ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DEPARTMENT OF PHYSICS AND ASTRONOMY "A. RIGHI"

SECOND CYCLE DEGREE

PHYSICS

First measurement of atmospheric ν_μ and $\bar{\nu}_\mu$ flux with the KM3NeT/ARCA telescope

Supervisor

Prof. Annarita Margiotta

Co-supervisor

Dr. Rasa Muller
Dr. Francesco Filippini

Defended by

Martino Bernuzzi

Spring Session / March / 2026

Academic Year 2024/2025

Abstract

This thesis focuses on the high-energy atmospheric neutrino flux and its measurement with the second-generation neutrino telescope KM3NeT/ARCA, located in the depths of the Mediterranean Sea. Above 100 GeV, the atmospheric flux is known with an uncertainty of the order of 30%. Its precise measurement is crucial, as it also represents the main background for astrophysical neutrino analysis. The aim of this work is to perform the first ever measurement of the atmospheric ν_μ and $\bar{\nu}_\mu$ charged current flux in the energy range between 100 GeV and 100 TeV with KM3NeT/ARCA, with 21 active strings. To achieve this goal, a dedicated event selection strategy based on machine learning techniques is implemented. The flux is then measured through an unfolding procedure. This approach allows the deconvolution of detector effects from the reconstructed energy spectrum, providing an estimate of the true neutrino flux. The final result is compared with different theoretical curves and measurements performed by other neutrino telescope, such as ANTARES and IceCube.

Contents

List of Figures

List of Tables

Introduction

Neutrino plays a fundamental role in the framework of the Standard Model of particle physics, which currently represents our most successful description of the microscopic world. Several open questions and possibilities, however, still surround this topic, generating ongoing scientific interest. It is within this landscape that the KM3NeT experiment was conceived and developed. In particular, this thesis focuses on one of its two detectors, KM3NeT/ARCA, an acronym for *Astroparticle Research with Cosmics in the Abyss*. As the name suggests, over the past decades neutrinos start to play a relevant role also in the so called multi-messenger astrophysics. In fact, because of their elusiveness, neutrinos can travel through the Universe escaping from dense regions and avoiding to interact with the interstellar medium, reaching the Earth unaffected, unlike electromagnetic radiation. Together with photons and gravitational waves, neutrinos act as probes of the most energetic astrophysical phenomena. Beyond astrophysics, such detectors also offer the opportunity to investigate fundamental neutrino properties, such as their masses, which are not embedded within the Standard Model. In particular, the study of the atmospheric neutrino flux constitutes the primary focus of this thesis.

To provide a clearer overview of the structure of this work, the organization of the thesis is outlined below;

- Chapter One: an overview of the Standard Model is provided, with particular focus on neutrino properties. Subsequently, the relation between cosmic rays and atmospheric neutrinos is described. The Chapter ends with neutrino production within astrophysical phenomena;

- Chapter Two: a detailed description of KM3NeT/ARCA detector is given, including also a discussion on neutrino event topologies and background sources;

- Chapter Three: focused on Monte Carlo simulations, trigger conditions and the reconstruction chain adopted by the KM3NeT Collaboration;

- Chapter Four: the event selection strategy to achieve the required purity of the sample is presented. A preliminary selection based on reconstructed variables precedes an original development of a *Boosted Decision Tree* algorithm;
- Chapter Five: an unfolding method involving TUnfold is applied to the final sample. The full pipeline is previously developed and validated on a Monte Carlo pseudo-experiment. The final measurement of the atmospheric ν_μ and $\bar{\nu}_\mu$ flux is obtained using the same procedure with the data;
- Chapter Six: an overview of the analysis and of the result is presented, focusing on criticism and future improvements.

Historical note

At the dawn of the 1930s, particle physicists were convinced that matter was composed of protons and electrons, with the latter existing both within the nucleus and in the external atomic shell. Among the many theoretical issues associated with the presence of such light, negatively charged particles within the nucleus, the study of the beta decay further confirmed the criticalities of this interpretation. Indeed, measurement of the energy spectrum of electrons emitted in the radioactive β-decay of atomic nuclei revealed a continuous distribution, incompatible with a simple two-body decay process. It was in this context that, in December 1930, Pauli proposed, describing it as a "desperate remedy", the existence of a new elementary particle, to which he initially gave the name of neutron [1]. Only a few years later, in 1933, Fermi formalized this intuition by developing a mathematical theory capable of explaining β-decay as a point-like interaction involving four fermions. He also renamed this neutral half-integer spin particle as *neutrino*, in order to avoid confusion with the particle now known as neutron, discovered in 1932 by Chadwick [2]. This elegant theory also predicted a process known as inverse β-decay, providing a possible experimental mechanism to prove the existence of neutrinos. In this interaction, an electron antineutrino annihilates with a proton, leading to the production of a positron and a neutron. After more than twenty years of anticipation, the first experimental observation of a neutrino was achieved by Cowan and Reines [3]. Exploiting the intense neutrino flux produced by a nuclear reactor near the Savannah River in South Carolina, they used two tanks containing approximately 200 liters of water doped with $CdCl_2$ as neutrino target. Following the interaction with protons in the hydrogen atoms, the produced positron quickly annihilates with an electron, emitting two photons with energies equal to the electron mass. The neutron, on the other hand, can be detected through its capture by a cadmium nucleus, releasing an additional gamma ray several microseconds after the first two.

Chapter 1

Neutrino Physics and Astrophysics

In the theoretical framework of the Standard Model (SM) of particle physics [4], neutrinos are described as massless fermions with half-integer spin. Furthermore, neutrinos interact only via the weak interaction, as they do not have electric or color charge. Each electrically charged lepton, the electron (e), muon (μ) and tau (τ), has a corresponding neutrino flavour ν_e, ν_μ and ν_τ. These pairs form three doublets, which are known as families. Similarly, each charged antilepton is associated with its corresponding antineutrino. Contrary to the theoretical expectation, experimental evidences for extremely small neutrino masses are provided by the observation of neutrino oscillations between flavour eigenstates [5, 6]. As a consequence, neutrinos can be described as quantum mechanical superposition of mass eigenstates, labeled as ν_1, ν_2 and ν_3.

The Lagrangian of the SM of particle physics is obtained requiring the invariance under the local gauge symmetry group $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, where $\mathrm{SU}(3)_C$ describes strong interactions, while $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ describes electroweak interactions [7]. The Lagrangian describing neutrino interactions involves only the latter symmetry group and, as for the other fermions, neutrinos are described by Dirac spinors:

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{1.1}$$

The left-handed ψ_L and right-handed ψ_R components of the field can be obtained by applying projector operators to the Dirac spinor:

$$\psi_L = P_L \psi = \frac{1-\gamma_5}{2}\psi, \qquad \psi_R = P_R \psi = \frac{1+\gamma_5}{2}\psi \tag{1.2}$$

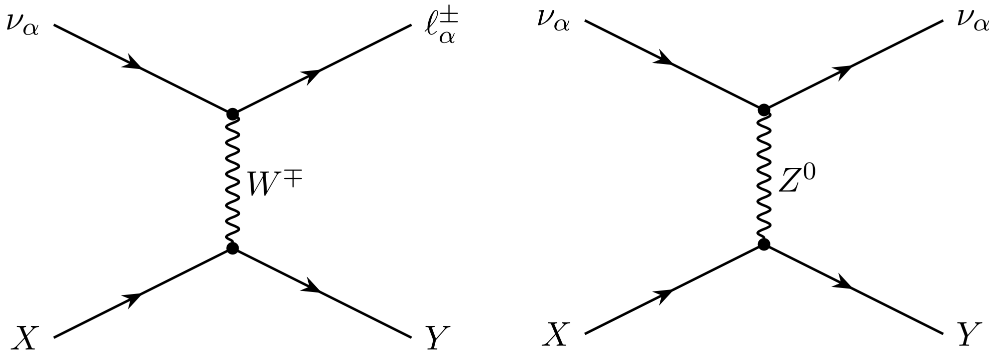
Moreover, as a consequence of parity violation in weak interactions [8], only left-handed components for particles and right-handed components for antiparticle can interact. The neutrino Lagrangian, written as the sum of charged current (CC) and neutral current (NC) terms, which respectively represent the exchange of a $W^{\pm}$ and Z^0 gauge boson field, take the following form:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{CC} + \mathcal{L}_{NC} \\ &= -\frac{g}{\sqrt{2}} \sum_{\alpha} \left(W_{\mu}^{+} \bar{\nu}_{\alpha L} \gamma^{\mu} l_{\alpha L} + W_{\mu}^{-} \bar{l}_{\alpha L} \gamma^{\mu} \nu_{\alpha L} \right) - \frac{g}{2\cos\theta_W} \sum_{\alpha} \left(Z_{\mu}^{0} \bar{\nu}_{\alpha L} \gamma^{\mu} \nu_{\alpha L} \right) \end{aligned} \tag{1.3}$$

where α runs over the three lepton families, g is the weak interaction coupling constant and θ_W is the Weinberg angle.

1.1 Neutrino's interactions

Considering the two terms in the Lagrangian (1.3), the Feynman diagrams for charged current (left) and neutral current (right) interactions are shown in Fig.1.1. In this context, it is convenient to represent neutrino interacts with a generic target rather than with a single lepton, the differences between specific cases will be discussed subsequently.



(a) $\nu_\alpha + X \rightarrow \ell^{\pm}_{\alpha} + Y$

(b) $\nu_\alpha + X \rightarrow \nu_\alpha + Y$

Figure 1.1: Feynman tree-level diagrams for charged-current (a) and neutral-current (b) neutrino interactions.

In both cases, a neutrino interacts with the target matter X. For CC case, the mediator is a $W^{\mp}$ gauge boson and in the final state is produced a lepton with opposite charge of the boson and the same flavour of the interacting neutrino. Moreover, a hadronic system Y is generated after the interaction. In the NC case, the intermediate boson is a Z^0 and in the final state the lepton is replaced by a neutrino of the same flavour of the interacting one.

The dominant process is the neutrino-nucleon scattering and it is important to note that both the interaction cross-section and the final state products strongly depend on the neutrino energy [9]. For energies below 100 MeV, the dominant process is inverse β-decay, $\bar{\nu}_e + p \to e^+ + n$. In the energy range between approximately 0.1 and 20 GeV, initially elastic and quasi-elastic scattering processes dominate, while above 2 GeV the main contribute is given by resonance production in which a neutrino can excite a nucleon. At higher energies, above 20 GeV, the dominant process is deep inelastic scattering (DIS). Given that the minimum detectable energy by KM3NeT/ARCA is of the order of 100 GeV, the majority of detected events undergo this process, therefore, DIS will be discussed in greater detail.

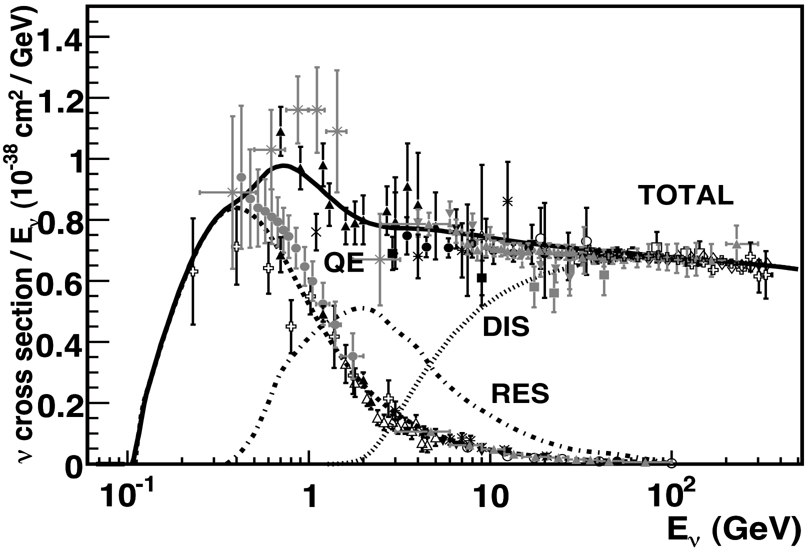


Figure 1.2: Neutrino-nucleon CC cross section as a function of the neutrino energy: total cross-section (black line), quasi-elastic scattering (dashed), resonance production (dot-dashed) and deep inelastic scattering (dotted). Figure taken from [9]

In Fig.1.2 the total neutrino-nucleon interaction cross section is shown, highlighting the different contributions that identify the energy ranges in which each process is dominant.

Within the DIS regime, the neutrino energy is sufficiently high to allow the interaction with the quarks composing the target nucleon. In particular, a neutrino scatters off a quark inside the nucleon through the exchange of a virtual $W^\pm$ or Z^0 boson, producing a lepton and an hadronic system in the final state, as shown in Fig.1.3.

In order to write the inclusive cross section for this process, it is useful to introduce four kinematic variables, based on the four-momenta of the initial and final particles (defined as in Fig.1.3), which are invariant under Lorentz transformations:

$$s = (k_1 + p)^2 \quad \text{(centre of mass energy)} \tag{1.4}$$

$$Q^2 = -q^2 = -(k_1 - k_2)^2 \quad \text{(four-momentum transferred)} \tag{1.5}$$

$$y = \frac{p \cdot q}{p \cdot k_1} \quad \text{(inelasticity)} \tag{1.6}$$

$$x = \frac{Q^2}{2p \cdot q} \quad \text{(Bjorken scaling variable)} \tag{1.7}$$

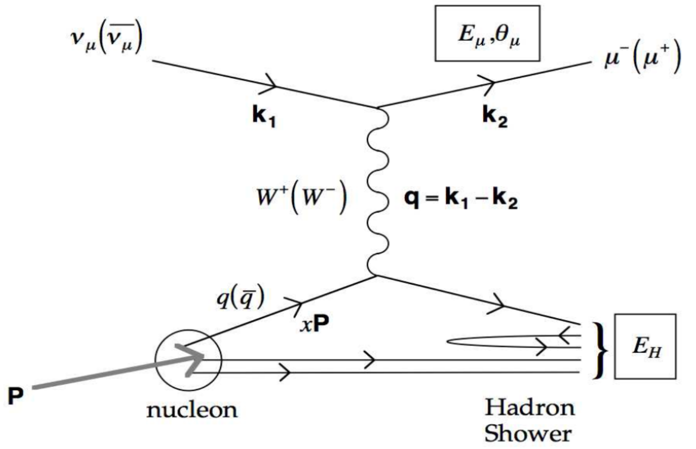


Figure 1.3: Feynman diagram for a CC neutrino DIS process involving a muon neutrino. Figure taken from [9]

Here, the inelasticity represents the fraction of the neutrino energy transferred to the target in the laboratory frame, while the Bjorken scaling variable represents the fraction of the nucleon momentum carried by the quark involved in the interaction. The differential cross section for DIS of neutrinos and antineutrino is given by:

$$\frac{\mathrm{d}^2\sigma^{\nu,\bar{\nu}}}{\mathrm{d}x\,\mathrm{d}y} = \frac{G_F^2\, M\, E_\nu}{\pi\left(1 + Q^2/M_{W,Z}^2\right)^2}\left[\frac{y^2}{2}\,2xF_1(x,Q^2) + \left(1 - y - \frac{Mxy}{2E_\nu}\right)F_2(x,Q^2) \pm y\left(1 - \frac{y}{2}\right)xF_3(x,Q^2)\right] \tag{1.8}$$

where M is the nucleon mass, $M_{W,Z}$ is the mass of the exchanged gauge boson, the $(-)+$ sign refers to (anti)neutrino scattering, and $F_i\left(x, Q^2\right)$ with $i = 1,\, 2,\, 3$ are the nucleon structure functions, which encode the probability of finding a quark carrying a fraction x of the nucleon momentum.

In principle, neutrino can also scatter off atomic electrons, however this process is

strongly suppressed due to the small mass of the electron. Nevertheless, it is worth mentioning that at very high energies the so called Glashow resonance is observed, peaked at 6.32 PeV [10, 11], in which an electron antineutrino resonantly scatters off an electron.
From Eq.(1.8), it is relevant to note that the cross-section increases with the neutrino energy. Moreover, integrating over neutrino energies above 100 GeV, the following ratio can be extracted:

$$\frac{\sigma_{CC}}{\sigma_{NC} + \sigma_{CC}} \sim 0.7 \tag{1.9}$$

which indicates that charged current interactions account for approximately 70% of all neutrino-nucleon interactions.

1.2 Neutrino oscillations

The first, and so far only, experimental evidence of Physics beyond the Standard Model (BSM) is provided by neutrino masses, first observed through neutrino oscillations by the Super-Kamiokande and MACRO experiments [5, 6]. Both were focused on the detection of atmospheric neutrinos, which revealed a deficit of ν_μ-like events compared to expectations from Monte Carlo simulations, as depicted in Fig.1.4, showing the results of SK.

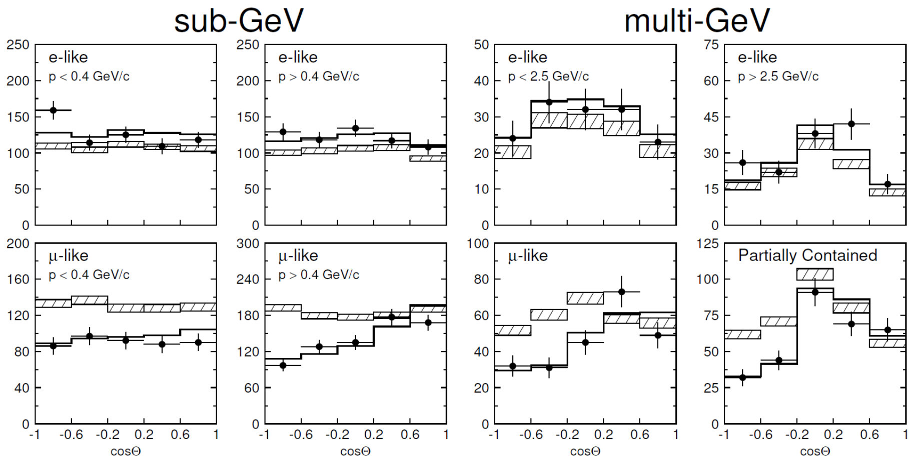


Figure 1.4: Zenith angle distributions of μ-like and e-like events for sub-GeV and multi-GeV data sets measured by SK. Upward-going particles have $\cos\Theta < 0$ and downward-going particles have $\cos\Theta > 0$. The hatched region shows the Monte Carlo expectation for no oscillations normalized to the data livetime with statistical errors. The bold line is the best-fit expectation for $\nu_\mu \leftrightarrow \nu_\tau$ oscillations. Figure taken from [5]

In particular, this discrepancy appears in the upward-going region (i.e. neutrinos originating from the opposite side of the Earth and traversing it), corresponding to the distance required for $\nu_\mu \to \nu_\tau$ oscillations to occur.

The idea that neutrinos can oscillate between $\nu - \bar{\nu}$ was first proposed by Pontecorvo [12], in analogy of the oscillation in the $K_0 - \bar{K}_0$ system [13]. Lately, this concept was extended to oscillations that arise from the interference between different massive neutrino states, implying oscillations in flavour states. To describe oscillation phenomena, a neutrino produced with flavour α is described as a superposition of three mass eigenstates $|\nu_k\rangle$ [14]:

$$|\nu_\alpha\rangle = \sum_{k=1}^{3} U^*_{\alpha k}\, |\nu_k\rangle \quad (\alpha = e,\, \mu,\, \tau) \tag{1.10}$$

The matrix $U^*_{\alpha k}$ is known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix and rules the mixing between flavour and mass eigenstates. It can be represented as:

$$\begin{aligned} U &= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \qquad (1.11) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\alpha_1} & 0 & 0 \\ 0 & e^{i\alpha_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

The most common representation involves three mixing angles, θ_{12}, θ_{13} and θ_{23} defined through $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$, and a single complex phase δ which takes into account CP-violation. Are also present two complex Majorana phases, α_1 and α_2, to include Majorana neutrinos [15].

A neutrino created with a definite flavour α at time $t = 0$ evolves as:

$$|\nu_\alpha(t)\rangle = \sum_{k=1}^{3} U^*_{\alpha k}\, e^{-iE_k t}\, |\nu_k\rangle \tag{1.12}$$

and, substituting $|\nu_k\rangle$ using Eq.(1.10), one obtains:

$$|\nu_\alpha(t)\rangle = \sum_{\beta=e,\mu,\tau} \left(\sum_{k=1}^{3} U^*_{\alpha k}\, e^{-iE_k t} U_{\beta k} \right) |\nu_\beta\rangle \tag{1.13}$$

As a consequence, the superposition of mass eigenstates becomes a superposition of different flavour eigenstates for $t > 0$.

Then, the transition amplitude for $\nu_\alpha \to \nu_\beta$ is therefore:

$$A_{\nu_\alpha \to \nu_\beta}(t) \equiv \langle \nu_\beta | \nu_\alpha(t) \rangle = \sum_{k=1}^{3} U^*_{\alpha k} U_{\beta k}\, e^{-iE_k t} \tag{1.14}$$

and the corresponding transition probability is given by:

$$P_{\nu_\alpha \to \nu_\beta}(t) = |A_{\nu_\alpha \to \nu_\beta}(t)|^2 = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \, e^{-i(E_k - E_j)t} \tag{1.15}$$

After appropriate manipulations (more on [14]) which take into account ultrarelativistic neutrinos, the oscillation probability in vacuum can be written as:

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta}(L, E) = \delta_{\alpha\beta} - 4 \sum_{k>j} \mathfrak{Re} \left(U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \right) \sin^2 \left(\frac{\Delta m^2_{kj} L}{4E} \right) \\ +2 \sum_{k>j} \mathfrak{Im} \left(U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \right) \sin \left(\frac{\Delta m^2_{kj} L}{2E} \right) \end{aligned} \tag{1.16}$$

where L is the distance traveled by the neutrino (called baseline) in time t, E its energy and $\Delta m^2_{kj} \equiv m^2_k - m^2_j$ is the squared mass difference between two mass eigenstates. It is evident that the oscillation probability in vacuum at a certain energy strongly depends on the mass splitting and the oscillation length. For neutrinos propagating in matter, an additional potential term must be included in the Hamiltonian.

Oscillation experiments are sensitive only to mass squared differences and not to the absolute neutrino masses. Moreover, only two independent mass splitting exist, since $\Delta m^2_{kj} + \Delta m^2_{ji} = \Delta m^2_{ki}$. Solar neutrinos observations have established that Δm^2_{21} is positive [16], while the sign of Δm^2_{32}, or equivalently Δm^2_{31}, remains unknown. This leads to two possible neutrino mass orderings, the normal ordering ($m_1 < m_2 < m_3$) and the inverted ordering ($m_3 < m_1 < m_2$), as shown in Fig.1.5.

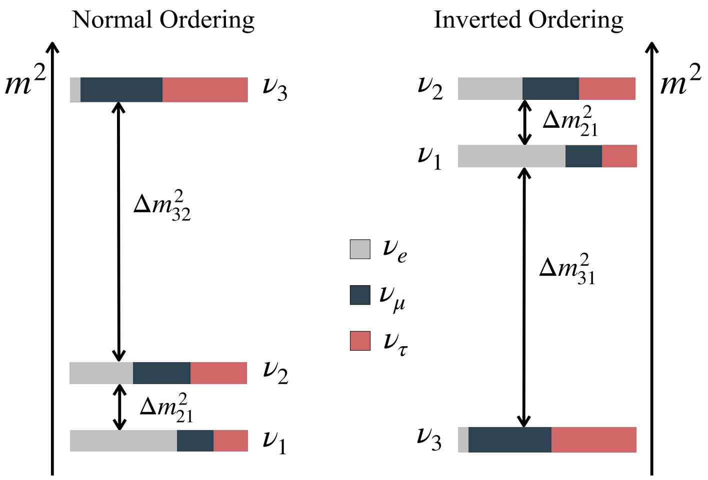


Figure 1.5: Normal Ordering (left) and Inverted Ordering (right) of neutrino mass eigenstates. The color code represents the relative contribution of the flavour eigenstates to the mass eigenstates. Figure taken from [17]

This is one of the biggest open questions in neutrino physics, known as neutrino mass hierarchy, which is among the main goals of the KM3NeT/ORCA experiment.

1.3 Cosmic Rays

Cosmic Rays (CRs) are high-energy protons, nuclei and a small fraction of electrons, produced in different astrophysical phenomena, that travel through the Universe. They hit the top of the Earth's atmosphere at the rate of about 1000 particles per square meter per second. Their composition is dominated by hadrons, such as protons (87%), helium nuclei (12%) and heavier nuclei (1%), while leptons constitute only a very small fraction (about two orders of magnitude less than hadrons).

The CR flux spans several orders of magnitude in energy and can be described by a power law [18] as:

$$\Phi(E) = K \cdot E^{-\alpha} \quad [\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}] \tag{1.17}$$

where K is a normalization factor and α is the spectral index, which characterizes the different regions of the spectrum (see Fig.1.6).

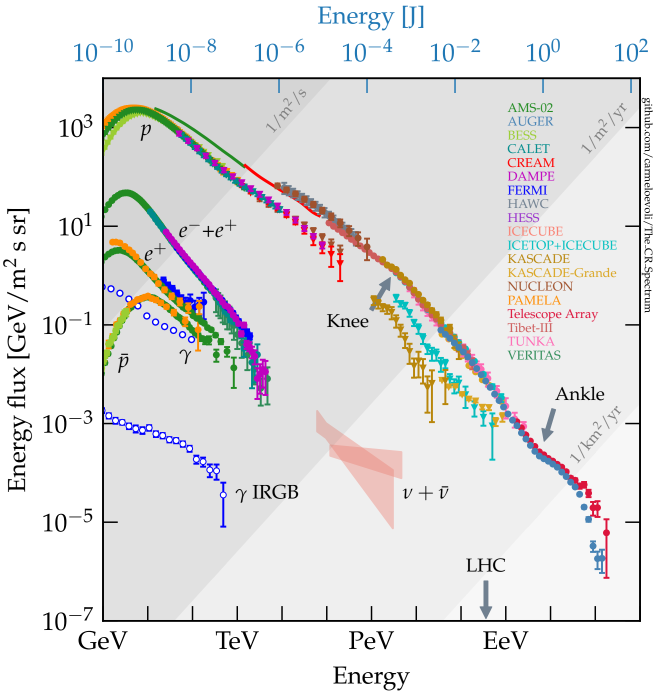


Figure 1.6: CRs differential energy spectrum, measured by multiple experiments. Figure taken from [19]

At least three regions with different spectral index can be identified;

- $E \lesssim 3 \cdot 10^{15}$ eV, with $\alpha \sim 2.7$. CRs up to this energy, known as the knee region of the spectrum, are generally believed to be produced by sources located within our Galaxy;
- $3 \cdot 10^{15}\,\text{eV} \lesssim E \lesssim 3 \cdot 10^{19}\,\text{eV}$, with $\alpha \sim 3.1$. In this energy range, which ends in the so called ankle region, a transition from Galactic to extra-Galactic sources is expected;
- Above $3 \cdot 10^{19}$ eV, with $\alpha \sim 2.6$. In this region the spectrum flattens and the sources are considered to be extra-Galactic.

Moreover, a cutoff is observed at the highest energies, around 50 EeV. The most widely accepted explanation for this phenomena is the so called Greisen-Zatsepin-Kuzmin (GZK) effect [20]. With increasing energy, the dominant contribution to the cosmic ray flux shifts from Galactic to extra-Galactic sources. This transition arises because only particles with sufficiently high energy can escape confinement by the Galactic magnetic field and propagate beyond the Galaxy. The propagation of CRs at different energies in the presence of a Galactic magnetic field is illustrated in Fig.1.6.

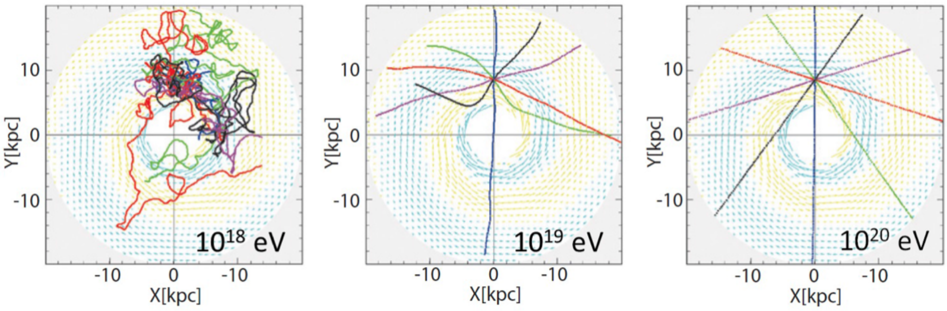


Figure 1.7: Simulated trajectory of CRs in the Galactic magnetic field. The three energy regimes are chosen to highlight how CRs at energies above $\sim 10^{19}$ eV travel along nearly straight lines and are not trapped inside the Galaxy. Figure taken from [21]

Currently, the Fermi acceleration mechanism [22] is the most commonly used model to explain the acceleration of cosmic rays. However, it is unable to explain the entire energy spectrum and the exact mechanisms responsible for the origin of CRs remain an open question.

1.4 Atmospheric neutrinos

When CRs enter the Earth's atmosphere, they propagate until they collide with nuclei present in the atmosphere, at an altitude of approximately $10 - 20$ km, producing cascades of secondary particles known as extensive air shower. In order to describe the production of atmospheric neutrinos, it is necessary to examine the composition of air shower, which can be divided into three components (see Fig.1.8);

- Hadronic: This forms the core of the air shower and consists of protons, neutrons, pions, kaons and heavier hadrons. These particles feed the electromagnetic component, for example through the decay of neutral pions into photons. Moreover, charged pions and kaons decay into muons, contributing to the muonic part;

- Muonic: This component is composed of muon and muon neutrinos produced through the decay of charged pions and kaons;

- Electromagnetic: This consists of electrons, positrons and photons that are produced mainly through the decay of muons and neutral pions. Photons can also be produced via lepton bremsstrahlung.

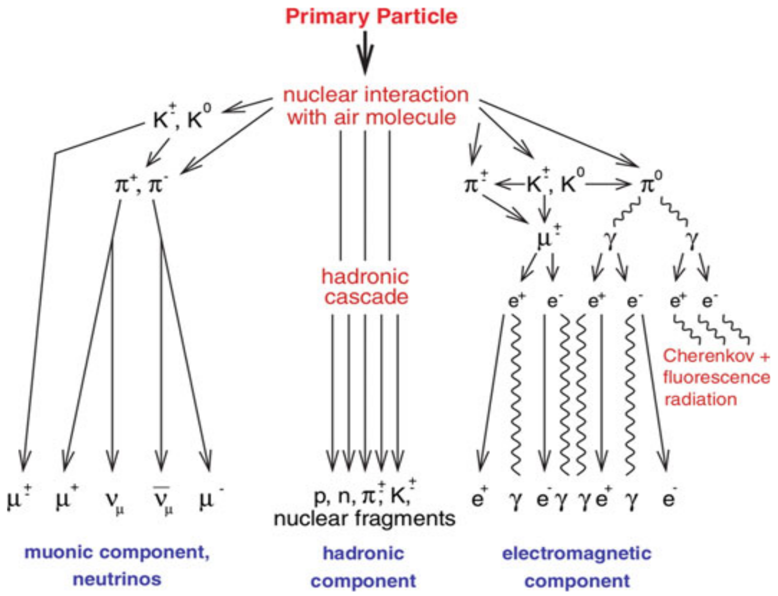


Figure 1.8: Schematic representation of an extensive air shower induced by a cosmic ray particle. The secondary particles in the shower can be grouped into three components: the muonic component including neutrinos, the hadronic and the electromagnetic component. Figure taken from [23]

The most relevant channels for the atmospheric neutrino productions, with the relative branching ratio [24], are:

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \quad [BR \sim 100\%] \tag{1.18}$$

$$K^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \quad [BR \sim 63.5\%] \tag{1.19}$$

and subsequently also muon decay produces neutrinos:

$$\mu^{\pm} \rightarrow e^{\pm} + \bar{\nu}_{\mu}(\nu_{\mu}) + \nu_{e}(\bar{\nu}_{e}) \quad [BR \sim 100\%] \tag{1.20}$$

The production of ν_{τ} is mainly due to τ decays, but its contribute to the atmospheric flux is negligible compared to the other two flavours. These processes allow a simple but still accurate description of the neutrino flavour production ratio, $\nu_e : \nu_{\mu} : \nu_{\tau} = 1 : 2 : 0$. Considering a generic meson 2-body decay into a muon and a neutrino (as in Eq.(1.18) and Eq.(1.19)), limits on the energy of the secondary muons and neutrinos, in the laboratory frame, can be respectively derived:

$$\frac{m_{\mu}^2}{M^2} \cdot E \leq E_{\mu} \leq E \quad \text{and} \quad 0 \leq E_{\nu} \leq (1 - \frac{m_{\mu}^2}{M^2}) \cdot E \tag{1.21}$$

where E is the lab energy of the decaying meson and M its mass. Therefore, numerically:

$$\langle E_{\mu} \rangle / E_{\pi} = 0.79 \quad \text{and} \quad \langle E_{\nu} \rangle / E_{\pi} = 0.21 \tag{1.22}$$

$$\langle E_{\mu} \rangle / E_{K} = 0.52 \quad \text{and} \quad \langle E_{\nu} \rangle / E_{K} = 0.48 \tag{1.23}$$

which explain why at high energy the pions contribution to neutrino events is suppressed and kaons decay start to be relevant.

Moreover, when the muon energy exceeds approximately 2.5 GeV, its decay length becomes larger than the typical production height in the atmosphere (about 15 km). As a consequence, the number of electron neutrino decreases at high energies (see Fig.1.10), and their production is dominated by kaon decays:

$$K_L^0 \rightarrow \pi^{\pm} + e^{\mp} + \bar{\nu}_e(\nu_e) \tag{1.24}$$

$$K^{\pm} \rightarrow \pi^0 + e^{\pm} + \nu_e(\bar{\nu}_e) \tag{1.25}$$

Fig.1.9 shows different measurements of neutrino flux for ν_e and ν_{μ} components provided by different experiments.

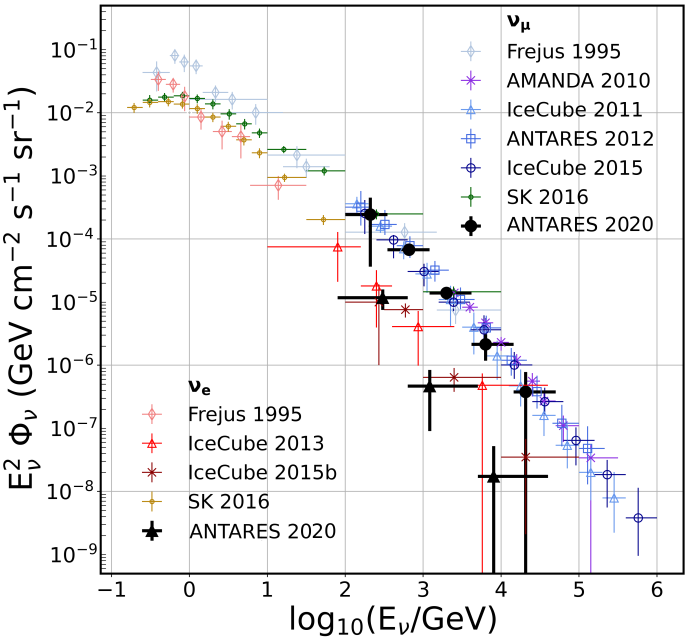


Figure 1.9: Measurement of the atmospheric ν_e and ν_μ fluxes reported by various experiments. Figure taken from [25]

The objective of this thesis is to perform the first measurement of the muon neutrino (ν_μ) component using KM3NeT/ARCA data in the energy range from 100 GeV to 100 TeV.

The atmospheric flux consists of two components, the conventional and the prompt components. The conventional component originates from pion and kaon decays, while the prompt component is mainly produced by the decay of charm mesons as D mesons. From low energies up to several tens of TeV for ν_e and up to a few PeV for ν_μ, the conventional flux dominates the spectrum. At higher energies, the prompt flux becomes significant. Since atmospheric neutrinos are produced by CRs interaction, both components follow a power law behavior similar to Eq.(1.17). The conventional flux has a spectral index of approximately $\alpha \sim -3.7$, while the prompt flux has $\alpha \sim -2.7$. As a result, at high-energy the flux is strongly suppressed, making the construction of detectors with large volume necessary to detect the limited number of incident events.

The flux is also modulated as a function of the zenith angle. Due to the differences in path length through the atmosphere, high-energy muons produced in vertical directions are more likely to reach the ground before decaying, while muons produced near the horizon have a longer path and therefore decay more frequently, as described in Eq.(1.20). As a result, the neutrino flux near cosine of zenith angle equal to ±1 is lower compared to that near the horizon (cosine of zenith angle equal to zero) (see Fig.1.10).

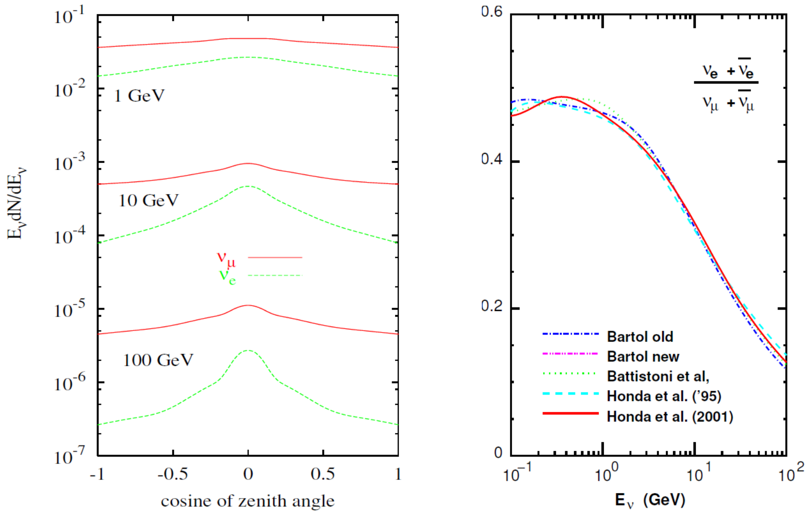


Figure 1.10: *Left:* Zenith-angle dependence of ν_e and ν_μ. *Right:* $(\nu_\mu + \bar{\nu}_\mu)/(\nu_e + \bar{\nu}_e)$ ratio as a function of energy. Figures taken from [26]

To complete the description of the atmospheric neutrino flux, it is necessary to consider neutrino oscillations. From Eq.(1.16), it is evident that the oscillation probability depends on the ratio between the traveled distance and the neutrino energy. Considering the disappearance of ν_μ into ν_τ, downward-going neutrino do not have a sufficiently long baseline (L) to undergo significant oscillations. In contrast, upward-going neutrinos traverse a baseline of approximately 13000 km (the Earth's diameter), leading to a maximum oscillation probability at energies of the order of tens of GeV.

1.5 Neutrinos in the Universe

The Universe is permeated by neutrinos which, due to their extremely small interaction cross-section with matter, are highly elusive particles. To illustrate how weakly they interact, consider electron anti-neutrinos produced in nuclear reactors with ~ 1 MeV energy, which have a cross-section of $\sim 10^{-44}$ cm^2. This implies that only one neutrino out of 10^{11} would interact while traveling along the Earth's diameter [27].

As will be discussed, detecting neutrinos with appreciable event rates requires detectors with huge interaction volumes.

However, this small interaction cross-section also represents an extraordinary opportunity: high-energy neutrinos produced within astrophysical objects can propagate through the Universe essentially unmodified. By combining neutrino observations with traditional astronomical messengers such as γ-rays, cosmic rays and gravitational waves, it becomes possible to study the same astrophysical event or system from multiple perspectives. This approach is known as multimessenger astronomy.

Is assumed that most of the observable cosmic neutrinos comes from astrophysical mechanism which involves cosmic ray collisions [28]. The main processes involved in neutrino production in the Universe are;

- Hadronic interaction (pp): in which nucleon-nucleon collisions produce pions $\left(\pi^0,\, \pi^+,\, \pi^-\right)$ or kaons that eventually can decay into neutrinos, similarly to the atmospheric case (see Eq.(1.24) and Eq.(1.25)). Experimental observations shows that the leading pion carries on average 1/5 of the initial kinetic energy of the proton;
- Pion photo-production ($p\gamma$): in which CRs interact with high-energy gamma rays producing pions which decay into neutrinos. Also in this case the secondary pions carry about 1/5 of the initial kinetic energy.

This latter process is expected to occur in the environments of powerful astrophysical systems. Several classes of astrophysical objects are currently considered as potential sources of CRs and, consequently, cosmic neutrinos. We can briefly mention some of the most relevant;

- Supernovae: explosive deaths of massive stars that release enormous amounts of energy and can accelerate charged particles through shock waves;
- Active Galactic Nuclei (AGN): extremely energetic central regions of galaxies, capable of accelerating particles in relativistic jets;
- Gamma Ray Bursts: short and intense flashes of gamma rays generated by collision of compact objects or collapse of massive stars into Black Holes;
- Galaxy Clusters: the largest gravitationally bound structures in the Universe, made of thousands of galaxies and plasma. The matter inside these environments can accelerate protons and produce hadronic cascades.

Chapter 2

The KM3NeT neutrino telescopes

In this Chapter, a first introduction to the concept behind the construction of massive underwater neutrino detectors is given. Then, the detection method is described, in which an interacting neutrino produces a charged lepton that, traveling in water, gives rise to Cherenkov radiation. An exhaustive description of the KM3NeT/ARCA detector is subsequently provided. The Chapter ends with the presentation of the various types of signals and the description of the background.

2.1 Underwater detector

As discussed in Section 1.5, neutrinos have a very small cross-section compared to other particles, therefore, in order to achieve a statistically significant number of events, a huge volume of target material is required. Moreover, another delicate aspect of atmospheric and cosmic neutrino detection is that cosmic rays constitute the major source of background from which the detector must be shielded as much as possible. For these reasons, the deep sea represents an ideal location, since it provides an enormous, low-cost target, and effective shield from cosmic rays and a natural Cherenkov radiator.

The first neutrino telescope built in the deep sea was ANTARES [29], which operated in the Mediterranean Sea between 2008 and 2022. Based on the experience gathered, the KM3NeT (*Cubic Kilometer Neutrino Telescope*) experiment [30] was established. It currently comprises a network of Cherenkov detectors located at different locations in the Mediterranean Sea with distinct scientific goals (see Fig.2.1);

- KM3NeT/ARCA: stands for "*Astroparticle Research with Cosmics in the Abyss*". It is located ∼ 100 km off-shore from the coasts of Sicily at 3500 meters under the sea level. Its final configuration will cover a volume larger than one cubic kilometer. Its main goal is the detection of astrophysical neutrinos in the TeV−PeV energy range and the identification of neutrino sources;

- KM3NeT/ORCA: stands for "*Oscillation Research with Cosmics in the Abyss*". It is located ∼ 40 km off-shore from the coasts of France, near Toulon, at a depth of 2500 meters. Its final configuration will instrument ∼ 7 Mton of seawater. Its main goal is the detection of atmospheric neutrinos in the few-GeV energy range, in order to determine the neutrino mass hierarchy.

The two detectors exploit the same construction technology (see Section 2.3), namely a three-dimensional array of photomultiplier tubes. The main differences lies in the horizontal spacing between the strings and the vertical distance of optical modules within the string, which are optimized for the energy range of interest. Currently, both detectors are under construction, their completion is foreseen by the end of decade.

Figure 2.1: Detector sites of the KM3NeT Collaboration. The KM3NeT-Gr indicates the possible place of a future third detector. Figure taken from [31]

2.2 Detection method: the Cherenkov effect

The detection principle at the base of the KM3NeT detectors is the detection of the Cherenkov light produced by charged secondary particles generated in neutrino interactions. Precisely, the Cherenkov effect occurs when a charged particle propagates in a medium with refractive index n at a speed greater than the speed of light in that medium ($v = c/n$, where c is the speed of light in vacuum) [21]. In this case, the charged particle polarizes the molecules of the medium along its trajectory, whose subsequent de-excitation generates a coherent radiation. The emitted photons form a characteristic conical shape (see Fig.2.2) with an angle that depends on the particle velocity and the refractive index:

$$\cos(\theta_c) = \frac{c/n(\omega)}{\beta c} = \frac{1}{\beta n(\omega)} \tag{2.1}$$

For relativistic particles ($\beta \sim 1$) propagating in sea water ($n \sim 1.35$), the Cherenkov angle is approximately $\theta_c \sim 42°$, almost independent of the particle energy.

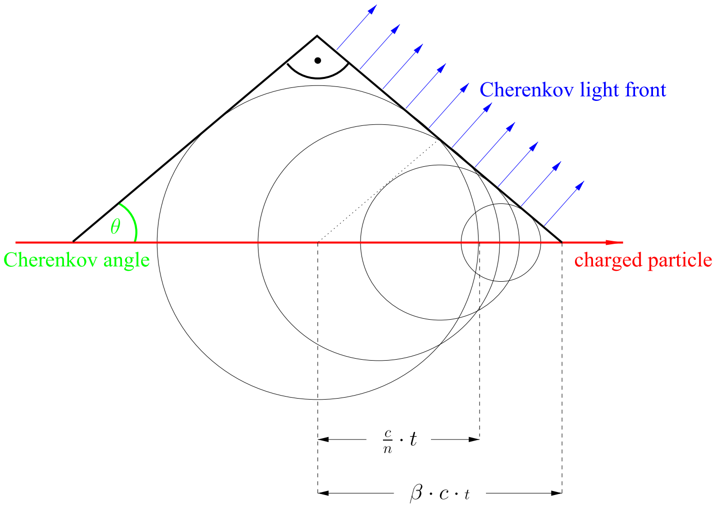


Figure 2.2: Schematic representation of Cherenkov light emission, where θ is the characteristic Cherenkov angle. The red arrow marks the propagation of the charged particle in the medium, while the blue arrows represent the propagation of the coherent wavefront of Cherenkov light. Figure taken from [32]

Between the emission point along the charged particle track and the detection point at the PMT, photons mainly undergo two processes: absorption and scattering. For an initial radiation flux ϕ_0, the light reaching the detector is attenuated as a function of the absorption length and the path length, resulting in:

$$\phi' = \phi_0 \cdot e^{-d/\lambda_{abs}} \tag{2.2}$$

Scattering changes the photon path and consequently alters the arrival time at the PMT. It is essential to distinguish between photons that propagate without scattering, referred as direct, and those that undergo through single or multiple scattering processes. Scattered photons, due to multiple changes in direction and partial energy loss, induce a smearing of time arrival reconstruction, unlike direct photons that bring important information for particle direction reconstruction. This distinction is crucial and will be at the base of the event selection developed in the context of this thesis to reach the required purity of the data sample.

2.3 Detector descriptions

The KM3NeT detectors consist of a three-dimensional array of Digital Optical Modules (DOMs) [33]. The DOMs are arranged in groups of 18 along two supporting ropes, forming vertical structures anchored to the seabed and kept taut by a buoy at the top. These vertical structures are called Detection Units (DUs). Each DU has a Base Module (BM) located at the anchor and is connected to Junction Boxes (JBs), which are connected to the electro-optical cable responsible for supplying power and connectivity with the shore station via optical fibers.

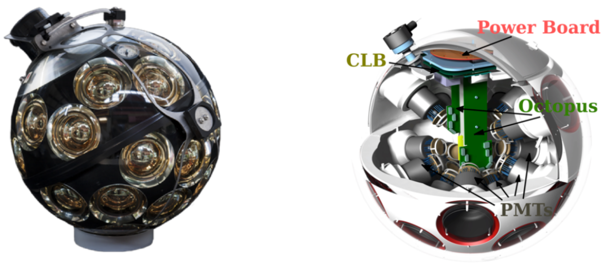


Figure 2.3: *Left:* picture of the KM3NeT Digital Optical Module. *Right:* schematic representation of the components inside a DOM. Figure taken from [31]

Thanks to the innovative design developed by the KM3NeT Collaboration (see Fig.2.3), the DOM contains 31 three-inch PMTs within a glass sphere of 0.44 m diameter. The PMTs are distributed over the DOM surface, providing nearly 4π solid angle coverage. They are arranged in circles: 19 PMTs are located in the lower hemisphere and 12 in the upper one. Moreover, all electronics for power supply and data acquisition are housed inside the module. An aluminum cap is placed on the top of the DOM, designed to dissipate the heat generated inside

the module. Since sea currents affect the DOM positions, each module has a calibration devices measuring tilt, pitch and yaw as well as a piezoelectric sensor capable of detecting signals generated by beacons, located at the DU base. Together with hydrophones anchored in the DU base, these elements form the acoustic positioning system, which can determine the positions of the DOMs with an accuracy better than 10 cm. A Central Logic Board (CLB) housing a Kintex-7 FPGA is the electronic board housed inside each DOM responsible for signal digitisation, data transmission and sensor control.
This multi-PMT configuration offers several advantages over the traditional single-PMT designs. Indeed, the photocathode area is comparable to the three optical modules of ANTARES that make up a detector floor, providing a more detailed directional information about the detected photons and the possibility to correlate signals of different PMTs within the same module greatly improves the event trigger.

Detection units are deployed using a Launcher of Optical Modules (LOM) [34], an aluminum structure that contains the rolled-up DU, anchor and buoy. Once the LOM reaches the seabed, the DU is released and the launcher is recovered in order to be used in next installations.

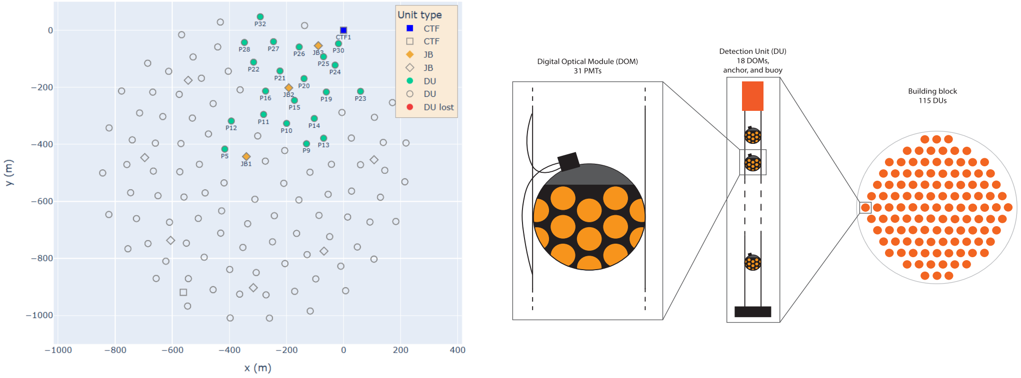


Figure 2.4: *Left:* Green dots represent the KM3NeT/ARCA21 detector footprint. The not colored ones represent the missing DU in order to complete the first ARCA building block. Figure taken from [35]. *Right:* Schematic representation of DOM, DU and an ARCA building block. Figure taken from [36]

Focusing on KM3NeT/ARCA layout, which is the detector used in this thesis, it will consist of two Building Block with 115 DUs each. The vertical distance between DOMs in a DU is ~ 36 m and the horizontal spacing between DUs is ~ 100 m. In particular, the present analysis focuses on the ARCA21 configuration, which comprises 21 DUs (see Fig.2.4). The total period of data-taking for this specific configuration is approximately 287.4 days.

One of the main experimental challenges for KM3NeT is to maintain and operate many optical modules for more than 20 years at thousands of meters below the sea level. Therefore, the probability of failures must be minimized. For this reason, only essential instrumentation is included within the DOM. The “all data to shore” approach is adopted for data transmission. This means that all signals are transmitted to shore station, where a PC farm processes the hits, applying dedicated trigger algorithms in order to reduce the data flow, storing only potentially interesting information.

2.4 Signal and background topologies

The light pattern observed in the detector depends on the neutrino interaction type. As discussed in Section 1.1, neutrinos undergo either neutral or charged current interactions. Event signatures in water can be classified into two main categories (see all interactions in Fig.2.5);

- Track-like events: are generated by a muon passing through the detector along a straight line. Due to the information left by the Cherenkov light emission along the straight line, its direction can be reconstructed with sub-degree resolution. At the same time, the energy reconstruction is more uncertain because part of the produced light can be deposited outside the detector. These events are mainly dominated by ν_μ CC interaction with a very small contribute from ν_τ CC;

- Shower-like events: all the other neutrino interactions produce secondary particles that propagate only for a few tens of meters, generating electromagnetic and hadronic cascades. Considering the average distance of the KM3NeT/ARCA optical modules, the light emitted in the cascade development can be approximated as an expanding spherical surface. For this reason direction reconstruction is less precise. On the other hand, almost all the light is deposited inside the detector, improving the accuracy of energy reconstruction.

The ν_μ CC interaction represents the golden channel for neutrino source identification. Thanks to the precise direction reconstruction of the produced muons also the neutrino direction can be reconstructed with high precision. The production angle θ_{ν_μ} between the neutrino and muon direction is approximately:

$$\theta_{\nu_\mu} \sim \frac{0.6^\circ}{\sqrt{E_\nu(\mathrm{TeV})}} \tag{2.3}$$

and above few TeV, the angular resolution is dominated by reconstruction algorithm uncertainties.

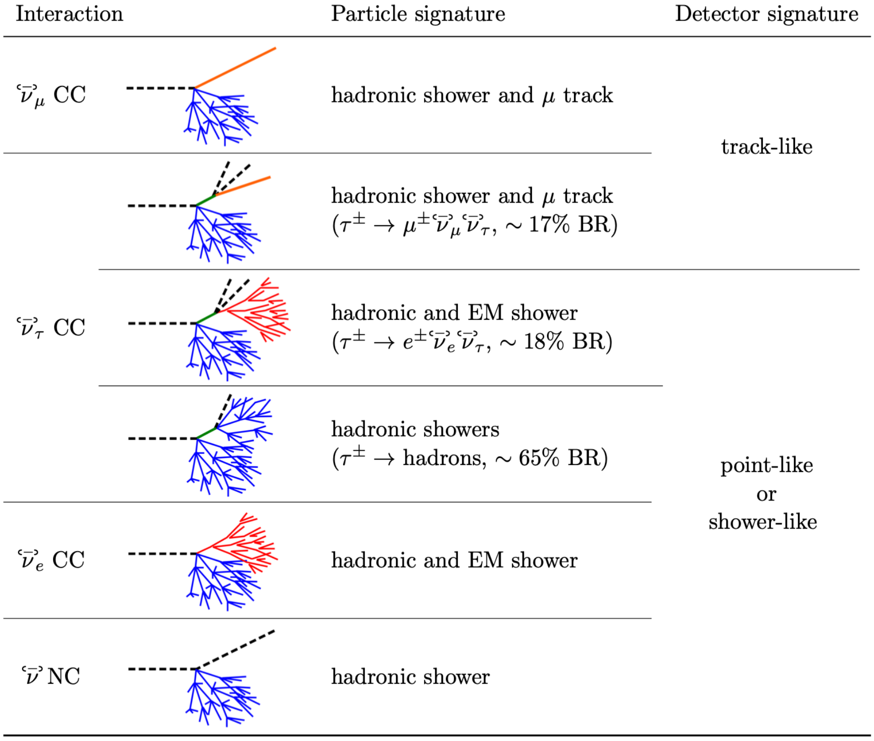


Figure 2.5: Comprehensive representation of all the event topologies occurring inside a neutrino telescope. Black dashed lines represent neutrinos, orange solid lines muons, red electromagnetic showers, blue hadronic showers, and green τ leptons. Figure taken from [21]

For shower-like events, as in the case of ν_e CC interaction, the radiation length X_0 of the high-energy electron released is typically about 36 cm. The hadronic interaction length in water is compatible, about 83 cm, and for this reason these two showers can be treated similarly. Therefore, the light emission can be approximated as point-like relative to DOM and DU spacing. Moreover, as long as the charged particles produced in the cascade exceed the Cherenkov threshold, they continue to produce light signals.

In addition to these topologies, there is another one that it is worth mentioning, the so called double bang event. It occurs when a ν_τ CC interaction occurs, producing a first hadronic shower and a τ lepton, which in turn decays producing a second hadronic or electromagnetic shower.

Background

In the context of underwater neutrino telescope it is appropriate to divide the background into environmental and physics background.

In the deep-sea the environmental background is mainly composed by radioactive decay and bioluminescence. Sea water contains about ~0.04% of radioactive isotopes, such as Potassium-40, which decay as:

$$^{40}K \rightarrow {}^{40}Ca + e^- + \bar{\nu}_e \quad [BR = 89.3\%] \tag{2.4}$$

$$^{40}K + e^- \rightarrow {}^{40}Ar + \nu_e + \gamma \quad [BR = 10.7\%] \tag{2.5}$$

In the case of (2.4) decay, the emitted electron can reach up to 1.33 MeV and produce Cherenkov light. Instead, for (2.5) decay, the photon has an energy of 1.46 MeV and, via Compton scattering, can generate an electron over the Cherenkov threshold. This background is continuous and, with a PMT threshold of 0.25 photo-electrons, the average ^{40}K background rate is about ~5 kHz.

Also living organisms can produce visible light, the so-called bioluminescence. This background event rate can be efficiently suppressed requiring for casual-connected hits over distant optical modules.

The most challenging source of physical background is represented by atmospheric muons which, at sea level, have a flux about 10^{11} times larger than that of atmospheric neutrinos.

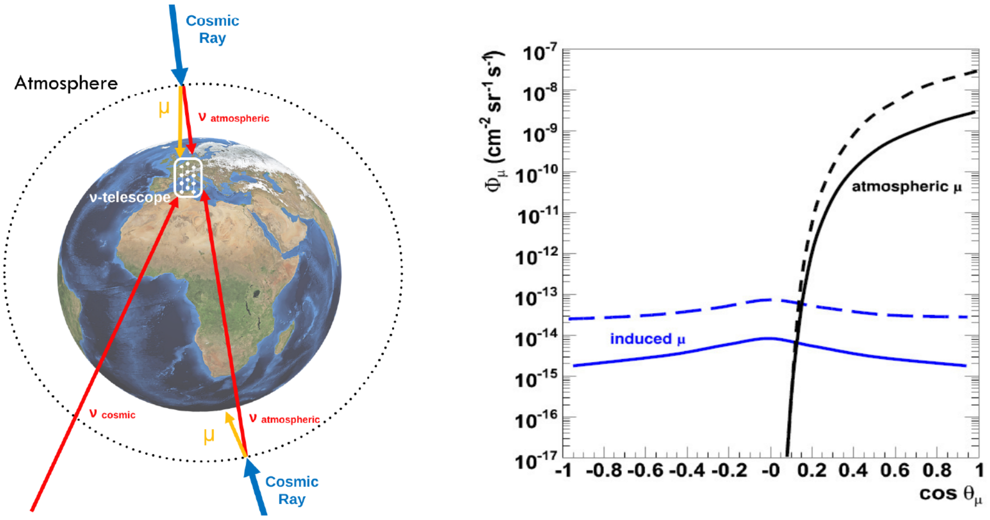


Figure 2.6: *Left:* Sketch of the muon and neutrino flux that can/cannot reach the location of a neutrino telescope. Figure taken from [37] *Right:* Different contributions (as a function of the cosine of the zenith angle) of the atmospheric muons for two different energy regimes (dashed: $E_\mu > 100$ GeV, while solid: $E_\mu > 1$ TeV) at a depth of 2400 m. Figure taken from [21]

As already mentioned in Section 2.1, deep water helps to shield the detector from primary and secondary cosmic rays, however, at the seabed the flux of atmospheric muons is still $\sim 10^6$ times larger than the neutrino flux. For this reason, a widely-used selection technique in neutrino telescopes is based on the fact that all the muons reaching the detector originate in the portion of atmosphere above it. Indeed, unlike neutrinos, muons cannot travel through the $\sim$ 13000 km of the Earth without interact (see Fig.2.6). Based on this consideration, all events leaving a signature from the bottom to the top of the detector, which we will refer to from now on as upward-going events, should be generated only by neutrinos.

For clarity, Fig.2.7 provides a pictorial representation of both signal and background events.

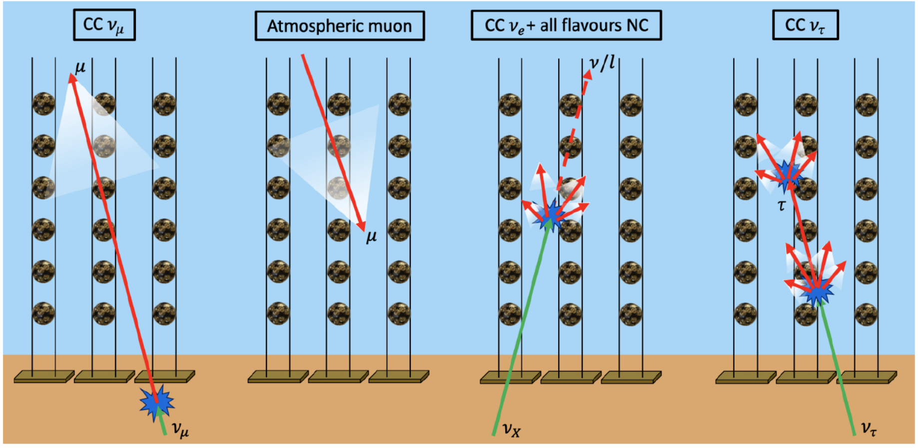


Figure 2.7: From left to right: ν_μ CC event (track-like); atmospheric muon event (track-like); NC neutrino or ν_e CC event (shower-like); ν_τ CC double bang event. Figure taken from [38]

Chapter 3

Monte Carlo simulations and event reconstruction

In this Chapter, an overview of the basic ingredients to perform Monte Carlo simulations in the context of KM3NeT/ARCA neutrino telescope is provided. In general, there are four fundamental steps: generation of neutrino and muon events, particle propagation through rock under the telescope and water, light production and propagation, detector response and event reconstruction.

3.1 Simulations

The starting point of Monte Carlo simulations lie on the algorithms used to generate particles. The gSeaGen code [39] is used to simulate events of all neutrino flavours undergoing both CC and NC interactions. gSeaGen is an open source library, developed by the KM3NeT Collaboration, based on GENIE package [40] and generating neutrino interaction from few GeV up to EeV scale. Muons events are simulated using the MUPAGE generator [41, 42], which is based on parametric formulas obtained from a full Monte Carlo simulation of cosmic ray showers and from measurements done with the MACRO experiment. MUPAGE generates downward-going single and bundles of muons, with the zenith angle ranging from 0 to 85 degrees.

Events are generated in proximity of the detector in order to avoid wasting of computational resources. This makes it necessary to introduce the concepts of instrumented volume, detector can and interaction volume. The instrumented volume corresponds to the region containing all the DOMs and is represented as a blue cylinder in Fig.3.1. The radius that includes all the instrumentation is defined as R_{det}. The same figure shows in yellow the detector can, defined as a cylinder whose bottom is at the seabed, with a height of Z_{can}^{max} which can be determined depending on the type of event involved, and a radius equal to R_{det} plus n times the light absorption length ($R_{can} = R_{det} + n \cdot L_a$). A typical value used by the KM3NeT Collaboration is $n = 5$. All events leading to Cherenkov radiation inside the detector can are stored in the output file.

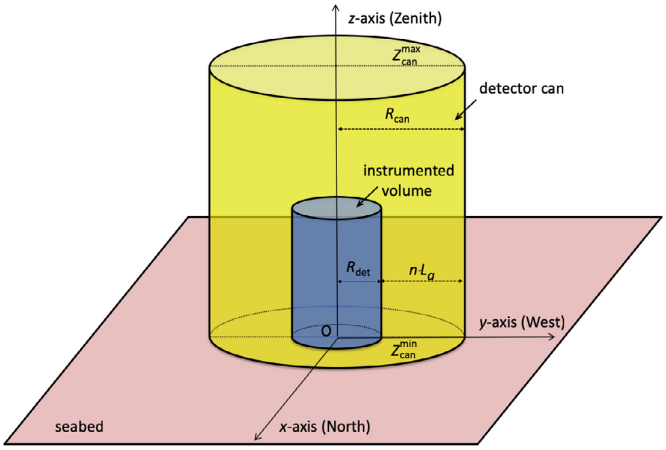


Figure 3.1: Definition of the detector can. Figure taken from [39]

The interaction volume (see Fig.3.2) is the region surrounding the detector in which a neutrino interaction can produce detectable particles. Unlike the detector can, the interaction volume is not fixed, as it depends on the topology and on the energy of the neutrino events involved.

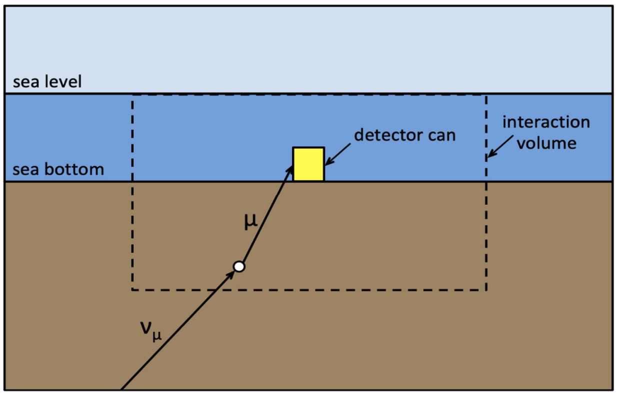


Figure 3.2: Schematic view of the interaction volume for track-like events. Figure taken from [39]

For ν_e CC and neutral current interactions due to all neutrino flavours, generating shower-like events, the interaction volume coincides with the detector can. In the case of ν_μ CC interactions, the leading muon can be produced outside the can and still reach the detector. For this reason, track-like events require an interaction volume that depends on the maximum muon range in water and rock, evaluated at the highest simulated neutrino energy. For ν_τ CC interactions, the contribution of muons produced by tau decays is also taken into account.

A key feature of the simulation framework is the *run-by-run* approach, developed for the ANTARES experiment [43] which takes into account variations of deep-sea environmental conditions and the detector status. All this information is extracted from recorded data and used as input for the simulations.

Considering the neutrino interaction probability and the steeply falling flux of atmospheric and cosmic neutrinos, in order to populate all the parameter space with enough statistics, neutrino events are oversampled and generated according to a $E^{-1.4}$ spectrum. In order to recover the physical event rate $f(E, \theta, \phi)^{phys}$ (both atmospheric or cosmic event rates can be derived from the same generated events), for each neutrino event a generation weight is associated according to:

$$w_{evt} = \frac{f(E, \theta, \phi)^{phys}}{f(E, \theta, \phi)^{gen}}$$

This procedure allows events to be generated efficiently and subsequently reweighted to any desired physical flux.

The simulation of the atmospheric muon background follows a different approach. The MUPAGE package provides kinematical details of muon bundles reaching the can surface according to parametric formulae [41, 42]. In this case the weight applied to each event takes into account only the livetime corresponding to each data-taking run.

Once the particles are generated on the can surface, they are propagated through the active volume of the detector using different propagators available on the market, such as MUSIC and PROPOSAL. At this point the light generation and propagation is simulated through a specific code, JSirene, developed by the KM3NeT Collaboration, based on pre-compiled tabulated probability distribution functions (PDFs) accounting for particle distance, PMT orientation, particle energy and direction.

3.2 Trigger and events reconstruction

The final step required to reproduce a complete event as recorded by the telescope, is the simulation of the detector response. To this end, trigger conditions based on space-time correlations between the registered photon hits must be defined. This procedure represents the first stage in the selection of physical events and in the reduction of the environmental background (see Section 2.4).

A hit is defined as the signal produced when a PMT collects photons that generate a voltage amplitude above a given threshold. The time duration of this pulse is called Time-over-Threshold (*ToT*) and is proportional to the number of photoelectrons. Within KM3NeT experiment, three different types of hit conditions are considered in the trigger system;

- L0 filter: each hit recorded with a typical threshold of 0.3 photoelectrons;
- L1 filter: coincidence of two or more L0 hits recorded by different PMTs of the same DOM within a given time window, typically ΔT equal to 10 ns;
- L2 filter: considering only PMTs that registered a L1 signal, a constraint is set on the angular separation between PMT orientations within the same DOM.

Two trigger algorithms have been developed and optimized for the two possible event topologies, track-like and shower-like events. The trigger condition for track-like event is satisfied with at least five correlated L2 hits. In shower-like case, since the light emission is only weakly dependent on the initial direction of the interacting particle, the trigger condition requires a minimum number of L2 hits and activated DOMs within a spherical volume, with radius related to the absorption length.

Events reconstruction

All triggered events are subsequently processed through the event reconstruction chain. Two reconstruction algorithms, optimized for the track-like and shower-like topologies, are employed. Since the event topology is not known a priori, both algorithms are applied to the recorded signals.

The track reconstruction aims at determining the direction and energy of muons, assuming their trajectory to be a straight line. It starts by performing a fixed number of pre-fits over an equally spaced grid of direction covering the full sky, in order to estimate the time and position of the muon track that minimize a χ^2 function which compares the observed hit times with the expected ones under the assumption of direct Cherenkov hit (see Section 2.4). Starting from this track and position hypothesis, a maximum likelihood method is applied to extract the final values of these parameters. The reconstruction concludes with an estimate of the muon energy based on the number and spatial distribution of the hit PMTs.

The shower reconstruction begins with the estimation of the position of the shower vertex, obtained by requiring coincidences of hits in the same DOM within a very narrow time window. This vertex is then used as the starting point for the reconstruction of the shower energy and direction, performed by minimizing a likelihood function based on the probability of observing or not a hit.

Chapter 4

Event Selection

In this Chapter, my work done on the event selection strategy is presented. The goal is to obtain a high purity sample of atmospheric ν_μ and $\bar{\nu}_\mu$ charged current events. The main source of background is represented by atmospheric muons, therefore different strategies to remove these events from the sample are developed. Initially, a fixed-cut selection based on a set of reconstructed variables is applied to the dataset. This serves also as a common starting point for all the KM3NeT analysis. In order to reach the required level of atmospheric muon contamination a machine learning algorithm has been trained.

4.1 Preselection

As mentioned in Section 2.4, the most relevant source of background is composed by atmospheric muons; during the full livetime of KM3NeT/ARCA21, one neutrino over fifteen thousands atmospheric muons is expected. The event selection procedure is organized into two steps. The first step is carried out through the use of cuts on reconstructed variables. These cuts have been developed by the KM3NeT Collaboration with the aim of removing environmental noise and part of the atmospheric muon background. In the following, we will examine these cuts in detail. The next section will focus on the second step, which consists in the development of a *Boosted Decision Tree* (BDT).

In the following, the two fundamental quantities that will drive the selection process are defined;

- Efficiency of the atmospheric neutrino signal ($\nu_{atm}\,$sgl): ratio between the number of atmospheric neutrino events at a specific cut level and at preselection level;

- Muon contamination (μ cont): percentage of muon events over the whole number of events at a specific cut level.

For the purpose of this thesis, the objective is to achieve a muon contamination level below 1%, trying also to retain the highest possible signal efficiency.

Data quality and run selection

A first selection is performed by identifying a sample of so called *good data-taking runs*. This run list includes only those runs that satisfy a list of specific criteria and thresholds define by the Collaboration such as;

- livetime between the upper (61200 s) and the lower (1200 s) limits;
- upper and lower limit on the trigger rate, appropriately tuned based on each specific detector geometry;
- upper and lower limit on the mean and root mean squared rate of PMTs, in order to avoid issues related to malfunctioning PMTs and bioluminescence activity;
- good and stable conditions for the electronic and time synchronization of the active elements in the detector.

About 10% of the total runs are discarded at this step. Once the data quality criteria have been applied, the dataset is ready for analysis.

	Preselection criteria
1	The fit of the reconstruction algorithm converge and complete all the steps
2	All reconstructed variables are generated
3	Positive track length
4	Positive angular error estimation
5	Likelihood greater than 50
6	More than 20 hits used in the reconstruction
7	Reconstructed energy above 100 GeV
8	Upward-going event: $\theta_{zenith} < 90°$
9	At least two DOMs triggered by hits compatible with Cherenkov hypothesis
10	Distance between first and last Cherenkov hits above 100 meters

Table 4.1: KM3NeT/ARCA21 preselection criteria.

The starting point for KM3NeT analyses is the application of what we will, from now on, refer to as *preselection*: a combination of physics driven cuts as well as criteria to select only the events that have been successfully reconstructed by the track algorithm (see Table 4.1).

Cuts (1) and (2) ensure that the events have been successfully reconstructed by the track algorithm, while cuts (3) and (4) guarantee that variables such as track length and angular uncertainty are physically meaningful, in particular, that they take values greater than zero. These criteria are designed to select events with a reliable reconstruction and good fit quality, meaning that the algorithm provides a stable and physically consistent description of the observed data.

A recent study by the Collaboration introduced new variables based on PMT hits whose arrival times are consistent with Cherenkov light propagation. Part of my work involved adapting the processing workflows to incorporate these new variables into the standard analysis files used within KM3NeT, as well as refining the variable definitions to account for the specific geometry of the ARCA21 detector. At this point, cuts (9) and (10) are introduced to select events that are more likely to be well reconstructed. Cuts (5) and (6) help to reject events generated by random hit coincidences over different optical modules, mainly generated by the environmental background. Cut (7) requires that the reconstructed energy must be greater than one hundred GeV. Finally a selection on the reconstructed arrival direction: since only neutrinos can traverse the Earth without interacting, an effective way to reject atmospheric muons is to select only reconstructed upward-going event (8). The survival downward-going atmospheric muons are therefore particles that have been wrongly reconstructed as upward-going.

Table 4.2 reports the number of events before and after the preselection. This step is clearly necessary, as it reduces the atmospheric muon background by approximately five orders of magnitude.

KM3NeT/ARCA21 - 287.4 days							
	ν_{atm}	ν_{cos}	μ_{atm}	DATA	μ cont	ν_{atm} sgl	DATA/MC
Full set	13868	–	$1,95 \cdot 10^8$	$1,91 \cdot 10^8$	~100%	–	0.98
Preselection	1236	12.9	7165	8039	85.2%	100%	0.96

Table 4.2: KM3NeT/ARCA21 number of events in the full data set and at preselection level.

However, these cuts are not sufficient to achieve the required level of muon contamination.

Another quantity that shows the quality of the sample is the ratio between the total number of data and Monte Carlo events. In order to perform a proper data analysis, it is necessary to ensure that the simulations faithfully reproduce real data.

4.2 Selection through *Boosted Decision Tree*

To improve the event selection, a *Boosted Decision Tree* technique is implemented. It belongs to the category of *supervised machine learning* algorithms [44, 45]. Decision trees are used to classify the events in the sample as either signal or background. Each event is defined by label, signal (s_i) or background (b_j), its corresponding weight (w_i^s, w_j^b) and by a set $\vec{x}_i$ of reconstructed variables. For each event is found the splitting value which gives the best separation for all the variables in $\vec{x}_i$. Starting from one variable, which gives the name of root node, the algorithm makes a binary classification on the base on the best separation, splitting the sample into two new nodes. At each node, all variables can be considered, even if they have been used in a previous iteration. This process continue since a stopping condition is reached and the node is turned into a leaf with a specific value which goes from -1 (event identified totally as background) to +1 (signal). Using the *Adaptive Boost* boosting technique, the algorithm is extended to several trees (forest) that, one after the other, re-weight the incorrect classification. Our signal events are defined by requiring the angle between the MC muon track and the best reconstructed track to be smaller than one degree. Since after preselection all the surviving atmospheric muons are misreconstructed as upward-going, this signal definition is designed to remove them. It is important to note that this choice also excludes badly reconstructed neutrino events.

A two-fold cross-validation technique is adopted: each training file is composed of two separate trees (Tree A and Tree B), each representing a subset of approximately the same livetime. To achieve this, the Monte Carlo sample is split according to even and odd run numbers. Each tree is used to train an independent BDT, which is then applied to the other tree. This procedure is essential to prevent overtraining and also allows the use of all available MC events in the analysis.

Several tests were done to optimize the algorithm, until reaching the following hyperparameters configuration;

- Number of trees is fixed to 500;

- Each tree can include at most 5 nodes;
- 20 cut points are tested for each variable in order to find the best separation value;
- Shrinkage (i.e. learning rate), the contribute that each new tree gives to the output score, is fixed to 30%.

Below are listed the sixteen variables used for the training, mainly focused on the distributions of Cherenkov hits. Their selection follows a work of the KM3NeT Collaboration, where a similar procedure was applied to the atmospheric neutrino analysis for the full detector (KM3NeT/ARCA115). Within the scope of this thesis, the efficiency of these variables in rejecting atmospheric muons was tested for the first time for the KM3NeT/ARCA21 detector. The motivation behind this choice is that Cherenkov hits carry the most relevant information, as they are mostly direct hits, i.e. photons that reach the PMT without undergoing scattering in sea water.

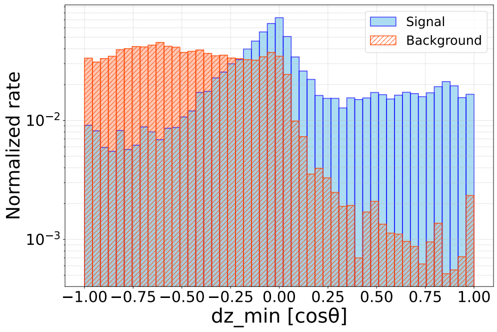
- Likelihood normalized to the number of hits: `nlik`;
- Average distance from Cherenkov hit to closest PMT: `avgclosestdist`;
- Number of DOMs with Cherenkov hits: `chits_doms`;
- Distance between the first and the last Cherenkov hit along the best track: `clen`;
- Distance to the detector hull: `dist_hull`;
- Maximum distance between any two successive Cherenkov hit along the track: `dist_maxdiff`;
- Smoothness of Cherenkov hits (hits homogeneously distributed for smoothness = 0): `dist_smooth`;
- Zenith angle of the most downgoing track of an event: `dz_min`;
- Reconstructed zenith angle: `dz`;
- Fraction of hits in the upper two rows of DOMs (Floor 17 and 18): `frac_border`;
- Ratio between Cherenkov hits and total number of hits: `frac_chits`;
- Ratio of hits in the lower hemisphere of the DOM over the total number: `frac_down`;
- Normalized difference between the upgoing and downgoing maximum likelihood: `lik_diff`;

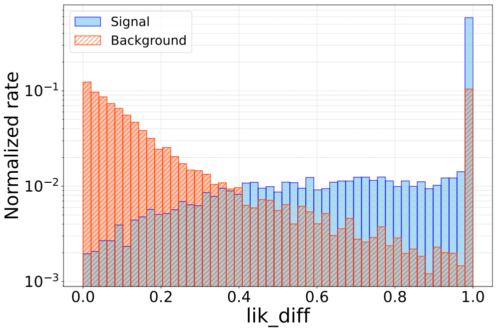
- Common logarithm of the angular error β_0 estimated by the reconstruction algorithm: `log10beta`;
- Number of reconstructed tracks within one degree of the best track: `n_good`;
- Largest angular separation between the best track and any other track whose likelihood is within 10% of the best track's likelihood: `angle`.

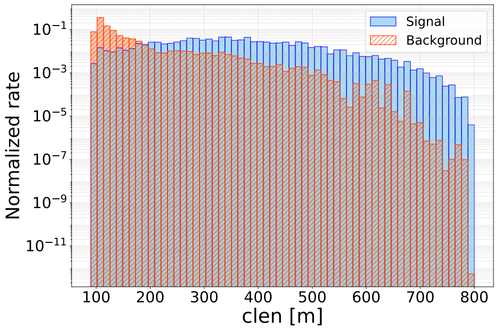
In Fig.4.1 are represented all the BDT variables, showing their rate normalized to signal and background events. Below, Table 4.3, shows the variables importance derived from the trained BDT model.

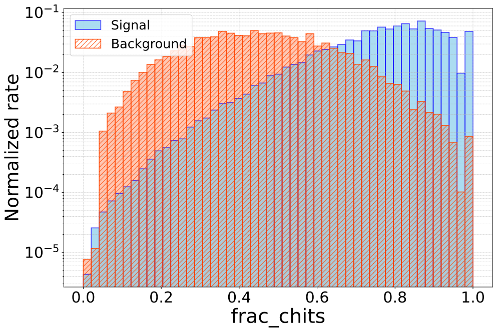
Rank	**Variable**	**Variable importance**
1	dz_min	8.4%
2	lik_diff	8.1%
3	clen	7.9%
4	frac_chits	7.4%
5	nlik	7.0%
6	log10beta	6.8%
7	dist_smooth	6.5%
8	frac_down	6.4%
9	n_good	6.2%
10	angle	5.8%
11	avgclosestdist	5.5%
12	dist_maxdiff	5.4%
13	dz	4.9%
14	dist_hull	4.8%
15	chits_doms	4.6%
16	frac_border	4.3%

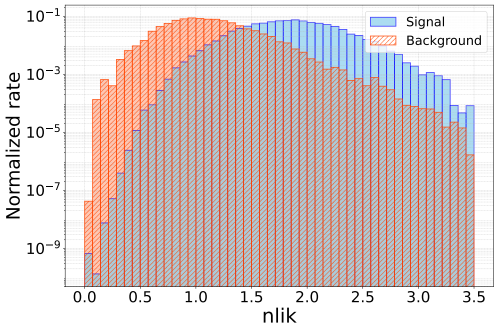
Table 4.3: Ranking result (top variable is best ranked).

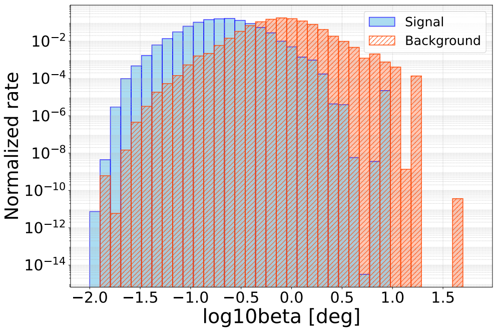

Signal
Background
Normalized rate
10^{-2}
10^{-3}
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
dz_min [cosθ]

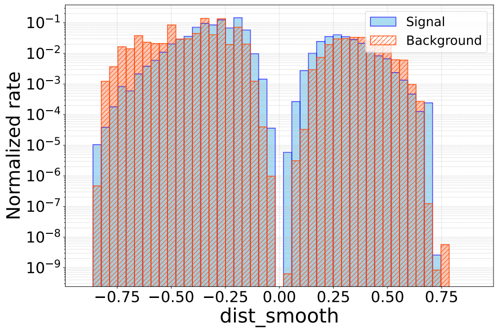

Signal
Background
Normalized rate
10^{-1}
10^{-2}
10^{-3}
0.0 0.2 0.4 0.6 0.8 1.0
lik_diff

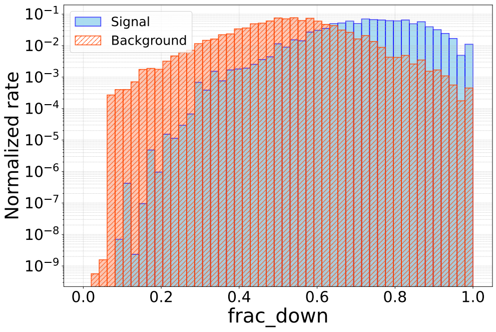

Signal
Background
Normalized rate
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
100 200 300 400 500 600 700 800
clen [m]


Signal
Background
Normalized rate
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0.0 0.2 0.4 0.6 0.8 1.0
frac_chits


Signal
Background
Normalized rate
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5
nlik


Signal
Background
Normalized rate
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
−2.0 −1.5 −1.0 −0.5 0.0 0.5 1.0 1.5
log10beta [deg]


Signal
Background
Normalized rate
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
−0.75 −0.50 −0.25 0.00 0.25 0.50 0.75
dist_smooth


Signal
Background
Normalized rate
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
0.0 0.2 0.4 0.6 0.8 1.0
frac_down

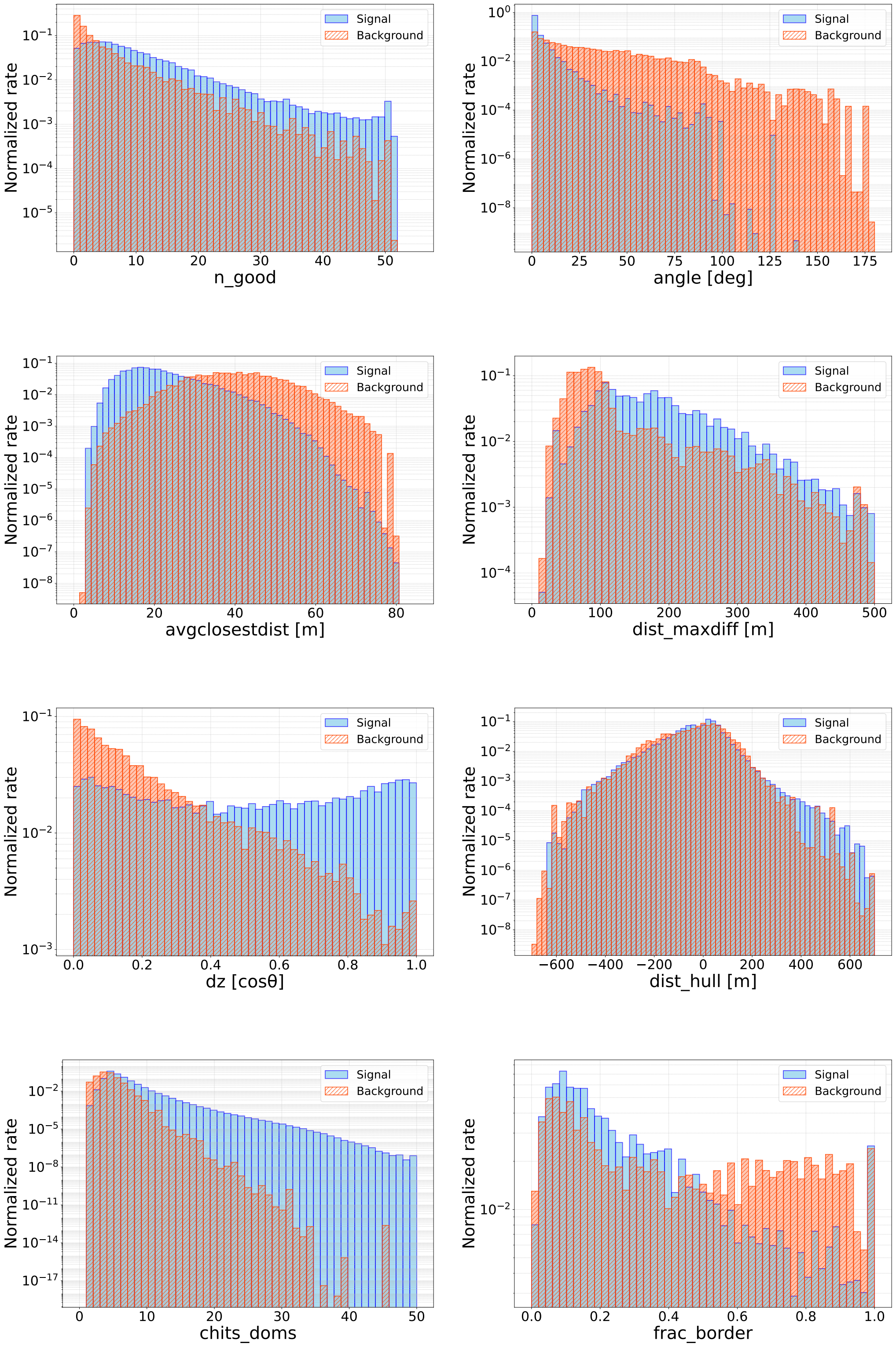


Figure 4.1: Normalized signal and background distribution of variables involved into the BDT's training. The distributions are listed from most to least important.

After the training, each of the two independent BDT is applied to the opposite MC sample. The results are shown in Fig.4.2 and in Fig.4.3.

The first plot shows the BDT score distribution for atmospheric neutrinos, cosmic neutrinos and atmospheric muons.

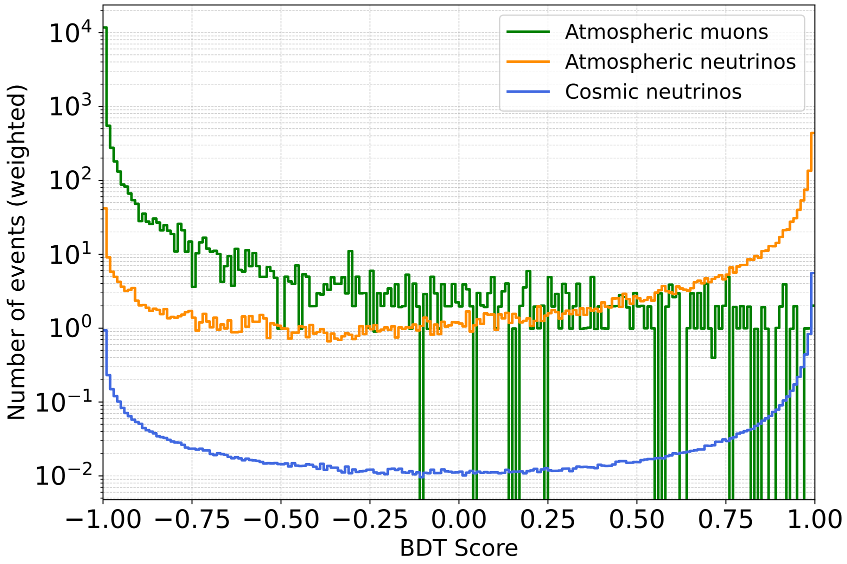


Figure 4.2: BDT score distributions for Monte Carlo atmospheric muon (green line), atmospheric neutrino (orange line) and cosmic neutrino (blue line) events.

The number of events is weighted according to the corresponding flux, as discussed in Section 3.1. Only neutrino events exhibit a peak in the region of scores close to 1, whereas muon events show a peak at scores close to -1 and then decrease as the score increases. However, neutrino events also exhibit a peak near a score of -1. This is about the 10% of the total number of atmospheric muon neutrino events, that have been wrongly classified as background.

In the second plot is shown the *Receiver Operating Characteristic* (ROC) curve which is a standard procedure to illustrate the performance of a binary classifier model. The integral of this curve is called Area Under the ROC Curve (AUC). It is a scalar metric that represents the probability that the classifier ranks a randomly chosen positive example (signal event) higher than a randomly chosen negative example (background event). In this case, weights are not applied to the events since with this metric we want to know how the algorithm classifies the raw events. Along the y-axis is shown the true positive rate, interpreted as how many of the events labeled as signal are correctly classified, while the x-axis shows the amount of wrongly classified signal (false positive) which are rejected. A AUC value of 0.9927, as obtained in our case, is highly promising for achieving the selection of a high-purity sample with a small loss of efficiency.

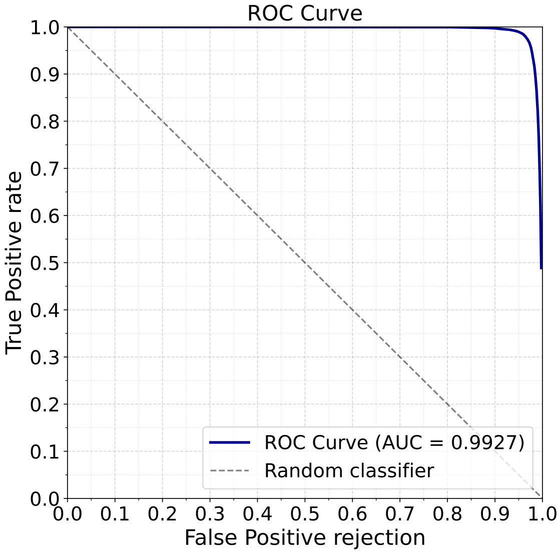


Figure 4.3: *Receiver Operating Characteristic* (ROC) curve: it is a metric for the performance of a binary classifier.

To gain a deeper understanding of the algorithm's performance, we consider additional quantities. Among the most relevant are the signal efficiency and the background rejection as functions of BDT score. An ideal algorithm would achieve a complete rejection of the background while leaving the signal efficiency unaffected.

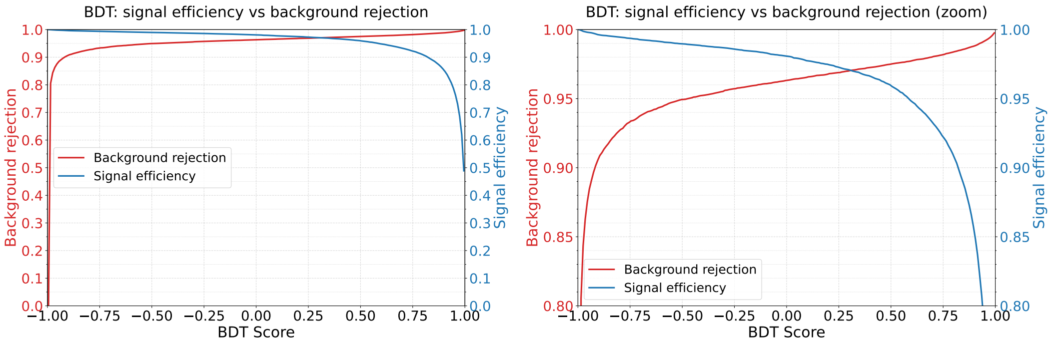


Figure 4.4: *Left:* Shape of the background rejection and the signal efficiency as functions of the BDT score. *Right:* Zoom of the left plot including only y-axis values grater than 0.8

However, as shown in Fig.4.4, a trade-off is observed in the choice of the optimal

BDT score. Around a score of 0.3, the two curves intersect at a common value of approximately 0.97. This condition is not sufficient to reach the desired requirement of a muon contamination below 1%, therefore, a higher score threshold must be selected. The good performance of the algorithm is demonstrated by the fact that, at about 99% background rejection, more than 70% of the signal efficiency is retained.

The choice of the optimal BDT score cut is driven by the priority of obtaining the purest dataset possible. By selecting a BDT score of 0.915, we are able to achieve a muon contamination well below the percent level, while retaining about 64.5% of the signal efficiency for atmospheric neutrinos.
To conclude the performance analysis, it is important to examine the confusion matrix at this BDT cut value (see Fig.4.5). A confusion matrix is a specific representation that allows to compare the algorithm's prediction with respect to the true labels of the Monte Carlo events. This analysis is performed without considering the physical weights of the events, since the goal is to infer how often the algorithm produces false positive or false negative outcome. The perfect scenario corresponds to a matrix populated only along the diagonal entries. The matrix entries correspond to true negative (TN), false positive (FP), false negative (FN) and true positive (TP) events.

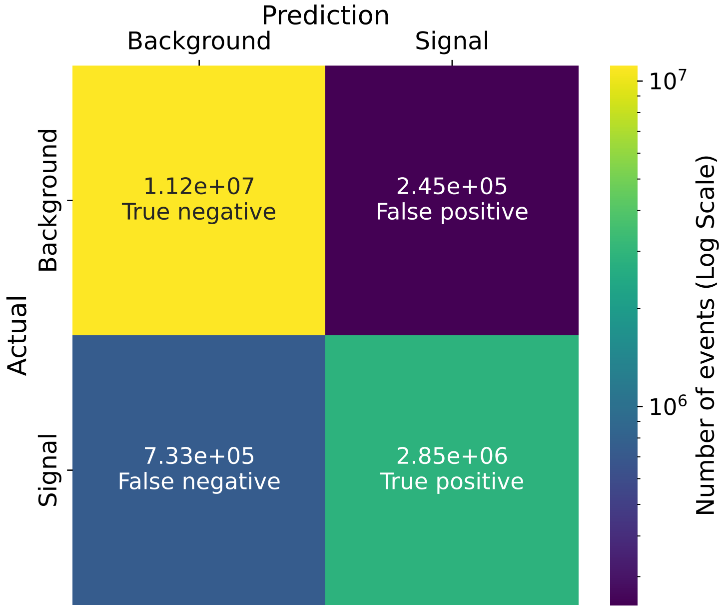


Figure 4.5: Confusion matrix for events with BDT score grater than 0.915.

We now introduce a set of metrics, based on the confusion matrix, that provide a clearer understanding of the algorithm prediction;

- Accuracy:
  $$\frac{TP+TN}{TP+TN+FP+FN}$$
  represents the amount of correct prediction of the algorithm (for BDT 0.915 the value is about 0.932);

- Precision:
  $$\frac{TP}{TP+FP}$$
  represents the purity of the signal, i.e. the amount of predicted signal which is also the actual signal (for BDT 0.915 the value is about 0.930);

- Specifity:
  $$\frac{TN}{TN+FP}$$
  represents the background rejection, i.e. the amount of actual background identified by the algorithm (for BDT 0.915 the value is about 0.982);

- Recall:
  $$\frac{TP}{TP+FN}$$
  represents the efficiency of the signal, i.e. the amount of actual signal identified by the algorithm (for BDT 0.915 the value is about 0.777).

These values, overall, confirm that the trained model achieves satisfactory performance.

4.3 Final Samples

Table 4.4 shows the most relevant quantities for the analysis of the atmospheric ν_μ CC flux. The implementation of a BDT algorithm drastically reduces the muon contamination to a value of about 0.7%, which can be considered optimal for the purpose of the analysis. The drawback of achieving such a high sample purity is a reduction in the signal efficiency for atmospheric neutrino events. However, the statistics of the final number of events is adequate to perform the measurement. Moreover, the majority of the atmospheric neutrinos surviving the BDT selection are ν_μ CC events, while neutral current interactions and other neutrino flavours are present in such a limited number that they can be considered negligible as background.

KM3NeT/ARCA21 - 287.4 days			
		Preselection	+ BDT > 0.915
Atmospheric $\boldsymbol{\nu}$ events	Total	1236	797
	ν_μ CC	1208	796
	ν_μ NC	17.9	0.32
	ν_e CC	8.6	0.13
	ν_e NC	0.81	0.01
Cosmic $\boldsymbol{\nu}$ events	Total	12.9	7.8
Atmospheric $\boldsymbol{\mu}$ events	Total	7165	6.0
Muon contamination	$\mu/(\nu_{atm} + \nu_{cos} + \mu)$	$\sim$ 85.2%	$\sim$ 0.7%
$\boldsymbol{\nu_{atm}}$ signal efficiency	$\nu^i_{atm}/\nu^{pres}_{atm}$	100%	$\sim$ 64.5%
DATA		8039	1429
DATA/MC		0.96	1.76

Table 4.4: Table summarizing the number of events, detailed composition of atmospheric ν and their signal efficiency and muon contamination at preselection and BDT level (with the particle symbol, e.g. ν_μ, are indicated both particle and anti-particle).

One final aspect that should not be ignored is the ratio between data and Monte Carlo events: the transition from the preselection level (with a ratio of 0.96) to the BDT sample (with a ratio of 1.76) suggests a degradation in the agreement. Nevertheless, this issue is well known and is not related to the performance or possible bias introduced in the training of the algorithm itself. When neutrinos become the dominant contribution, the agreement worsens due to uncertainties in the simulation of neutrino interactions, particle propagation in water, its optical properties and the PMT quantum efficiency, which are currently under intense study by the KM3NeT Collaboration, as well as uncertainties in the flux normalization of atmospheric neutrinos, today known with an accuracy up to 40%.

Fig.4.6 and Fig.4.7 show the reconstructed energy and zenith angle distributions before and after the application of the BDT cut, respectively, together with the corresponding data/Monte Carlo ratios. To account for simulation uncertainties, ± 40% variations are applied to both neutrino and atmospheric muon MC samples (see [46] for details). As expected, the distributions after the BDT selection show a worst agreement with respect to the ones at preselection. However, looking better at Fig.4.7 it is evident that, even at preselection, for events with cosine values close to one, where the neutrino component is dominant, the ratio between data and Monte Carlo increases. The Appendix reports the same distributions of MC events and data for all the variables used in the BDT training.

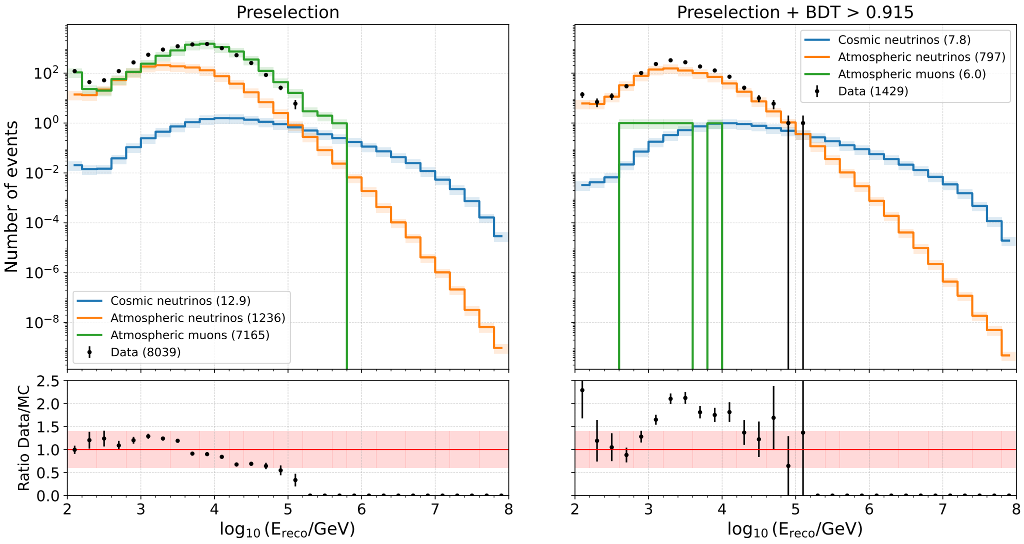


Figure 4.6: Reconstructed energy distributions of Monte Carlo events and data at the preselection and BDT cut levels. The corresponding data/MC ratio is shown in the bottom panel. The statistical error is applied to data. Variations of ± 40%, corresponding to a conservative uncertainty estimation, are applied to all the MC distributions and it is also shown in the ratio plot.

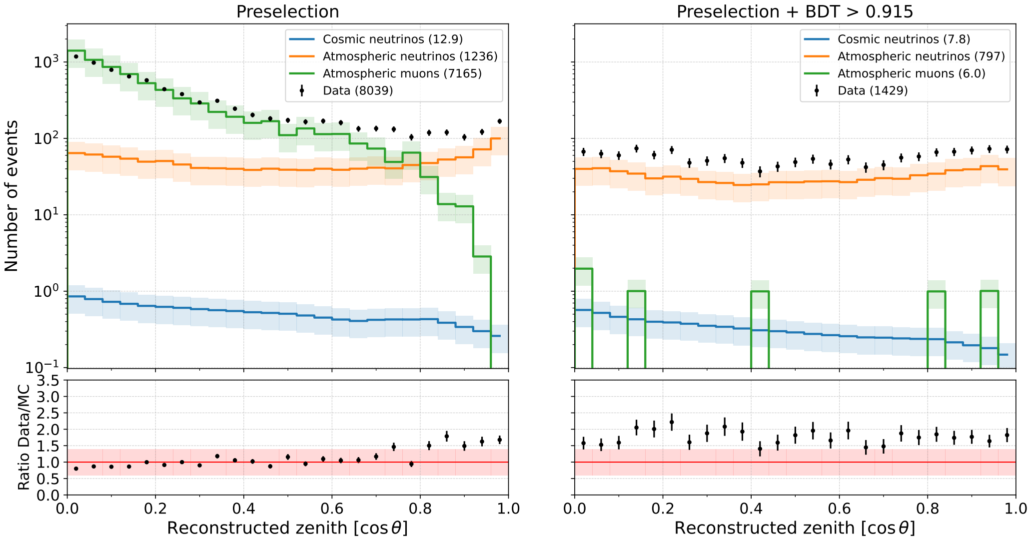


Figure 4.7: Reconstructed zenith angle distributions of Monte Carlo events and data at the preselection and BDT cut levels. The corresponding data/MC ratio is shown in the bottom panel. The statistical error is applied to data. A cosine value of zero corresponds to horizontal events, while a value of one corresponds to completely upward-going events. Variations of ± 40%, corresponding to a conservative uncertainty estimation, are applied to all the MC distributions and it is also shown in the ratio plot.

Chapter 5

Atmospheric flux measurement using an unfolding method

In this Chapter, the entire procedure developed in order to perform the first measurement of the atmospheric ν_μ and $\bar{\nu}_\mu$ flux with KM3NeT/ARCA is explained. The first part of the chapter is focused on the theoretical description of the unfolding procedure. Then a complete analysis pipeline which starts from the sample obtained in Chapter 4 and ends with the unfolded atmospheric flux is developed. A preliminary analysis is performed using only Monte Carlo simulation, in order to verify the correctness of the procedure comparing the obtained results with theoretical expectations. Once the validity of all steps has been verified, the full pipeline has also been applied to data.

5.1 The TUnfold method

In high-energy physics, most experiments are performed by counting and grouping events into fixed regions of phase-space. In this context, the response of the detector plays a crucial role: an ideal experiment would perfectly reconstruct the true phase-space of the measured variables. However, experiments are always affected by detector resolutions, which cause migration of events in the reconstructed phase-space. This problem can be translated in mathematical language as:

$$\tilde{y}_i = \sum_{j=1}^{m} A_{ij}\tilde{x}_j \quad 1 \leq i \leq n \tag{5.1}$$

where the m bins $\tilde{x}_j$ represent the average true distribution of the signal, the n bins $\tilde{y}_i$ represent the average reconstructed distribution at detector level and A_{ij} is a matrix describing the probability of migration of an event from a true bin j to one of the n reconstructed bins.

The unfolding technique [47] is the statistical method in which, starting from a Monte Carlo simulation, the response matrix A_{ij} is built and inverted in order to estimate the true distribution from the detector measurement. This process is represented in Fig.5.1.

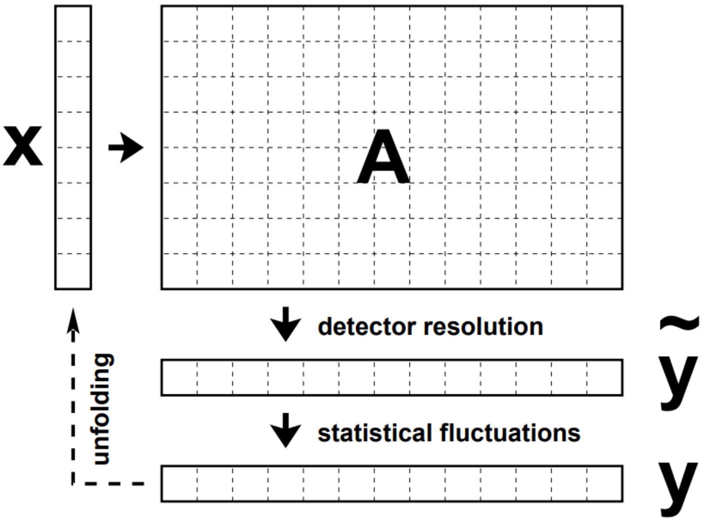


Figure 5.1: Schematic view of the unfolding approach. Migration effects and statistical fluctuations are highlighted. Figure taken from [47]

If background contributions are also present in the detector measurements, the situation becomes more complex and an additional term must be included (see Eq.(5.1)):

$$\tilde{y}_i = \sum_{j=1}^{m} A_{ij}\tilde{x}_j + b_i \quad 1 \leq i \leq n \tag{5.2}$$

Unfolding becomes fundamental when the results need to be compared with other experiments, since the deconvolution of the detector effects is then crucial.

There are two main classes of unfolding algorithms. The first is based on matrix inversion, like TUnfold [47], which will be used in this analysis. The second class are algorithms based on iterative methods and on the use of the Bayes' theorem [48].

In Eq.(5.2), one might consider just to replace $\tilde{y}_i \rightarrow y_i$ and $\tilde{x}_j \rightarrow x_j$ (where y_i and x_j are specific distributions) and find x_j simply by inverting the matrix. However, in this way the statistical fluctuations of y_i are amplified once multiplied by the elements of the inversed matrix. The TUnfold algorithm, developed within the ROOT analysis package [49], avoid this problem estimating $\tilde{x}_j$ using a least square method and introducing a term which controls the fluctuation through

Tikhonov regularization [50].

In particular, the algorithm gives an estimator of a set of truth parameters described by the vector $\boldsymbol{x}$, using a single measurement of a set of observables $\boldsymbol{y}$. These observables are treated as random variables described by a multivariate Gaussian distribution with mean $\tilde{\boldsymbol{y}} = \boldsymbol{A}\tilde{\boldsymbol{x}}$. The algorithm works only if the dimension of $\tilde{\boldsymbol{x}}$ is less or equal to the dimension of $\tilde{\boldsymbol{y}}$ $(m \leq n)$. The estimator $\boldsymbol{x}$ of the truth parameters $\tilde{\boldsymbol{x}}$ is obtained by finding the stationary point of the equation:

$$\mathcal{L}(x,\tau) = \mathcal{L}_1 + \mathcal{L}_2 \qquad \text{where;} \tag{5.3}$$

$$\mathcal{L}_1 = (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x})^T \boldsymbol{V}_{\boldsymbol{yy}}^{-1} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}) \tag{5.4}$$

$$\mathcal{L}_2 = \tau^2 (\boldsymbol{x} - \boldsymbol{x_0})^T (\boldsymbol{L}^T \boldsymbol{L})(\boldsymbol{x} - \boldsymbol{x_0}) \tag{5.5}$$

Starting from Eq.(5.4), $\boldsymbol{V}_{\boldsymbol{yy}}$ is the covariance matrix of $\boldsymbol{y}$ and $\mathcal{L}_1$ is what we expect from a least square minimization.
Then, Eq.(5.5), describes the regularization which damps fluctuations in $\boldsymbol{x}$. The parameter τ^2 controls the strength of the regularization . If it is too small, the unfolding result often has large fluctuations, if it is too large, the result is biased toward $\boldsymbol{x_0}$. The matrix $\boldsymbol{L}$ has n columns and n_R rows, corresponding to n_R regularization conditions.
There are different methods to choose the strength of the regularization, in this particular case the L-curve method is used [51]. It consists in studying the parametric plot of two variables defined as:

$$L_x^{curve} = \log \mathcal{L}_1 \tag{5.6}$$

$$L_y^{curve} = \log \frac{\mathcal{L}_2}{\tau^2} \tag{5.7}$$

The optimal τ^2 value is chosen in correspondence of the maximal curvature of this specific curve, called L-curve.

5.2 Monte Carlo analysis pipeline

The first step in order to develop an unfolding procedure is to define the energy range of the flux measurement. As discussed in Chapter 2, the geometry of the KM3NeT/ARCA detector implies a minimum detectable energy of about 100 GeV and it is optimised to measure TeV−PeV neutrinos. However the steeply falling conventional atmospheric neutrino spectrum is overtaken around 100 TeV by the prompt atmospheric and by cosmic neutrino fluxes. Moreover, the par-

tial configuration of KM3NeT/ARCA21 do not yet allow to populate with high statistics this region of the phase space. Therefore, the energy range is restricted to 100 TeV.
To perform the unfolding, the reconstructed track energy is used as the experimental observable (see Section 3.2). A first look at the response matrix for reconstructed versus true energy, considering only the atmospheric ν_μ and $\bar{\nu}_\mu$ charged current MC events is shown in Fig.5.2.

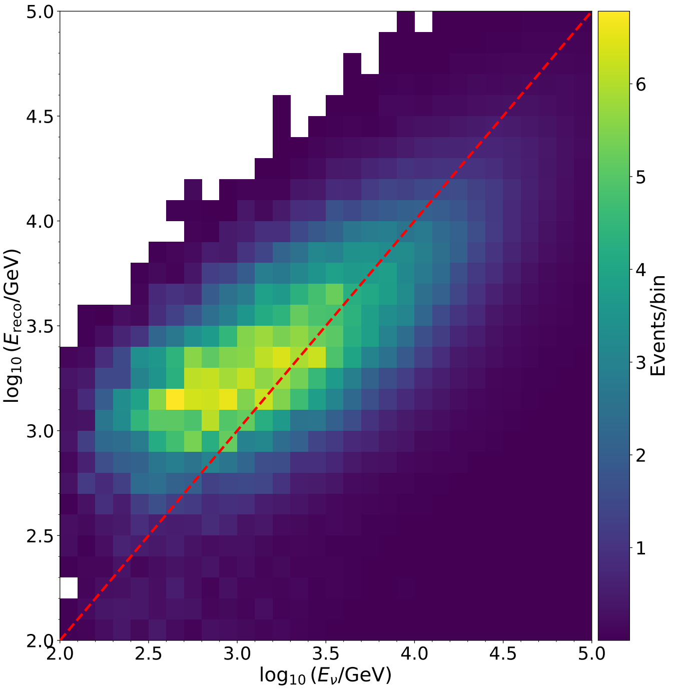


Figure 5.2: Response matrix of the reconstructed track energy versus the true energy for atmospheric ν_μ and $\bar{\nu}_\mu$ CC MC simulated events surviving the selection criteria.

Moreover, for the algorithm to work properly, sufficient statistics are required in each bin.

Therefore, the second step is to define the binning of the true phase-space. There are different motivations that lead to the final configuration selected in the analysis. As just said, sufficient number of events in each true energy bin is necessary to build a non-empty response matrix. Furthermore, this binning determines the number of points in the final spectrum: finer bins correspond to smaller energy intervals.
The binning optimization procedure can be schematize as follow;

1. A binning scheme is defined for the true energy;
2. For each bin of reconstructed energy, the fraction of events that fall in the different true energy bins is computed;
3. The purity is defined as the portion of events in a reconstructed energy bin for which the true energy belongs to the same bin, corresponding to the entries of the anti-diagonal bins.

This computation can be interpreted as the number of reconstructed events that do not migrate from the corresponding true bin to the others. For this reason, the higher the purity, the lower the event migration along the true bins. This aspect is crucial since the algorithm wants exactly to infer the true distribution starting from the reconstructed one. Different binning schemes have been tested: for each of them the bin purity and the statistical error for each bin has been measured. The optimal solution found corresponds to $\log_{10}(E_\nu/\text{GeV}) : \{2.0, 2.7, 3.4, 4.2, 5.0\}$.

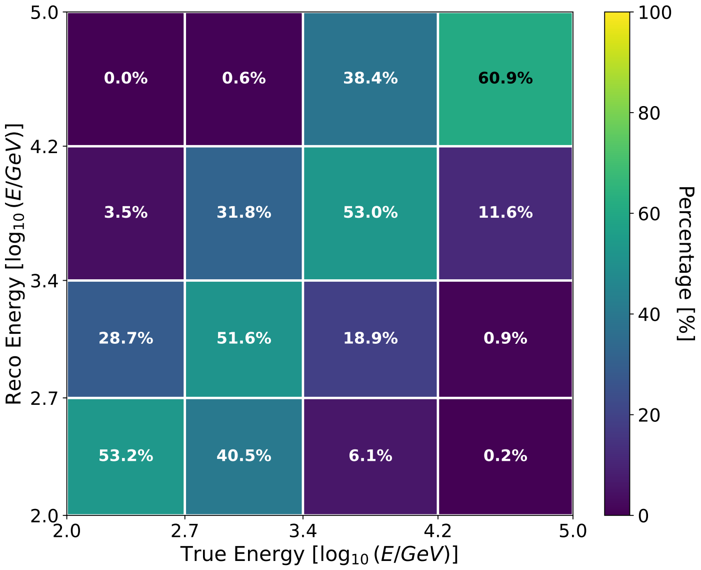


Figure 5.3: Migration matrix for ν_μ and $\bar{\nu}_\mu$ CC MC simulated events. The percentages are normalized along the total number of events in each reconstructed energy bin. The values on the anti-diagonal bins indicate the purity.

A commonly adopted threshold for purity is 50%. In this case all selected bins satisfies this condition (see Fig.5.3).

The binning on the reconstructed phase-space is less demanding and can be safely set as $\log_{10}(E_{reco}/\text{GeV}) : \{2.0, 2.1, 2.2, ..., 4.9, 5.0\}$. The response matrix with the final binning setup can be then generated and is presented in Fig.5.4.

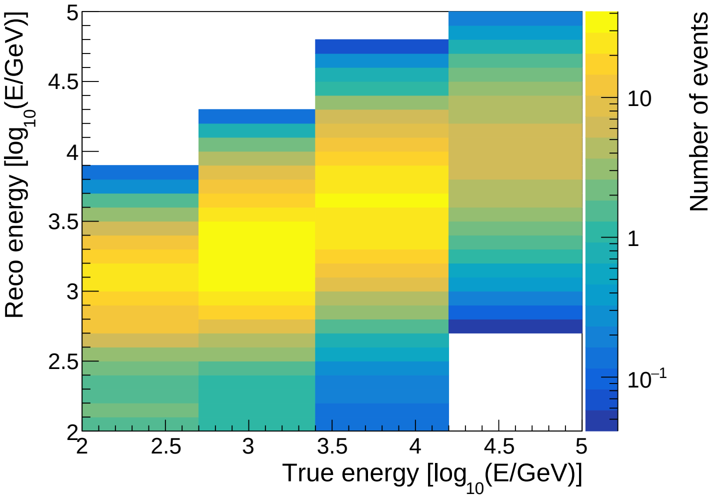


Figure 5.4: Response matrix used in the unfolding procedure, four bins are fixed in the true phase-space and thirty in the reconstructed phase-space.

Before performing the unfolding, statistical and systematic errors must be estimated. Since the reconstructed events distribution is simulated, each events carries a generation weight which is different from the unitary weight that a counted data event distribution should have. For this reason a pseudo-experiment is built following the procedure below;

1. An event is randomly selected from the true distribution;
2. This event is propagated through the response matrix, generating a reconstructed event with unit weight. This process simulates the detector response to real events;
3. This process is repeated until the total number of selected events is reproduced.

Then the Poissonian error in each bin is set as statistical error and it is propagated through the response matrix.

Systematic uncertainties are treated in a simplified manner. A more comprehensive study of the impact of systematic uncertainties on the final atmospheric neutrino spectrum will be conducted in the future and falls outside the scope of the present thesis. A deliberately conservative 40% energy independent systematic uncertainty is adopted, in order to estimate its impact on the final result. This choice is motivated by a previous study performed by the KM3NeT Collaboration that identified the dominant systematic contributions with PMT quantum efficiency, light absorption in water and the uncertainty on the theoretical atmospheric flux normalization. Therefore, two additional pseudo-experiments are generated with ±40% variations and propagated through the response matrix.

At this stage, all the necessary inputs for the algorithm are prepared. The signal is given by the atmospheric muon neutrino CC events reported in Table 4.4, while the background sources come from all the other events (all ν NC, ν_e CC, cosmic ν and badly reconstructed atmospheric μ). A first consistency check has been performed applying the unfolding algorithm to a pseudo-experiment generated from MC sample as described above. The statistical and systematic uncertainties have also been propagated. The pseudo-experiment unfolded energy spectrum is then presented in Table 5.1 and in Fig.5.5.

$\Delta\log_{10}(E_\nu/\mathrm{GeV})$	True MC	Unfolded	Stat. Err. (%)	Sys. Err. (%)
2.0 – 2.7	143.4	179.4	± 13.3	± 46.9
2.7 – 3.4	324.8	279.5	± 7.0	± 35.9
3.4 – 4.2	265.7	278.5	± 9.2	± 38.9
4.2 – 5.0	59.8	53.6	± 31.5	± 46.1

Table 5.1: Number of events for the unfolded energy spectrum, statistical and systematic unfolding errors.
The number of true MC events after the event selection for each bin has been also added as a reference.

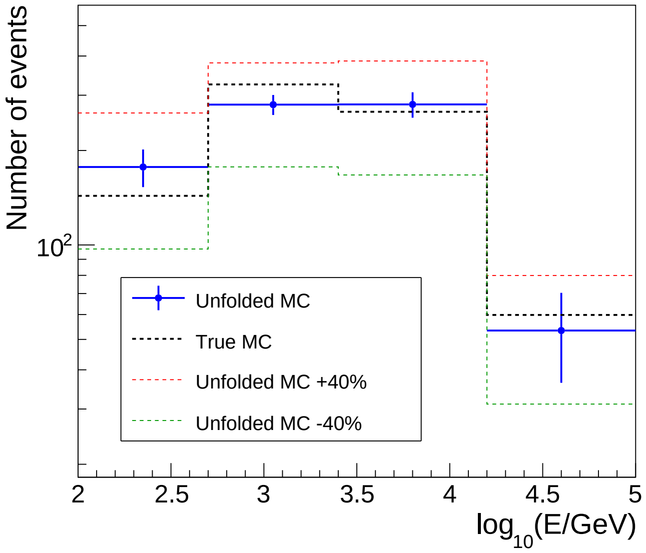


Figure 5.5: Distribution of the unfolded energy spectrum. The blue points represent the unfolded number of events with the statistical error propagated. The red and green lines represent respectively the pseudo-experiment with the addition and subtraction of 40% of the statistics. The black line represents the true MC number of events.

The consistency between the unfolded spectrum and the true Monte Carlo distribution provides initial evidence that the algorithm is performing well. Indeed, starting from a pseudo-experiment with an events distribution affected by detector effects, the unfolding procedure is able to reconstruct a spectrum compatible with the underlying true distribution.

The final check to verify the quality of the pipeline is the chi-squared test. In this context, this method is used to assess the compatibility of the resulting unfolded spectrum and the theoretical one extracted from Monte Carlo simulations. There are two contributions to the chi-squared, one from least squares term and one from the regularization term, coherently as reported in Eq.(5.3). The reduced chi-squared is presented in Table 5.2 and it shows that the overall agreement between the unfolded spectrum and the theoretical one is very good.

χ^2_{LS}	χ^2_{Re}	ndf	χ^2_{tot}/ndf
14	2	23	0.7

Table 5.2: χ^2 corresponding to the least squares term and the regularization term are shown. The reduced χ^2 is computed as the sum of these two contributes divided by the degrees of freedom.

The value of the optimal τ^2 parameter $\left(\tau^2_{opt} = 9.4 \cdot 10^{-3}\right)$ also suggests that the regularization damps statistical fluctuations without introducing a bias [47, 51].

The last step of the analysis is to convert the number of unfolded events into flux values. Fluxes are normally expressed as the number of events normalized to energy, time, solid angle and area, with units $\left[\mathrm{GeV}^{-1} \cdot \mathrm{s}^{-1} \cdot \mathrm{sr}^{-1} \cdot \mathrm{cm}^{-2}\right]$. However, it is common to multiply the flux by a factor E^2, which flattens its overall slope for a clearer graphical representation. The contributions involved into the computation of the flux in each bin are shown in Eq.(5.8):

$$E^2_\nu \phi^{unf}_\nu \left[\mathrm{GeV} \cdot \mathrm{s}^{-1} \cdot \mathrm{sr}^{-1} \cdot \mathrm{cm}^{-2}\right] = \frac{N^{unf}_{evt} \cdot \langle E_\nu \rangle^2}{\Delta E_\nu \cdot t_{liv} \cdot 2\pi \cdot A_{eff}} \tag{5.8}$$

which can be described as follows;

1. Bin width (ΔE_ν): the unfolded number of events are binned as a function of $\log_{10}(E_\nu)$. Since the flux must be normalized to E_ν then, the number of events is divided by the corresponding bin width in linear scale $\Delta E_\nu = \langle E_\nu \rangle \cdot \ln(10) \cdot \Delta \log_{10}(E_\nu)$, where $\langle E_\nu \rangle$ is the weighted mean energy within the bin;

2. Livetime (t_{liv}): the data-taking period of our sample is approximately 287.4

days. Dividing by the livetime expressed in seconds provides the number of events per bin per second;

3. Half of the solid angle (2π): since the flux is isotropic at the Earth's surface, we divide by the portion of the sky considered into the analysis (only upward-going events);

4. Effective area (A_{eff}): this parameter represents the size of a virtual target that the detector presents to a given simulated neutrino flux. By definition, it is used to estimate the number of detected events for a given flux $\phi(E_\nu)$ within a time window T:

   $$N_\nu = \int dE_\nu \; \phi(E_\nu) \cdot A_{eff} \cdot T \tag{5.9}$$

   It has dimension of an area and also provides a metric to evaluate and compare the performances and construction status of neutrino telescopes. In our specific case, it can be computed in the four unfolded bins for atmospheric ν_μ and $\bar{\nu}_\mu$ CC that survive the selection (see Fig.5.6 and the Appendix for details);

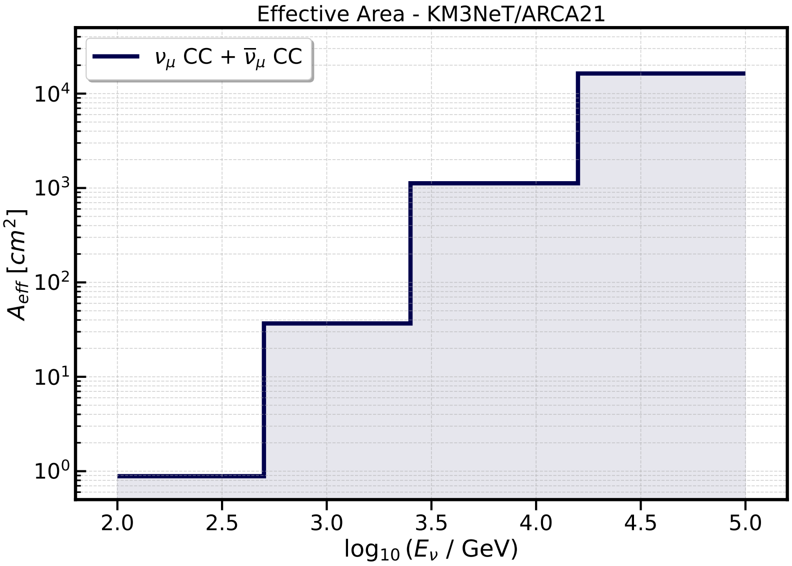


Figure 5.6: Effective area computed in the unfolded bins for atmospheric ν_μ and $\bar{\nu}_\mu$ CC that survive the selection.

5. Mean squared energy in each bin ($\langle E_\nu \rangle^2$): the E^2 factor is chosen, as in the case of bin width, as the squared weighted mean energy in each bin.

All these contributions are constants within each bin. Therefore, using Eq.(5.8), the following relation between the true MC and the unfolded spectrum can be established:

$$\frac{E_\nu^2 \phi_\nu^{MC}}{N_{evt}^{MC}} = \frac{\langle E_\nu \rangle^2}{\Delta E_\nu \cdot t_{liv} \cdot 2\pi \cdot A_{eff}} = \frac{E_\nu^2 \phi_\nu^{unf}}{N_{evt}^{unf}}$$

$$\Rightarrow \phi_\nu^{unf} = \phi_\nu^{MC} \cdot \frac{N_{evt}^{unf}}{N_{evt}^{MC}} \tag{5.10}$$

The MC flux values corresponding to the mean energy in each bin are extrapolated from the theoretical curve. The final unfolded flux is then computed at these energy values.

The unfolded MC simulated atmospheric ν_μ and $\bar{\nu}_\mu$ CC flux is presented in Table 5.3 and Fig.5.7.

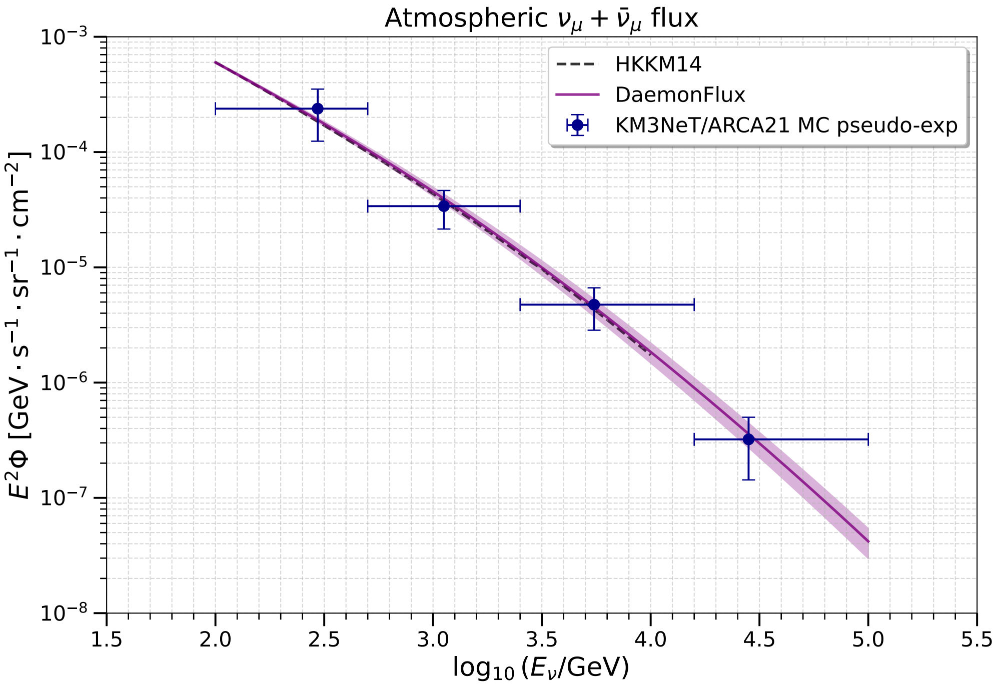


Figure 5.7: Flux of the unfolded MC atmospheric ν_μ and $\bar{\nu}_\mu$ CC events compared to the theoretical prediction of HKKM14 [52] and Daemonflux [53]. The x-error is given by the bin width.

In Fig.5.7, the unfolded MC flux is compared with two theoretical curves. The first is the Honda (HKKM14 [52]) flux, which was used to generate the MC simulation. The second one, Daemonflux, is a newer data-driven flux [53] recently exploited by the KM3NeT Collaboration, which also accounts for model uncertainties.

$\Delta\log_{10}(E_\nu/\mathrm{GeV})$	$\log_{10}(\langle E_\nu\rangle/\mathrm{GeV})$	N_{evt}^{unf}	$E_\nu^2\phi_\nu^{unf}$	Stat.err.(%)	Sys.err.(%)
2.0 – 2.7	2.47	179.4	$2.38\cdot10^{-4}$	± 13.3	± 46.9
2.7 – 3.4	3.05	279.5	$3.39\cdot10^{-5}$	± 7.0	± 35.9
3.4 – 4.2	3.74	278.5	$4.74\cdot10^{-6}$	± 9.2	± 38.9
4.2 – 5.0	4.45	53.6	$3.22\cdot10^{-7}$	± 31.5	± 46.1

Table 5.3: Flux values in each bin with the corresponding statistical and systematic errors for the unfolded MC events. Also the number of unfolded MC events and the mean energies are reported.

It is evident that this pseudo-measurement is in agreement with the predictions. This is expected as the pseudo-experiment is built from a distribution generated exactly from these models. This step is necessary to assess the correctness of the full pipeline and check the consistency of the results with the Monte Carlo.

5.3 Unfolding KM3NeT/ARCA21 data

After an internal revision from the Collaboration, the same unfolding procedure has been applied to the KM3NeT/ARCA21 data. The parameters optimised on the Monte Carlo sample are kept fixed, such as the energy range, the binning choice, the regularization parameter and the response matrix. Regarding the uncertainties, the statistical errors must be recalculated based on the data distribution in each bin, while the systematics uncertainties remain the same as in the Monte Carlo analysis. Since the distributions in each bin change, the mean energy at which the flux is computed must also be recalculated.

In Fig.5.8, the unfolded number of data events is shown in red (see also Table 5.4) with their total uncertainties, computed as the quadratic sum of statistical and systematic errors. They are compared with both the true Monte Carlo and the unfolded MC.

$\Delta\log_{10}(E_\nu/\mathrm{GeV})$	$\log_{10}(\langle E_\nu\rangle/\mathrm{GeV})$	N_{evt}^{data}	$E_\nu^2\phi_\nu^{data}$	Stat.err.(%)	Sys.err.(%)
2.0 – 2.7	2.41	145.8	$1.93\cdot10^{-4}$	± 18.0	± 46.9
2.7 – 3.4	3.16	586.5	$7.12\cdot10^{-5}$	± 3.9	± 35.9
3.4 – 4.2	3.69	631.3	$1.07\cdot10^{-5}$	± 4.9	± 38.9
4.2 – 5.0	4.41	58.4	$3.50\cdot10^{-7}$	± 34.9	± 46.1

Table 5.4: Flux values in each bin with the corresponding statistical and systematic errors for the unfolded data events. Also the number of unfolded data events and the mean energies are reported.

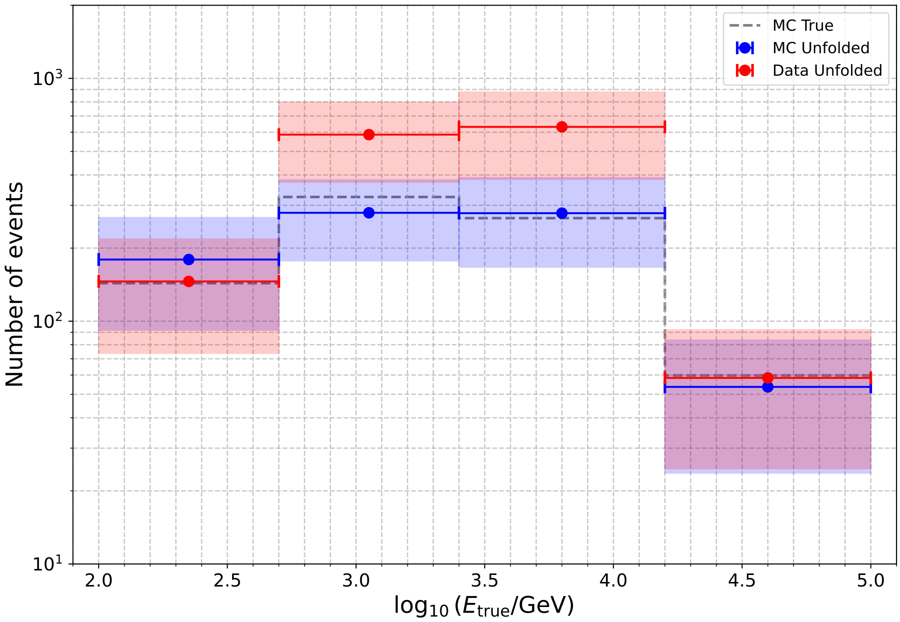


Figure 5.8: Unfolded number of events for data compared to the unfolded pseudo-experiment and the true MC distribution. The colored regions show their respective errors as quadrature sums of statistical and systematic uncertainties.

A good agreement between data and MC can be observed in the first and last bins. In contrast, large discrepancies in the central bins are evident. However, as shown in Fig.4.6, the region between 1−10 TeV is where the data significantly exceed the Monte Carlo prediction. Therefore, this behavior of the unfolded data is expected and does not indicate an issue with the unfolding procedure applied. In addition, when considering the ratio between the total number of data and MC events, it is important to note that this quantity is preserved through the unfolding procedure (DATA/MC is equal to 1.76 at selection level and 1.79 after unfolding).

The unfolded atmospheric ν_μ and $\bar{\nu}_\mu$ CC data flux is presented in Table 5.4 and Fig.5.9. In this final plot, the unfolding results for the KM3NeT/ARCA21 data are compared to the two theoretical curves and with two other experimental measurements performed by ANTARES [25] and IceCube [54].

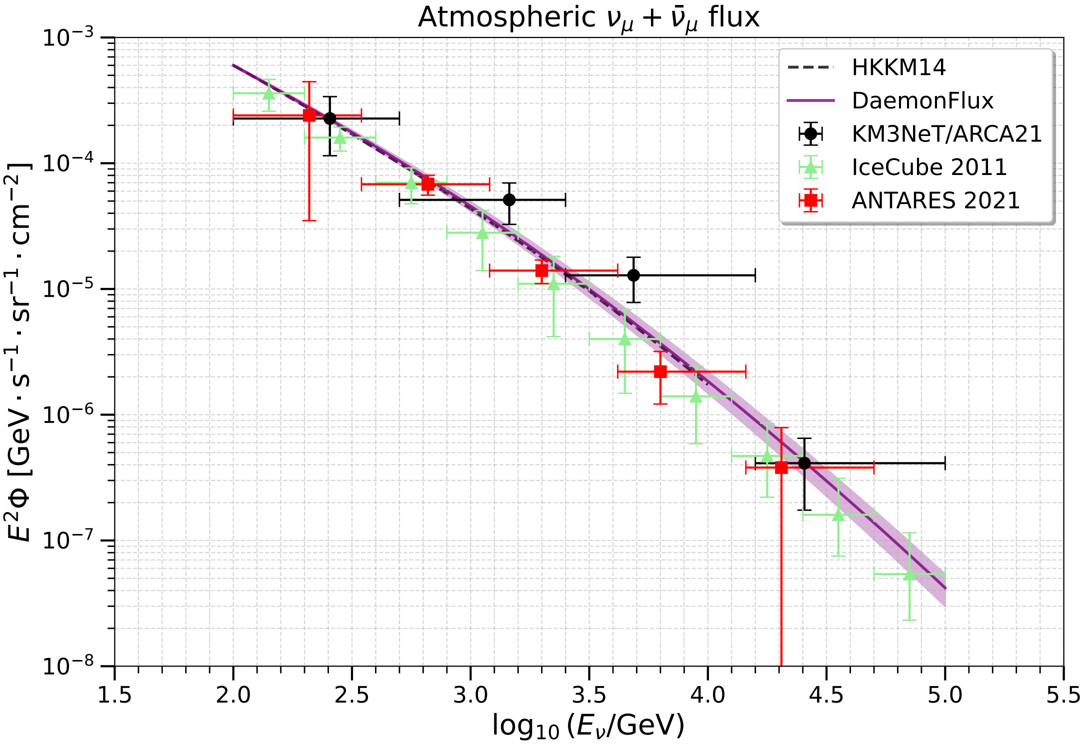


Figure 5.9: Flux of the unfolded atmospheric ν_μ and $\bar{\nu}_\mu$ CC data events compared to the theoretical prediction of HKKM14 [52] and Daemonflux [53] and also to other experimental measurements like ANTARES [25] and IceCube [54]. The x-error is given by the bin width.

Chapter 6

Conclusions

In this thesis, the first measurement of the atmospheric ν_μ and $\bar{\nu}_\mu$ charged current flux in the energy range between 100 GeV and 100 TeV with the KM3NeT/ARCA, with 21 active strings, neutrino telescope is presented. The data-taking period corresponds to approximately 287.4 days.

Several studies have been conducted to develop an optimal event selection strategy. A preselection based on reconstructed variables, combined with a newly trained *Boosted Decision Tree* (BDT) classifier, enables the extraction of a high-purity sample of atmospheric muon neutrinos, with residual atmospheric muon contamination reduced to below the percent level. The proposed event selection is currently being evaluated by the Collaboration for application in other analyses, further demonstrating the robustness and effectiveness of the developed procedure.

As no previous measurement of the atmospheric neutrino flux has been performed with KM3NeT/ARCA, a dedicated analysis pipeline, including the optimization of the unfolding algorithm, was developed, tested, and validated using Monte Carlo simulations. After verifying the consistency of the reconstructed spectrum with theoretical expectations, the validated procedure was subsequently applied to experimental data.

The final measurement shows good agreement with theoretical expectations and with the results of other experiments, particularly in the two energy ranges $10^2 - 10^{2.7}$ GeV and $10^{4.2} - 10^5$ GeV. In the other two energy bins, the flux appears to be overestimated. However, these discrepancies can be explained by the poor agreement between the Monte Carlo used and the data in those regions.

Beyond the final results, this work aims to develop a complete analysis chain for this type of measurements. Indeed, KM3NeT/ARCA21 currently represents less than one fifth of the final detector configuration and future expansions will provide significantly higher statistics. Moreover, a new version of the Monte Carlo simulation is being developed within the KM3NeT Collaboration, with the expectation that it will address the issues mentioned above. By combining these improvements with a proper implementation of systematic uncertainties, a more precise measurement of the atmospheric flux can be achieved in the future.

Appendix: further plots and tables

In this Appendix, additional plots and tables complementing Chapter 4 and Chapter 5 are presented.

BDT variables distributions

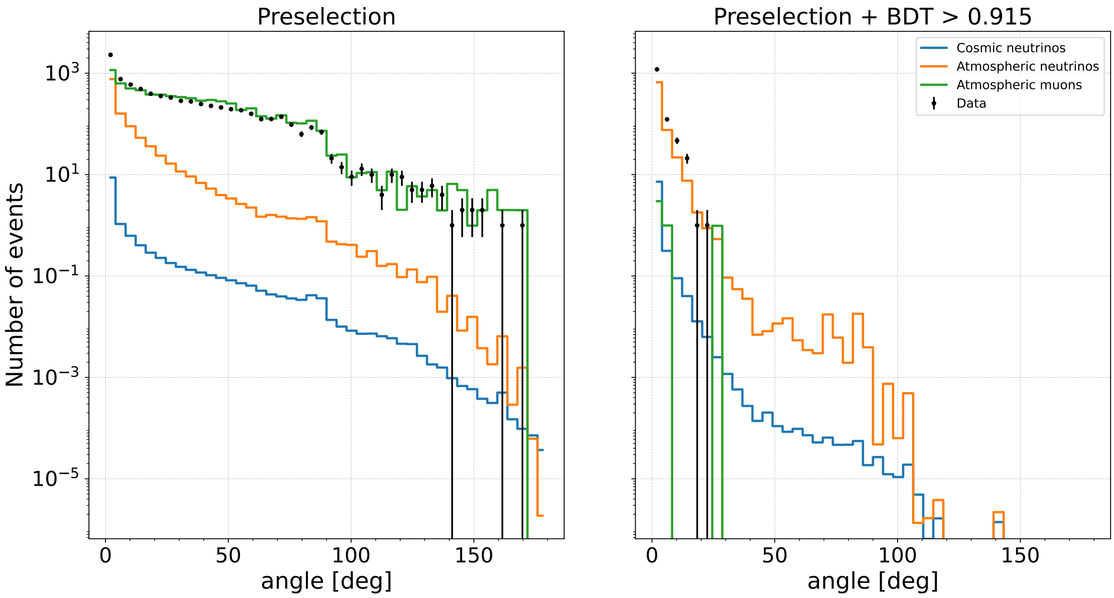


Figure 6.1: *BDT variables:* angle distribution before and after BDT selection.

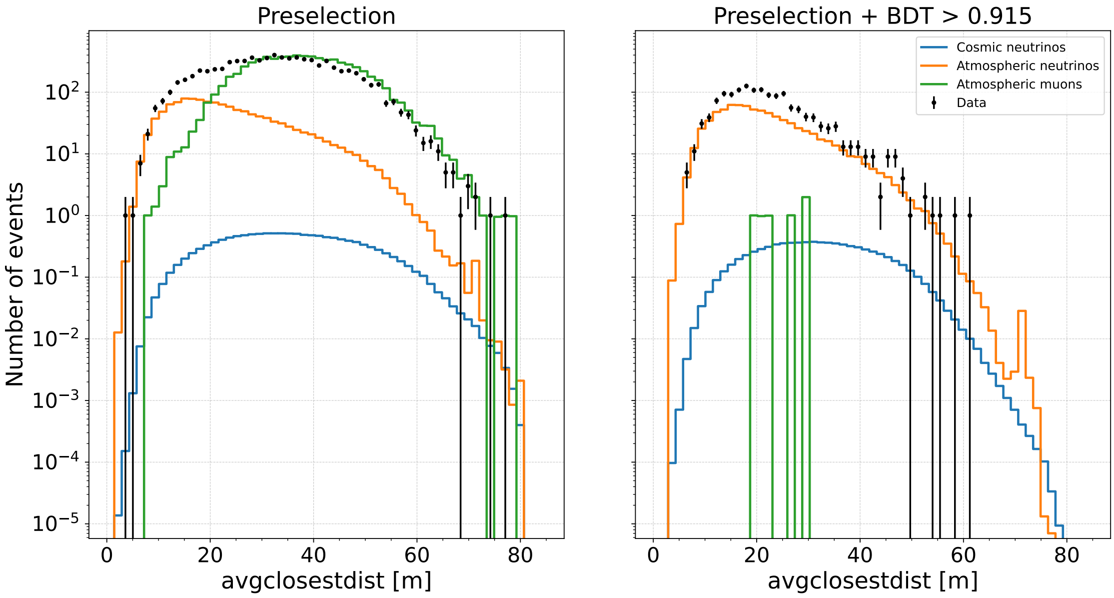


Figure 6.2: *BDT variables:* avgclosestdist distribution before and after BDT selection.

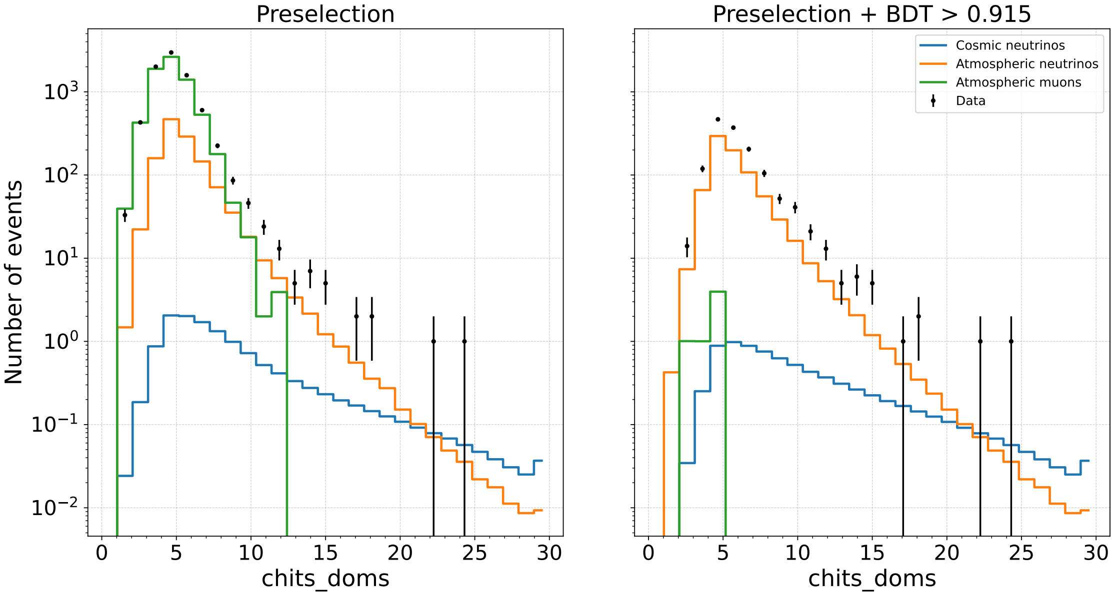


Figure 6.3: *BDT variables:* chits_doms distribution before and after BDT selection.

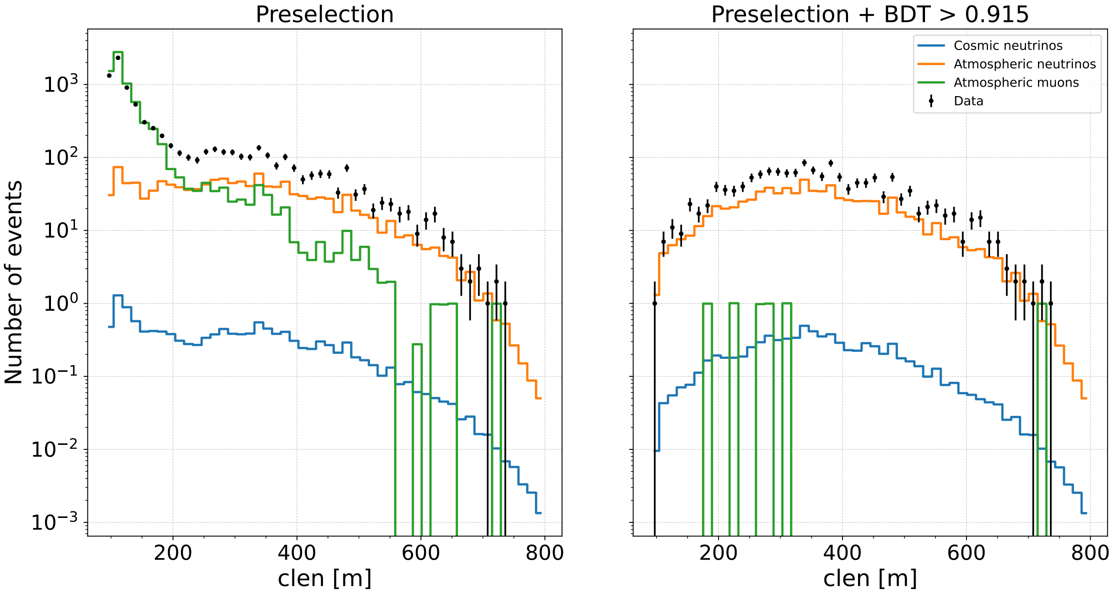


Figure 6.4: *BDT variables:* clen distribution before and after BDT selection.

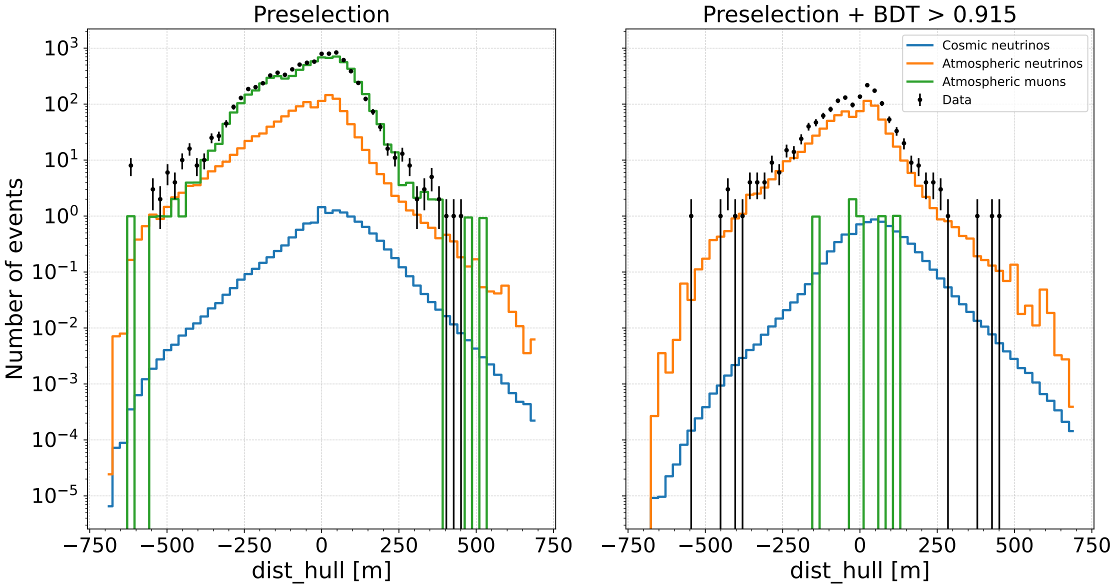


Figure 6.5: *BDT variables:* dist_hull distribution before and after BDT selection.

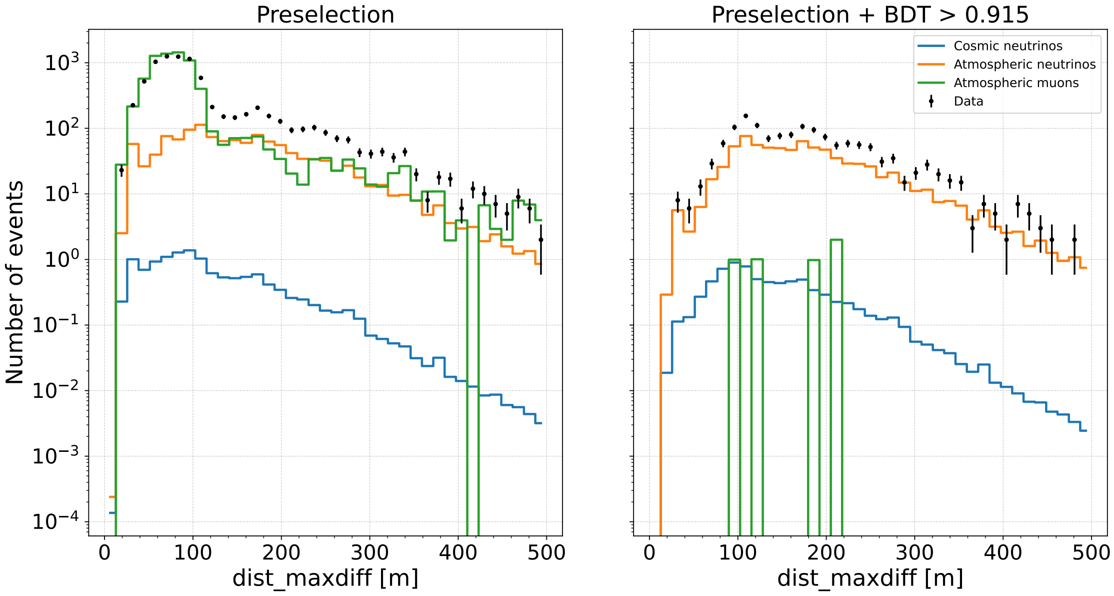


Figure 6.6: *BDT variables:* dist_maxdiff distribution before and after BDT selection.

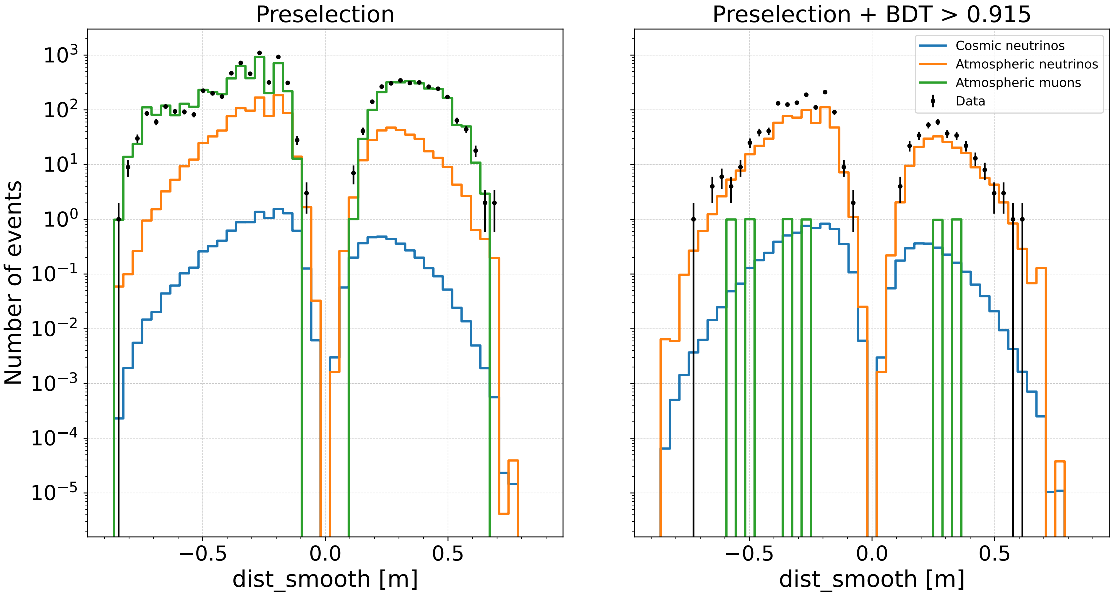


Figure 6.7: *BDT variables:* dist_smooth distribution before and after BDT selection.

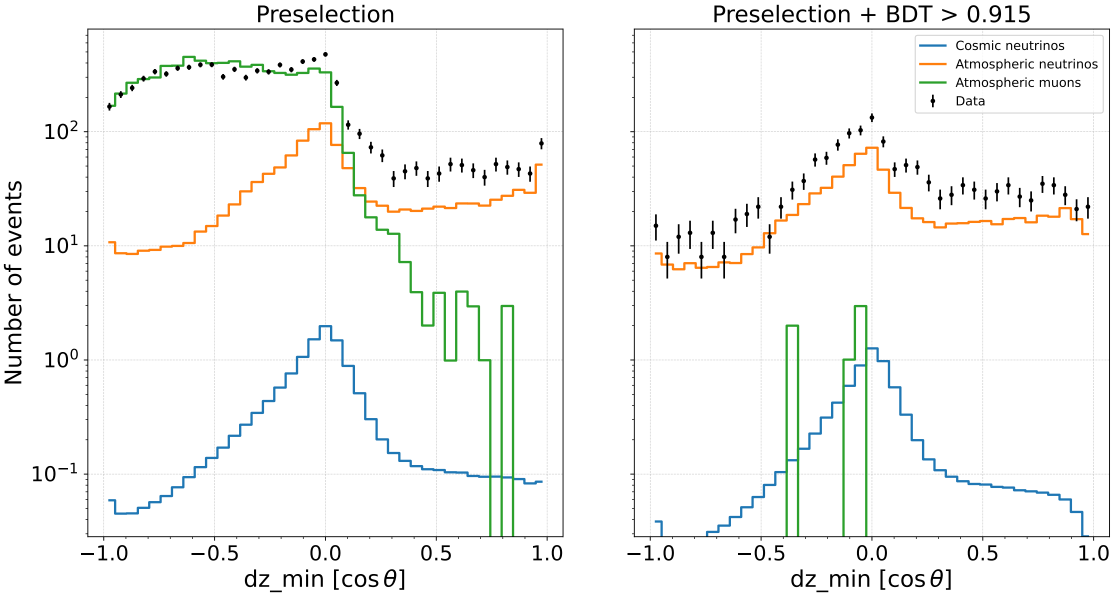


Figure 6.8: *BDT variables:* dz_min distribution before and after BDT selection.

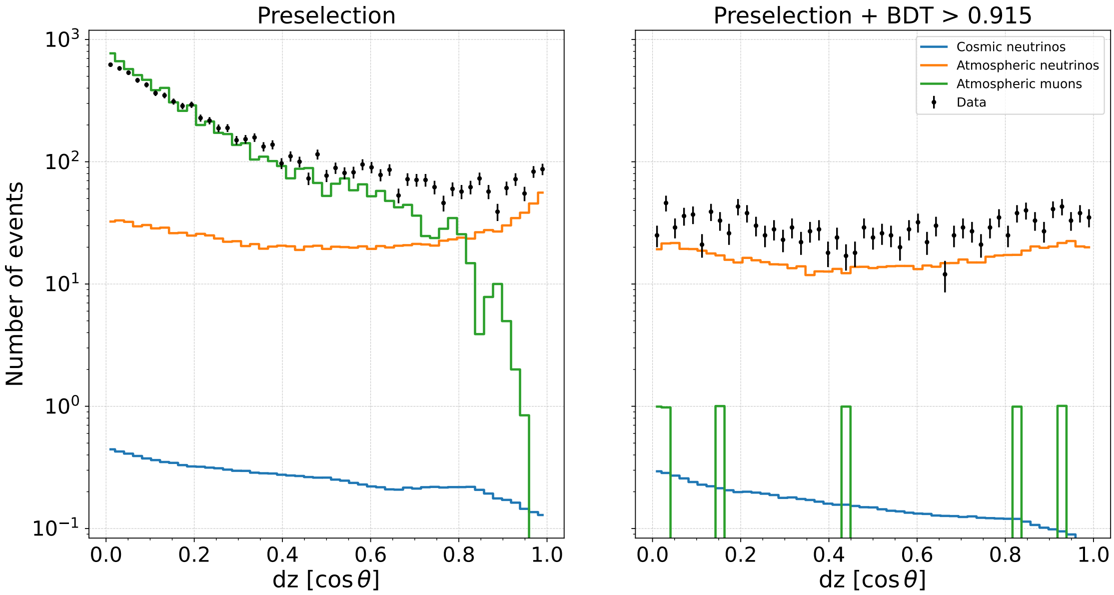


Figure 6.9: *BDT variables:* dz distribution before and after BDT selection.

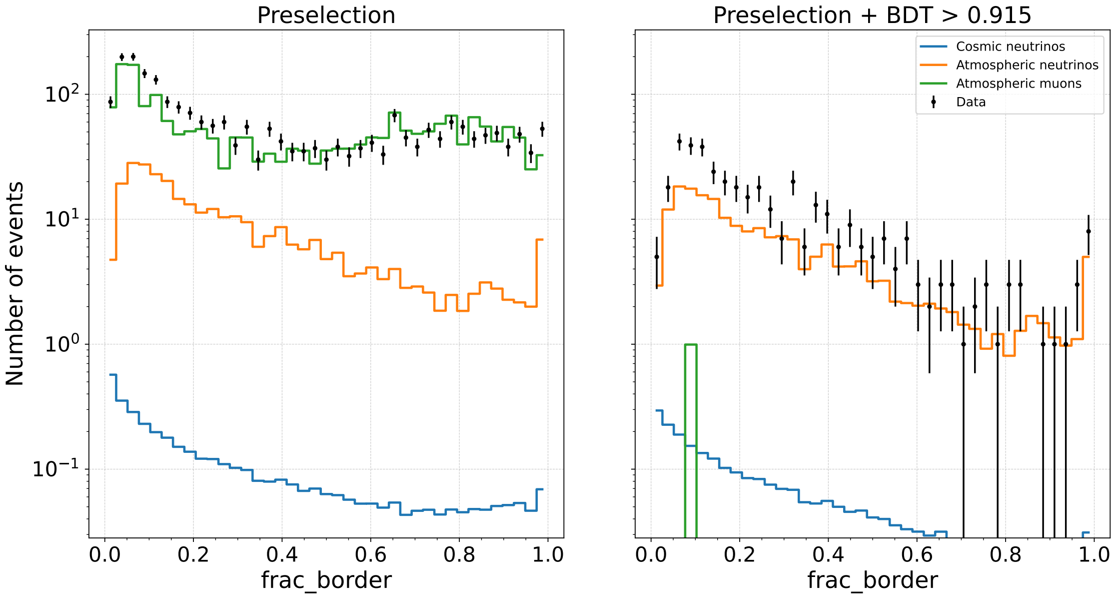


Figure 6.10: *BDT variables:* frac_border distribution before and after BDT selection.

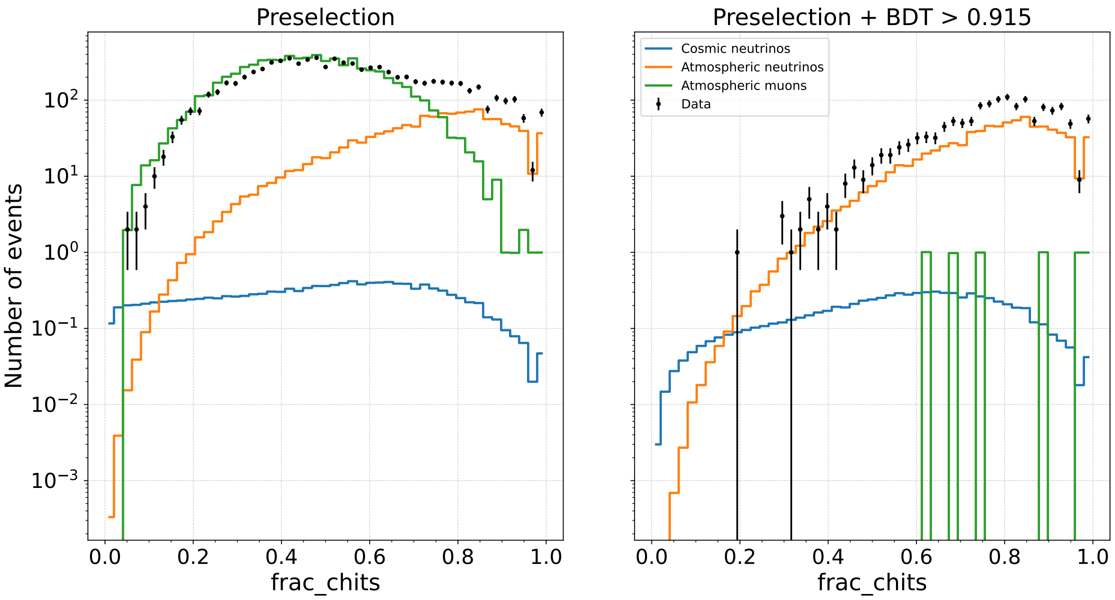


Figure 6.11: *BDT variables:* frac_chits distribution before and after BDT selection.

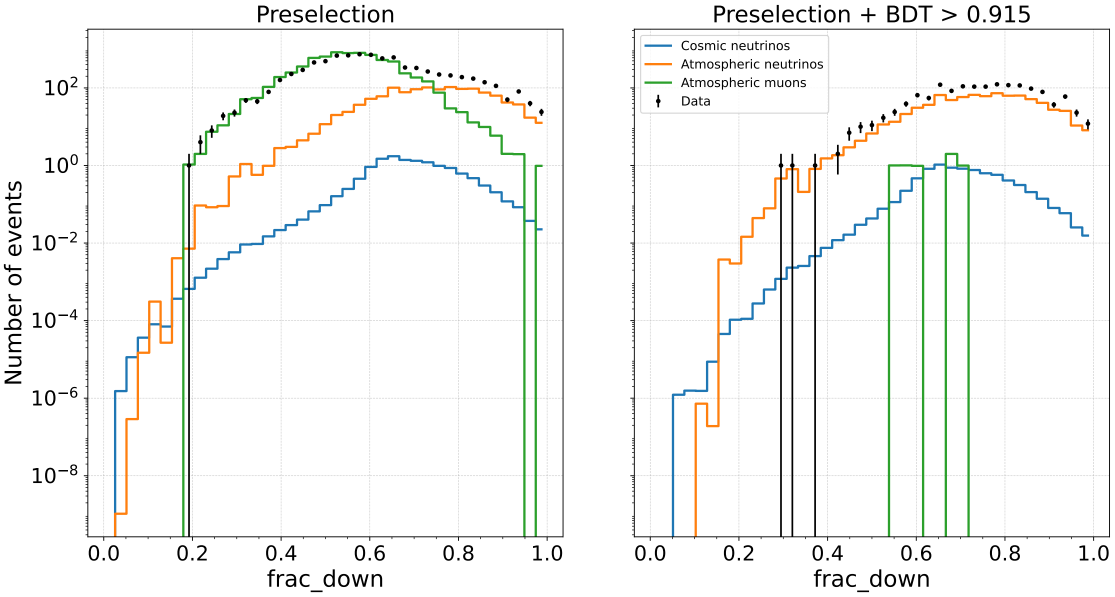


Figure 6.12: *BDT variables:* frac_down distribution before and after BDT selection.

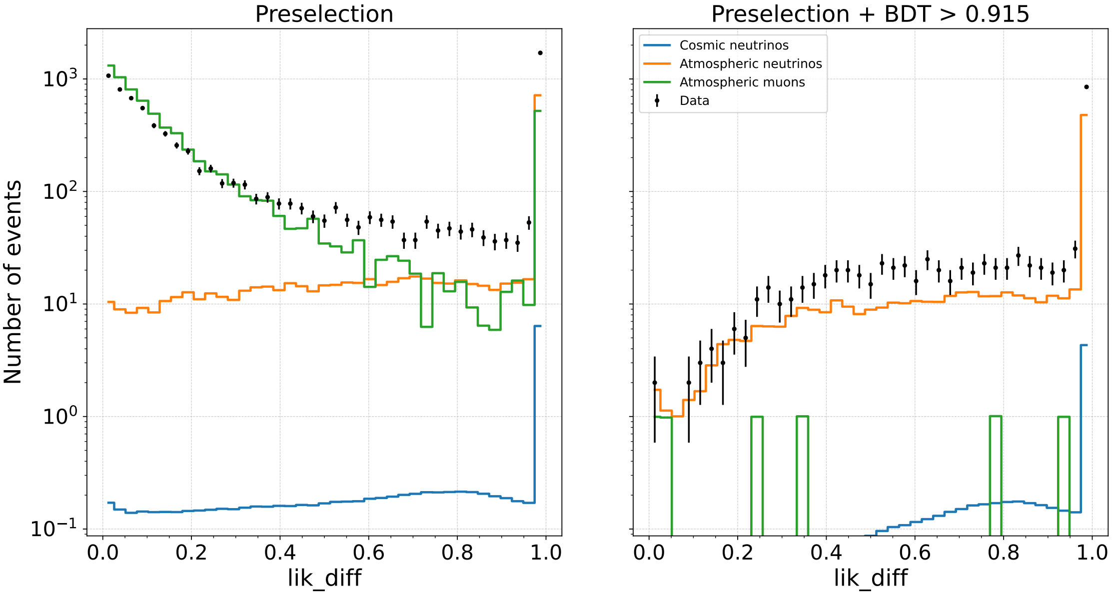


Figure 6.13: *BDT variables:* lik_diff distribution before and after BDT selection.

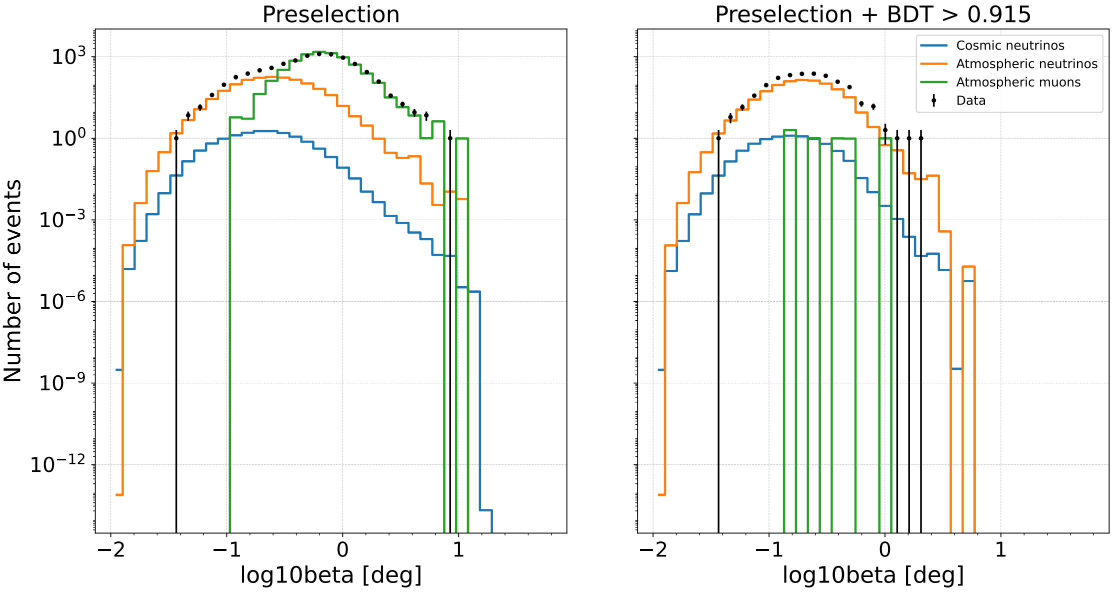


Figure 6.14: *BDT variables:* log10beta distribution before and after BDT selection.

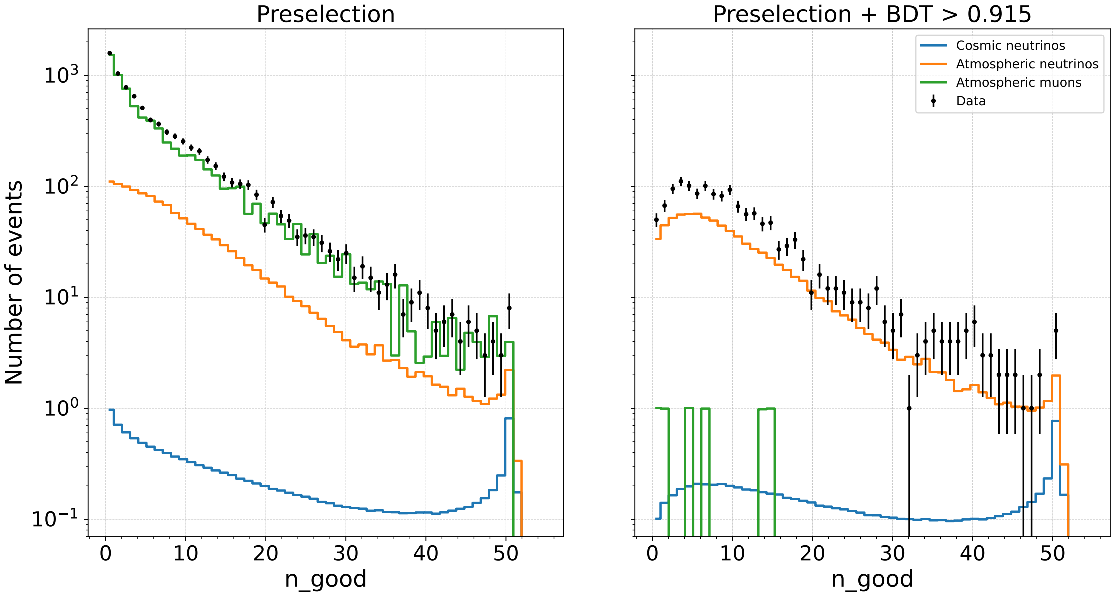


Figure 6.15: *BDT variables:* n_good distribution before and after BDT selection.

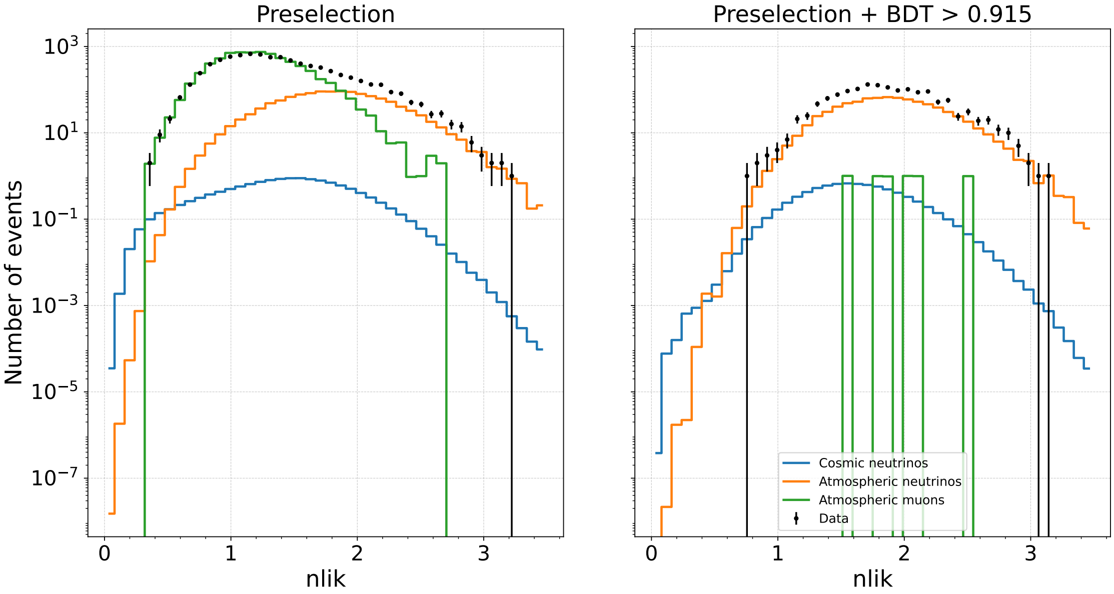


Figure 6.16: *BDT variables:* nlik distribution before and after BDT selection.

Effective area

In Section 5.2, the concept of effective area is introduced to convert the number of neutrino events into a flux. However, this quantity encodes more physical information than it may initially appear. Indeed, it accounts for the probability that the neutrino survives absorption in the Earth and interacts sufficiently close to the detector, that the secondary particle reaches the instrumented volume, and that it produces enough light to trigger the detector and allow track reconstruction. Therefore, an alternative definition of the effective area can be expressed as [18]:

$$A^{\nu}_{\text{eff}}(E_\nu) = A \cdot P_{\nu_\mu}(E_\nu, E^{\text{thr}}_\mu) \cdot \epsilon \cdot e^{-\sigma(E_\nu)\rho N_A Z(\theta)} \tag{6.1}$$

where;

- A $\left[\text{cm}^2\right]$ is the geometrical projected detector surface;
- $P_{\nu_\mu}(E_\nu, E^{\text{thr}}_\mu)$ represents the probability that a neutrino with energy E_ν produces a muon arriving with a residual threshold energy E^{thr}_μ at the detector;
- ϵ corresponds to the fraction of muons with energy E^{thr}_μ that are detected. It includes the trigger efficiency, the reconstruction efficiency and the probability to pass the analysis selection cuts;
- the term $e^{-\sigma(E_\nu)\rho N_A Z(\theta)}$, where $\sigma(E_\nu)$ is the total neutrino cross section, ρN_A is the target nucleon density and $Z(\theta)$ is the neutrino path as a function of the zenith angle, takes into account the neutrino absorption.

In Table 6.1 are reported the computed values of KM3NeT/ARCA21 effective area for atmospheric ν_μ and $\bar{\nu}_\mu$ charged current events that survive the event selection.

$\Delta\log_{10}(E_\nu/\text{GeV})$	Effective area $\left[\text{cm}^2\right]$
$2.0 - 2.7$	$8.85 \cdot 10^{-1}$
$2.7 - 3.4$	$3.67 \cdot 10^{1}$
$3.4 - 4.2$	$1.12 \cdot 10^{3}$
$4.2 - 5.0$	$1.63 \cdot 10^{4}$

Table 6.1: Effective area of KM3NeT/ARCA21 for ν_μ and $\bar{\nu}_\mu$ charged current events that survive to the final event selection. The values are computed within all energy ranges involved in the unfolding.

Ringraziamenti

In this last Chapter, ...

Volevo innanzitutto ringraziare la Prof.ssa Margiotta, che mi ha permesso di lavorare a questo progetto, dandomi anche l'opportunità di due bellissime trasferte al CERN e ai LNS di Catania. La ringrazio per avermi aperto le porte del mondo della ricerca e per avermi fatto capire quale è la mia strada. Volevo anche ringraziare tutti i membri del gruppo KM3NeT di Bologna, che mi hanno fatto sentire da subito parte del gruppo e mi hanno stimolato durante tutti gli incontri. In particolare, un immenso grazie va a Rasa e Francesco. Ho visto in loro la passione per questo lavoro e una grande volontà di passarmi tutte le loro conoscenze. È stato davvero bello lavorare con voi! {In het bijzonder gaat een enorme dank uit naar Rasa en Francesco. Ik heb in jullie de passie voor dit werk gezien en een grote bereidheid om al jullie kennis met mij te delen. Het was echt heel fijn om met jullie te werken!}

Un grazie va a tutti i regaz di "Per fare robe" con cui ho trascorso due anni bellissimi a Bologna, sperando di poter continuare le balotte anche dopo che inizieremo a disperderci tutti in giro per il mondo.
A tutti i miei amici fisici con cui ho trascorso prima i tre anni a Parma e poi (quasi con tutti) i restanti due a Bologna. È stato bellissimo condividere questo percorso con voi e sono convintissimo che se siamo arrivati qua ora, gran merito è dovuto a ciò che abbiamo costruito insieme in questi anni.
Agli amici di sempre, Fara, Pot e Guate. Con voi ho condiviso qualsiasi tipo di esperienza e di momento, dai più belli ai più brutti. Avete sicuramente un posto speciale in questi ringraziamenti. Siamo cresciuti insieme, siamo cambiati insieme, abbiamo intrapreso strade diverse eppure, anche se nel nostro modo alquanto discutibile di comunicare, ci siamo sempre stati l'uno per l'altro.

A Laura. Anche noi ci siamo incontrati la prima volta tra le mura di fisica a Parma. Chi l'avrebbe mai detto che dopo quasi 6 anni saremo stati qua, lo stesso identico giorno, ancora a condividere insieme questo traguardo? Tutto ciò che abbiamo vissuto insieme è stato davvero speciale. Sei stata un punto di riferimento per me in questi anni e mi hai fatto crescere veramente tanto. Grazie di tutto e che questo sia solo l'inizio di un nuovo periodo bellissimo insieme.

A Seba, Jess e ai miei nipotini Johannes e Emiliana. Anche se siamo lontani e ci sentiamo poco, vi voglio un mondo di bene.
Alle mie zie Luciana e Francesca, ai miei zii Andrea e Renato, ai miei cugini Riccardo e Letizia, alle mie nonne Marcellina e Lucia, ai miei nonni Bruno e Amadio. Grazie a tutti dell'amore che mi avete dato.

Infine, sicuramente, il grazie più grande va ai miei genitori. Grazie per tutto l'amore che mi date, grazie di non avermi mai fatto mancare niente e grazie di essere sempre stati di supporto. Vi ammiro per come siete e per quello che fate, se sono così oggi gran merito è sicuramente vostro. Anche se non lo dico spesso, vi voglio tantissimo bene, grazie!!

2005, Dipartimento di Fisica Unipr.
In primo piano mio fratello, neolaureato magistrale in fisica. Appena dietro, io.
Volevo chiudere così questo ciclo, che forse era già iniziato 21 anni fa. In futuro chissà, magari spunterà fuori un articolo con due Bernuzzi tra gli autori ;-P

Bibliography

[1] W. Pauli. “The idea of the neutrino”. In: *Physics Today* (Sept. 1978). DOI: `10.1063/1.2995181`. URL: `https://physicstoday.aip.org/features/the-idea-of-the-neutrino`.

[2] J. Chadwick. “The Existence of a Neutron”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 136.830 (1932), pp. 692–708. ISSN: 09501207. URL: `http://www.jstor.org/stable/95816`.

[3] C. L. Cowan et al. “Detection of the Free Neutrino: A Confirmation”. In: *Science* 124.3212 (1956), pp. 103–104. ISSN: 00368075, 10959203. URL: `http://www.jstor.org/stable/1751492`.

[4] Mark Thomson. *Modern particle physics.* New York: Cambridge University Press, Oct. 2013. ISBN: 978-1-107-03426-6, 978-1-139-52536-7. DOI: `10.1017/CBO9781139525367`.

[5] Y. Fukuda et al. “Evidence for oscillation of atmospheric neutrinos”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: `10.1103/PhysRevLett.81.1562`. arXiv: `hep-ex/9807003`.

[6] G. Giacomelli and A. Margiotta. “Macro results on atmospheric neutrinos”. In: *Nucl. Phys. B Proc. Suppl.* 145 (2005). Ed. by P. Bernardini, G. Fogli, and E. Lisi, pp. 116–119. DOI: `10.1016/j.nuclphysbps.2005.03.047`. arXiv: `hep-ex/0504029`.

[7] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model.* Cambridge University Press, Mar. 2014. ISBN: 978-1-107-03473-0, 978-1-107-03473-0. DOI: `10.1017/9781139540940`.

[8] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: `10.1103/PhysRev.105.1413`. URL: `https://link.aps.org/doi/10.1103/PhysRev.105.1413`.

[9] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino Cross Sections Across Energy Scales”. In: *Rev. Mod. Phys.* 84 (2012), pp. 1307–1341. DOI: `10.1103/RevModPhys.84.1307`. arXiv: `1305.7513 [hep-ex]`.

[10] Sheldon L. Glashow. “Resonant Scattering of Antineutrinos”. In: *Phys. Rev.* 118 (1 Apr. 1960), pp. 316–317. DOI: `10.1103/PhysRev.118.316`. URL: `https://link.aps.org/doi/10.1103/PhysRev.118.316`.

[11] M. G. Aartsen et al. “Detection of a particle shower at the Glashow resonance with IceCube”. In: *Nature* 591.7849 (2021). *[Erratum: Nature 592, E11 (2021)]*, pp. 220–224. DOI: `10.1038/s41586-021-03256-1`. arXiv: `2110.15051 [hep-ex]`.

[12] S. M. Bilenky. “Bruno Pontecorvo — pioneer of neutrino oscillations”. In: *Il Nuovo Cimento* 37.3 (2014). *Published online 29 July 2014*, pp. 21–28. DOI: `10.1393/ncc/i2014-11760-0`.

[13] J. H. Christenson et al. “Evidence for the 2π Decay of the K_2^0 Meson”. In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: `10.1103/PhysRevLett.13.138`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.13.138`.

[14] Carlo Giunti and Chung W. Kim. *Fundamentals of Neutrino Physics and Astrophysics.* Oxford: Oxford University Press, 2007. ISBN: 9780198508717. DOI: `10.1093/acprof:oso/9780198508717.001.0001`.

[15] Boris Kayser. “Majorana Neutrinos and their Electromagnetic Properties”. In: *Phys. Rev. D* 26 (1982), p. 1662. DOI: `10.1103/PhysRevD.26.1662`.

[16] G. Bellini et al. “Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy”. In: *Phys. Rev. D* 89 (11 June 2014), p. 112007. DOI: `10.1103/PhysRevD.89.112007`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.89.112007`.

[17] Brian O Fearraigh. “Following the light - Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA”. PhD thesis. Amsterdam U., 2024.

[18] Maurizio Spurio. *Particles and Astrophysics. A Multi-Messenger Approach.* Astronomy and Astrophysics Library. Springer, 2015. ISBN: 978-3-319-34539-0, 978-3-319-08051-2. DOI: `10.1007/978-3-319-08051-2`.

[19] Carmelo Evoli. *The Cosmic-Ray Energy Spectrum.* Dec. 2020. DOI: `10.5281/zenodo.4396125`. URL: `https://doi.org/10.5281/zenodo.4396125`.

[20] A A Watson. “High-energy cosmic rays and the Greisen–Zatsepin–Kuz’min effect”. In: *Reports on Progress in Physics* 77.3 (Feb. 2014), p. 036901. DOI: `10.1088/0034-4885/77/3/036901`. URL: `https://doi.org/10.1088/0034-4885/77/3/036901`.

[21] Francesco Filippini. “Search for a diffuse astrophysical neutrino emission from the Galactic Ridge with the first KM3NeT/ARCA data and strategies for real-time follow-up searches of rapid transient phenomena”. PhD thesis. Bologna U., 2024. DOI: `10.48676/unibo/amsdottorato/11156`.

[22] Matthew G. Baring. “Diffusive shock acceleration: The Fermi mechanism”. In: *32nd Rencontres de Moriond: High-Energy Phenomena in Astrophysics*. 1997, pp. 97–106. arXiv: `astro-ph/9711177`.

[23] Anne Zilles. *Emission of Radio Waves in Particle Showers: Validation of Microscopic Simulations with the SLAC T-510 Experiment and their Potential in the Future Square Kilometre Array.* Jan. 2017. ISBN: 978-3-319-63410-4. DOI: `10.1007/978-3-319-63411-1`.

[24] S. Navas et al. “Review of particle physics”. In: *Phys. Rev. D* 110.3 (2024), p. 030001. DOI: `10.1103/PhysRevD.110.030001`.

[25] A. Albert et al. “Measurement of the atmospheric ν_e and ν_μ energy spectra with the ANTARES neutrino telescope”. In: *Phys. Lett. B* 816 (2021), p. 136228. DOI: `10.1016/j.physletb.2021.136228`. arXiv: `2101.12170 [hep-ex]`.

[26] T. K. Gaisser and M. Honda. “Flux of atmospheric neutrinos”. In: *Ann. Rev. Nucl. Part. Sci.* 52 (2002), pp. 153–199. DOI: `10.1146/annurev.nucl.52.050102.090645`. arXiv: `hep-ph/0203272`.

[27] Maurizio Spurio. *Probes of Multimessenger Astrophysics. Charged cosmic rays, neutrinos, γ-rays and gravitational waves.* Astronomy and Astrophysics Library. Springer, 2018. ISBN: 978-3-319-96853-7, 978-3-319-96854-4. DOI: `10.1007/978-3-319-96854-4`.

[28] Andrea Palladino, Maurizio Spurio, and Francesco Vissani. “Neutrino Telescopes and High-Energy Cosmic Neutrinos”. In: *Universe* 6.2 (2020), p. 30. DOI: `10.3390/universe6020030`. arXiv: `2009.01919 [astro-ph.HE]`.

[29] M. Ageron et al. “ANTARES: The first undersea neutrino telescope”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 656.1 (2011), pp. 11–38. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2011.06.103`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900211013994`.

[30] S. Adrian-Martinez et al. “Letter of intent for KM3NeT 2.0”. In: *J. Phys. G* 43.8 (2016), p. 084001. DOI: `10.1088/0954-3899/43/8/084001`. arXiv: `1601.07459 [astro-ph.IM]`.

[31] *KM3NeT - opens a new window on our universe.* KM3NeT. URL: `https://www.km3net.org/news-media/blogs/`.

[32] Jerzy Mikolaj Manczak. “Search for Neutrino Non-Standard Interactions with the KM3NET/ORCA 6 Lines Detector”. PhD thesis. U. Valencia (main), 2022.

[33] S. Aiello et al. “The KM3NeT multi-PMT optical module”. In: *JINST* 17 (2022), P07038. DOI: `10.1088/1748-0221/17/07/P07038`. arXiv: `2203.10048 [astro-ph.IM]`.

[34] E. de Wolf et al. “A launching vehicle for optical modules of a deep-sea neutrino telescope”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 725 (2013). *VLVT 11, Erlangen, Germany, 12 - 14 October, 2011*, pp. 241–244. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2012.11.146`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900212014933`.

[35] *nikhef.nl/˜dveijk/ARCA.html.* URL: `https://www.nikhef.nl/~dveijk/ARCA.html`.

[36] S. Aiello et al. “Astronomy potential of KM3NeT/ARCA”. In: *Eur. Phys. J. C* 84.9 (2024), p. 885. DOI: `10.1140/epjc/s10052-024-13137-2`. arXiv: `2402.08363 [astro-ph.HE]`.

[37] Federico Versari. “Measurement of the atmospheric electron and muon neutrino flux with the ANTARES neutrino telescope”. PhD thesis. Bologna U., Bologna U., 2021. DOI: `10.6092/unibo/amsdottorato/9664`.

[38] Dimitrios Stavropoulos. “Measurement of the atmospheric muon neutrino flux with the KM3NeT/ORCA6 detector”. PhD thesis. Natl. Tech. U., Athens, 2024. DOI: `10.12681/eadd/57393`.

[39] S. Aiello et al. “gSeaGen: The KM3NeT GENIE-based code for neutrino telescopes”. In: *Computer Physics Communications* 256 (Nov. 1, 2020), p. 107477. ISSN: 0010-4655. DOI: `10.1016/j.cpc.2020.107477`. URL: `https://www.sciencedirect.com/science/article/pii/S0010465520302241`.

[40] C. Andreopoulos et al. “The GENIE Neutrino Monte Carlo Generator”. In: *Nucl. Instrum. Meth. A* 614 (2010), pp. 87–104. DOI: `10.1016/j.nima.2009.12.009`. arXiv: `0905.2517 [hep-ph]`.

[41] Y. Becherini et al. “A parameterisation of single and multiple muons in the deep water or ice”. In: *Astroparticle Physics* 25.1 (Feb. 2006), pp. 1–13. ISSN: 09276505. DOI: `10.1016/j.astropartphys.2005.10.005`. arXiv: `hep-ph/0507228`. URL: `http://arxiv.org/abs/hep-ph/0507228`.

[42] G. Carminati, A. Margiotta, and M. Spurio. “Atmospheric MUons from PArametric formulas: a fast GEnerator for neutrino telescopes (MUPAGE)”. In: *Computer Physics Communications* 179.12 (Dec. 2008), pp. 915–923. ISSN: 00104655. DOI: `10.1016/j.cpc.2008.07.014`. arXiv: `0802.0562[physics]`. URL: `http://arxiv.org/abs/0802.0562`.

[43] A. Albert et al. “Monte Carlo simulations for the ANTARES underwater neutrino telescope”. In: *JCAP* 01 (2021), p. 064. DOI: `10.1088/1475-7516/2021/01/064`. arXiv: `2010.06621 [astro-ph.HE]`.

[44] Yann Coadou. “Boosted decision trees”. In: Mar. 2022, pp. 9–58. DOI: `10.1142/9789811234033_0002`. arXiv: `2206.09645[physics]`. URL: `http://arxiv.org/abs/2206.09645`.

[45] *ROOT: TMVA::MethodBDT Class Reference.* URL: `https://root.cern.ch/doc/v636/classTMVA_1_1MethodBDT.html`.

[46] S. Aiello et al. “Atmospheric muons measured with the KM3NeT detectors in comparison with updated numeric predictions”. In: *Eur. Phys. J. C* 84.7 (2024), p. 696. DOI: `10.1140/epjc/s10052-024-13018-8`. arXiv: `2403.11946 [astro-ph.HE]`.

[47] Stefan Schmitt. “TUnfold: an algorithm for correcting migration effects in high energy physics”. In: *JINST* 7 (2012), T10003. DOI: `10.1088/1748-0221/7/10/T10003`. arXiv: `1205.6201 [physics.data-an]`.

[48] G. D’Agostini. “A multidimensional unfolding method based on Bayes’ theorem”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/0168-9002(95)00274-X`. URL: `https://www.sciencedirect.com/science/article/pii/016890029500274X`.

[49] Rene Brun and Fons Rademakers. “ROOT — An object oriented data analysis framework”. In: *Nucl. Instrum. Meth. A* 389.1-2 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: `10.1016/s0168-9002(97)00048-x`.

[50] Daniel Gerth. “A new interpretation of (Tikhonov) regularization”. In: *Inverse Problems* 37.6 (June 2021), p. 064002. DOI: `10.1088/1361-6420/abfb4d`. URL: `https://doi.org/10.1088/1361-6420/abfb4d`.

[51] Per Christian Hansen. "The L-Curve and Its Use in the Numerical Treatment of Inverse Problems". In: vol. 4. Jan. 2001, pp. 119–142.

[52] M. Honda et al. "Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model". In: *Physical Review D* 92.2 (July 7, 2015), p. 023004. DOI: `10.1103/PhysRevD.92.023004`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.92.023004`.

[53] Juan Pablo Yañez and Anatoli Fedynitch. "Data-driven muon-calibrated neutrino flux". In: *Phys. Rev. D* 107.12 (2023), p. 123037. DOI: `10.1103/PhysRevD.107.123037`. arXiv: `2303.00022 [hep-ph]`.

[54] R. Abbasi et al. "Measurement of the atmospheric neutrino energy spectrum from 100 GeV to 400 TeV with IceCube". In: *Phys. Rev. D* 83 (2011), p. 012001. DOI: `10.1103/PhysRevD.83.012001`. arXiv: `1010.3980 [astro-ph.HE]`.