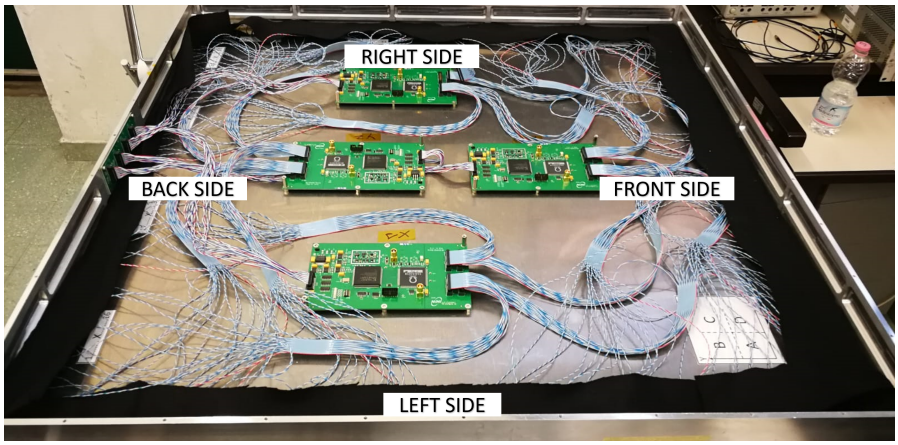


Figure 2.9: Finished module with electronics boards mounted. The connections between these and the SiPMs are visibile.

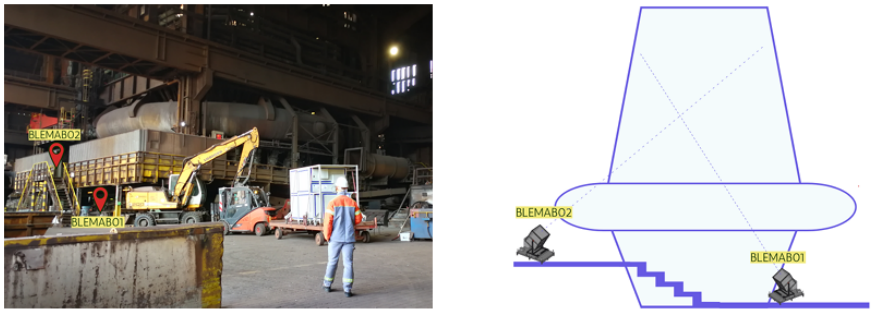


Figure 2.10: Left: installation points to the blast furnace underneath the walkway, corresponding to BLEMAB01 and at the tuyere level corresponding to BLEMAB02. Right: schematic representation of the position of the two detectors pointing at 47° in elevation in the direction of the blast furnace.

Once this is done, the module is completed by inserting the electronic boards and connecting the SiPMs to them, as shown in Figure 2.9. There are 4 electronic boards in total.
Each detector consists of 3 modules, each consisting of two planes with 63 bars each and 126 SiPMs. A total of two detectors were built. The total construction time was approximately one year. Once assembled and configured, they were shipped to the city of Bremen, Germany for installation near a blast furnace in the ArcelorMittal steel plant [39].

2.2.3 Installation

The installation of BLEMAB detectors, which was completed in August 2023, constituted a pivotal phase of the project. The BLEMAB01 and BLEMAB02 detectors were installed around blast furnace No. 2 at the ArcelorMittal site in Bremen, Germany, with the objective of studying the internal structure and dynamics of the furnace. Given the furnace's height of approximately 40 meters and its radius of approximately 16 meters at the tuyere level, precise positioning of the detectors is of paramount importance for accurate measurements.

Figure 2.10 a) illustrates the location of the two detectors, while Figure 2.10 b) presents a schematic representation of their installation relative to the blast furnace. During the installation process, BLEMAB01 was situated at the base of the furnace, oriented at an elevation angle of 47 towards the center of the furnace. This positioning was selected to obtain a comprehensive view of the cohesive zone of the furnace, where the material begins to soften and melt. BLEMAB02, on the other hand, had to be installed at a higher level, just below the furnace, oriented also with an elevation angle of 47, due to spatial limitations and operational

considerations. Despite the difference in elevation, both detectors were aligned to point approximately at the same level, ensuring consistent data collection.

Furthermore, the installation process entailed the utilisation of adjustable feet to guarantee that the detectors were aligned with the ground's surface, a procedure that had to be conducted on-site. The orientation and positioning of the detectors were designed to be controlled remotely, thus allowing adjustments to be made without the need for manual access to the detectors.

The data collection process from these detectors concluded in July 2024. This includes the collection of baseline data during a planned furnace shutdown, which will serve as a reference point for comparison with operational data and facilitate a more accurate understanding of the blast furnace's internal processes. The results are currently being processed.

This constitutes a comprehensive account of the detectors that may be employed in muon radiography, with a specific focus on those based on scintillator bars. The Florence group has extensive experience with this technology, particularly with the development of the MIMA detector, which is currently the main detector used in the group's measurement campaigns. Moreover, the group was instrumental in the development of the European BLEMAB project, thanks to their knowledge in the construction of these types of detectors. In conclusion, it can be stated that detectors based on scintillator bars are particularly well-suited for the investigation of large structures (such as volcanoes, mines, and hills) or highly dense materials in challenging environments, such as blast furnaces. This constituted a significant component of the thesis, as it permitted the critical issues associated with the construction and subsequent installation of a detector for a muographic measurement to be elucidated.

Chapter 3

Monitoring water infiltration inside tailing dams

Tailing dams are engineering structures in the mining industry, designed to contain tailings, the byproducts of mineral mining and processing. These tailings often consist of finely ground materials and process waters and may contain hazardous chemicals such as heavy metals, reagents, and acids. Tailing dams play a critical role in preventing the release of hazardous materials into the environment and their failure can have severe environmental, social, and economic consequences for local communities and ecosystems. The primary risks associated with the failure of a tailing dam include flooding of populated areas, contamination of surface and groundwater, and destruction of natural habitats.

In recent decades, numerous incidents of tailing dam collapse across disparate global locations have underscored the necessity for secure and sustainable management of these structures, alongside the imperative for rigorous regulatory supervision and meticulous risk evaluation. The recurrence and gravity of these incidents demand a profound examination of current standards for tailing dam design, construction, and maintenance to prevent future catastrophes and ensure the safety of communities and the environment.

In this context, the application of advanced monitoring technologies is of paramount importance to prevent tailing dam failures. Among the most promising emerging techniques is muon radiography, which enables the acquisition of detailed images of the interiors of dense structures like tailing dams. As previously mentioned, this non-invasive technique can detect structural anomalies, density variations, and potential weaknesses within these dams. Continuous monitoring of dams through muon radiography can, therefore, provide valuable information for implementing effective preventive maintenance strategies, allowing early intervention to correct structural defects before they escalate into critical risks. Integrating this technology with other monitoring practices could be a significant step toward

safer and more sustainable management of tailing dams, helping mitigate the risks associated with potential failures.

3.1 The study case of the tailing dams

In the context of the global mining industry, case study analysis of tailing dams provides crucial insights into the dynamics that contribute to the success or failure of these structures. By examining specific instances of collapse or high-risk episodes, it is possible to identify the underlying causes, evaluate the efficacy of the monitoring technologies employed, and delineate optimal practices for preventing future disasters.

One illustrative example is the Brumadinho disaster in Brazil in January 2019 [40]. The collapse of the tailing dam, owned by the Vale mining company, resulted in the deaths of more than 250 people and caused significant environmental damage along the Paraopeba River. Analysis of this event revealed several critical issues, including inadequate drainage, insufficient monitoring, and an inadequate assessment of the structural risks, particularly given the construction technique used known as "upstream" which is generally considered less safe than alternative methods.

Another notable incident is the Mount Polley tailing dam disaster in Canada in 2014 [41]. This failure occurred in a tailings storage facility with a capacity of approximately 44 million cubic meters. The dam was around 40 meters in height and 4 kilometers in length. The failure resulted in the release of approximately 25 million cubic meters of contaminated water and residues into the surrounding environment. This incident underscored the imperative for enhancements to tailings dam design and monitoring.

These cases highlight the necessity for proactive and technologically advanced management and monitoring practices. Implementing novel technologies, alongside fostering a comprehensive safety culture at all operational levels, is crucial to mitigate the risks associated with these structures. Insights from these case studies can be applied in various contexts to improve the design, construction, and management of tailing dams worldwide. The objective of this study is to determine whether muography can effectively detect water infiltration within tailing dams and, if so, to establish the timeframe required for detection.

3.2 Simulation of the infiltration inside the dam

To determine if the muography method could detect infiltration in the dam, we make use of our fast simulation tool. We used a simplified model with four zones

(Figure 3.1, left). Zones one and two were simulated with a dry density of 2.7 g/cm^3, while zone four always contained water, with a density of 3.0 g/cm^3. Zone three was initially modeled as dry density and subsequently with water infiltration, which corresponds to average density of 3.0 g/cm^3 to test whether muography could detect this change. We summed the opacity of each individual simulated piece, to obtain an equivalent opacity composed of areas with different densities for each of the two cases. These were then converted into simulated transmissions. As a result, we obtained two distinct maps, whose difference arises from the difference in the average density used to simulate zone three. We will refer to the simulation conducted with zone three at a density of 2.7 g/cm^3 as the dry (or standard) case, while the simulation conducted with zone three at a density of 3.0 g/cm^3 will be referred to as the infiltrated (or anomalous) case.

By calculating the transmission ratio between the two cases (dry versus infiltrated), we can identify the infiltrated zone. The term "relative transmission" will be used henceforth to denote this ratio. On the right side of Figure 3.1, this ratio from the simulations is shown in elevation-azimuthal coordinates. Using Figure 1.11 as a reference, the azimuthal angle (φ) is plotted on the X-axis, while the elevation angle (e) is on the Y-axis. White-colored areas are outside of the dam profile, where the opacity is zero. Purple-colored areas indicate regions with equal density (ratio equal to one), while yellow-colored areas show regions with infiltration (ratio greater than one). This pattern arises because the relative transmission is defined as the ratio between the dry and infiltrated cases, with the latter simulated at a higher density, resulting in lower expected counts than the dry case. This is an ideal measurement obtained from our simulation software, as described in Section 1.3. Detector performances and systematic effects were not included.

Since the measured transmission is the ratio between the counts with the target present and the counts in free-sky conditions, to replicate a real measurement, we multiplied the simulated transmission by the MIMA free-sky obtained at a 20° elevation inclination after 30 days of acquisition, $N_{i,j}^{\text{target}} = t_{i,j}^{\text{simu}} \cdot n_{i,j}^{\text{free-sky}}$. The choice of 20° is due to the fact that the dam is located at very low elevations. Therefore, we decided to simulate our data acquisition campaign by pointing the detector at this elevation. The value of the *i-th, j-th* bin in the free-sky histogram is represented by $n_{i,j}^{\text{free-sky}}$, while the value of the *i-th, j-th* bin in the Simulated Transmission histogram is represented by $t_{i,j}^{\text{simu}}$. Thus, $N_{i,j}^{\text{target}}$ represents the value of the *i-th, j-th* bin referenced to a data acquisition period of 30 days of expected counts. We did this procedure for both the standard and infiltrated configurations. This procedure is visually demonstrated in Fig. 3.3. The histogram of the free-sky obtained with the MIMA detector at a 20° inclination is displayed on the left. Since the simulations show the dam for elevation values between zero and approximately 25°, a cropped version of the histogram is presented. For a comprehensive view,

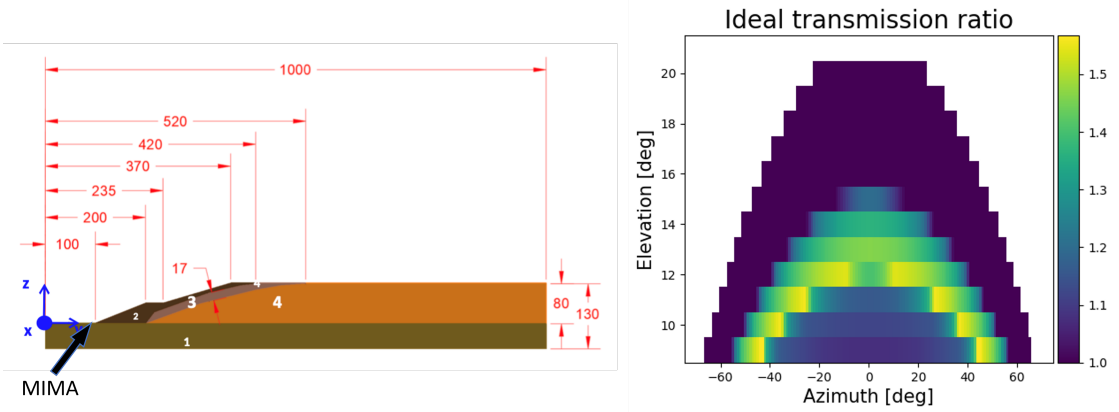


Figure 3.1: On the left is the bi-dimensional projection of the *computer-aided drafting* (CAD) of the dam used for the simulations; zone three is where we simulated the infiltration by varying its density. The numbers in red represent distances expressed in meters. On the right is the ideal ratio between the transmission measured with and without infiltration.

the full version is provided in Fig. 3.2. Due to the extensive acceptance of the detector and the decline in muon flux with decreasing elevation, the region with the highest number of counts appears between 50° and 60° even at a 20° inclination. The center of Fig. 3.3 displays the histogram of simulated transmission obtained with the simulation software for the non-infiltrated case, and the histogram of expected counts after 30 days for the non-infiltrated case is shown on the right.

To best reflect a genuine measurement, we incorporated statistical error. This was obtained by fluctuating the expected counts within each bin according to Poisson statistics:

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{3.1}$$

where k is the extracted random value fluctuated from the Poisson distribution having mean value λ equal to the bin content. This process replicates a true measurement by multiplying the simulated transmission by the free-sky (a measurement conducted with MIMA), which introduces measurement errors. Additionally, the inclusion of statistical fluctuations contributes to the measurement's statistical independence. This procedure was applied solely to the infiltrated case, as the objective was to ascertain whether muography can detect the presence of water infiltration. The dry case serves as a baseline for comparison, representing a continuous measurement of the dam under these conditions and therefore for a long datataking period of time statistical error are negligible.

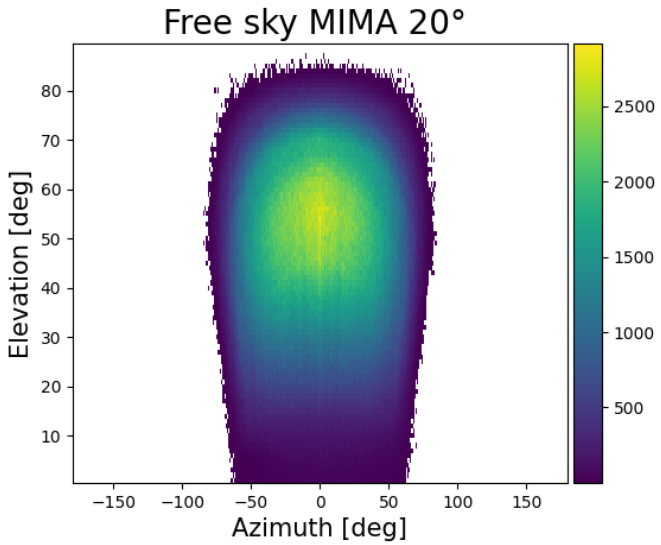


Figure 3.2: Full-size free-sky of MIMA after 30 days with an elevation of 20°.

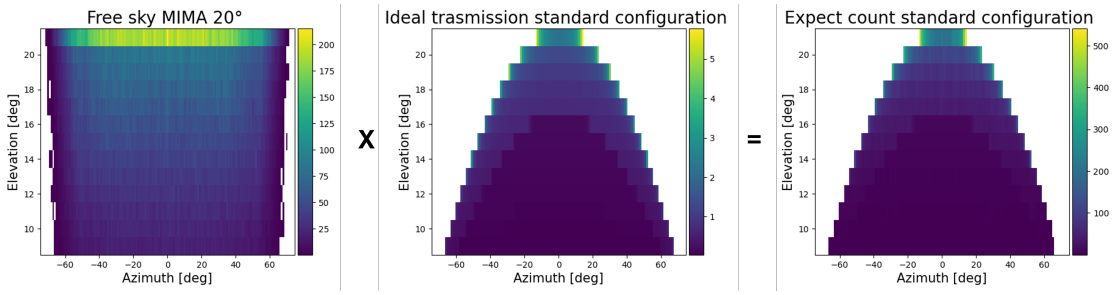


Figure 3.3: This graph illustrates the methodology employed in calculating the expected counts. On the left is the histogram of the MIMA sky at 20° elevation after thirty days of data acquisition. In the center is displayed the simulation transmission map, extracted from the simulation software for the non-infiltrated case. The number of expected counts for the non-infiltrated case after thirty days of data acquisition is presented on the right.

3.2.1 First result

To assess whether these densities allow for the detection of potential infiltration within the dam structure, data were simulated with varying acquisition statistics, effectively increasing or decreasing the free-sky acquisition time to simulate longer or shorter measurement periods. The initial results are shown in Figure 3.4. In the upper left of the figure, we see the results with a 20-fold increase in acquisition time, which corresponds to 20 months of monitoring—an impractically long duration for timely detection of infiltration in critical areas of the dam. The upper right figure shows a histogram with a 40-fold increase in statistics, where the infiltrated zone becomes somewhat discernible.

Since these extended measurement periods do not suit the study's objectives, it was decided to alter the bin size from a 1x1 to a 3x5 configuration. The lower left histogram in Figure 3.4 shows a 3x5 bin arrangement with a 4-fold increase in acquisition time, while the lower right histogram shows the same binning with a 20-fold increase. Comparing these with the initial 1x1 bin size histograms (e.g., the top left figure versus the bottom right figure, both with a 20-fold increase), it is evident that the 3x5 bin arrangement enhances the visibility of the infiltrated zone. However, in all cases, the presence of the infiltrated zone remains faint. Additionally, even the promising results would require prolonged acquisition times, delaying potential corrective actions on the dam structure.

The results confirmed that the standard binning method alone does not provide a sufficiently clear view of the infiltrated area. To refine the analysis, we decided to implement an adaptive binning method known as Voronoi tessellation. The following section outlines this approach and how it improves signal resolution.

3.3 Use of adaptive binning

This section presents a detailed account of the Voronoi tessellation [42] of a two-dimensional histogram. To better understand the operation of this tessellation, histograms are expressed in polar instead of alt-azimuth coordinates. The X-axis represents the projection $\theta \cos \varphi$ and the Y-axis $\theta \sin \varphi$. When collecting particle counts, such as muons, with a detector like MIMA, the data are affected by the detector's geometry and by the nature of the flow which decreases as the square cosine of the zenith angle ($cos^2\theta$). Using MIMA's free-sky counts at $\theta = 0°$ as an example (Fig. 3.5), it can be observed that counts are higher at the center of the detector and decrease toward the sides, resulting in a variation of the signal-to-noise ratio (S/N) across the detector, with some spatial elements showing reduced accuracy given their low S/N.

Binning techniques allow for local averaging of data. Given the high S/N

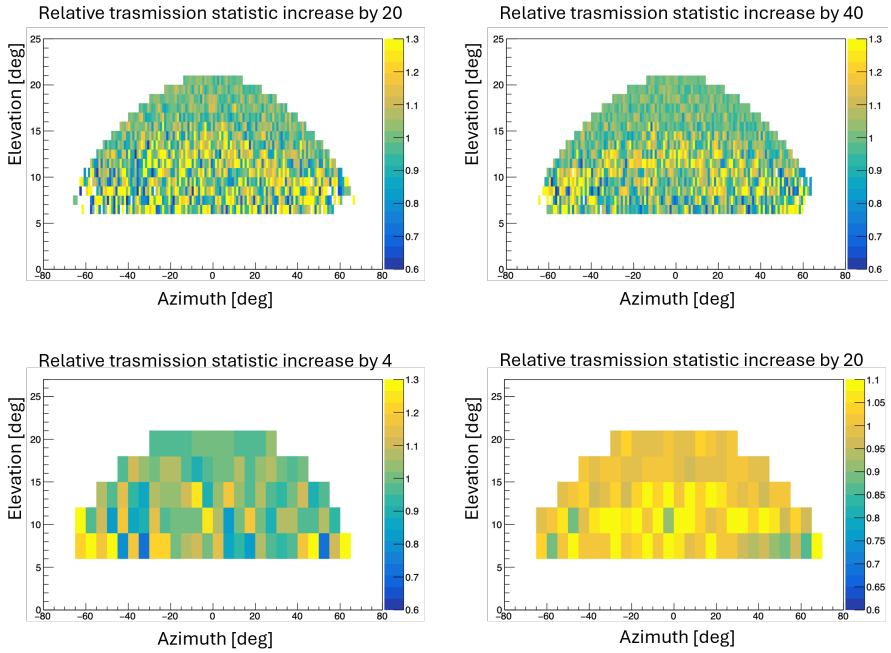


Figure 3.4: Plots of relative transmission (ratio of transmission in the dry case versus the case with infiltration). Top left: statistic increased by 20. Top right: statistic increased by 40. Bottom left: statistic increased by 4 with bins changed from 1x1 to 3x5. Bottom right: statistic increased by 20 with bins changed from 1x1 to 3x5.

variability across the detector, it is useful to apply an adaptive binning scheme where bin size is adjusted to the local S/N: larger bins will appear in regions of low S/N, while smaller bins appear in high S/N areas, offering higher resolution.

In this section, we refer to "pixel" as a specific spatial data element (e.g., X and Y points in the original histogram) and "bin" as the elements in the new histogram obtained by merging the original pixels.
Each pixel i has its own associated signal S_i, with the noise defined as $N_i = \sqrt{S_i}$ (where the signal is the bin's count). The resulting S/N when merging two pixels is defined as:

$$\left(\frac{S}{N}\right)_{1+2} = \frac{S_1 + S_2}{\sqrt{N_1^2 + N_2^2}}. \tag{3.2}$$

A good binning scheme should satisfy the following criteria:

- **Topological requirement:** Bins should tessellate the region of the plane (Ω), creating a partition without overlaps or gaps.

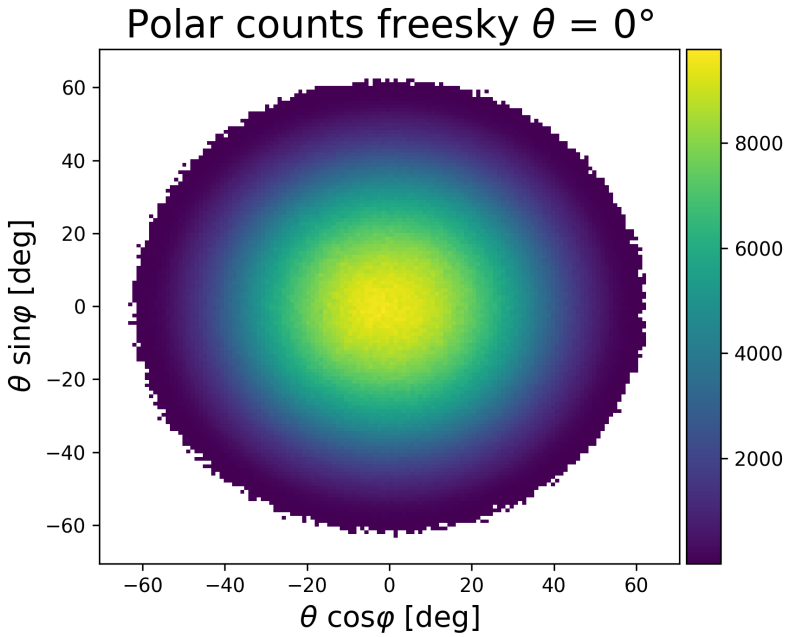


Figure 3.5: MIMA free-sky at $\theta = 0°$ after one month of acquisition in polar coordinates.

- **Morphological requirement:** Bins should be as "compact" (or "round") as possible so that pixels within a bin are close together, giving the best spatial resolution in all directions.
- **Uniformity requirement:** The S/N scatter among bins should be minimal, centering around the target value.

3.3.1 Voronoi Tessellation

A suitable method [42] for generating bins that satisfy these criteria is the *Voronoi Tessellation* (VT), defined as follows:

Given a region Ω and a set of points $\{z_i\}_{i=1}^N$ called generators, a Voronoi tessellation (VT) of Ω partitions it into regions $\{V_i\}_{i=1}^N$, where each region encloses points closer to z_i than to any other generator.

Voronoi tesselation has several favorable features for binning:

1. It satisfies the topological requirement by default.
2. It is efficiently described by the coordinates of its generators alone.

3. It is straightforward to implement in the discrete case: given generator positions, it is sufficient to locate the generator closest to a given pixel to determine its bin.

In the discrete case, to determine whether a pixel belongs to a region $\{V_i\}$ or $\{V_j\}$ (where $\{V_i\}$ is the area of the bin containing generator $\{z_i\}$), we calculate the distance $d_i = ||z_i - x_i||$ (where x_i is the pixel) and assign the pixel to the region where this distance is minimized, as illustrated in Fig. 3.6.

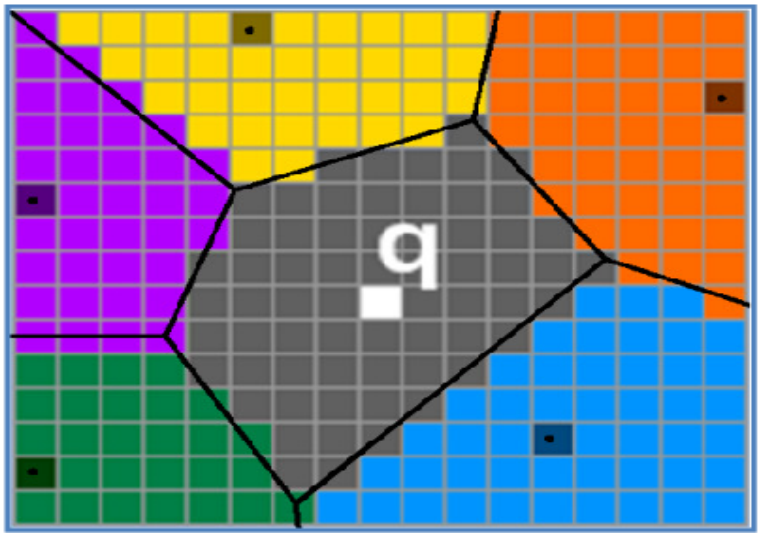


Figure 3.6: Voronoi tessellation in the discrete case. Each generator (white square) calculates the distance to all pixels. Pixels are associated with region V_i (in gray) when this distance is minimal.

VT alone does not impose the morphological requirement: bins may have sharp angles, and uniformity is also unmet with standard VT. These requirements must be addressed by custom-distributing the Voronoi generators. The following section outlines this distribution process.

Centroidal Voronoi Tessellation

A density distribution $\rho(\mathbf{r}) = (S/N)^2(\mathbf{r})$ can be defined, so binning with constant S/N reduces to creating a tessellation that encloses equal mass according to ρ. Centroidal Voronoi Tessellation (CVT) [42] is a technique that can generate optimally smooth and uniform VT in the continuous case or when the number of pixels is large. Given a distribution $\rho(\mathbf{r})$ over Ω, a CVT of Ω is a VT where each z_i generator aligns with the mass barycenter:

$$z_i^* = \frac{\int_{V_i} \mathbf{r}\rho(\mathbf{r})d\mathbf{r}}{\int_{V_i} \rho(\mathbf{r})d\mathbf{r}} \tag{3.3}$$

of the corresponding Voronoi region V_i. In the 2D case, CVT creates bins that vary in size according to density distribution but tend asymptotically toward a hexagonal lattice. Lloyd's (1982) algorithm [43] can calculate CVT in practice. Although CVT bins shrink with greater density, the bins do not maintain constant enclosed mass directly, but a modified Lloyd's algorithm allows it to converge to an equal-mass 2D-CVT, as reported in [42].

With this preparation, the scheme proceeds as follows:

Given Ω, ρ, and N generators $\{z_i\}_{i=1}^N$, the modified Lloyd method creates an equal-mass 2D CVT by:

1. Selecting an initial random set $\{z_i\}_{i=1}^N$ of generator positions, with the probability of each point based on $\rho(\mathbf{r})$.

2. Performing VT of Ω around $\{z_i\}_{i=1}^N$: the Voronoi regions $\{V_i\}_{i=1}^N$ average equal mass, but with irregular shape and S/N scatter.

3. Calculating the center of mass of each $\{V_i\}_{i=1}^N$ according to density $\rho' = \rho^2$, creating a new set of generators $\{z_i\}_{i=1}^N$.

4. Repeating steps 2-3 until VT generator coordinates converge.

Some practical challenges remain:

1. The algorithm produces bins of equal mass but not equal S/N.

2. It works less effectively when bins have very few pixels, as Lloyd's method assumes continuity and may not reach the global minimum in discrete cases.

For discrete cases, selecting initial generators is crucial for Lloyd's algorithm to converge globally, producing bins with a S/N in a neighborhood of the region with a chosen S/N threshold ($(S/N)_T$) and achieving efficiency. The following section introduces the generator-selection method used with the modified Lloyd method.

Voronoi Growth Algorithm and Generator Choice

The Voronoi 2D binning algorithm is as follows [42]:

1. Choose the first bin from the pixel with the highest S/N in the starting histogram.

2. Calculate the bin's center of mass and select the nearest unbinned pixel as a candidate for addition.

3. Add the candidate pixel if these criteria are met:

   i Topological: The pixel is adjacent to the current bin.

   ii Morphological: Adding the pixel keeps roundness R below a threshold.

   iii Uniformity: Adding the pixel improves S/N.

4. End bin growth. If S/N exceeds 80% of $(S/N)_T$, mark as "success bin"; otherwise, as "failure bin".

5. Calculate the center of all binned pixels and start a new bin from the nearest unbinned pixel.

6. Reassign pixels from "failed bins" to the nearest successful bin center.

7. Recalculate the center of each bin, using these as CVT starting points with the modified Lloyd method.

Quantity R in step $3ii$ measures bin roundness as:

$$R = \frac{r_{max}}{r_{eff}} - 1, \tag{3.4}$$

where r_{max} is the maximum distance between the bin's center and any pixel in it, and r_{eff} is the radius of a disk of the same area as the bin. A threshold of $R_{max} = 0.3$ is used. Fig. 3.7 shows an example of steps 1-5.

Given the complexity of this technique, before showing its application in the field of muography, we briefly summarize it here.

The adaptive Voronoi tessellation method is used to enhance the analysis of images obtained from muography techniques by optimizing data binning. Starting from a two-dimensional histogram, this method redistributes bins so that they are proportional to the local signal-to-noise ratio (S/N). The bins become smaller in areas with high S/N, where higher resolution is needed, and larger in low S/N areas to reduce statistical fluctuations.

Voronoi tessellation divides a plane into regions according to a series of generators, assigning each pixel to its nearest generator. To meet uniformity and shape compactness requirements, a variant called *Centroidal Voronoi Tessellation* (CVT) places generators at the centroids of their respective regions. Lloyd's method is then used iteratively to move the generators towards the center of mass of their assigned regions, ensuring a more uniform and compact bin distribution.

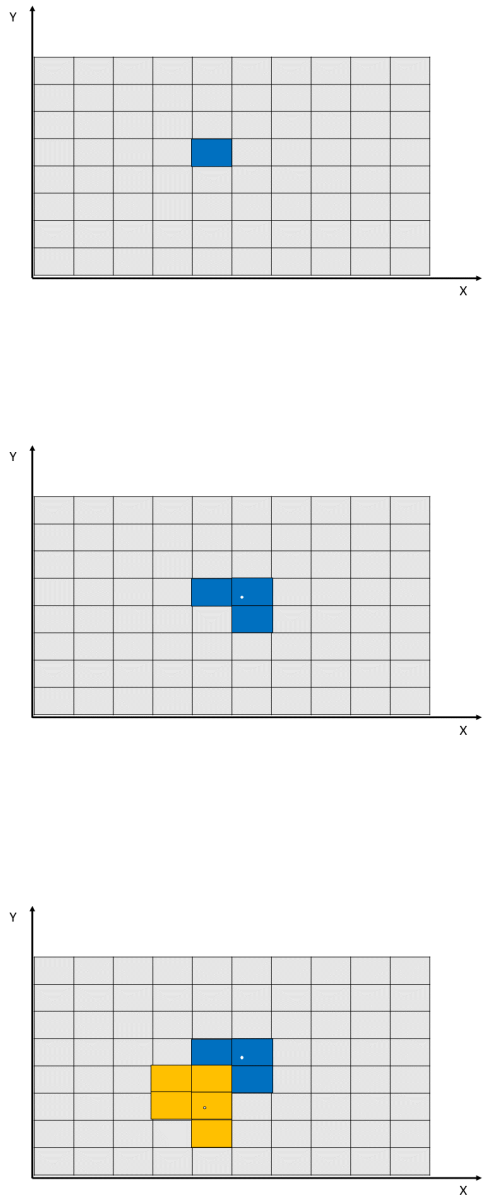


Figure 3.7: 1) The pixel with the largest S/N is chosen as the starting bin. 2) The first bin is created by aggregating adjacent pixels. 3) The center of mass (white dot) of the first bin (blue) is calculated, and the nearest unbinned pixel becomes the starting pixel for the second bin (yellow). The process is repeated for subsequent bins, each starting from the nearest unbinned pixel to the center of all prior binned pixels.

To optimize the method in discrete cases, the process is initialized with a bin "growth" algorithm: starting from the pixel with the highest S/N, the bin expands by adding adjacent pixels that satisfy topological, morphological, and S/N uniformity criteria. Finally, any unsuccessful bins are reassigned to the nearest successful bin, and the resulting centroids are used as starting points for the final CVT. This procedure provides an adaptively optimized resolution, enhancing the analysis of muographic images.

3.4 Use of the Voronoi tessellation in the Tailing dam

An application of the accretion algorithm using Voronoi tessellation is presented here. Taking Figure 3.5 as a reference, and applying the previously described procedure, the resulting tessellation is shown on the left in Figure 3.8. For this analysis, we chose a target signal-to-noise ratio defined as the square root of the maximum histogram count, scaled by 1.5:

$$\left(\frac{S}{N}\right)_T = \sqrt{N_{max}} \times 1.5.$$

Given the scarcity of expected counts due to the low elevation values, we decided to choose this S/N value to ensure that all bins were rebinned and could achieve a signal-to-noise ratio as free as possible from statistical fluctuations. We also experimented with larger values, but at that point, we lost resolution because the bins became too large. This choice allows us to strike the right balance between reducing statistical uncertainties and maintaining precision.

As seen in the center of the histogram, the areas with initially higher counts yield smaller-sized bins, while the outer regions, where the initial counts are lower, have larger bins. This outcome achieves the goal of adaptive binning, maintaining high resolution in areas with abundant data while reducing statistical fluctuations by merging bins in regions with lower counts. The histogram on the right in Figure 3.8 compares the distribution of the signal-to-noise ratio before (in blue) and after (in orange) the application of Voronoi tessellation. It is evident that the initial random S/N distribution transitions to a Gaussian distribution centered around the target S/N value. An example of the successful use of this technique in muon radiography is provided in [44].

Now we will demonstrate the application of this variable tessellation to the study of tailing dams.

After understanding the operation and benefits of Voronoi tessellation, we decided to apply it to our tailings dam analyses. In particular, we processed the

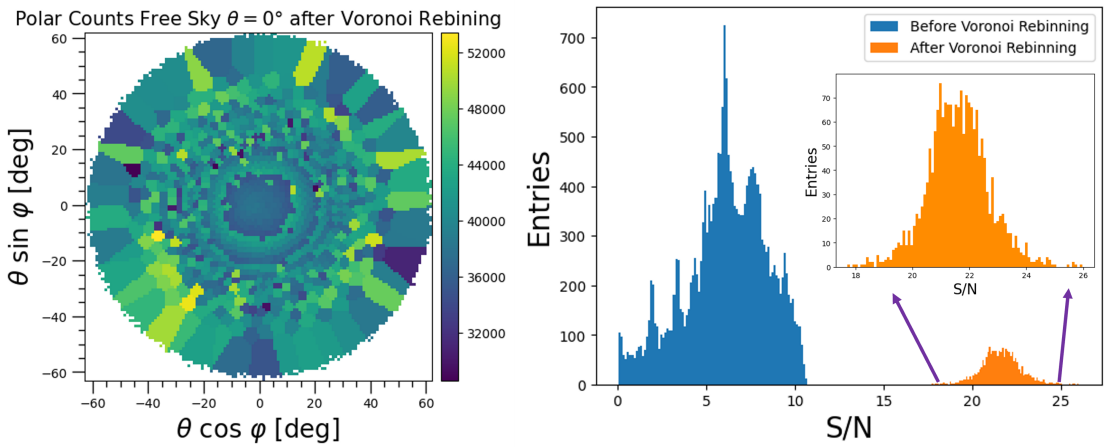


Figure 3.8: On the left, the free-sky in polar representation after applying Voronoi transformation is shown. On the right, the S/N distribution before and after Voronoi tessellation is illustrated.

histograms, using Voronoi tessellation for both the expected counts and the simulated infiltration data. The results are shown in Figure 3.9. The first four images show the results obtained with different acquisition statistics, while the last one represents the ideal transmission ratio, i.e., without applying fluctuations. The latter represents a visual benchmark for our algorithm. The first four images were obtained with an increased statistic of 4, 16, 20, and 40. Compared to the previous cases (without Voronoi tessellation), the infiltrated area is more visible with the same statistics for 16, 20 and 40. A preliminary analysis sets a lower limit of 16, which equals four months of data collection with a detector four times larger than MIMA, for example BLEMAB. These results are certainly more satisfactory and better compared to those previously obtained with a fixed bin, although not yet entirely exhaustive. Four months using a detector like BLEMAB would be far too long to detect an infiltration. Our goal is to identify it within a timeframe that allows for prompt intervention ideally, within just a few days.

This first result nonetheless provides evidence that the use of this method, to rebin the bins according to the signal-to-noise ratio, is efficient and allows for better visualization of sensitive areas.

The previous assumed densities are averages of different dams. In particular cases, however, these densities can be significantly lower. In fact, collaborators of the Geology department of the University of Florence, have encountered case were dry areas can reach $\rho = 1.4, \mathrm{g/cm}^3$ and $\rho = 1.9, \mathrm{g/cm}^3$ for the areas with the presence of water. Therefore, this density change can reduce the acquisition times even by a factor of 4-5. In fact, the total numbers of counts go from 18386 to 278169 for the dry case, while for the infiltration case, the number of counts went

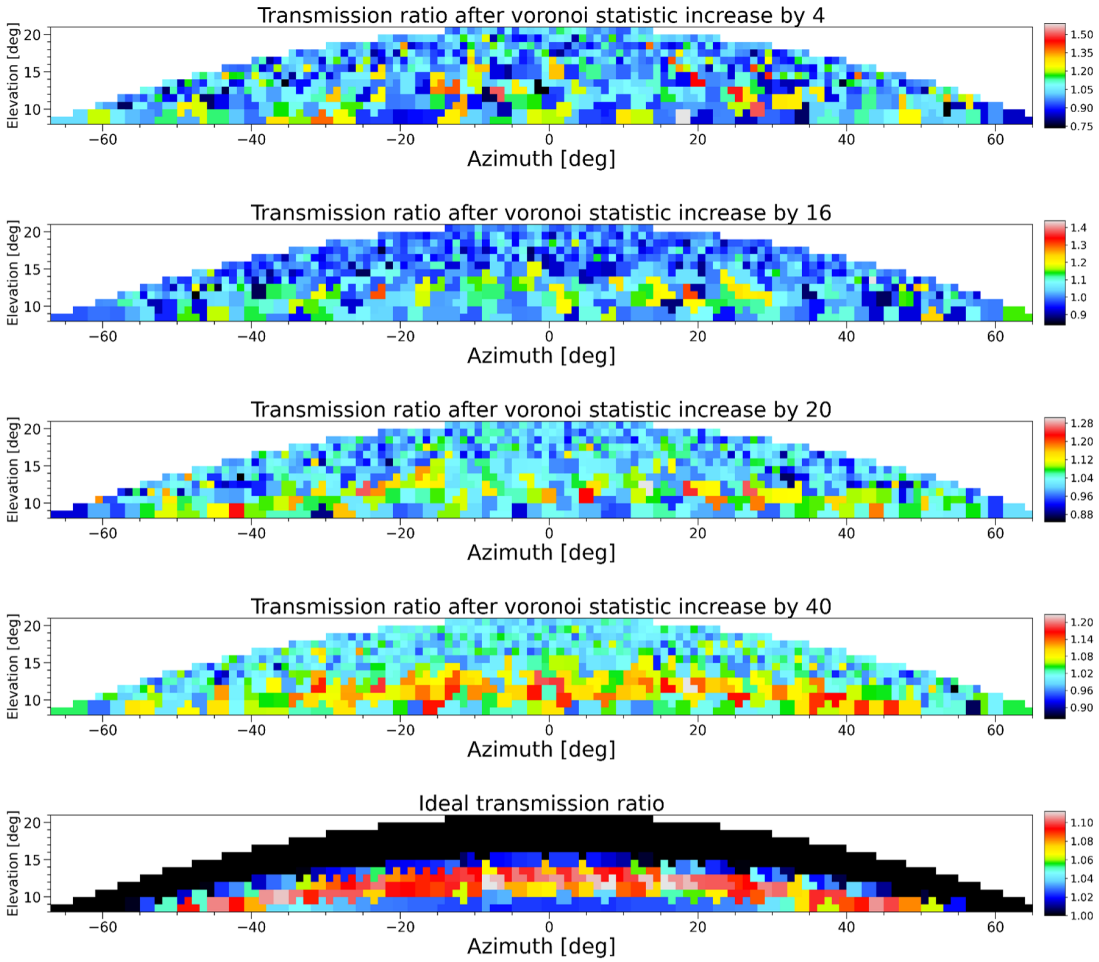


Figure 3.9: The relative transmission (dry/infiltration) is shown for various statistics: 4, 16, 20, 40 months of acquisition. The bottom figure represents the ratio between the dry situation and the infiltrated one for the ideal case, i.e., without adding statistical fluctuations.

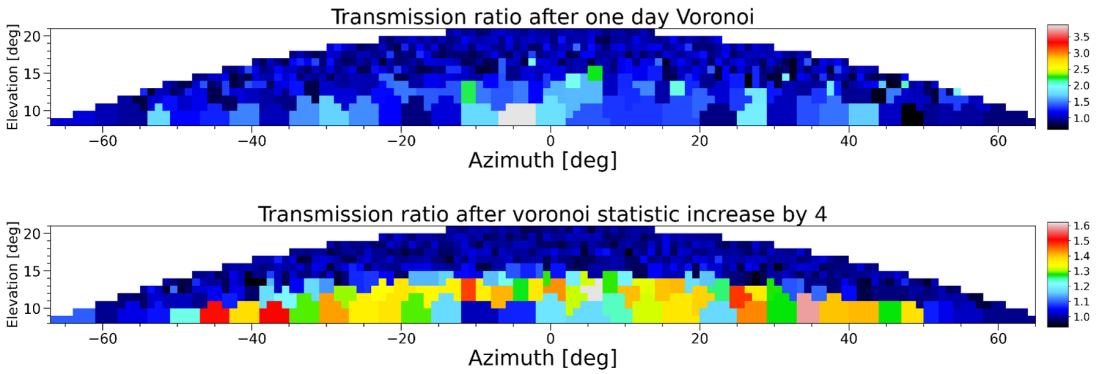


Figure 3.10: At the top is the relative transmission map, with a simulated statistic of one day using a detector four times the size of MIMA. In the bottom figure is the relative transmission map with the statistic simulating one month of data acquisition with a detector the size of MIMA. Both refer to the new densities used.

from 18275 to 275941. Both refer to acquisition times of one month with MIMA.

In Figure 3.10, the results with the new simulated densities can be seen. The first corresponds to a relative transmission map with a data acquisition statistic of one day, using a detector four times the size of MIMA. The second shows the relative transmission map with a four-month statistic using a MIMA-sized detector. It is interesting to note that, in the case of four months, the evidence of the infiltrated area is clear. Compared to the same statistic in the previous figure (Fig. 3.9), it is noticeable how the change in density plays a fundamental role in this study. Observing the first transmission map in Figure 3.10, we can see that even with a detector four times the size of MIMA, we are unable to observe the infiltrated area in one day.

Although Voronoi tessellation improved the data resolution, there was still room for improvement in accurately identifying infiltrated zones. To address this challenge, we introduced the *Random Forest* an advanced classification method. This algorithm allows to increase the accuracy of predictions and further reduce classification errors.

3.5 Use of the Random Forest classifier in the tailing dams case

Random Forest (RF) is a prediction and classification method using multiple decision trees (500-1000) [45] that is widely used across research fields, including

healthcare. Before examining RF, it's important to understand decision trees, which form its core structure. Given its popularity in healthcare, we will use this field as an example for our introduction.

3.5.1 Introduction to Decision Trees

Decision trees classify different outcomes (e.g., tumor subtype A vs B) by evaluating all observations and identifying the best feature and split value. For instance, if the predictors include age, BMI, sex, diabetes, and hypertension, the tree determines the optimal feature and cutoff point for distinguishing between tumor types, a flexibility that is a key strength of decision trees.

For example, the possible cutoff for patient age could be any year observed in the dataset (e.g., 29-81 years). The decision tree examines every possible cutoff (e.g., age < 30 vs age ≥ 30, age 31 vs age ≥ 31, etc.) to determine which cutoff divides the dataset to produce the best two child nodes. How does the decision tree identify age 60 as the best way to separate tumor type A from tumor type B? As with most algorithms, the term "best" can vary; here, we assume "best" denotes finding a split that yields two purer child nodes in terms of composition, aiming to place patients with subtype A in one node and patients with subtype B in another. After the initial split, this process continues independently within each child node until the tree is fully grown, at which point further splits do not improve purity. In the context of decision trees, purity refers to how homogeneous the observations in a node are with respect to the target variable. A pure node contains observations that all belong to a single class (e.g., all patients have tumor subtype A or all have subtype B). The goal of splitting a node is to increase the purity of the resulting child nodes compared to the parent node. Decision trees can also be pruned or limited in depth to avoid overfitting to the training data. The final tree assigns a predicted outcome (e.g., tumor subtype A or B) based on the majority of observations within that node. Predictions for new patients are generated by running their data through the tree until reaching a terminal node associated with a prediction.

Understand when stop splitting the tree is complex, as theoretically, we could keep splitting until purity improvements cease. Such a large tree might fit the dataset well, but its performance would likely decline on a different dataset, indicating overfitting. Overfit decision trees are high-variance classifiers; even with samples from the same population, large, unpruned decision trees tend to overfit the dataset in question, leading to significant variation in predictions.

Now that we understand how and why a decision tree is created, its strengths and weaknesses, we will examine how Random Forest can be used to improve decision tree performance.

3.5.2 Introduction of Random Forest Classifier

Random Forest is a collection of 500-1000 decision trees. To construct each tree to be as independent as possible, RF perturbs the dataset using bootstrapping (i.e., randomly sampling, with replacement, members of the original dataset, resulting in a dataset of the same size but consisting of a constantly perturbed version of the data) and randomly selects only a subset of the total available features to evaluate at each node as potential splitters. We can predict whether a new patient will have tumor subtype A by obtaining the prediction from each decision tree and allowing the final prediction to be based on the proportion of trees that predict subtype A (if 800/1000 trees all predict subtype A, the final prediction will be that the patient has subtype A). The advantage of Random Forest is that, by counting the "votes" of each member tree, the prediction tends to be much more stable compared to the prediction from a single decision tree.

Let's analyze some of these aspects to understand better.
Suppose we have a training set composed of n individuals. When we perform bootstrap on this set, we sample n individuals from this set with replacement, so for a particular bootstrap sample, some individuals will appear in the sample more than once and others (about $\frac{1}{3}$ of the data) will not appear in the sample at all. When performing bootstrapping multiple times, we essentially create several new training sets that are similar but all slightly perturbed from each other and from the original dataset. This allows Random Forest to build a series of different trees. If we had only one training set, every tree would be exactly the same.

Recall that every time we try to split a parent node into two child nodes in a decision tree, we evaluate all the variables at all possible cut points. Random Forest selects only a random sample of predictors from which the tree can choose (instead of the entire set of predictors). We do this because, like bootstrapping, it is another way to allow the resulting trees to be as different as possible from each other.

By taking all these high-variance trees and averaging their predictions, the resulting classification tends to have much lower variance. In summary, Random Forest creates a collection of decision trees that are forced to be as different as possible through bootstrap sampling and random selection of predictors. It is this averaging across the trees that makes Random Forest so effective at reducing the high variance of a decision tree.

3.5.3 Data preprocessing

As mentioned, the goal of this work is to develop algorithms for muography that are capable of providing both numerical and visual results. It was therefore decided to use a *Random Forest* (RF) classification algorithm [46] to identify the areas

where infiltration is present. This algorithm complements the Voronoi tessellation, as having bins with a similar S/N ratio throughout the histogram will allow for better training of the algorithm itself. As described in the previous subsections, it is important to have as many variables as possible in order to generate trees with different nodes. New variables useful for training the algorithm have thus been created. The first is the *Significance*, defined as:

$$S = \frac{|t_{standard}^{no-flut} - t_{inf}^{flut}|}{\Delta t_{inf}^{flut}} \tag{3.5}$$

where $t_{standard}^{no-flut}$ is the transmission of the dry area without fluctuations. This is because, as mentioned, the dry configuration is assumed to be the standard one, which would represent a measurement taken over a long period of time, and thus statistical fluctuations are negligible. t_{inf}^{flut} is the transmission in the infiltrated case after applying the fluctuations, while Δt_{inf}^{flut} is a weight and represents the ratio between the square root of the counts in the infiltrated case compared to the counts in the free-sky. The meaning of significance is to find how much the infiltrated case deviates from the standard (dry) case.
For each bin, the significance and relative transmission values of neighboring bins were then considered. Specifically, for each bin, 4 variables were added (in addition to transmission and significance), which are the sum of the relative transmission and the significance of all bins enclosed within an ellipse with semi-axes (6°,3°) and a circle of radius 3°, both centered on the bin itself. A graphical example of what is described is shown in Figure 3.11. Each dot represents the center of the Voronoi bin.

To train the Random Forest algorithm, we considered different infiltration shapes into the dam. Some examples are shown in Figure 3.12. The figure shows the ground truth maps of the infiltration shapes used to train the model. The white bins represent the presence of infiltration, while the black bins represent the background, i.e., where no infiltration is present. The first map in Figure 3.12 shows the baseline ground truth map for complete infiltration, meaning the entire Zone Three is considered infiltrated. For the other images, however, we only considered portions of this infiltration. The second shows the ground truth map where a small part of the infiltration is missing. The third, instead, shows the ground truth map with only a small part of infiltration present. Five configurations were created where only a piece of infiltration is missing, and another five where only a piece of infiltration is present. In total, there are 11 configurations. Each of these was then fluctuated 100 times. By fluctuating, we mean that the counts were fluctuated, as described previously in Section 3.1. The purpose of this is to generate for each configuration 100 cases of study, which helps generalize the model as much as possible and reduce statistical fluctuations. To test the model,

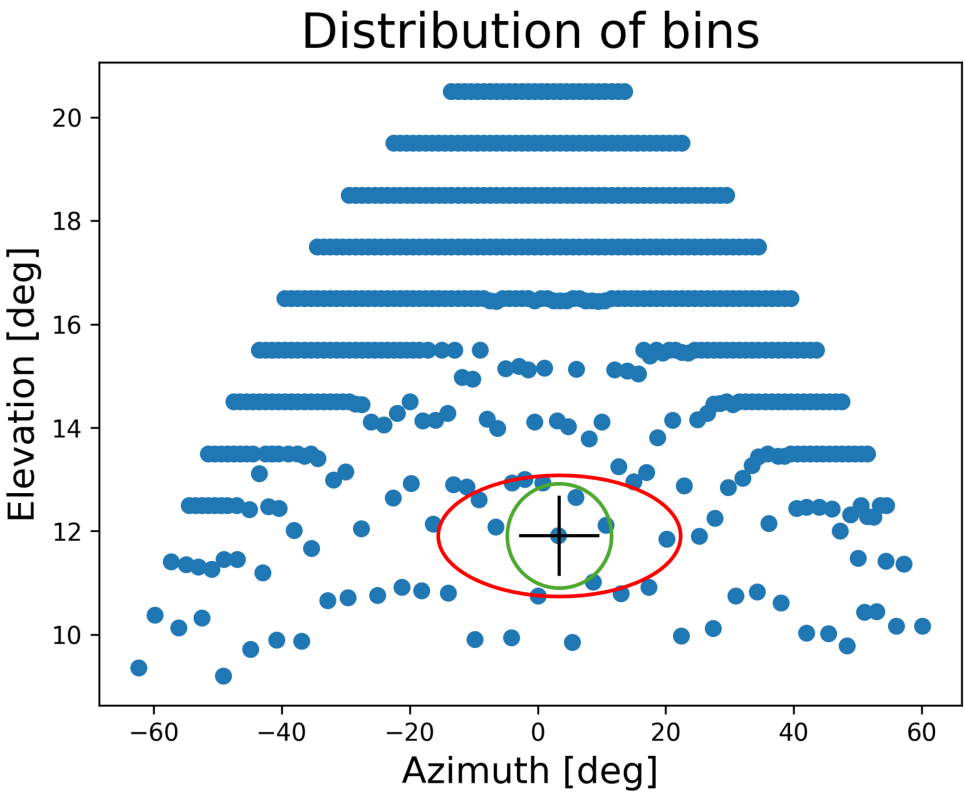


Figure 3.11: Plot of the center of each bin after transformation through Voronoi Tessellation. The circle (with a radius of 3°) and the ellipse (with semi-axes of (6°,3°)) represent an example to understand how neighboring bins are selected from which the sums of significance and transmission are derived.

three additional configurations were also created, in which the infiltration regions were modified by applying random alterations to their shape and extent within the dam structure. One of these is shown at the bottom of Figure 3.12. Each of these three configurations was also fluctuated 100 times. Each of these experiments was generated with statistics equivalent to one day of data collection with a detector four times larger than MIMA. Therefore, in a situation where, as we have seen in Figure 3.10, we are not able to observe visually the infiltration.

At this point, we are ready to train the *Random Forest* (RF) algorithm. One premise is that it does not analyze the entire image but examines each bin individually, evaluating the variables provided to it and deciding whether to classify that bin as 0 (absence of infiltration) or 1 (presence of infiltration).

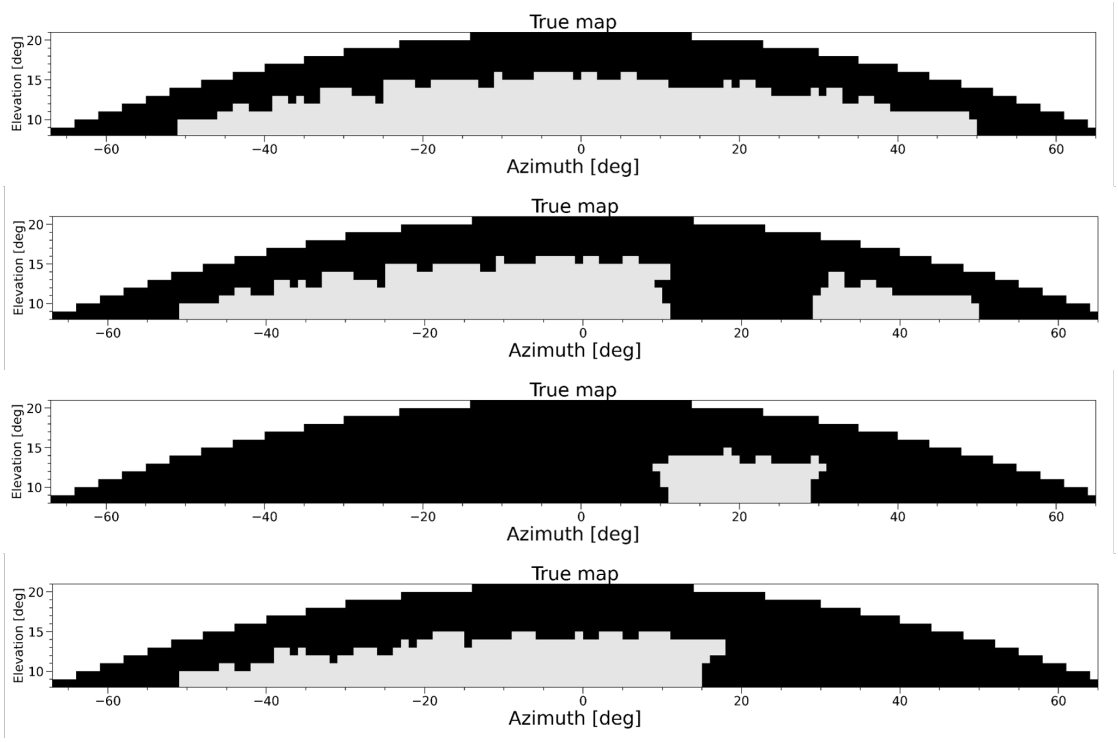


Figure 3.12: All the images show the ground truth map of the infiltration location (white bins) relative to the background (black bins), i.e., the non-infiltrated area of the dam. At the top, the dam with the complete infiltration is shown. The second image shows a configuration where a piece of infiltration is missing. The third shows a configuration where only a piece of infiltration is present. The fourth image, instead, shows a configuration with random cuts, used to test the algorithm.

3.5.4 Results of the Random Forest

We trained the RF algorithm on a total of 594,000 events, with each event corresponding to a bin. In fact, we generated a total of 1100 experiments, and each experiment contains 540 events (bins). Of these 594,000 events, 60% were used for training and 40% for validation. For the test, we had 162,000 bins available. For each bin, the algorithm was provided with the following variables: transmission of the single bin, significance of the single bin, number of bins inside the ellipse centered on that bin, sum of the significance of the bins inside the ellipse, sum of the transmission inside the ellipse, number of bins inside the circle centered on that bin, sum of the significance of the bins inside the circle, sum of the transmission inside the circle.
To avoid overfitting, we imposed a maximum forest depth of 10. Additionally, we instructed the algorithm to give more importance to bins with a value of 1, i.e.,

those with the presence of infiltration, compared to those with a value of 0. This means that bins with a value of 1 will have a greater weight compared to bins with a value of 0 during the training phase.

In Figure 3.13, the importance percentages of each feature within the RF model are shown. As can be seen, the two most important features are *number neighboring bins (3, 3)* and *number neighboring bins (6, 3)*. (3, 3) and (6, 3) refer to the value in degrees of the semi-axes of the ellipse centered on the considered bin. Semi-axes (3, 3) therefore refer to a circle with a radius of 3°. This tells us that the two most important features are the number of neighboring bins for each bin, as these two indicate the number of bins within the ellipse and circle centered on each bin. On the other hand, the individual values of significance and transmission for each bin are of little importance. This is quite expected. In fact, if the transmission value of each bin were important, it would simply be enough to set a threshold on this value, and we would notice it even visually. The fact that the next two most important features are *sum neighboring transmission (3, 3)* and *sum neighboring transmission (6, 3)* indicates that the model heavily weighs the information from neighboring bins.

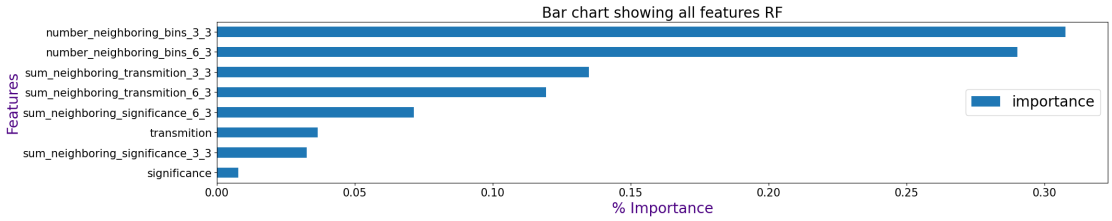


Figure 3.13: Bar chart representing the feature importance within the Random Forest model.

To quantify and evaluate the performance of the RF we used to start the *Receiver Operating Characteristic* (ROC) curve. The ROC curve is a graphical tool used to evaluate the performance of a binary classifier. It plots the relationship between the *True Positive Rate* (TPR) and the *False Positive Rate* (FPR) at different classification threshold levels, providing an overall view of the model's discrimination capabilities. To fully understand the ROC curve, it is important to clarify some key concepts related to classifier performance:

- **True Positive (TP)**: Cases where the model correctly predicts the positive class (predicts infiltration and it is actually present).
- **False Positive (FP)**: Cases where the model incorrectly predicts the positive class (predicts infiltration but it is not present).

- **True Negative (TN)**: Cases where the model correctly predicts the negative class (predicts no infiltration and it is not present).

- **False Negative (FN)**: Cases where the model incorrectly predicts the negative class (predicts no infiltration when it is actually present).

From these values, two fundamental metrics are derived:

True Positive Rate (TPR): measures the model's ability to correctly identify positives. It is defined as:

$$\text{TPR} = \frac{TP}{TP + FN} \tag{3.6}$$

False Positive Rate (FPR): measures the proportion of false positives relative to the total number of negative observations. It is defined as:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{3.7}$$

The ROC curve is constructed by varying the model's classification threshold. In a binary classifier, for each prediction, the model assigns a probability or score of belonging to the positive class. Depending on the chosen threshold value, an observation is classified as positive or negative. With a very high threshold, the model tends to classify as positive only cases with high probabilities, which reduces false positives but may also reduce true positives (high precision, low sensitivity). With a very low threshold, the model tends to classify more cases as positive, increasing sensitivity but also the number of false positives (low precision, high sensitivity).

Varying this threshold generates a set of pairs of values (FPR, TPR), which are plotted on the graph. The result is a curve that starts at the point $(0, 0)$ (no positive classifications) and ends at the point $(1, 1)$ (all classified as positive).

The perfect point on a ROC curve would be the point $(0, 1)$, which indicates a classifier with FPR = 0 (no false positives) and TPR = 1 (all positives correctly identified). In practice, the classifier's performance is evaluated based on how close the curve gets to this point.

The dashed diagonal line that runs from point $(0, 0)$ to point $(1, 1)$ represents a random classifier, which has no discriminative ability. Any classifier that approximately follows this line is no better than a random decision.

A quantitative way to summarize the ROC curve is to calculate the *Area Under the Curve (AUC)*. This value ranges from 0.5 (random classifier) to 1 (perfect classifier). The higher the AUC, the greater the model's ability to discriminate between positive and negative classes.

- AUC = 1: The classifier is perfect.

- $0.5 < \text{AUC} < 1$: The classifier has some discriminative ability.
- $\text{AUC} = 0.5$: The classifier is equivalent to random guessing.
- $\text{AUC} < 0.5$: The model performs worse than random, implying that it likely reverses the classes.

The AUC provides a single numerical value that summarizes the model's discriminative ability. In binary classification contexts, the ROC is particularly useful in the presence of imbalanced datasets, as in our case, where the number of bins with infiltration is smaller than the bins without infiltration.

At this point, we can look at our ROC curve and describe it. This is shown in Figure 3.14 and displays the performance of the *Random Forest* model on three distinct datasets: *training*, *validation*, and *test*, with very similar AUC values. The graph highlights the model's excellent predictive ability and generalizability. Below, we analyze each part of the ROC curve.

1. **AUC of the training set (99.1%)**:
   - The AUC value of 99.1% for the training set indicates that the model has learned the characteristics of the training data very well, managing to correctly separate most of the observations.
   - Although very high, this value is not perfect, which means there is no obvious overfitting. The model has acquired good discriminative abilities without excessively memorizing the training data.
2. **AUC of the validation set (98.9%)**:
   - The AUC value for the validation set is 98.9%, almost identical to that of the training set. This indicates that the model was able to generalize very well on the validation data, with a minimal difference compared to the training data.
   - This similarity between the curves suggests that the model has not overfitted the training set, but is able to maintain similar performance on unseen data, even though they are similar.
3. **AUC of the test set (99.0%)**:
   - The AUC value for the test set, which measures the model's performance on entirely new data, is 99.0%, very close to the other values. This demonstrates that the model is able to generalize effectively on unseen data, maintaining a high capacity to correctly separate the positive and negative classes.

- The fact that the AUC on the test set is very similar to the other two sets confirms the absence of overfitting and suggests that the model is stable and reliable in correctly predicting new observations.

The ROC curves for the three sets (training, validation, and test) are almost completely overlapping, and all show AUC values higher than 98%. This indicates that the model has excellent discrimination ability between positive and negative classes, with highly stable performance across the different datasets. The model has been well-calibrated and is capable of handling new data without sacrificing accuracy.

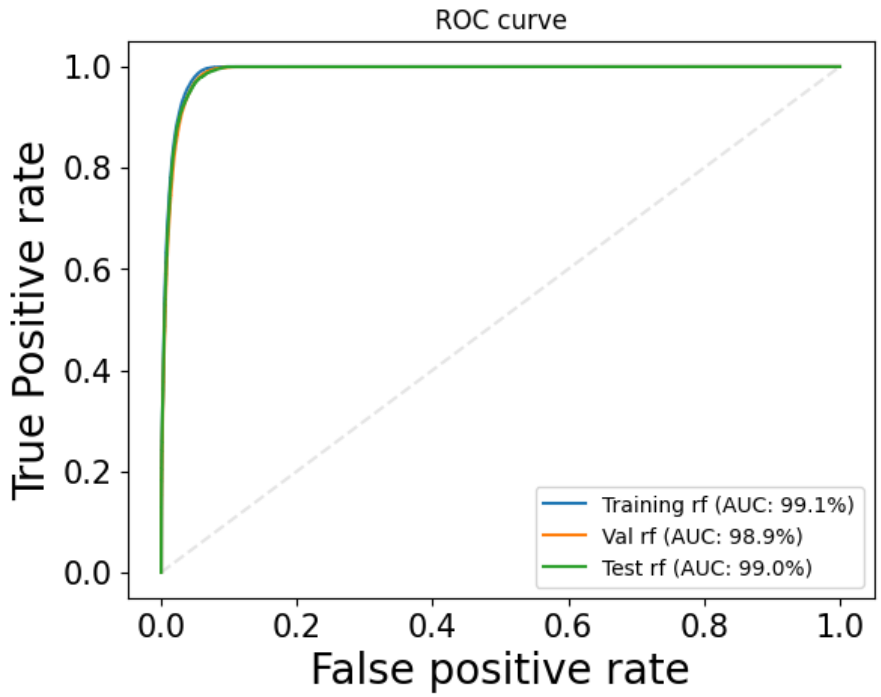


Figure 3.14: ROC curve showing the performance of the Random Forest model on the three datasets: training (AUC: 99.1%), validation (AUC: 98.9%), and test (AUC: 99.0%). The almost perfect overlap of the curves indicates an excellent discrimination ability of the model and good generalization without obvious signs of overfitting.

3.5.5 Confusion Matrix

In Figure 3.15, another estimator is presented to better understand the performance of our RF, the *confusion matrix*. This represents the results of the model

applied to the test dataset. The confusion matrix is a useful method to visualize the performance of a classifier, especially in binary classification problems. As for the ROC curve we can also define the TN, TP, FP and FN, for the confusion matrix which is applied only to the test data. At top left (TN) we have 140241, this value represents the number of instances where the actual class is 0, and the model correctly predicted 0. Therefore, the model correctly predicted 140241 non-infiltration examples. At the top right(FP) we have 8859, this value represents the number of instances where the actual class is 0, but the model incorrectly predicted 1. In other words, 8859 times infiltrations were predicted when there were none. At bottom left (FN) we have 267, this value represents the number of instances where the actual class is 1, but the model incorrectly predicted 0. Therefore, the model failed to correctly classify 267 infiltrations, classifying them as non-infiltrations. At the bottom right (TP) we have 12633, this value represents the number of instances where the actual class is 1, and the model correctly predicted 1. The model correctly predicted 12633 infiltrations.

From this matrix, we can calculate several evaluation metrics to better understand the model's performance:

- **Accuracy**: Accuracy measures the proportion of correct predictions over the total predictions:

  $$\text{Accuracy} = \frac{TN + TP}{TN + FP + FN + TP} = \frac{140241 + 12633}{140241 + 8859 + 267 + 12633} \approx 0.94$$

  Therefore, the model has an accuracy of about 94% on the test data.

- **Precision (Positive)**: Precision indicates the proportion of positive predictions that are actually correct, i.e., when infiltration is present and correctly predicted:

  $$\text{Precision} = \frac{TP}{TP + FP} = \frac{12633}{12633 + 8859} \approx 0.59$$

  The model has a precision of 59% in correctly classifying positive cases.

- **Sensitivity (True Positive Rate)**: Sensitivity measures the proportion of positive examples that were correctly identified:

  $$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{12633}{12633 + 267} \approx 0.98$$

  The model has a sensitivity of 98%, indicating that it is very effective at capturing most positive cases.

- **F1-score**: The F1-score combines precision and sensitivity, offering a balanced measure:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} = 2 \times \frac{0.59 \times 0.98}{0.59 + 0.98} \approx 0.73$$

  The F1-score is about 0.73, indicating a good balance between precision and sensitivity.

The model has good overall accuracy, with an accuracy of 94%. However, it shows moderate precision at 59%, suggesting that while it correctly captures many true positives, it also generates a significant number of false positives (8859). The sensitivity is very high (about 98%), indicating that the model is very effective at recognizing most positive observations. The trade-off between precision and sensitivity is reflected in the F1-score, which is 0.73, suggesting that the model performs well but could be improved in terms of precision.

In summary, the model appears well-balanced, with a good ability to correctly identify positive cases, although it has some tendency to generate false positives. This is probably due to the fact that we have imbalanced classes, meaning the number of bins with infiltration is much smaller than the number of bins without infiltration. Even when balancing it during the training phase, the algorithm still has fewer "examples" to train on.

3.5.6 Evaluation of test data

An additional test was conducted to understand the potential of this RF model. For each of the three test configurations, 100 experiments were created as mentioned earlier. An example of the ground truth of these three configurations is shown in the fourth image of figure 3.12. We decided to separately evaluate the values of Average Accuracy, Average Precision, Average sensitivity, and Average F1-Score for each experiment. We then calculated the average for each configuration. Essentially, it is as if we computed the *Confusion Matrix* for each experiment in a configuration and then averaged the results. This ensures that each experiment is evaluated individually and the global average is computed on the scores from each experiment. In contrast, the Confusion Matrix shown earlier in Figure 3.15 was calculated on all the bins from all experiments of all configurations combined. The average results for each configuration are reported below.

Experiment 1

- **Average Accuracy**: 94.26%
- **Average Precision**: 58.92%

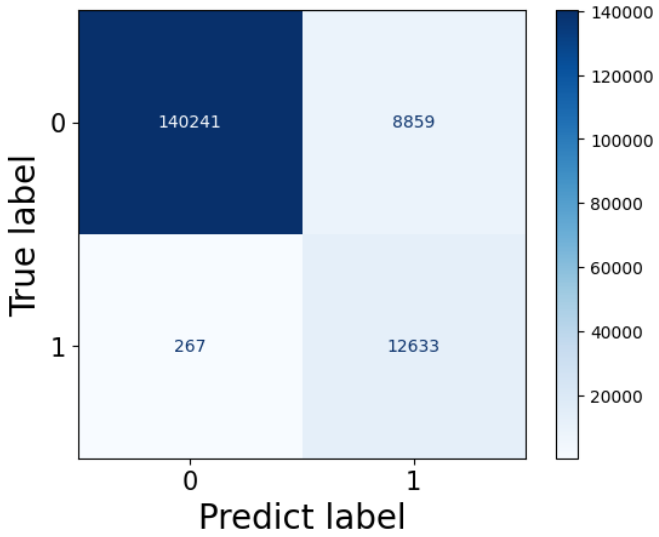


Figure 3.15: Confusion matrix of the *Random Forest* model on the test data. The model shows good discriminative ability between positive and negative classes, with an overall accuracy of 94%. However, the precision is moderate (59%), indicating a significant number of false positives. The sensitivity is high (98%), suggesting that the model successfully identifies most positive instances.

- **Average Sensitivity**: 99.39%
- **Average F1-Score**: 73.92%

In the first experiment, the model showed high *accuracy* of 94.26%, but relatively low *precision* (58.92%). This suggests that the model generates a considerable number of false positives, mistakenly classifying areas without infiltration as areas with infiltration. However, the *sensitivity* is very high (99.39%), indicating that the model is able to capture almost all areas with infiltration, with an overall *F1-score* of 73.92%.

Experiment 2

- **Average Accuracy**: 94.78%
- **Average Precision**: 62.38%
- **Average Sensitivity**: 99.07%
- **Average F1-Score**: 76.49%

In the second experiment, a slight improvement is observed both in terms of *accuracy* (94.78%) and *precision* (62.38%), with a slight reduction in *sensitivity* (99.07%). This indicates that the model reduced the number of false positives compared to the previous experiment, while still maintaining a high ability to capture true positives. The *F1-score* increased to 76.49%, demonstrating a better balance between precision and sensitivity compared to the first experiment.

Experiment 3

- **Average Accuracy**: 95.50%
- **Average Precision**: 69.03%
- **Average Sensitivity**: 98.79%
- **Average F1-Score**: 81.22%

The third experiment produced the best overall results, with an accuracy of 95.50%, a significantly higher precision (69.03%), and still a very high *sensitivity* (98.79%). This experiment further reduced false positives and improved the trade-off between precision and sensitivity, as evidenced by the F1-score of 81.22%, the highest among the three experiments.

The results show that the Random Forest model is capable of effectively distinguishing areas with infiltration from areas without infiltration.

3.5.7 Visual Test

To provide a visual example of what has been discussed so far, let's take an experiment from one of the three cases and compare the predicted data with the ground truth. In Figure 3.16, we see one of these comparisons. The white bins indicate *TP*, the green bins indicate *FN*, and the orange bins indicate *FP*. As we can see in this case, the RF model predicts the presence of infiltration very well, as the white bins are the majority and the largest in size. Indeed, we also see that there are some green bins of very small size. However, these bins, when computing Precision, count the same as the larger white bins, but to the human eye, it is clear that they do not represent a bin with infiltration. Overall, we can say that in most cases, the algorithm predicts the presence of infiltration well and makes mistakes in only a few instances.

At this point, we must ask ourselves what score we aim to achieve to consider ourselves satisfied. To answer this, we calculated the Confusion Matrix with three levels of statistics, as shown in Figure 3.17. We evaluated the statistics corresponding to one day, two days, and three days of data collection with a detector

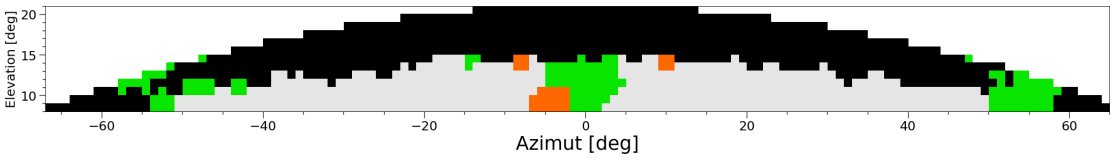


Figure 3.16: This represents a visual test comparing the model's predictions with the ground truth, with statistic equal at one month with a detector like BLEMAB. The white bins indicate *TP*, the green bins indicate *FP*, and the orange bins indicate *FN*.

four times the size of MIMA. For example, if Precision is critical for us, it is evident that one day of statistics is not sufficient, we have a value less then 70%, and is no longer our lower limit. On the other hand, with three days of statistics, the Precision would reach 79%. However, if the goal is simply not to miss any infiltration parts, the sensitivity value tells us that we already have 98% with just one day. If we want to consider both aspects, focusing on the F1-Score, it has a value of 82% with two days of statistics.

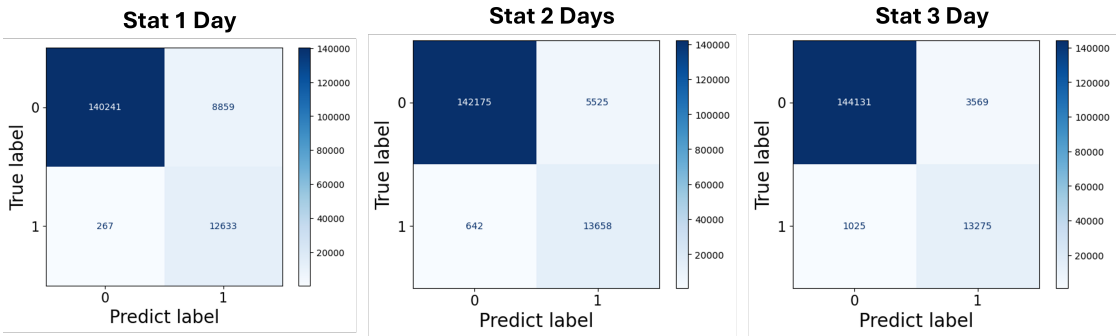


Figure 3.17: Three confusion matrices compared, each referring to a detector four times the size of MIMA. On the left, we find the one related to one day of data collection, in the center the one for two days of data collection, and on the right the one for three days of data collection.

However, it can be stated that the use of the RF algorithm has drastically helped in identifying the infiltrated area in a very short time. In fact, the combination of human expertise and the algorithm allows for the detection of infiltration within 1 or 2 days with a detector 4 time the size of MIMA. In contrast, simply visually observing the relative transmission map, even with the help of Voronoi tessellation, does not allow us to distinguish the infiltration with one day of statistics using a detector four times the size of MIMA.

3.6 Tailing Dams Conclusion

To conclude this chapter, we can affirm that the application of *Random Forest* classification for detecting infiltrations within the dam has shown promising results. One important factor that affects this technique is certainly the choice of the appropriate density for the dry and wet areas. This could yield higher expected counts given the same acquisition time.

However, this choice alone was not sufficient to visually identify the presence of an anomaly over a period of one or two days with a detector four times the size of MIMA. Therefore, the help of a classification model such as Random Forest was needed. The model proved capable of effectively distinguishing areas subject to infiltration from dry areas, as evidenced by the evaluation metrics, including *accuracy*, *precision*, *sensitivity*, and *F1-score*.

The Random Forest model has proven to be a robust tool for managing the binary classification problem, particularly due to its ability to balance overall accuracy with high sensitivity in detecting infiltrations. Additionally, the use of *Voronoi Tessellation* further improved the quality of the input data, reducing statistical noise and enhancing prediction accuracy.

Although the results are very promising, there is still room to improve the model's performance, especially in further reducing false positives without compromising the infiltration detection rate (sensitivity).

Therefore, we can conclude that adopting this approach in real cases could represent an important step forward in monitoring critical structures such as earthfill dams, providing a non-invasive and highly accurate tool for early detection of water infiltrations, and enabling timely interventions to prevent potential disasters.

Chapter 4

Use of a UNet algorithm for cavity identification and segmentation within mines

In muon radiography, anomaly identification is typically conducted through visual inspection of radiographic maps. This method, however, can sometimes lack objectivity, making it challenging to accurately identify or segment anomalies, particularly when they are near the instrument's resolution limits. Developing a systematic and objective approach free from human bias is therefore essential. The goal of this work is to develop a neural network capable of identifying cavities within muon radiographies of mining environments by classifying each bins inside the histogram, to determine whether or not it corresponds to an anomaly.

This chapter discusses the application of a UNet algorithm [47] for identifying cavities within mining structures, focusing specifically on measurements taken at the Temperino mine [10]. In this context, anomalies are represented by cavities within a rocky environment. Few studies have explored the application of machine learning, such as a UNet network, for anomaly segmentation in muon radiography, making this one of the first approaches in this field. However, this type of network is widely used in fields like medical imaging, as illustrated in [47].

Training neural networks like UNet for this purpose introduces several challenges that can affect both efficiency and accuracy. The primary challenge is the need for a large amount of labeled data to ensure robust model training, yet no labeled database of muographic images currently exists for geological or mining contexts. Another significant challenge is parameter optimization. Neural networks, particularly deep architectures such as UNet, involve numerous parameters that require tuning during training. Selecting appropriate hyperparameters, including the learning rate, network depth, and number of convolutional filters, is crucial, as these choices can greatly influence the model's performance.

Despite these challenges, implementing UNet for mining cavity segmentation is a promising approach to increase the reliability of muon radiography. Therefore, the primary aim of this study is to enhance muon radiography's reliability in identifying geological irregularities, such as cavities and concealed structures, through deep learning techniques.

4.1 The measure at Temperino mine

To evaluate the effectiveness of the developed UNet network, we applied it to a case study by the Florence muography group at the Temperino mine. This site is located within the San Silvestro Archaeo-Mining Park, in the rural area of Campiglia Marittima (LI), Tuscany. Mining at Temperino is believed to date back to Etruscan times, as evidenced by small, winding tunnels and vertical shafts. Further mining activity continued through medieval and post-medieval periods until the late 19th century, when operations ceased completely. Mining efforts focused primarily on skarn, a high-density silicate rock containing sulfides of Cu, Fe, Zn, Pb, and Ag [48].

The mine consists of numerous tunnels arranged across six levels, reaching depths of approximately 200 meters. Fig. 4.1 a) shows the drone-acquired point cloud of the Temperino Mine area, along with a 3D laser scanner reconstruction of the tourist access tunnel, which extends approximately 360 meters and is situated 40 meters below the surface. The image also includes a large cavity known as "Gran Cava", located 20 meters above the tunnel.

Since 2018, several data collection campaigns have been conducted within the Temperino Mine. This study examines results from three measurement locations, shown in fig. 4.1b) and labeled as position 1, 3, and 5. These numbers indicate the sequence of the measurements over time. The Florence muography group conducted these measurements with two main objectives: to validate muography's geological applications by locating and imaging known cavity and to investigate previously unmapped cavity and mineral deposits [10]. The detector was installed in the tourist tunnel, oriented vertically ($\theta = 0°$), to observe the Gran Cava quarry, the primary test cavity. This cavity has a length of approximately 40 meters, a width of 6-7 meters, and a height of 6-7 meters.

The images in fig. 4.1 b were obtained from point clouds acquired with a laser scanner and processed in the CloudCompare software [49]. The Gran Cava has two external entrances: the main entrance, located to the northwest and accessible via a well-worn road, and a second entrance, known as the "downfall zone", located to the southeast and of uncertain origin, possibly resulting from a collapse. In addition to these entrances, the Gran Cava is connected to the upper and lower mine levels by internal tunnels, which are now inaccessible.

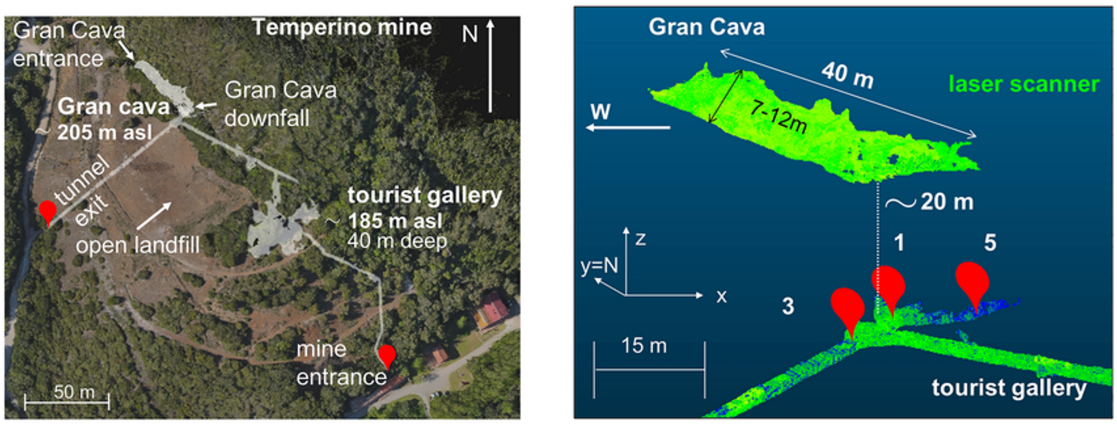


Figure 4.1: (a) the top view of the area of the Temperino mine (LI) in Tuscany obtained from a drone point cloud acquired for this study. The point cloud is displayed in CloudCompare [49] software. In transparency are shown the gallery opened to the public (located about 40 m deep) and the Gran Cava, a known cavity, 20 m higher than the tunnel. (b) localization of the measurement positions inside the tourist gallery. The figure shows the adopted three-dimensional Cartesian reference system with the positive direction of the y axis coinciding with the North direction [10] and the z axis pointing to the upper vertical direction.

4.2 Generation Dataset

Our goal is to train a UNet-type neural network to identify and segment cavities within mining environments. As previously mentioned, one of the most critical aspects in training such networks is the dataset used for training. Since no database of muographic images in mining environments currently exists, we created a custom dataset for this purpose. A key requirement for this dataset is that it should be as realistic as possible and capable of simulating all measurements within a reasonably short time frame, ideally around one week. To achieve this, we opted to modify the software currently in use by the Florence muography group, as described above in section 1.3.

4.2.1 Generation of measured transmission maps

Using CloudCompare software [49], we created 2422 artificial cavities, following realistic shapes from the Temperino mine or other mining environments. These cavities can be single or grouped together. The objective is to create transmission maps with these artificial cavities to closely replicate real measurements. We will therefore refer to these transmission maps as “measured transmission maps”,

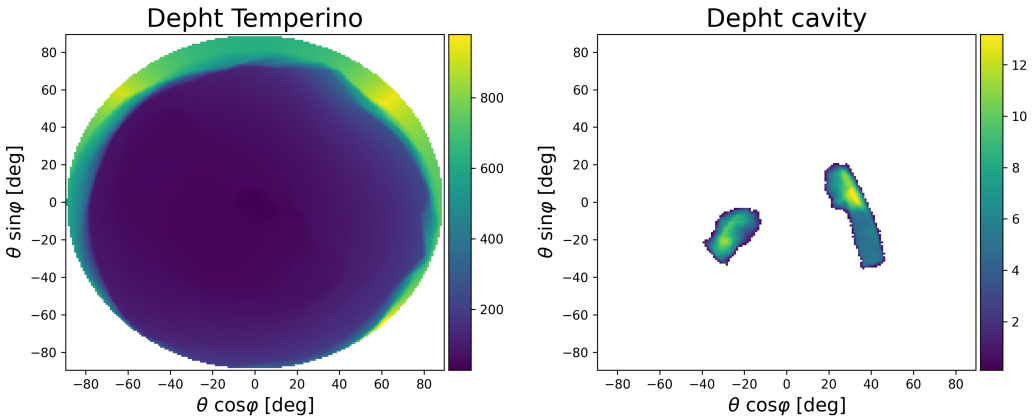


Figure 4.2: On the left is the depth map of the Temperino, while on the right is the depth map of one of the cavities used. Both maps are in polar coordinates. The color scale indicates depth in meters. At this stage of the process for creating the measured transmission maps, the depth of the cavity is subtracted from that of the Temperino mine.

which have been artificially generated. For this study, we chose to use maps in a polar configuration ($\theta \cos\varphi$ vs. $\theta \sin\varphi$). This approach is advantageous, as the detector's vertical orientation ($\theta = 0$) facilitates measurement interpretation. Additionally, this method allows us to obtain square images, which is beneficial for the subsequent development of the network.

The following subsection describes the procedure used to develop these maps. Fig. 4.2 illustrates the initial step, which involves calculating the depth of both the Temperino mine and the cavities. Additionally, to obtain a greater number of images, a process of data augmentation is employed. This involves shifting the cavity's position by a random azimuth angle (φ), effectively rotating it within the image. Each cavity is rotated to ten random positions. At this point, the cavity is added to the Temperino map by subtracting the cavity's depth from the depth of the Temperino mine. This process is shown in fig. 4.2. The depth map of the Temperino mine is displayed on the left, and that of one of the cavities used is displayed on the right. The color scale represents depth in meters.

Once this new depth map is obtained, we proceed to calculate the opacity. As described in section1.3, opacity is defined as the product of depth and density, $X = \rho \cdot l$. In our simulation software, density is fixed. To make the measurement as realistic as possible, we decided to modify this part of the program. Instead of multiplying the depth histogram by a constant, we multiplied it by the histogram describing the density maps of the Temperino extrapolated from measurements 3 and 5 [10]. This gives us the true opacity of the Temperino and cavity configu-

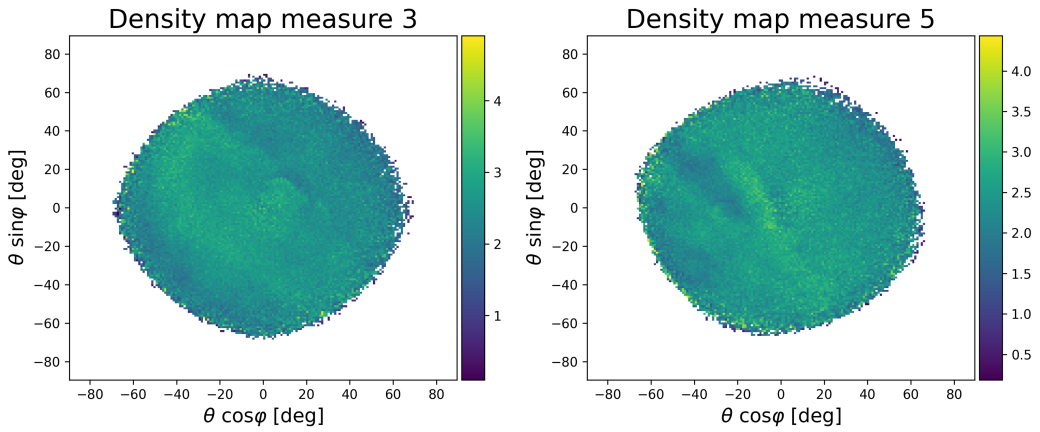


Figure 4.3: Density maps in polar coordinates obtained for measurements 3 and 5 conducted at the Temperino mine in polar configuration ($\theta \cos\varphi$ vs. $\theta \sin\varphi$) [10]. The color bar indicates density in g/cm^3.

ration. The density maps in polar configuration are shown in fig. 4.3. The color bar indicates density in g/cm^3. Since these density maps were obtained from real measurements, the instrumental error of the MIMA detector is already present. To further augment the data, both the density map of measurement 3 and that of measurement 5 were randomly rotated before being multiplied by the depth. The last step in this process is to convert the opacity map to a transmission map using the conversion tables, as described in section 1.3.4.

In summary, to create transmission maps we have generate 2422 cavities, each rotated to a random position 10 times and then subtracted from the Temperino mine depth. Each map was then multiplied by the density map from measurement 3 once, and from measurement 5 once, with a random rotation applied each time. Finally, the measured transmission maps were generated using the conversion tables.

4.2.2 Preprocessing of data

The measured transmission maps we obtained were created using conversion tables. To replicate a real measurement as closely as possible, we derived the expected counts of the simulated measurement and applied fluctuations according to the Poisson distribution.

As described in Section 1.10, the measured transmission is defined as the ratio between the flux in the target configuration and the flux in the open sky. If we consider the same detector and the same measurement time for acquiring the

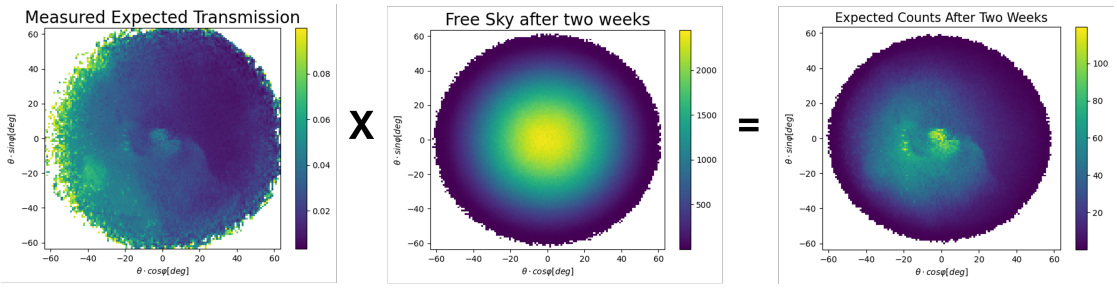


Figure 4.4: On the left is the measured transmission map extracted from the simulation software. In the center is the free-sky map after two weeks of data collection. On the right are the expected measured counts after two weeks.

fluxes, we can write the measured transmission as:

$$T_{\text{measured}} = \frac{N_{\text{measure}}}{N_{\text{freesky}}}. \tag{4.1}$$

We can use the inverse process to obtain the expected measured counts. Additionally, to fully assess the network's functionality, even with shorter measurement times, we conducted simulations with varying acquisition periods, namely 2 months, 1 month, and 2 weeks. Fig. 4.4 illustrates the methodology employed. The simulated measured transmission, displayed on the left, is multiplied by the free-sky counts (middle figure) at $\theta = 0°$ for a chosen acquisition time. In the figure, the free-sky counts are shown after two weeks of data collection. This product yields the expected counts for the measurement (right figure), specifically after two weeks in this example.

At this stage, it is essential to introduce the statistical fluctuations inherent to the measurement process. Each count within each bin was subjected to fluctuations according to the Poisson distribution, represented by the equation $P_\lambda(n) = \frac{\lambda^n e^{-\lambda}}{n!}$. Here, the variable λ represents the expected count, while n represents collected the number of events. At this point, we obtain a realistic measurement, as both the instrumental errors of the MIMA instrument and the statistical fluctuations are present.

We then recalculated the measured transmission and divided it by the expected transmission of the Temperino with a constant density, thus obtaining a relative transmission map. As with a real measurement, the transmission map of the Temperino (without added cavities) was obtained with a constant density, specifically the standard rock density of 2.65 g/cm^3.

Each relative transmission map is paired with a corresponding truth map, representing the cavity's position and shape. Fig. 4.5 illustrates an example of an image provided to the network for training. The accompanying truth map depicts

the shape and position of the cavity. The image representing the measurement can be interpreted as a grayscale image, with only one channel. A color image has three channels: one for red, one for blue, and one for green, while a grayscale image has only one channel. Each image thus has a size of (128,128,1), where the color scale represents the relative transmission value. When this value exceeds one, it indicates an area with an average density less than the expected density (2.65 g/cm^3). The first two values in the size represent the pixel resolution of the image, where the pixel size is (1x1)deg^2. Conversely, areas with values less than one represent regions with a higher average density than expected. The truth map, in contrast, is a binary image with only two possible values: 0 and 1. A value of 1 (white) indicates the presence of a cavity, while a value of 0 (black) indicates its absence. The goal is to have an accurate segmentation in order to determine, pixel by pixel, whether a cavity is present.

Given the difficulty of accurately generating cavities within the MIMA acceptance cone, we opted to assign a value of 0 to all bins corresponding to directions forming an angle greater than 55° with respect to the central axis of the detector's field of view. This is because the simulation software treats MIMA as a point-like detector, effectively giving it infinite acceptance. However, from field measurements and more precise simulation software, such as Geant4, we know that its acceptance is approximately 60°. These simulated cavities do indeed exist, but they are located in a region outside MIMA's acceptance cone and, consequently, would not be detected by the detector. This strategy is highly beneficial for network training, as it helps the network learn to predict empty outcomes.

4.3 Neural Networks

Machine learning (ML) is among the most innovative and rapidly expanding areas of computer science and artificial intelligence (AI). This discipline aims to enable computers to learn from data without explicit programming for specific tasks. The primary goal of ML is to develop algorithms and models capable of identifying patterns, making predictions, and autonomously making decisions based on large datasets. This approach is particularly useful across various fields, including medicine, business, and robotics. In the specific case of this work, ML is applied for anomaly detection in muon radiography.

Artificial neural networks (ANNs), a fundamental aspect of machine learning, are loosely modeled after the functioning of the human brain. As shown in fig. 4.6, these networks are composed of interconnected units called neurons that collaboratively process information. The neurons are arranged into layers: an input layer that captures raw data, one or more hidden layers where computations take place, and an output layer that delivers the final outcome. Connections between neurons

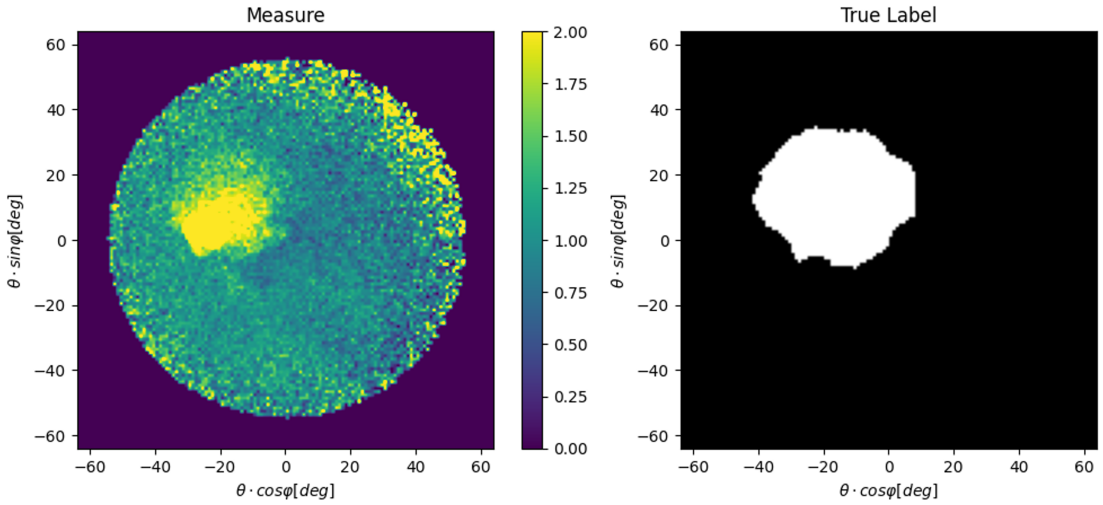


Figure 4.5: Left: the relative transmission map used for network training. Right: the truth map with discrete values of 0 and 1, showing the cavity's position and shape in white.

have weights, which are adjusted during the training process to determine how strongly one neuron influences another. In recent years, a subset of AI known as Deep Learning (DL), which utilizes neural networks with multiple hidden layers, has become increasingly important.

Neural networks are particularly powerful due to their ability to learn and adapt dynamically. By receiving feedback on the accuracy of their outputs, they can modify and refine their internal parameters, effectively "learning" from past experiences to improve over time (fig. 4.7). Many machine learning models leverage neural network architectures to encode knowledge in the form of weights distributed across input, hidden, and output layers. By continuously adjusting these weights, the network becomes increasingly skilled at identifying complex patterns, such as those found in images, audio, or text. This capability makes neural networks a cornerstone of modern deep learning, a highly specialized area within machine learning.

The ability to learn directly from data without explicit human intervention makes these technologies promising, paving the way for a future where automation and AI are increasingly integral to daily life and global industries.

4.3.1 Introduction to image recognition

Image recognition is a key area within machine learning, aimed at enabling computers to "see" and interpret images similarly to the human eye. The appli-

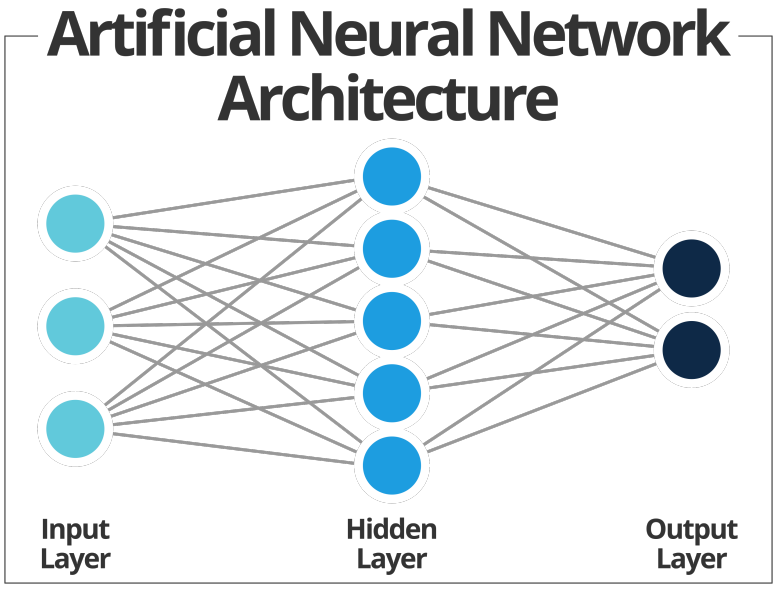


Figure 4.6: Architecture of a simple "feed forward" neural network (NN), illustrating the input, hidden, and output layers. A NN may contain multiple hidden layers.

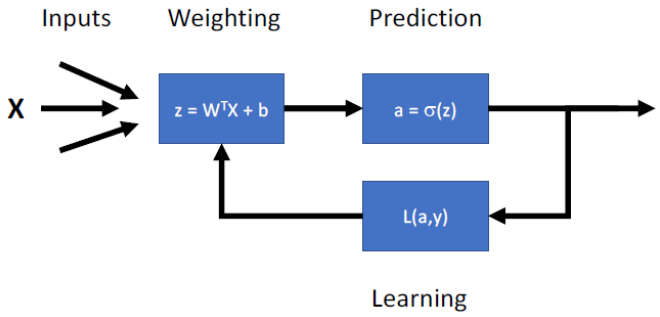


Figure 4.7: Architecture of a single neuron in a neural network. Inputs (X) are weighted (W) and a bias (b) is added, before passing the output (z) through an activation function (σ) to produce a prediction (a). The result is then compared to a known target (y) to compute a Loss Function (L), which quantifies prediction quality. When L is large, weights adjust significantly for the next iteration. As learning progresses, L diminishes, and the weights approach optimal values for a given dataset.

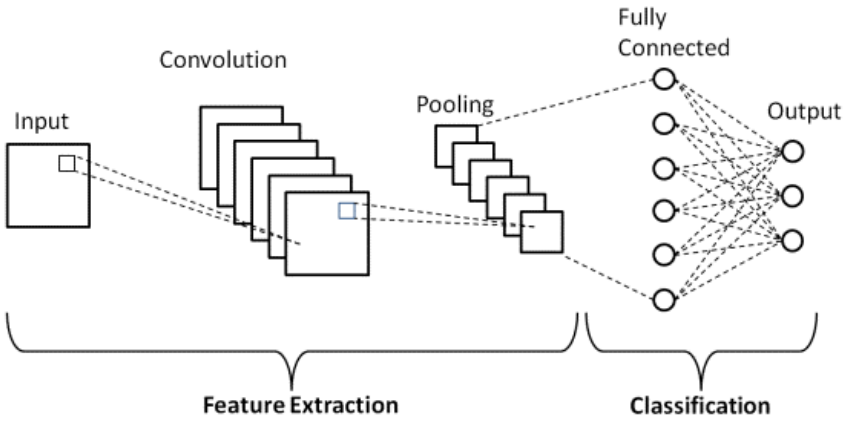


Figure 4.8: Architecture of a Convolutional Neural Network (CNN), illustrating key features: convolution layers and pooling layers for feature extraction, and final classification layers.

cations of this discipline are extremely broad, spanning facial recognition, object classification in photos and videos, and automatic diagnosis in the medical field through radiological image analysis. Recent advances in artificial neural networks, particularly *Convolutional Neural Networks* (CNNs), have driven rapid progress in image recognition techniques [50].

CNNs are specifically designed neural network architectures optimized for image processing and recognition. Traditional neural networks, or feedforward networks, treat each image pixel as an independent input, disregarding spatial relationships between neighboring pixels. While effective in some contexts, this approach is not ideal for handling complex images. In contrast, CNNs exploit the hierarchical structure of images, where simpler patterns (e.g., edges or shapes) are detected in the first layers of the network, while more complex patterns (e.g., faces or objects) emerge in the deeper layers.

A CNN comprises three main types of layers: convolutional layers, pooling layers, and fully connected layers, as shown in fig. 4.8. The convolutional layers are the core of the CNN, responsible for extracting the key features of an image. These layers apply a series of filters, or kernels, that move over the image to detect details such as edges, corners, and textures. Each filter generates a feature map representing a particular aspect of the image [51].

Pooling layers reduce the spatial dimensions of feature maps, condensing information while retaining only the most relevant details. This operation decreases the number of parameters to manage, making the model more efficient and helping to avoid overfitting. Finally, the fully connected layers process the extracted features and connect all the neurons, allowing the model to make final predictions, such as

in image classification.

One of the major advantages of CNNs is their capacity for abstraction and generalization. In the network's initial layers, CNNs identify low-level features, such as contours or textures; as the network progresses to deeper layers, it can recognize more complex shapes, even to the point of identifying specific objects. For instance, in a network trained to recognize human faces, the first layers may detect lines or edges, while the final layers recognize facial features such as eyes, nose, mouth, and their arrangement. For a more detailed explanation of CNNs, see ref. [52].

4.3.2 UNet

Our objective is not just to classify an entire image as belonging to a specific category. Instead, we aim to go further by identifying both the shape and the exact location of each cavity within muon radiographies. This process is known as *image segmentation*. Unlike simple classification, which assigns a single label to the whole image, image segmentation involves analyzing the image at a much finer level. Specifically, it assigns a category to every single pixel in the image. This approach, called *semantic segmentation*, allows us to determine which region each pixel belongs to, providing a detailed and precise understanding of the image's content. This high level of detail is crucial for accurately detecting and mapping cavities.

In this context, a standard CNN architecture, while effective for general feature extraction, may lack the spatial accuracy required for segmentation tasks. Unlike in a classification problem, where the primary goal is to recognize the main object in an image, segmentation demands preserving image resolution to detect even the smallest details. This is why *UNet*, a neural network architecture specifically designed for segmentation, is often preferred.

UNet, introduced in 2015 for biomedical image processing [47], features a symmetric "U-shaped" structure with two main paths:

1. Contraction Path (Encoder): This path, similar to a traditional CNN, is responsible for feature extraction using convolutional and pooling layers. The image progressively reduces in spatial size while relevant features are extracted. However, this process can lead to some loss of spatial information, crucial for accurate segmentation.

2. Expansion Path (Decoder): The decoder is tasked with recovering the spatial resolution of the image. Using upsampling operations and transposed convolutions, the network reconstructs the image, producing a segmented map at the original resolution. A unique feature of UNet is the *skip connections* linking corresponding layers of the encoder and decoder paths. These

connections transfer low-level information (e.g., edges or contours) directly from the contraction path to the expansion path, significantly enhancing segmentation accuracy.

A schematic structure of our neural network is shown in fig. 4.9. We used the Keras library to build our model [53]. The model consists of two main sections: the encoder and the decoder. The encoder (blue layers) extracts the primary features of the input image, starting from an input layer that receives tensor images, followed by a series of convolutional blocks. Each block applies a two-dimensional convolution, followed by a *Rectified Linear Unit* (ReLU) activation function. The number of filters doubles with each block (starting at 64) to capture increasingly complex patterns as depth increases. At the encoder's end is the neck layer (yellow), where *pooling* further reduces spatial dimension, facilitating deeper feature abstraction.

The decoder section (green layers) reconstructs the segmented image from extracted features. Through a series of transposed convolution blocks, the original image size is restored. Skip connections enable each decoder block to receive input from the corresponding encoder activations, preserving detailed spatial information and enhancing segmentation quality. The feature maps from the encoder (local details and low-scale information) are concatenated or combined with those from the decoder (global and high-level information). This concatenation creates a "mix" of information at different levels of abstraction, which helps the decoder reconstruct fine details in the image. Finally, the decoder output passes through an additional transposed convolution layer to produce the final segmented map, with a sigmoid activation function that helps distinguish between segmentation classes by returning values between 0 and 1. To prevent overfitting and improve generalization, we implemented several regularization techniques, including L2 regularization [53] to reduce model complexity and a SpatialDropout2D [53] layer that randomly excludes certain activations during training. This prevents the model from becoming overly dependent on specific filters or patterns.

Before proceeding, we provide a brief explanation of the activation functions used. For the hidden layers, we chose the $ReLU$ activation function, while for the decoder output we used *sigmoid*. These functions are represented by the following formulas:

$$\begin{aligned} f(x)_{RelU} &= \max(0, x), \\ f(x)_{sigmoid} &= \frac{1}{1 + e^{-x}}. \end{aligned} \tag{4.2}$$

ReLU is commonly used in hidden layers as it promotes faster learning and better feature representation. In addition, the use of the ReLU activation function

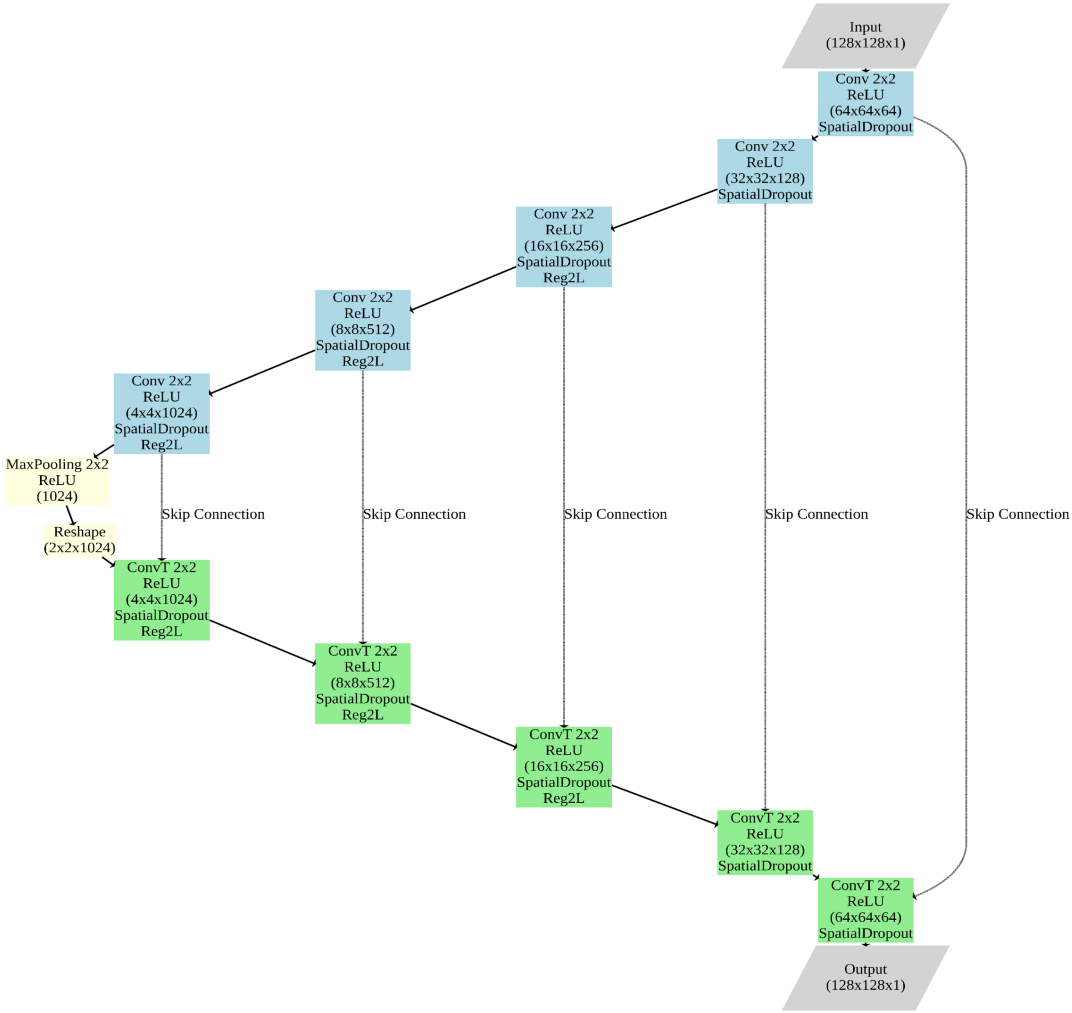


Figure 4.9: In blue, the encoder schematic; in yellow, the neck layer: in green, the decoder. The network takes a 128x128 one-channel image as input, reducing its size by half with each convolution step until reaching a vector in the neck layer, where pooling flattens and transforms the image into a vector. In the decoding phase, the image is reconstructed via transposed convolutions back to a 128x128 one-channel image. Encoder and decoder layers are linked by skip connections.

in the hidden layers of neural networks is motivated by its ability to mitigate the vanishing gradient problem, enabling models to learn more quickly and effectively [54]. Sigmoid, on the other hand, is useful for binary classification in neural network outputs, as it allows the output to be interpreted directly as probabilities [55]. The goal is to produce a probability map for each pixel that indicates the likelihood of that pixel belonging to a specific class (e.g., cavity vs. background). Because sigmoid outputs values between 0 and 1, it is well-suited for this purpose.

The strength of UNet lies in its ability to combine information at different detail levels. The contraction path captures global features and broader contexts, while the expansion path, enhanced by skip connections, recovers fine details lost during contraction. This makes UNet highly suitable for segmentation tasks, balancing global context with the precise localization of objects or regions within the image.

Compared to traditional CNNs, which tend to lose spatial detail through pooling and downsampling operations, UNet preserves high spatial resolution due to skip connections. This is particularly useful in our case, where accurately detecting the presence and shape of anomalies in muon radiographies involves classifying each pixel. Specifically, we aim to determine if each pixel represents part of a cavity.

UNet's efficiency is another advantage: despite being a powerful model, it is relatively lightweight compared to deeper architectures and can be trained with a limited number of labeled images. This is crucial in our case, where there is no large database of muon radiographies containing cavities in a mining context. As mentioned, we generated these images artificially, which was time-intensive, as each cavity was manually created using CloudCompare [49] software.

In summary, for segmentation tasks, UNet has significant advantages over traditional CNNs. Its architecture effectively combines global and local information while preserving the spatial resolution needed to segment fine image details. This balanced structure, with both contraction and expansion paths connected by skip connections, makes it particularly effective for applications requiring pixel-level accuracy.

4.3.3 Metrics and Loss

Before analyzing the results of training this neural network, it is useful to explain the type of metrics and loss used during training. First, let us clarify the concepts of loss and metrics. The *loss* function is a numerical indicator that measures how much the model deviates from its training objectives. The goal in training a neural network is to minimize this loss function, as a lower loss value indicates improved model predictions. On the other hand, *metrics* are quantitative measures used to assess the model's performance but do not directly influence weight updates during training. While the loss function guides the optimization

process, metrics offer a more intuitive understanding of performance, particularly when evaluating the model on validation or test data.

For image segmentation, the *Dice Similarity Coefficient* (DSC) and the *Intersection over Union* (IoU) are widely used metrics [56].

Fig. 4.10 shows the metrics we used. Intersection over Union is defined as the ratio between the intersection and the union of two regions (prediction and ground truth), mathematically expressed as:

$$IoU = \frac{A_{pred} \cap A_{true}}{A_{pred} \cup A_{true}}. \tag{4.3}$$

The IoU returns a value between 0 and 1, where 0 means no overlap and 1 indicates perfect overlap. IoU is stricter than the DSC in penalizing discrepancies between the prediction and ground truth; even small errors lead to a significantly reduced IoU.

The Dice Similarity Coefficient measures the similarity between two sets and is defined as the ratio between twice the intersection of the two regions (prediction and ground truth) and their disjoint union:

$$DSC = \frac{2 \times A_{pred} \cap A_{true}}{A_{pred} + A_{true}}. \tag{4.4}$$

Like the IoU, DSC also yields a value between 0 and 1, with 1 indicating a perfect match. The DSC is typically less strict than IoU with small differences between regions, as it places more emphasis on the intersection, making it slightly more forgiving of minor discrepancies. For our case A corresponds to individual pixels of the image.

The IoU metric tends to be more stringent in penalizing small errors, while the DSC is more lenient. This difference arises because IoU considers the total union of the areas, while DSC normalizes the outcome, balancing the weight between prediction and ground truth. For the same overlap level, DSC generally yields higher values than IoU, especially in cases of partial overlap. DSC is commonly applied in medical image segmentation or tasks where minor errors are tolerable.

In our case, we chose Dice Loss = 1−DSC as the loss function. This choice is motivated by DSC's ability to weigh the area of correspondence more heavily, making it suitable when predicted and ground truth regions differ significantly in size. Dice Loss is particularly effective in cases of class imbalance, as is typical in our dataset. Considering all possible cases, there may also be situations with relatively small cavities compared to the entire investigated area. In these cases, the Dice coefficient has even more advantages.

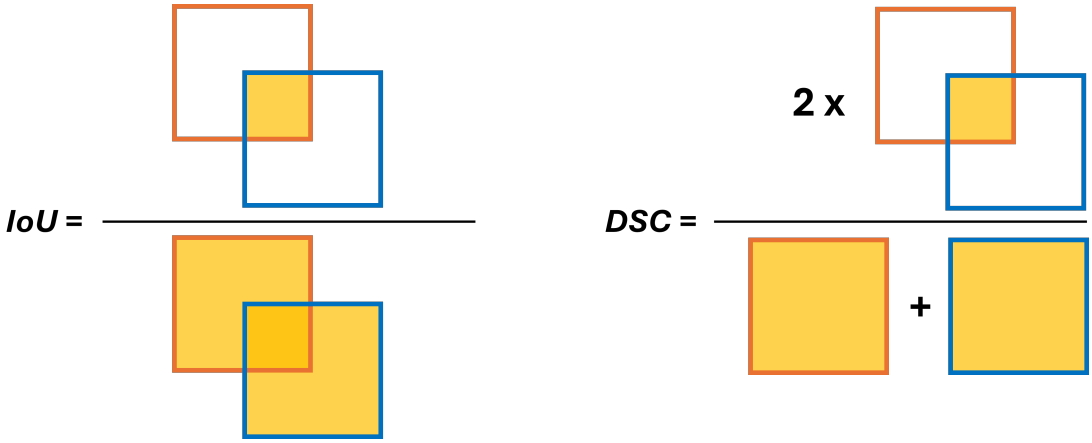


Figure 4.10: Left: the Intersection over Union metric. Right: the Dice Similarity Coefficient. IoU is defined as the ratio of the intersection of the truth map and model prediction to the union of the two, while the *Dice Similarity Coefficient* is the ratio of twice the intersection of the truth map and model prediction to their disjoint union.

4.4 First UNet Results

In this section, we present the results of training the UNet model for the identification and segmentation of cavity within muon radiographies. Below, we describe the model configuration and the rationale behind the methodological choices adopted to optimize segmentation performance and reduce the risk of overfitting. Before showing the results, we revisit our hyperparameter selection and the reasoning that led us to choose them.

The model architecture follows the classic encoder-decoder structure of UNet. The encoder consists of 5 convolutional blocks, which progressively reduce the image shape while extracting its salient features. The decoder restores the original shape using transpose convolution layers and employs skip connections to retain details lost during the encoding phase.

The implemented UNet model was designed with 64 initial filters in the first convolutional block and a dropout rate of 0.1, aiming to balance the model's learning capacity and the prevention of overfitting. L2 regularization, with a coefficient of 0.000001, was included in the last convolutional blocks to reduce the impact of excessive weights, promoting more stable and generalizable learning. This approach is particularly important given the limited number of segmented cavity images.

To prevent saturation, the ReLU activation function is used in the intermediate layers of the model, facilitating more efficient learning and enabling the model to

represent complex image features. The final output, however, uses a sigmoid activation function to produce a binary segmentation (cavity vs. background), with values between 0 and 1 that can be interpreted as class membership probabilities.

The Dice Loss function was chosen as it is particularly suitable for imbalanced datasets like the one under analysis, where cavity represent a small portion of the image. In addition to Dice Loss, the model was also evaluated using the DSC and the IoU metrics, that allow for precise overlap measurement (IoU) as well as structural similarity (DSC) between the model prediction and the ground truth.

4.4.1 Performance Visualization

This preliminary analysis of model settings and training techniques provides the necessary context to understand the results presented below, illustrating the model's behavior and its ability to segment cavity in muon radiographies. The following sections discuss the results of training the UNet model for cavity segmentation with data collected over different time periods: 2 months, 1 month, and 2 weeks. As shown in the learning curve plots in figure 4.11, the loss function was monitored for both the training and validation sets to evaluate the model's behavior and its generalization ability.

The plots in fig. 4.11 clearly show the trend of loss as a function of training epochs. Each epoch represents a complete cycle in which the entire training dataset is passed through the network. The loss decreases rapidly in the early stages of training, for both the training and validation sets, with a progressive slowdown as the model converges toward an optimal solution. Starting from 2 months, we observe that the curve for the training set loss shows a rapid decrease within the first 10 epochs, stabilizing at very low values. This indicates that the model effectively captures the dataset's main features. The mean of validation loss curve is 0.0789, indicating good model generalization on validation data without signs of overfitting. The model demonstrated excellent learning and generalization capabilities with this longer temporal dataset. The relatively low loss value suggests that the model captured cavity features effectively. For the one month, the trend of training loss curve is very similar to that of the 2-month dataset, with a rapid descent and stabilization around the tenth epoch. The mean validation loss value is slightly higher than in the 2-month case, reaching 0.0926. However, the training and validation curves remain close, suggesting a good balance between learning and generalization. As in the previous cases, for the 2-week dataset, the loss decreases rapidly in the early epochs and stabilizes afterward, with a value close to that of the longer sets. In this case, the validation loss is the highest among the three datasets, with an average value of 0.1096. This difference may be attributed to the shorter acquisition period, which provides the model with less information to effectively learn cavity features. Although the model still achieves

good results, the higher validation loss suggests that the model could benefit from an extended acquisition period or more data. The loss difference compared to longer sets may be related to lower statistical significance in the bins, resulting in greater statistical fluctuations.

Analyzing the three temporal scenarios, a clear correlation emerges between the data acquisition period duration and the model's performance in terms of loss. The average loss of the validation set increases as the temporal period decreases, from 0.0789 for 2 months to 0.1096 for 2 weeks. This trend suggests that a larger amount of data enables the model to generalize better and learn cavity features more effectively. Although the training set shows similar learning curves across the different temporal scenarios, the validation curves suggest that longer acquisition periods lead to a reduction in the discrepancy between training and validation loss, signaling greater generalization capability. It is noteworthy that no evident signs of overfitting are observed in the three cases analyzed, indicating that the model maintained a good balance between learning and validation, even with less data. In all three cases, model saturation is observed, indicating that further improvements are not achievable, as the training curve improves, but the validation curve remains constant. This suggests that the model has reached its maximum potential.

4.4.2 Test Networks

The results obtained with the IoU metric 4.12 are presented here, representing the degree of overlap between the model's prediction and the ground truth for cavity segmentation. The IoU compares bin-by-bin the cavity in the ground truth map with the model's prediction. In the three graphs shown in figure (2 months, 1 month, 2 weeks), we can observe the IoU value distribution for 686 test images in each case. We recall tha the IoU represents the fraction of pixel overlap between prediction and truth normalized to their sum. Therefore the correct 0 or 1 pixel predicted over the entire available pixels. An IoU score is calculated for each test image, and the results of each event are then displayed in the histogram shown in the figure. The IoU distribution is not uniform: there is a pronounced peak around 0, indicating that a significant portion of the images has an IoU value close to zero. This means that, in these images, the model's prediction does not significantly overlap with the ground truth. However, in all three periods, there is also a substantial number of images with IoU values above 0.6, indicating highly accurate predictions for these images.

For the 2 months the average IoU is 0.61, representing the model's best performance among the three acquisition periods. This suggests that with a larger amount of acquisition data, the model can segment cavity more accurately. For 1 month the average IoU decreases to 0.59, with similar to the result that come from the 2-month period. A good number of images still have IoU values above

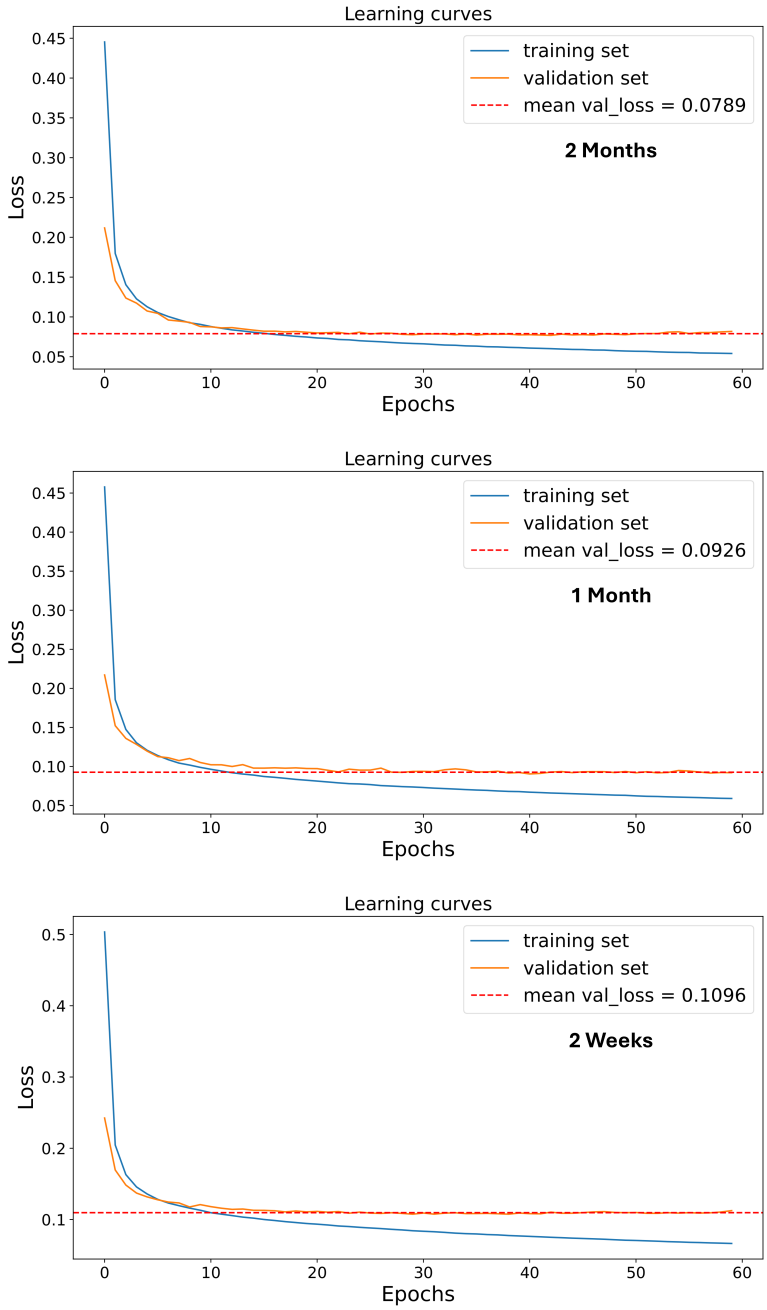


Figure 4.11: Loss function trend for the training and validation sets over the three acquisition periods: 2 months, 1 month, and 2 weeks. A rapid reduction in loss is observed in the first training epochs across all cases, followed by stabilization. The average validation set loss slightly increases as the acquisition period decreases, indicating that a greater amount of temporal data improves the model's generalization capability.

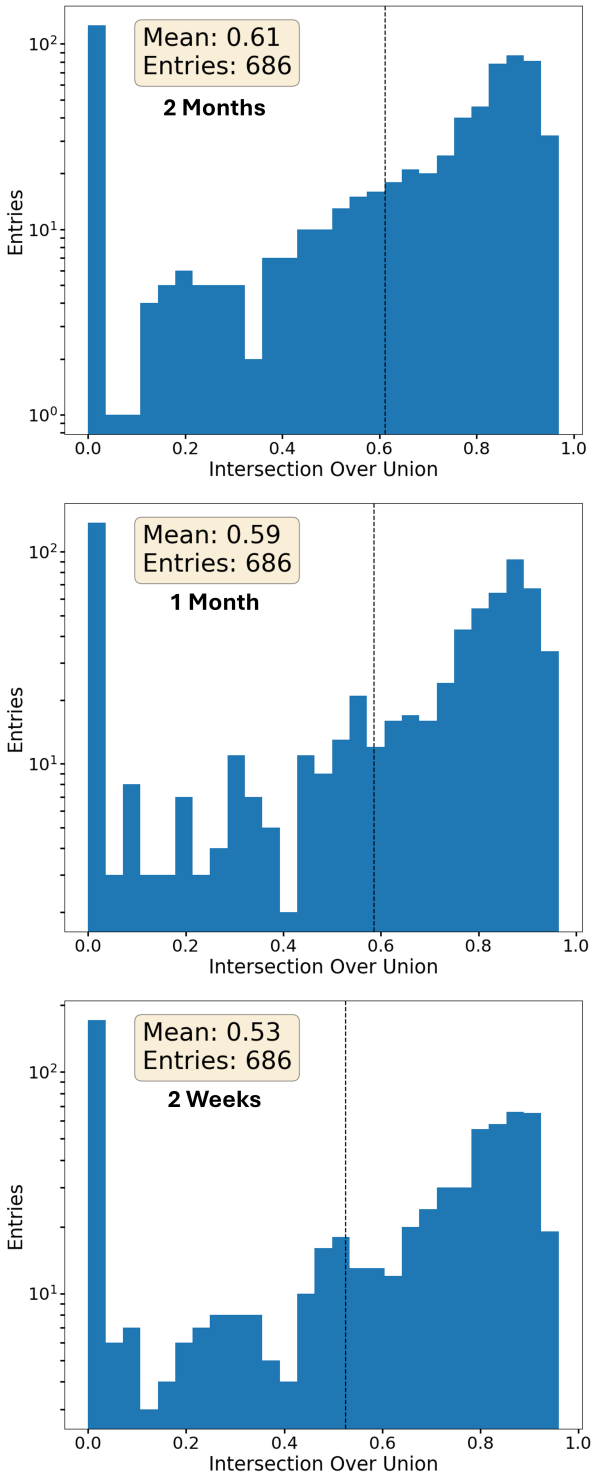


Figure 4.12: Distribution of the IoU metric across test sets with acquisition periods of 2 months, 1 month, and 2 weeks. The y-axis is on a logarithmic scale. The average IoU decreases with shorter acquisition periods: 0.61 for 2 months, 0.59 for 1 month, and 0.53 for 2 weeks. In each case, 686 test images not seen during model training were used. A significant number of images with IoU close to zero are observed, indicating model difficulties in correctly segmenting cavity in some cases.

0.6. In the end for 2 weeks the average IoU further decreases to 0.53, indicating weaker performance. The distribution shows more IoU values near zero compared to other periods, suggesting that with a shorter acquisition period, the model has less data to correctly learn cavity segmentation.

A significant number of images exhibit IoU values below 0.1, as shown by the pronounced peak at zero in the histogram of fig. 4.12. The different time periods in the figure correspond to simulated acquisition times, used to evaluate the effect of data accumulation on the model's segmentation performance. The mean IoU increases from 0.53 for a two-week equivalent acquisition, to 0.59 for one month, and reaches 0.61 for two months, suggesting that a larger dataset helps improve segmentation accuracy. However, despite this improvement, a non-negligible fraction of images still presents very low IoU values, which can be attributed to the difficulty of detecting small cavities, increased noise in certain regions, and the presence of cavities located near the MIMA acceptance boundary. Fig. 4.13 illustrates two representative cases. The top panel shows a well-segmented cavity with an IoU greater than 80%, corresponding to the right peak of the histogram in fig. 4.12. In contrast, the bottom panel represents a failure case, contributing to the zero-IoU peak in the histogram. In this instance, the cavity is positioned at the edge of the acceptance cone, has a small size, and is located in a high-noise region. The model struggles to distinguish such structures, leading to a complete miss in the predicted segmentation.

The model's decreasing performance with shorter acquisition periods (from 2 months to 2 weeks) is an expected result, as fewer data points provide reduced information for network training, limiting the model's ability to generalize. Longer acquisition periods allow the model to "see" more variations in the images and better capture cavity features, thereby improving segmentation accuracy.

Following this first promising test, fine-tuning was performed to find the best hyperparameters for our model using *Bayesian Optimization*.

4.5 Bayesian Optimization for Hyperparameter Optimization

Hyperparameter optimization is a crucial aspect in neural network training, as the correct selection of hyperparameters can have a significant impact on model performance. Among various methods used for hyperparameter optimization, *Bayesian Optimization* stands out for its efficiency in dealing with complex and expensive-to-evaluate problems, where the goal is to reduce the number of evaluations needed [57].

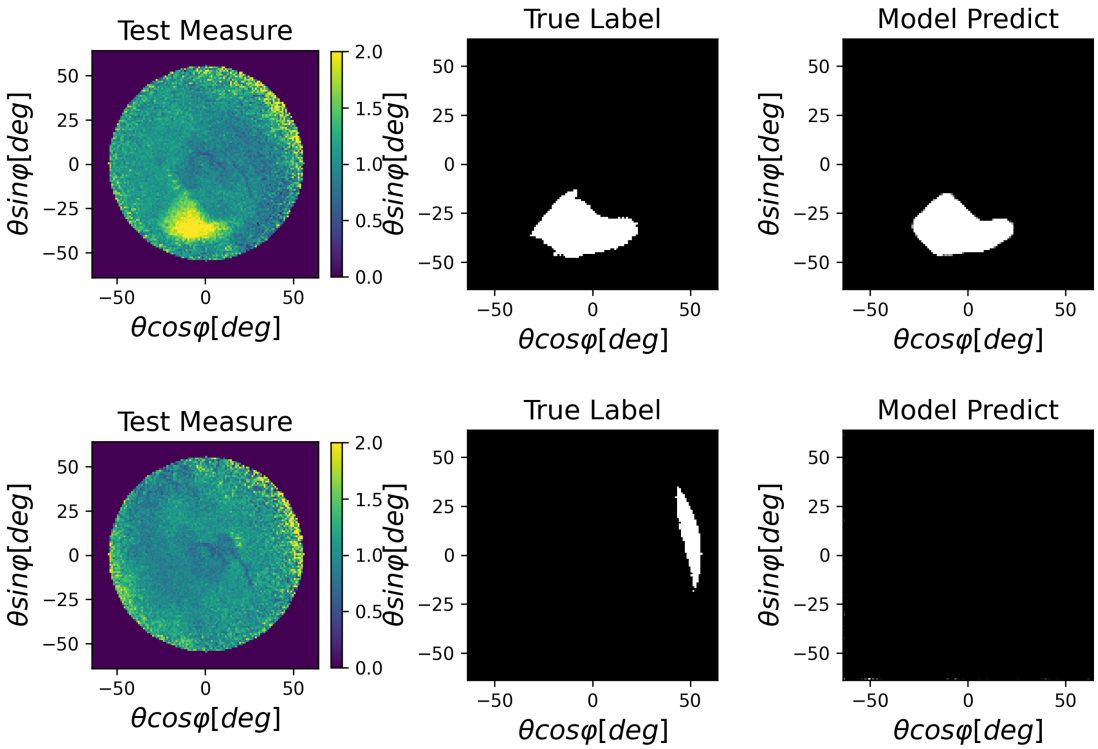


Figure 4.13: Two examples of test measurements, with 2 months of data acquisition simulated, are shown (on the left) along with the corresponding truth table (in the center) and the network prediction (on the right). At the top, a case where the network correctly segments the cavity is shown, while at the bottom, a case where the network fails to predict the cavity at all is displayed.

4.5.1 Theory of Bayesian Optimization

Bayesian Optimization is an iterative technique that aims to find the minimum of a costly objective function, as for the example te loss function. Instead of exploring the hyperparameter space randomly, as in *Grid Search* or *Random Search*, Bayesian Optimization builds a probabilistic model of the objective function behavior and uses it to decide where to sample next. This approach balances exploration of new regions and exploitation of known information [57].

This technique relies on two main components:

- **Surrogate Model**: A probabilistic model (e.g., a *Gaussian Process* [58] or a *Tree-Structured Parzen Estimator* [59]) is used to approximate the objective function. This model acts as a surrogate for the actual function, which is costly to evaluate.
- **Acquisition Function**: The acquisition function decides where to sample next, balancing exploration (testing new areas of the hyperparameter space) and exploitation (sampling near areas that already show good results). Common acquisition functions include:
  - *Expected Improvement (EI)*: maximizes the expected improvement over the best observed result.
  - *Probability of Improvement (PI)*: maximizes the probability that a future evaluation will improve the current result.
  - *Upper Confidence Bound (UCB)*: balances uncertainty and improvement in less explored areas.

For a more detailed description of *Acquisition Functions*, refer to [60].

4.5.2 Bayesian Optimization Process

The process proceeds in successive iterations:

1. An initial surrogate model is constructed based on a few initial evaluations of the objective function.
2. The acquisition function is used to select the next point to sample.
3. The objective function is evaluated at the new point (in our case, a new neural network training is performed with a specific set of hyperparameters).
4. The surrogate model is updated with the new sample, and the cycle is repeated until a stopping criterion is met (e.g., a maximum number of iterations).

A graphical example of the bayesian optimization process is shown in fig. 4.14.

This method is particularly useful when the objective function is costly to evaluate, as in the case of training deep learning models on large datasets, where each evaluation involves a lengthy training process.

4.5.3 Application to the studied case

For our work, the objective function calculated the average of the last 20 validation losses after training the model for which we want to find the best hyperparameters. The extracted value serves to choose the next hyperparameter values. In this work, we used *Optuna* [61], an open-source library for automatic hyperparameter optimization, which implements Bayesian Optimization. Optuna allows efficient exploration of the hyperparameter space, selecting the most promising configurations based on probabilistic models. Thanks to its flexibility, Optuna supports various sampling techniques, such as random search and Bayesian Optimization, to identify the best set of hyperparameters. It also offers advanced *pruning* functionalities, which allow early stopping of non-promising trials, saving computational resources. These features make Optuna an ideal tool for optimizing machine learning models, such as deep neural networks, minimizing the time needed to find optimal hyperparameters.

Specifically, the following UNet model hyperparameters were optimized: *Learning Rate*, *Initial Filter Number*, *Dropout Rate*, and *L2 Regularization Value*.

A detailed description of the intervals and reasoning behind each optimized hyperparameter is provided below.

The learning rate was optimized within a range between 10^{-6} and 10^{-4}. Optuna explored values on a logarithmic scale within this range. This scaling is crucial for the learning rate because small changes in its value can have a significant impact on model performance. The logarithmic scale enables the algorithm to explore values at different orders of magnitude, such as 10^{-6}, 10^{-5}, and 10^{-4}, without having to evaluate every possible intermediate value. Once an optimal range is identified, finer-grained testing can be conducted within this range to determine the best value. Choosing an appropriate learning rate is critical for effective training; if the value is too high, the model may diverge, while if it is too low, the training process will be slow and may fail to reach a minimum.

The initial filter number, which determines the number of filters in the first layers of the network, was optimized in the range of 32 to 128, with increments of 1. Unlike the learning rate, a linear distribution was used to explore this range, allowing Optuna to evaluate integer values uniformly. This parameter is important because a higher number of filters enables the network to capture more features, but it also increases the model's complexity. The chosen range is standard for networks dealing with problems of medium complexity, providing a balance between

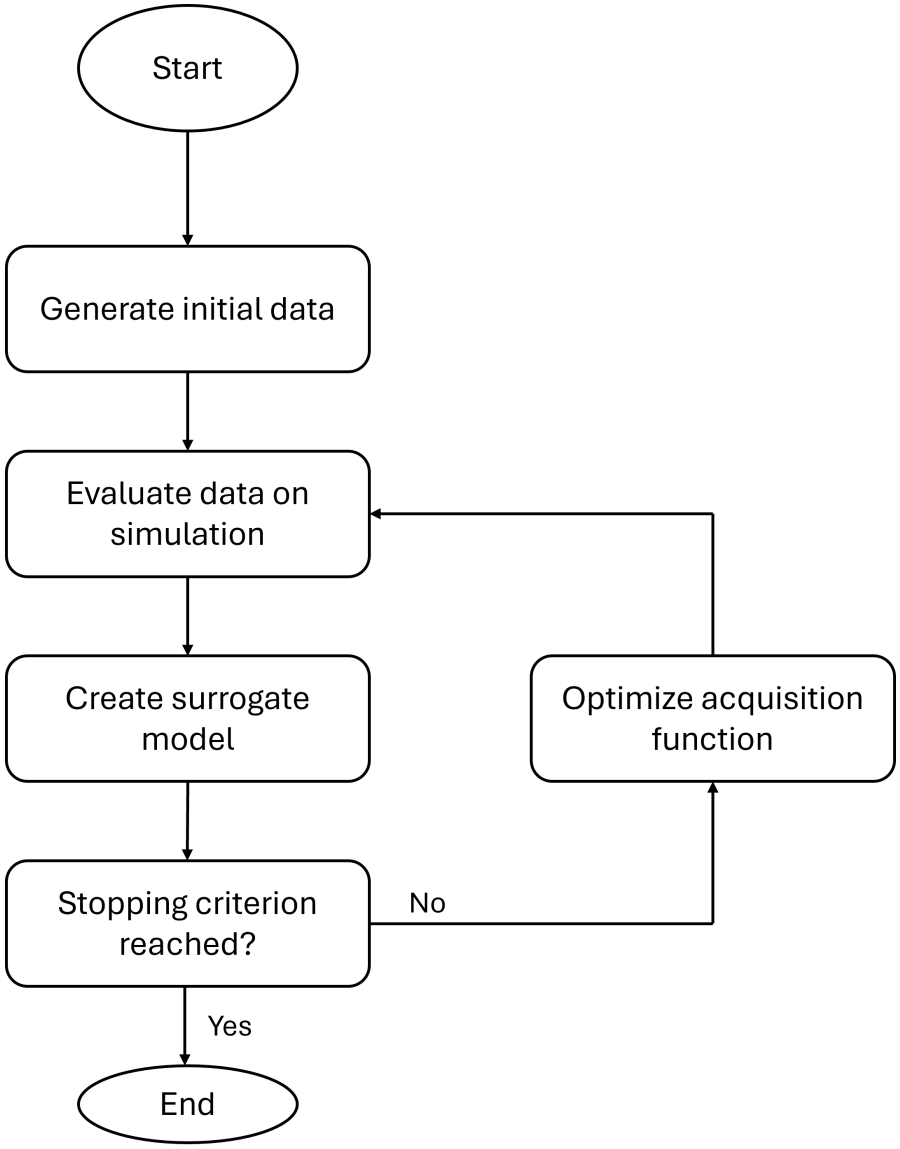


Figure 4.14: Diagram of the Bayesian Optimization iterative process. A surrogate model is built, the acquisition function is used to select the next point, the objective function is evaluated, and the surrogate model is updated with the new sample.

model flexibility and computational cost.

The dropout rate, a regularization parameter, was optimized within the range of 0.1 to 0.5 using a linear distribution. In this case, values were selected uniformly across the range without using a logarithmic scale. Dropout is a widely used regularization technique to reduce overfitting by randomly setting a fraction of layer outputs to zero during training. The selected range covers typical values, where 0.1 represents light dropout, allowing most connections to remain active, while 0.5 applies relatively strong regularization. This interval was chosen to balance the model's learning capacity and the need to prevent overfitting.

Finally, the regularization coefficient for L2 weight regularization was optimized over a range from 10^{-8} to 10^{-2}, using a logarithmic scale. This scaling allows the exploration of values across several orders of magnitude, enabling Optuna to test very small values, such as 10^{-8}, as well as larger values like 10^{-4} or 10^{-2}. Regularization helps to prevent overfitting by penalizing large weights, encouraging the model to find simpler solutions. The selected range covers scenarios ranging from minimal regularization to moderate levels of regularization, which are commonly effective. The logarithmic scale is particularly useful here, as small increments in lower values can lead to significant differences in the regularization effect.

This configuration balances linear and logarithmic explorations for key parameters:

- **Learning rate** and **regularization** use logarithmic scales to explore different orders of magnitude, focusing on significant variations while allowing small internal variations within the selected range.
- **Initial filters** and **dropout rate** follow a linear distribution, focusing more on discrete intervals and uniform distributions.

Essentially, the logarithmic distribution allows "zooming" into the most important orders of magnitude, offering a more balanced exploration for parameters distributed on a large scale. This combination is ideal for adapting a model to various training and regularization needs without excessive complexity.

After 100 iterations, the best hyperparameters found were:

- **Learning Rate**: 1.25e-05
- **Initial Filters**: 128
- **Dropout Rate**: 0.11
- **L2 Regularization Value**: 4.32e-08

These hyperparameter values were then used to retrain the network on the three datasets: 2 months, 1 month, and 2 weeks. The following section presents these results along with a comparison to the initial findings.

4.6 Results after Hyperparameter Optimization

After using Bayesian Optimization via the *Optuna* library to optimize the model's hyperparameters, the results show an improvement in the UNet model's performance evaluated as before in terms of: Dice Loss and IoU. Fig. 4.15 shows the new learning curves for the three acquisition periods: 2 months, 1 month, and 2 weeks.

4.6.1 Comparison with Previous Results

The graphs show an improvement in model performance compared to the results obtained before hyperparameter optimization, presented in the previous section.

As observed in fig. 4.15, the loss function for both the training and validation sets decreases more uniformly after hyperparameter optimization. After applying Bayesian optimization, the average validation loss showed only marginal improvements. For the two-month simulated acquisition, the validation loss decreased from 0.0789 to 0.0756, which, while measurable, does not represent a substantial numerical improvement. Similarly, in the one-month case, the validation loss decreased from 0.0926 to 0.0891, with a slightly faster stabilization of the validation curve. For the two-week simulation, the validation loss improved from 0.1096 to 0.1057, but again, the numerical difference is minimal. Although these changes in validation loss are not significant, an improvement is observed in the convergence behavior of the training process. The validation curve stabilizes earlier, and the gap between the training and validation loss curves has been reduced, indicating a better generalization of the model. While Bayesian optimization did not lead to a substantial numerical gain in validation loss, it was a necessary attempt to explore whether hyperparameter tuning could yield better performance. The results suggest that the model is already operating near an optimal regime.

4.6.2 Test Phase Results

After optimizing the hyperparameters, the model was tested using the same test set previously used to evaluate network performance. Fig. 4.16 shows the distribution of the IoU metric on test data for the acquisition periods of 2 months, 1 month, and 2 weeks. Over a period of 2 months, the average IoU improved from 0.61 to 0.63. The distribution reveals a greater concentration of images with IoU values above 0.6, indicating an enhancement in segmentation accuracy. For a duration of 1 month, the average IoU remained unchanged at 0.59, showing no notable improvement. However, within a shorter timeframe of 2 weeks, the average IoU increased slightly from 0.53 to 0.55. While the distribution suggests a tendency

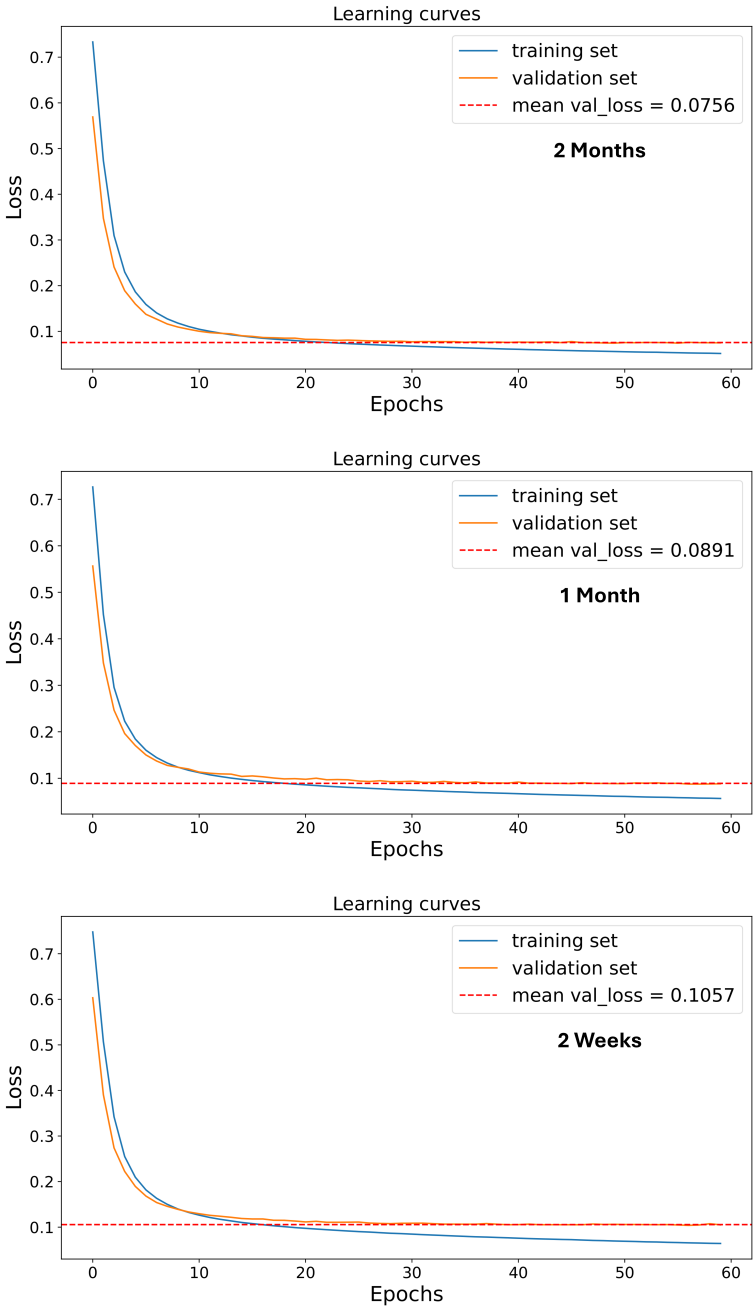


Figure 4.15: Learning curve after hyperparameter optimization for the datasets of 2 months, 1 month, and 2 weeks.

toward higher IoU values, a significant proportion of images still exhibited low IoU scores.

As previously mentioned, the peak at zero in all three histograms is due to the presence of cavity with very small shapes, which are challenging to detect or at the limits of acceptance. Ignoring this peak, it can be seen that the models segment the cavity within the muon radiography maps very well. Recall that the IoU metric compares bin-by-bin the cavity in the ground truth map with the model's prediction. Thus, an average value of 63% (which would rise to 76% if the zero peak is disregarded) indicates that the model can reconstruct a substantial portion of the cavity well. This is highly satisfactory for our purpose, as pixel-perfect accuracy for the cavity, given their typical size, is not essential. Our goal is to determine the number of cavities, their spatial positioning, and their size.

4.6.3 Discussion of Results

Hyperparameter optimization resulted in a small overall improvement in model performance, with a reduction in the loss function in all cases analyzed and an improvement in IoU values during the test phase. These results indicate that the optimized parameters allowed the model to learn more effectively, reducing the discrepancy between the training set and validation set. Regarding the test phase, it is observed that the dataset with a 2-month acquisition period shows the most notable improvements, with an increase in the mean IoU to 0.63.

The use of optimized L2 regularization and a well-calibrated dropout rate likely reduced overfitting, as evidenced by the more stable learning curve behavior. The optimized learning rate make sure that the model converges less quickly to the local minima. It takes about 10 more epochs. On the other hand we have an improvement in the mean validation values.

The increase in IoU values suggests improved precision in cavity segmentation by the model. Although the 2-week dataset showed a smaller improvement, the overall results indicate that the model benefited from hyperparameter optimization, especially for longer acquisition periods.

The results after Bayesian Optimization show only marginal improvements in segmentation accuracy. A comparison between Figures 4.12 and 4.16 reveals that the overall performance remains quite similar, with only slight variations in the mean IoU and validation loss. While the optimization process did not lead to substantial enhancements, it contributed to a better understanding of the model's learning dynamics. In particular, the optimization resulted in a slightly faster convergence and a more stable validation loss, suggesting a minor improvement in the generalization capability. However, the limited impact on overall segmentation accuracy indicates that the model was already operating in a near-optimal regime, and Bayesian Optimization was not a decisive factor in further improving

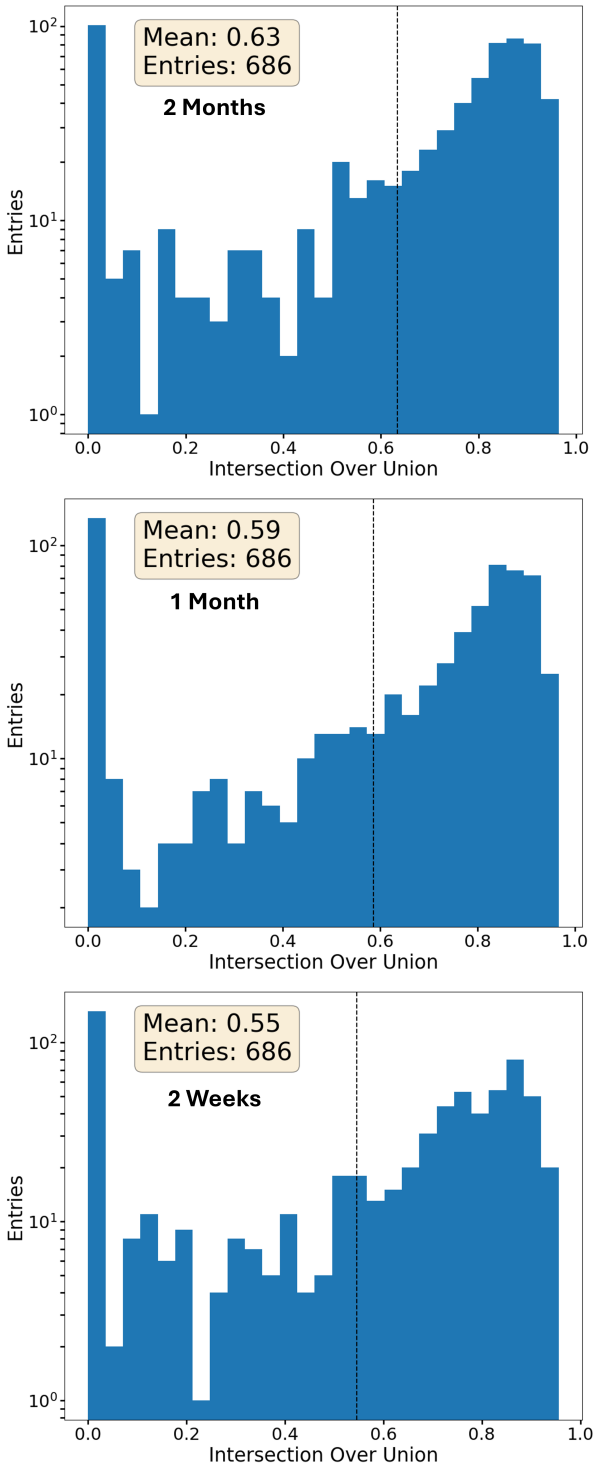


Figure 4.16: IoU distribution on test data after hyperparameter optimization, for acquisition periods of 2 months, 1 month, and 2 weeks.

performance.

4.7 Model behavior on real data

To understand the behavior of the UNet model in a real-world scenario, we tested it on the relative transmission maps of measurements 1, 3, and 5 conducted at the Temperino mines [10]. The results predicted by the model trained on the dataset from 2 months are initially shown, both before and after hyperparameter optimization.

Fig. 4.17 illustrates the transmission maps from measurements 1, 3, and 5 conducted by the Florence muography group at the Temperino mine [10]. The white areas indicate the model predictions, overlaid to verify their compatibility with the reference study [10]. The yellow ellipses highlight confirmed cavity, while the red ellipses represent potential cavity that are yet to be verified but align with the group's internal predictions.

For measurement 1, within the yellow ellipse, the model's prediction after optimization aligns excellently with the confirmed cavities, detecting and segmenting them accurately with a precise shape. In the non-optimized case, the cavity appears more angular and discontinuous than expected. The presumed cavities indicated by the red ellipses are consistent with the group's predictions, although they still need to be confirmed through field inspections. It is evident that the optimized model is more effective in identifying these presumed cavities compared to the non-optimized one. In measurement 3, the prediction within the yellow ellipse does not fully correspond to the actual cavity in either case, as it fails to capture the correct, wider, and elongated shape. Additionally, the extra prediction does not align with the presumed cavity hypothesized by the group, showing a broader area with less precise segmentation. Finally, for measurement 5, the prediction within the yellow ellipse shows perfect agreement with the confirmed cavity in the optimized case, accurately capturing its position and shape. The prediction in the area of the red ellipse suggests the possible presence of another cavity, though it remains unverified. The first model exhibits slightly worse segmentation in both the known cavity area and the presumed cavity area.

This comparison indicates that, albeit marginally, hyperparameter optimization has led to an improvement in the segmentation of both known and presumed cavities in real measurements.

In fig. 4.18, a comparison is shown between the predictions of the model trained using a simulated two-month dataset (on the left) and the model trained with two weeks of data (on the right). Both models underwent hyperparameter optimization. Starting with measurement 1, the predictions for the known cavities are very similar, but the primary difference lies in the prediction of presumed cavities.

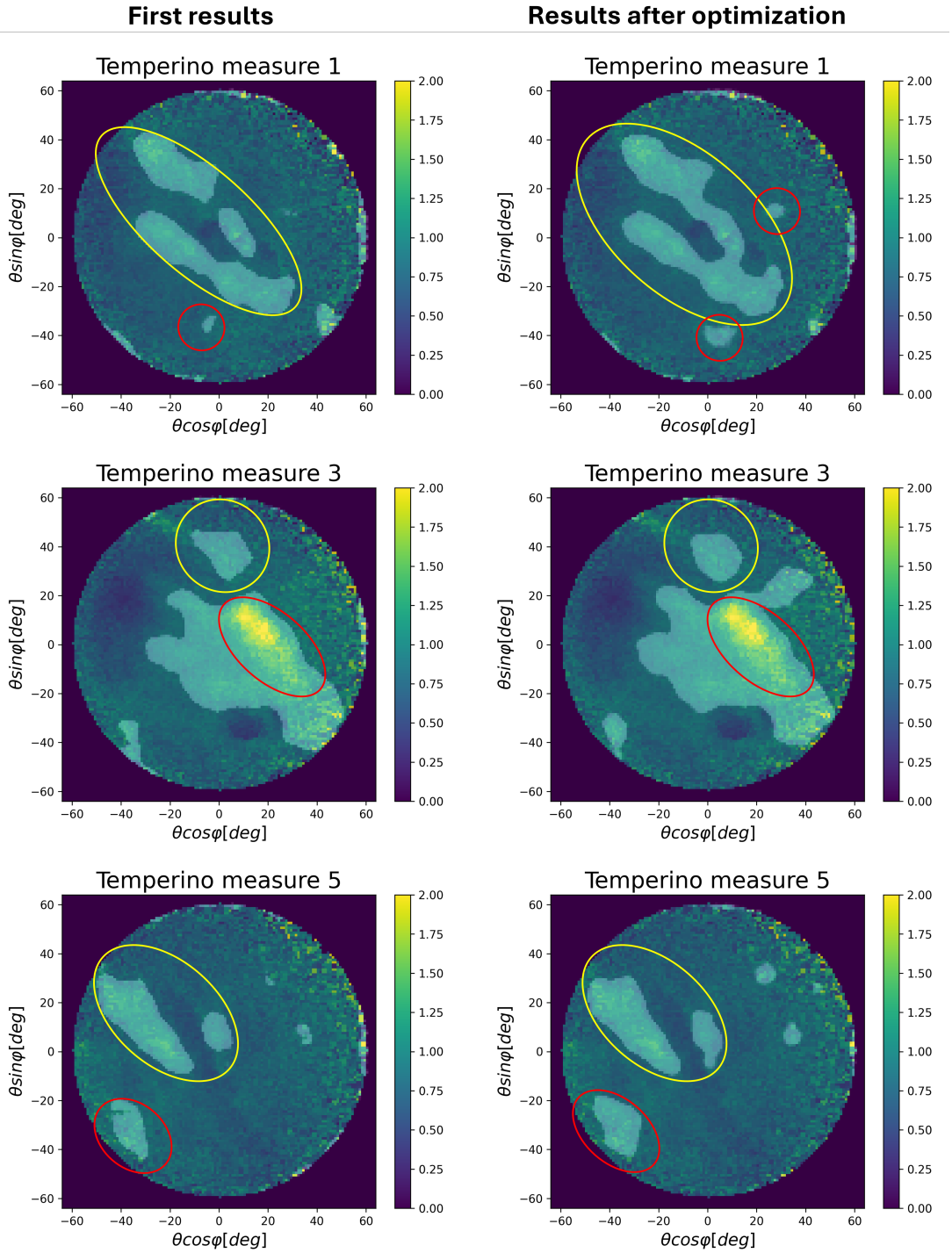


Figure 4.17: Relative transmission maps of measurements 1, 3, and 5 conducted at the Temperino mine by the Florence muography group [10]. Predictions for the models before and after hyperparameter optimization are shown. The white areas represent the UNet model predictions, overlaid on the transmission map to compare them with the group's observations. Yellow ellipses denote verified cavity, while red ellipses indicate potential cavity that are plausible according to the group's internal studies but not yet confirmed.

These presumed cavities are absent in the two-week model. For measurement 3, the predictions are nearly identical. Both models fail to correctly predict the known cavity (yellow ellipse), while for the possible cavities (red ellipse), both models predict a broad shape. For measurement 5, both models accurately predict the region of the known cavities (yellow ellipse), but the segmentation is slightly more precise in the model trained with two months of data. Regarding the prediction of possible new cavities (red ellipse), both models make similar predictions.

Based on these comparisons, and supported by the previously presented results for the loss curves and IoU histograms, it is evident that hyperparameter optimization has improved the network's performance. Specifically, the model trained with the simulated two-month dataset tends to perform better compared to the one trained with two weeks of data. The former predicts more cavities, as seen in measurement 1, and predicts the known cavities with greater precision, as observed in measurement 5. This suggests that training the model with a larger simulated dataset is preferable, as it reduces statistical fluctuations and provides clearer maps for training the network.

Thus, we conclude that the UNet network predicts the presence of cavity in all three measurements. In measurements 1 and 5, the model demonstrates high accuracy, correctly predicting both the position and shape of known cavity. The presumed cavity, highlighted in the red ellipses, are also in line with the predictions from the Florence muography group.

Regarding measurement 3, the model detects cavity but with less precision in shape and position, likely due to the presence of denser bodies that reduce cavity visibility. In the other two measurements, however, the dense bodies are not overlapping with the presumed cavity. Although the test statistics on real data are limited to three measurements, these results show that under favorable conditions, as in measurements 1 and 5, the UNet model is effective in accurately predicting both the position and shape of cavity.

4.8 UNet Conclusions

The purpose of this study was to develop software capable of assisting researchers in identifying and segmenting anomalies within muon radiography. In this context, these anomalies represent cavity within mines. The objective is thus to classify each pixel based on the presence or absence of cavity. The case study used was conducted by the Florence muography group at the Temperino mines [10].

To achieve this, we developed a UNet neural network, which is particularly suited for anomaly segmentation in the medical field, and adapted it to our needs. In addition to model development, the study addresses various challenges, includ-

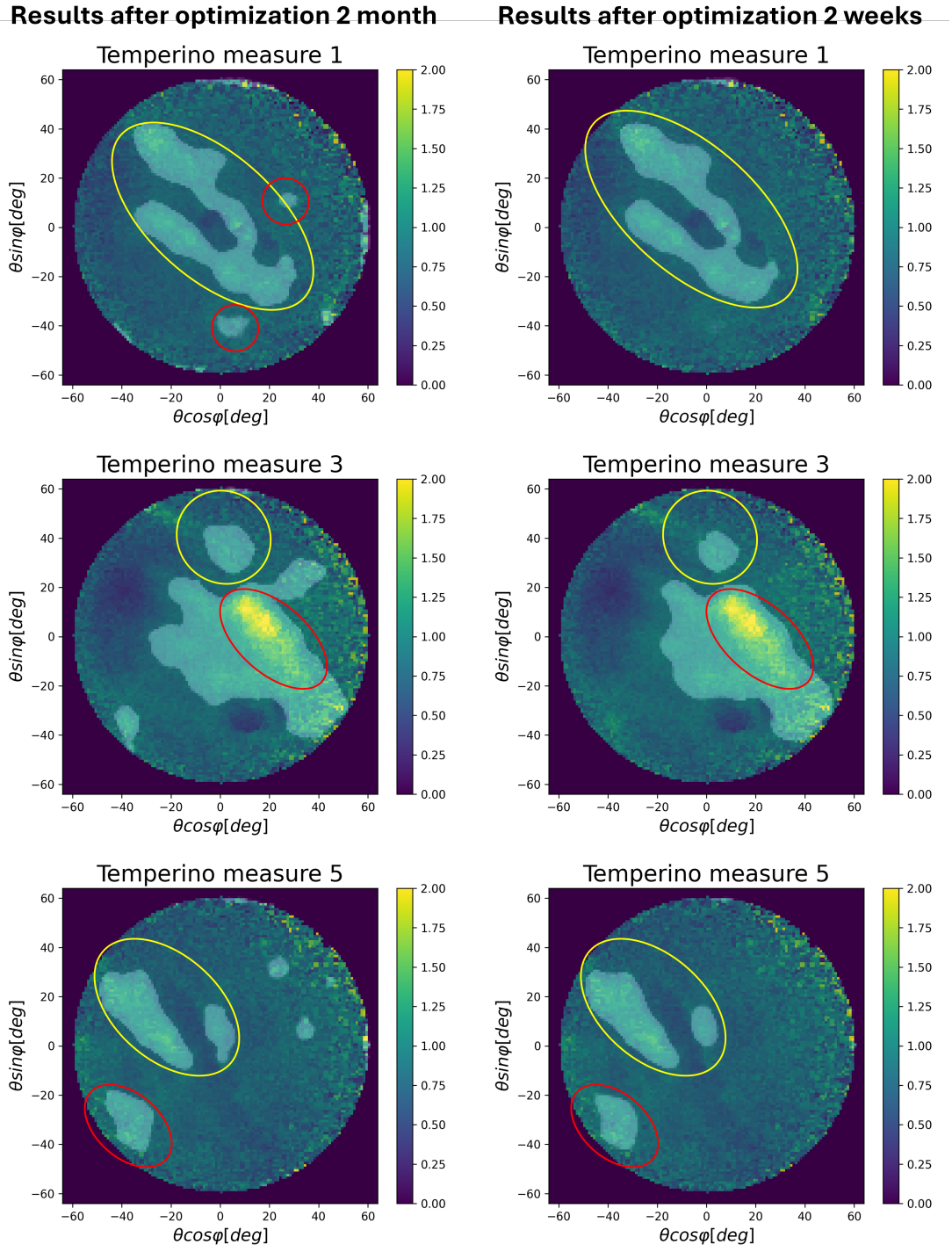


Figure 4.18: Relative transmission maps of measurements 1, 3, and 5 conducted at the Temperino mine by the Florence muography group [10]. The white areas represent the UNet model predictions, overlaid on the transmission map to compare them with the group's observations. On the left, the model predictions for a simulated two-month dataset are shown, while on the right, those corresponding to two weeks are displayed. Yellow ellipses denote verified cavity, while red ellipses indicate potential cavity that are plausible according to the group's internal studies but not yet confirmed.

ing the lack of a large training dataset. Therefore, the first part of this work focused on developing a synthetic and comprehensive dataset.

The network was initially trained with hyperparameters selected based only on our domain experience. These parameters still hinted at the potential of this approach. In fact, the average values for the validation loss and the test case metric calculated using Intersection over Union (IoU) were encouraging.

We subsequently fine-tuned the model by finding the best hyperparameters for training. These were determined using Bayesian Optimization, specifically targeting the best values for: learning rate, initial filter number, dropout rate, and L2 regularization. The subsequent results showed improvement, both in terms of Dice Loss and IoU, demonstrating that optimization allowed the model to learn more effectively and generalize better on test data.

Finally, we tested the model on real data obtained from the Temperino mine, where it was able to correctly identify known cavity and suggest the presence of possible new, yet unverified, cavity. These results, although evaluated on only three real cases, demonstrate the potential of the adopted approach for real-world applications, paving the way for further research and refinements.

This method is not intended to replace the researcher's expertise; rather, it aims to be a tool to complement their study. The key to accurately identifying anomalies within muon radiography is the combined use of this software and the scientist's experience. Previously, the analysis was based solely on visual observation; now, however, an additional tool provides an analytical and objective indication.

This methodology could be further improved by addressing two aspects: building a larger and more diverse dataset and using more advanced models.
The first aspect involves creating more artificial cavity and placing them at different points within the acceptance area. Another approach could be positioning MIMA at various locations.
The second aspect involves exploring more complex architectures, such as *Vision Transformers* [62]. This could involve using an already robust, pre-trained model and adapting it to our specific case.

Overall, it should be noted that the results obtained are highly satisfactory and meet our expectations. In summary, the work carried out demonstrated the effectiveness of using optimized UNet neural networks for cavity segmentation in complex geological contexts, contributing to the development of automated tools for muographic image interpretation and supporting the scientific community in analyzing subsurface structures.

Chapter 5

Summary and conclusions

This doctoral thesis is based on two main aspects. The first aspect focuses on the construction of a detector utilizing scintillating bar technology for the BLEMAB european project, with the aim of real-time monitoring of density variations within the cohesive zone of a blast furnace. The second aspect involves the development of an advanced approach for analyzing and monitoring geological structures through muon radiography, with applications in contexts such as tailing dams and mines. The main objective was to implement non-invasive imaging tools for identifying structural anomalies, optimizing detectors like MIMA and employing machine learning techniques to enhance analysis accuracy and effectiveness.

5.1 Summary of activities

The work carried out is organized into three main areas: participation in the BLEMAB project for monitoring internal density in blast furnaces, applying muon radiography to monitor infiltration in tailing dams, and identifying cavities in mines. The European BLEMAB project required the construction of two scintillating bar detectors larger than MIMA to monitor density variations within blast furnaces in real time. These detectors, designed to withstand challenging industrial environments, represent a significant advancement in muon radiography, as they can collect accurate and reliable data in contexts where traditional instruments would be ineffective. The ability to continuously and non-invasively observe these industrial structures demonstrates the potential of this technology in broader and more complex scenarios, with promising applications in heavy industry. The construction and testing of the detectors took approximately a year and a half. Data collection began in August 2023, with both detectors monitoring the same furnace. To explore the technique's potential, one of the detectors was relocated in November 2023 to monitor the operation of a smaller furnace. This data acqui-

sition phase concluded in January 2024, when we visited the site to carry out the disassembly procedures. The project was successfully completed, and preliminary results were delivered in the final report to the project manager in autumn 2024.

The application of muon radiography for detecting water infiltration in tailing dams demonstrated how this technology can improve the safety of critical structures. By using Voronoi tessellation in combination with a Random Forest classifier, it was possible to produce high-resolution density maps identifying potentially at-risk zones. This methodology has proven effective in delivering a detailed view of compromised structural areas, facilitating timely and accurate interventions.

The UNet network, optimized using Bayesian methods, was employed to identify and segment cavities in complex underground settings. Through the optimization of key hyperparameters, such as learning rate and initial filters, the network achieved accurate cavity mapping, overcoming some limitations posed by environmental complexity. This outcome highlights how advanced machine learning algorithms can significantly enhance muon radiography applications in challenging geological environments.

5.2 Conclusions and future developments

This thesis demonstrates the potential of muon radiography as an innovative tool for monitoring critical structures, with applications ranging from geological to industrial contexts. The results obtained open new perspectives for using muon radiography, especially when supported by advanced machine learning techniques, in sectors that require non-invasive and highly reliable solutions.

In the context of tailing dams, we have demonstrated that the combination of Voronoi tessellation and Random Forest classification effectively enhances the resolution of density maps, enabling the detection of water infiltration with an accuracy of approximately 60-70% in terms of density variations. Compared to conventional methods, this approach offers a significant improvement in response time, reducing the period required to detect infiltration zones to just 2 days when using a detector similar to the one employed for BLEMAB. However, we also showed that the final acquisition time strongly depends on the type of dam, particularly on the assumed density of the internal materials. For example, denser structures require longer acquisition times, underscoring the importance of preliminary material characterization.

The application of the UNet neural network has shown substantial advantages in cavity segmentation and identification in underground environments. In our case, we achieved a validation loss value of 0.0756, using a dataset simulating acquisition times of two months, demonstrating the model's ability to adapt very well without overfitting. The test values show an average IoU of 0.63, for the same

dataset, proving that the model can segment cavities with precision in most cases, even when using unseen data during the training phase.

In the specific case of the Temperino mine, the integration of muon radiography with the UNet network enabled us to develop software capable of autonomously and unbiasedly identifying cavities. The model successfully segments known cavities and predicts others in agreement with expert, forecasts for measurements 1 and 5. However, it shows greater difficulty in accurately addressing measurement 3.

The applications presented illustrate the potential of muon radiography to contribute to structural safety and environmental protection, enabling preventive and targeted interventions on strategic infrastructures. These findings suggest a broader and more diverse potential for muon radiography, which could become a reference method in non-invasive diagnostics and infrastructure protection.

Future developments in this field will involve the creation of increasingly precise and faster simulation techniques, both to accurately characterize the inspection environment and to generate transmission maps for model training. Future advancements in this field include the development of increasingly sophisticated and faster simulation techniques. These simulations aim to accurately model inspection environments by integrating more comprehensive geometric and physical interactions, ultimately enabling the creation of precise transmission maps that are essential for training advanced predictive models. Additionally, a parallel focus on the development of advanced machine learning techniques for anomaly detection will further refine these applications. At the same time, significant efforts will focus on refining machine learning approaches, particularly in anomaly detection, by leveraging large-scale datasets and improved algorithmic architectures.

For instance, in the Temperino case study, the detector could be simulated in position 3, identified as the most problematic. Pre-trained models could then be used, trained on features similar to ours but with a much larger dataset. Another option is to employ visual transformers, new models with promising results.

These improvements aim to achieve a higher degree of accuracy and reliability in identifying and analyzing subtle variations in dense and complex environments, broadening the scope of applications in both industrial and natural contexts. Additionally, integrating real-time data acquisition systems and adaptive learning will enable more dynamic and immediate responses to structural and environmental changes, showcasing the transformative potential of this technology across a wide range of disciplines.

Bibliography

[1] Juan Garzon et al. "TRAGALDABAS. First results on cosmic ray studies and their relation with the solar activity, the Earth magnetic field and the atmospheric properties". In: (Jan. 2017).

[2] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[3] Lorenzo Bonechi, Raffaello D'Alessandro, and Andrea Giammanco. "Atmospheric muons as an imaging tool". In: *Reviews in Physics* 5 (2020), p. 100038. ISSN: 2405-4283. DOI: https://doi.org/10.1016/j.revip.2020.100038. URL: https://www.sciencedirect.com/science/article/pii/S2405428320300010.

[4] Angelo Maggiora. *A.M.E.L.I.E. Apparatus for Muon Experimental Lifetime Investigation and Evaluation.* 2023. arXiv: 2306.10712 [physics.ed-ph]. URL: https://arxiv.org/abs/2306.10712.

[5] G. Bonomi et al. "Applications of cosmic-ray muons". In: *Progress in Particle and Nuclear Physics* 112 (2020), p. 103768. DOI: 10.1016/j.ppnp.2020.103768.

[6] P. Checchia. "Muon tomography applications and detector developments". In: *Journal of Instrumentation* 15 (2020), p. C04041. DOI: 10.1088/1748-0221/15/04/C04041.

[7] S. Pesente et al. "First results on material identification and imaging with a large-volume muon tomography prototype". In: *Nuclear Instruments and Methods in Physics Research Section A* 604 (2009), pp. 738–746. DOI: 10.1016/j.nima.2009.03.229.

[8] DONALD E. GROOM, NIKOLAI V. MOKHOV, and SERGEI I. STRIGANOV. "MUON STOPPING POWER AND RANGE TABLES 10 MeV–100 TeV". In: *Atomic Data and Nuclear Data Tables* 78.2 (2001), pp. 183–356. ISSN: 0092-640X. DOI: https://doi.org/10.1006/adnd.2001.0861. URL: https://www.sciencedirect.com/science/article/pii/S0092640X01908617.

[9] N. V. Mokhov D. E. Groom and S. Striganov (2001). *Muon stopping power and range tables 10 MeV–100 TeV*. URL: https://pdg.lbl.gov/2020/AtomicNuclearProperties/.

[10] D. Borselli et al. "Three-dimensional muon imaging of cavities inside the Temperino mine (Italy)". In: *Scientific Reports* 12 (2022). DOI: 10.1038/s41598-022-26393-7. URL: https://doi.org/10.1038/s41598-022-26393-7.

[11] Inc. Autodesk. *AutoCAD 2023: Computer-Aided Design Software*. Available at: https://www.autodesk.com/products/autocad. 2023.

[12] C. Frosin et al. "Optimizing sea level muon flux modeling: A 2D fit approach and integration into Geant4 generators". In: *Journal of Physics G* (2024). to be submitted.

[13] S. Gonzi et al. "A parametric analytical model of the atmospheric muon flux at sea-level and its application in the field of the muon transmission radiography". In: *38th International Cosmic Ray Conference*. Sept. 2024.

[14] L. Bonechi et al. "Development of the ADAMO detector: test with cosmic rays at different zenith angles". In: *29th International Cosmic Ray Conference (Pune)*. Vol. 9. Jan. 2005, p. 283.

[15] Lorenzo Bonechi. "Measurements of Cosmic Rays at Earth with the Experiment ADAMO". Other thesis. Apr. 2004.

[16] O. C. Allkofer et al. "Cosmic-ray muon spectra at sea-level up to 10 TeV". In: *Nucl. Phys. B* 259 (1985). [Erratum: Nucl.Phys.B 268, 747–747 (1986)], pp. 1–18. DOI: 10.1016/0550-3213(86)90270-1.

[17] M. Guan et al. "A parametrization of the cosmic-ray muon flux at sea-level". In: (2015). arXiv: 1509.06176 [hep-ex].

[18] T. K. Gaisser. *Cosmic rays and particle physics*. Cambridge University Press, 1990.

[19] Geant4 Collaboration. *Geant4 Software*. https://geant4.web.cern.ch/. Accessed on October 08, 2023. 2023.

[20] DONALD E. GROOM, NIKOLAI V. MOKHOV, and SERGEI I. STRIGANOV. "MUON STOPPING POWER AND RANGE TABLES 10 MeV–100 TeV". In: *Atomic Data and Nuclear Data Tables* 78.2 (2001), pp. 183–356. ISSN: 0092-640X. DOI: https://doi.org/10.1006/adnd.2001.0861. URL: https://www.sciencedirect.com/science/article/pii/S0092640X01908617.

[21] G. Baccani. "Development, testing and application to a real case studies of a threedimensional tomographic technique based on muon transmission radiography". Ph.D. Thesis. Faculty of Mathematical Physical and Natural Sciences: University of Florence, 2020.

[22] Peter K.F. Grieder. "Cosmic Rays at Earth". In: ed. by Peter K.F. Grieder. Amsterdam: Elsevier, 2001, pp. 1–53. ISBN: 978-0-444-50710-5. URL: https://shop.elsevier.com/books/cosmic-rays-at-earth/grieder/978-0-444-50710-5.

[23] Your citation for Paccagnella's work here. "Your title here". In: *Your journal here* (YYYY), XX–YY. DOI: YourDOIhere.

[24] F. Sauli. "Principles of operation of multiwire proportional and drift chambers". In: *CERN Yellow Reports: Monographs* 77-09 (1991), pp. 1–60. DOI: 10.5170/CERN-1977-009.

[25] C. Santoni et al. "Performance of resistive plate chambers for high rate applications". In: *Nuclear Instruments and Methods in Physics Research A* 515 (2003), pp. 545–561. DOI: 10.1016/S0168-9002(03)01892-3.

[26] F. Sauli. "GEM: A new concept for electron amplification in gas detectors". In: *Nuclear Instruments and Methods in Physics Research A* 386 (1997), pp. 531–534. DOI: 10.1016/S0168-9002(96)01172-2.

[27] F. Keizer et al. "A compact, high resolution tracker for cosmic ray muon scattering tomography using semiconductor sensors". In: *Journal of Instrumentation* 13.10 (Oct. 2018), P10028. DOI: 10.1088/1748-0221/13/10/P10028. URL: https://dx.doi.org/10.1088/1748-0221/13/10/P10028.

[28] Ignácio Lázaro Roche. "A Compact Muon Tracker for Dynamic Tomography of Density Based on a Thin Time Projection Chamber with Micromegas Readout". In: *Particles* 4.3 (Sept. 2021), pp. 333–342. DOI: 10.3390/particles4030028. URL: https://hal.science/hal-03276509.

[29] Ryuichi Nishiyama, Seigo Miyamoto, and Naotaka Naganawa. "Application of Emulsion Cloud Chamber to cosmic-ray muon radiography". In: *Radiation Measurements* 83 (2015). 26th International Conference on Nuclear Tracks in Solids, pp. 56–58. ISSN: 1350-4487. DOI: https://doi.org/10.1016/j.

radmeas.2015.04.013. URL: https://www.sciencedirect.com/science/article/pii/S1350448715300147.

[30] M. D'Errico et al. "The MURAVES muon telescope: a low power consumption muon tracker for muon radiography applications". In: *Journal of Physics: Conference Series* 2374.1 (Nov. 2022), p. 012190. DOI: 10.1088/1742-6596/2374/1/012190. URL: https://dx.doi.org/10.1088/1742-6596/2374/1/012190.

[31] D. Borselli et al. "Muography as a support technique for non-invasive research and three-dimensional localization of tombs in archaeological sites: a case study from Palazzone Necropolis (Perugia – Italy)". In: *Journal of Instrumentation* 19.02 (Feb. 2024), p. C02076. DOI: 10.1088/1748-0221/19/02/C02076. URL: https://dx.doi.org/10.1088/1748-0221/19/02/C02076.

[32] C. Frosin et al. "Exploring the potential of muon radiography for blast furnace assessments: advancements in non-invasive imaging and structural analysis". In: *Journal of Instrumentation* 19.02 (Feb. 2024), p. C02041. DOI: 10.1088/1748-0221/19/02/C02041. URL: https://dx.doi.org/10.1088/1748-0221/19/02/C02041.

[33] G. Baccani et al. "The MIMA project. Design, construction and performances of a compact hodoscope for muon radiography applications in the context of archaeology and geophysical prospections". In: *Journal of Instrumentation* 13.11 (Nov. 2018), P11001. DOI: 10.1088/1748-0221/13/11/P11001. URL: https://dx.doi.org/10.1088/1748-0221/13/11/P11001.

[34] F Ambrosino et al. "The MU-RAY project: detector technology and first data from Mt. Vesuvius". In: *Journal of Instrumentation* 9.02 (Feb. 2014), p. C02029. DOI: 10.1088/1748-0221/9/02/C02029. URL: https://dx.doi.org/10.1088/1748-0221/9/02/C02029.

[35] Stéphane Callier et al. "EASIROC, an Easy Versatile ReadOut Device for SiPM". In: *Physics Procedia* 37 (2012). Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011), pp. 1569–1576. ISSN: 1875-3892. DOI: https://doi.org/10.1016/j.phpro.2012.02.486. URL: https://www.sciencedirect.com/science/article/pii/S1875389212018688.

[36] Xilinx, Inc. *Spartan-II FPGA Family: Product Specification*. Version 1.0. 2000. URL: https://www.xilinx.com.

[37] BLEMAB collaboration et al. "BLEMAB European project: muon imaging technique applied to blast furnaces". In: *Journal of Instrumentation* 17.04 (Apr. 2022), p. C04031. DOI: 10.1088/1748-0221/17/04/C04031. URL: https://dx.doi.org/10.1088/1748-0221/17/04/C04031.

[38] URL: https://www.blemab.eu/.

[39] ArcelorMittal Bremen. *ArcelorMittal Bremen Website*. https://bremen.arcelormittal.com. Accessed: 2024-04-27. 2024.

[40] Vale S.A. *Brumadinho Dam Disaster*. Accessed: 2024-08-19. Jan. 2019. URL: https://en.wikipedia.org/wiki/Brumadinho_dam_disaster.

[41] Independent Expert Engineering Investigation and Review Panel. *Mount Polley Tailings Storage Facility Breach*. Accessed: 2024-08-19. Jan. 2015. URL: https://www.mountpolleyreviewpanel.ca/.

[42] Michele Cappellari and Yannick Copin. "Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations". In: *Monthly Notices of the Royal Astronomical Society* 342.2 (June 2003), pp. 345–354. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2003.06541.x. eprint: https://academic.oup.com/mnras/article-pdf/342/2/345/3407404/342-2-345.pdf. URL: https://doi.org/10.1046/j.1365-8711.2003.06541.x.

[43] S. Lloyd. "Least squares quantization in PCM". In: *IEEE Transactions on Information Theory* 28.2 (1982), pp. 129–137. DOI: 10.1109/TIT.1982.1056489.

[44] A. Paccagnella et al. "Adaptive Voronoi Binning in Muon Radiography for Detecting Subsurface Cavities". In: *Image Analysis and Processing - ICIAP 2023 Workshops*. Ed. by Gian Luca Foresti, Andrea Fusiello, and Edwin Hancock. Cham: Springer Nature Switzerland, 2024, pp. 170–178. ISBN: 978-3-031-51023-6.

[45] Gérard Biau and Erwan Scornet. *A Random Forest Guided Tour*. 2015. arXiv: 1511.05741 [math.ST]. URL: https://arxiv.org/abs/1511.05741.

[46] URL: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html.

[47] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. arXiv: 1505.04597 [cs.CV].

[48] Tommaso Beni et al. "Transmission-Based Muography for Ore Bodies Prospecting: A Case Study from a Skarn Complex in Italy". In: *Natural Resources Research* 32.4 (Aug. 2023), pp. 1529–1547. ISSN: 1573-8981. DOI: 10.1007/s11053-023-10201-8. URL: https://doi.org/10.1007/s11053-023-10201-8.

[49] CloudCompare Development Team. *CloudCompare: 3D point cloud and mesh processing software*. Open Source Project. 2022. URL: https://www.cloudcompare.org/.

[50] Xia Zhao et al. "A review of convolutional neural networks in computer vision". In: *Artificial Intelligence Review* 57.4 (Mar. 2024), pp. 99–99. ISSN: 1573-7462. DOI: 10.1007/s10462-024-10721-6. URL: https://doi.org/10.1007/s10462-024-10721-6.

[51] Neena Aloysius and M. Geetha. "A review on deep convolutional neural networks". In: *2017 International Conference on Communication and Signal Processing (ICCSP)*. 2017, pp. 0588–0592. DOI: 10.1109/ICCSP.2017.8286426.

[52] Keiron O'Shea and Ryan Nash. "An Introduction to Convolutional Neural Networks". In: *CoRR* abs/1511.08458 (2015). arXiv: 1511.08458. URL: http://arxiv.org/abs/1511.08458.

[53] Francois Chollet et al. *Keras*. 2015. URL: https://github.com/fchollet/keras.

[54] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. "Deep Sparse Rectifier Neural Networks". In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. Vol. 15. Proceedings of Machine Learning Research. PMLR, 2011, pp. 315–323. URL: https://proceedings.mlr.press/v15/glorot11a/glorot11a.pdf.

[55] Lars Nieradzik et al. *Effect of the output activation function on the probabilities and errors in medical image segmentation*. 2021. arXiv: 2109.00903 [eess.IV]. URL: https://arxiv.org/abs/2109.00903.

[56] Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. *Towards a Guideline for Evaluation Metrics in Medical Image Segmentation*. 2022. arXiv: 2202.05273 [eess.IV]. URL: https://arxiv.org/abs/2202.05273.

[57] A. Helen Victoria and G. Maragatham. "Automatic tuning of hyperparameters using Bayesian optimization". In: *Evolving Systems* 12.1 (2021), pp. 217–223. ISSN: 1868-6486. DOI: 10.1007/s12530-020-09345-2. URL: https://doi.org/10.1007/s12530-020-09345-2.

[58] *The Gaussian Process Website*. Accessed: 2024-10-25. URL: https://gaussianprocess.org/.

[59] Shuhei Watanabe. *Tree-Structured Parzen Estimator: Understanding Its Algorithm Components and Their Roles for Better Empirical Performance*. 2023. arXiv: 2304.11127 [cs.LG]. URL: https://arxiv.org/abs/2304.11127.

[60] Alexandros Kamperis. *Acquisition Functions for Bayesian Optimization*. Accessed: 2024-10-25. 2021. URL: https://ekamperi.github.io/machine%20learning/2021/06/11/acquisition-functions.html.

[61] *Optuna: A Hyperparameter Optimization Framework*. Accessed: 2024-10-25. URL: https://optuna.org/.

[62] Hans Thisanke et al. *Semantic Segmentation using Vision Transformers: A survey*. 2023. arXiv: 2305.03273 [cs.CV]. URL: https://arxiv.org/abs/2305.03273.

La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020, risorse FSE REACT-EU
Azione IV.4 “Dottorati e contratti di ricerca su tematiche dell’innovazione”
e Azione IV.5 “Dottorati su tematiche Green”