The Simulations Chain of the MURAVES Experiment

M. Moussawi,[1] S. Basnet,[1] L. Bonechi,[2] L. Cimmino,[3,4] R. D'Alessandro,[2,5] M. D'Errico,[3,4] A. Giammanco,[1] R. Karnam,[1] G. Macedonio,[6] C. Rendon,[7] A. Samalan,[7] G. Saracino,[3,4] and M. Tytgat[7]

[1]*Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université catholique de Louvain, Louvain-la-Neuve, Belgium*

[2]*INFN sez. di Firenze, Florence, Italy*

[3]*University of Naples Federico II, Napoli, Italy*

[4]*INFN sez. di Napoli, Naples, Italy*

[5]*University of Florence, Florence, Italy*

[6]*INGV, Osservatorio Vesuviano, Naples, Italy*

[7]*Department of Physics and Astronomy, Ghent University, Ghent, Belgium*

Corresponding author: M. Moussawi
Email address: marwa.almoussawi@uclouvain.be

Abstract

The MUon RAdiography of VESuvius (MURAVES) project is aimed at studying the summital cone of Mt. Vesuvius, an active and hazardous volcano near Naples, Italy. A detailed Monte Carlo simulation framework is necessary in order to investigate the effects of the experimental constraints and to perform comparisons with the actual observations. Our Monte Carlo setup combines a variety of Monte Carlo programs that address different aspects of cosmic muon simulation, from muon generation in the Earth's upper atmosphere to the response of the detector, including the interactions with the material of the volcano. We will elaborate on the rationale for our technical choices, including the trade-off between speed and accuracy, and on the lessons learned, which are of general interest for similar use cases in muon radiography.

1. INTRODUCTION

Muon radiography is an imaging method that uses natural cosmic-ray muons to characterize the density distribution of large-scale natural or human-made objects, complementary to conventional imaging techniques [1]. Imaging is performed from the measurements of the absorption profiles of muons as they pass through matter [2]. The MUon RAdiography of VESuvius (MURAVES) project [3] is aimed at studying the inner structure of the upper part of Mt. Vesuvius, an active volcano near Naples, Italy. The main expected outcome of MURAVES will be a muography of Mt. Vesuvius, i.e., to obtain a 2D image of the density structure by measuring the attenuation of the muons flux through the volcano cone. By combining sufficiently precise data from muography with those from gravimetric and seismic measurement campaigns, it may become possible to determine the stratigraphy of the lava plug at the bottom of the Vesuvius crater, in order to infer potential eruption pathways.

The detection setup of MURAVES consists of three identical and independent tracking hodoscopes, constituted of scintillator bars coupled to SiPMs, which are already installed on Mt. Vesuvius and fully operational. Each hodoscope includes four tracking stations, each with a surface of roughly $1\,\text{m}^2$, and a thick lead wall between the 3rd and the 4th station which is used for muon energy discrimination, as explained later in this document.

Comparing data with Monte Carlo (MC) simulations is very important for the imaging of a target. A popular approach in muography is to compare various density hypotheses in MC with the observed transmission map, seeking the simulated hypothesis that provides the best fit to real data. In other types of studies, one is interested in anomalies with respect to the expected map from MC. In all those cases, various effects can bias the expectations; thus, it is important to quantify their impact by testing various modeling assumptions in MC.

In this paper, we report on the development of an MC simulation chain based on the schematic steps illustrated in Figure 1, and we present a series of simulation studies that are being conducted to investigate the effects of the experimental constraints and to perform comparisons with the actual observations. Our simulation chain starts from the generation of cosmic muons with realistic angular and momentum distributions. Several tools for this task are available from *ab initio* simulations that simulate the development of the entire cosmic shower in the atmosphere (e.g., CORSIKA [4]), to parametric simulations (such as CRY [5] and EcoMug [6]), i.e., Monte Carlo programs that sample muon properties from pre-defined distributions. For the simulation of muon traversal through the volcano, we are investigating PUMAS [7, 8], a "backward MC" designed to optimize computational efficiency by only simulating the muons whose trajectories actually reach the detectors. A detailed simulation of the detector response and of its geometry becomes more and more important as the accumulation of statistics makes the muography of the volcano more

precise; for that, our work-horse is GEANT4 [9], a multipurpose tool very popular in particle, astroparticle, nuclear, and medical physics. To mimic the real data, GEANT4 raw hits are converted to clusters through simulated digitization. After the application of the tracking algorithm, we compare a few high-level quantities with data.
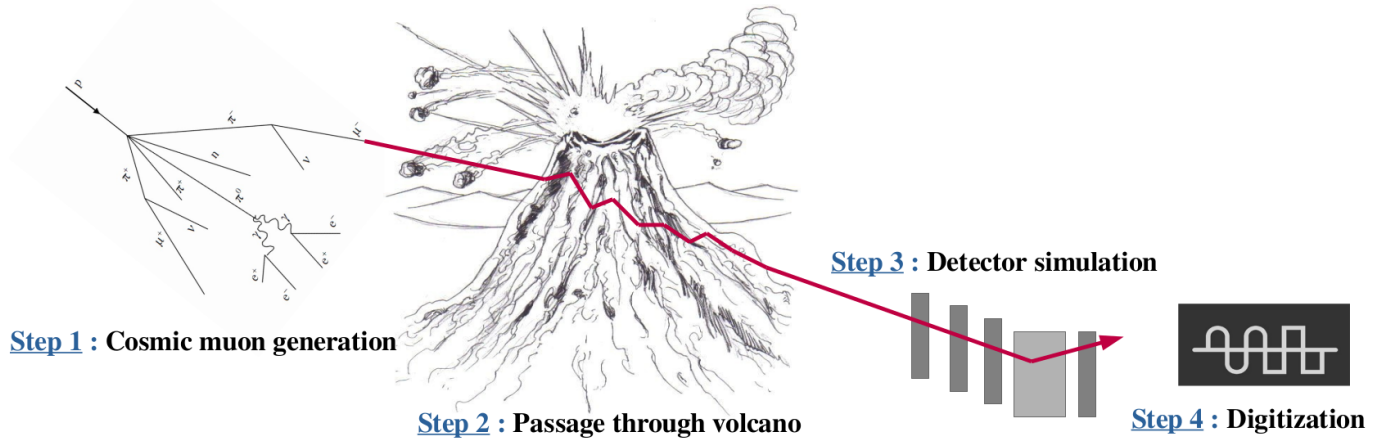

FIGURE 1: A schematic representation of the MURAVES simulation chain.

2. COSMIC-RAY MUON GENERATION

The first simulation step is the generation of muons, as well as background particles such as electrons, positrons, and hadrons, according to realistic models.

We examined and compared three Monte Carlo tools: CRY [5], CORSIKA [4], and EcoMug [6]. CRY and EcoMug are parametric simulations, while CORSIKA simulates the evolution of extensive air showers in the atmosphere step by step. EcoMug allows the generation of cosmic muons according to user-defined parametrizations of their differential flux, and a default one is provided based on data from the ADAMO experiment [10]. On the other hand, CRY's parameterization, based on MCNPX simulation [11], takes into account geomagnetic effects on the cosmic flux depending on time (solar cycle), latitude, and altitude (for the latter, CRY provides three options: 0, 2100, and 11300 m above sea level). Instead, in CORSIKA, the user can choose from various models of hadronic interaction at low and high collision energies.

Both absorption and scattering of the muons depend strongly on the particle momentum. Of particular importance in this context is the very high-energy end of the spectrum, as only the most energetic muons are able to pass through the deepest and most interesting regions of a volcano; e.g., it takes a bit less than 10 TeV for a muon to pass through 2.5 km of rock. In [12], we compared the muon flux as a function of kinetic energy for CRY and five different combinations of low- and high-energy hadronic interaction models available in CORSIKA in the range between 1 MeV and 10 TeV, finding them all in agreement concerning the peak position and the general shape but with a certain spread in predictions at high energy. These differences need to be taken into account as systematic uncertainties in the MURAVES measurements. To that end, the possibility in CORSIKA to vary the hadronic model can be exploited for assessing a physics-motivated modeling uncertainty.

In a particle generator, the user has to define the surface from where the particles start their trajectory, which is called the "generation surface." While most cosmic-particle generators, including CRY and CORSIKA, are limited to flat generation surfaces, EcoMug allows shooting muons also from cylindrical and hemispherical surfaces, ensuring the correct angular and momentum distributions. For the studies presented in this article, we set the generation surfaces at a very small distance from the detector, to minimize the fraction of muons outside of acceptance.

Figure 3 shows the distribution of zenith angle θ (a) and kinetic energy E (b) of muons simulated using CRY, CORSIKA (with the DPMJET [13] and GHEISHA [14] hadronic interaction models used, respectively, for high- and low-energy collisions in the atmosphere [15]), and the three EcoMug modes (as shown in Figure 2). It can be seen that the three generators yield consistent distributions for E; in θ, all generators are consistent when used with a flat generation surface, while the cylindrical configuration of EcoMug shows a substantial angular bias. Intuitively, this is explained by recalling that $\theta \sim 0°$ (very vertical muons) can never happen in this configuration as it lacks the top of the cylinder.
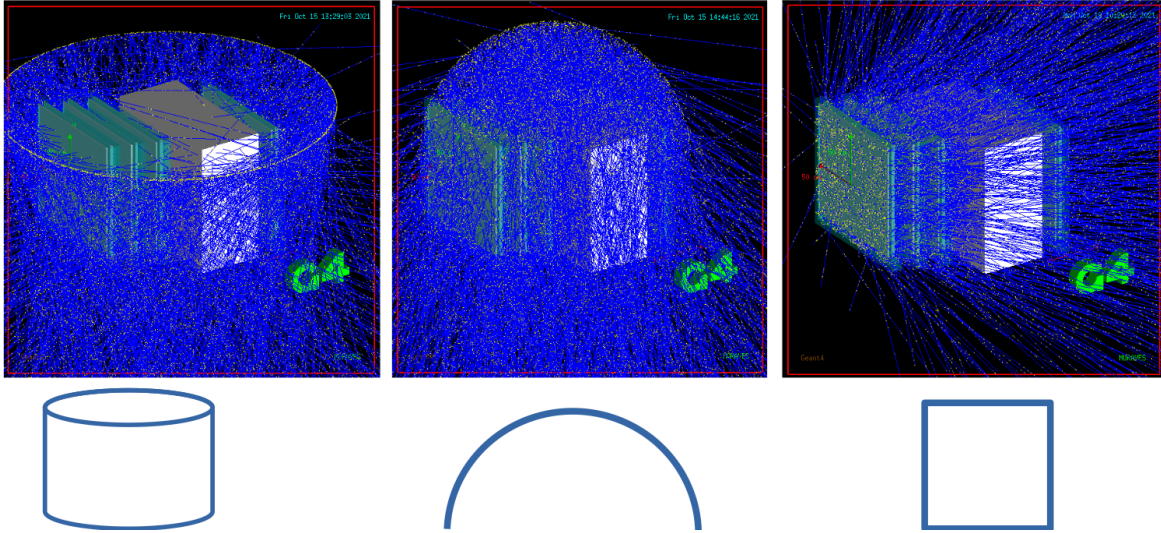
FIGURE 2: The three EcoMug modes: cylindrical (left), hemispheric (center), and flat (right) generation surfaces, here shown interfaced to the GEANT4 model of the MURAVES hodoscope described in Section 4.

An important practical consideration is the speed of execution, for which there are huge differences: we estimated that to generate 10^5 muons on a standard CPU, EcoMug employs O (sec) (depending on the mode: flat is faster, and hemisphere is slower, as thoroughly studied in [6]), and CRY O (min), while CORSIKA takes O (hours).
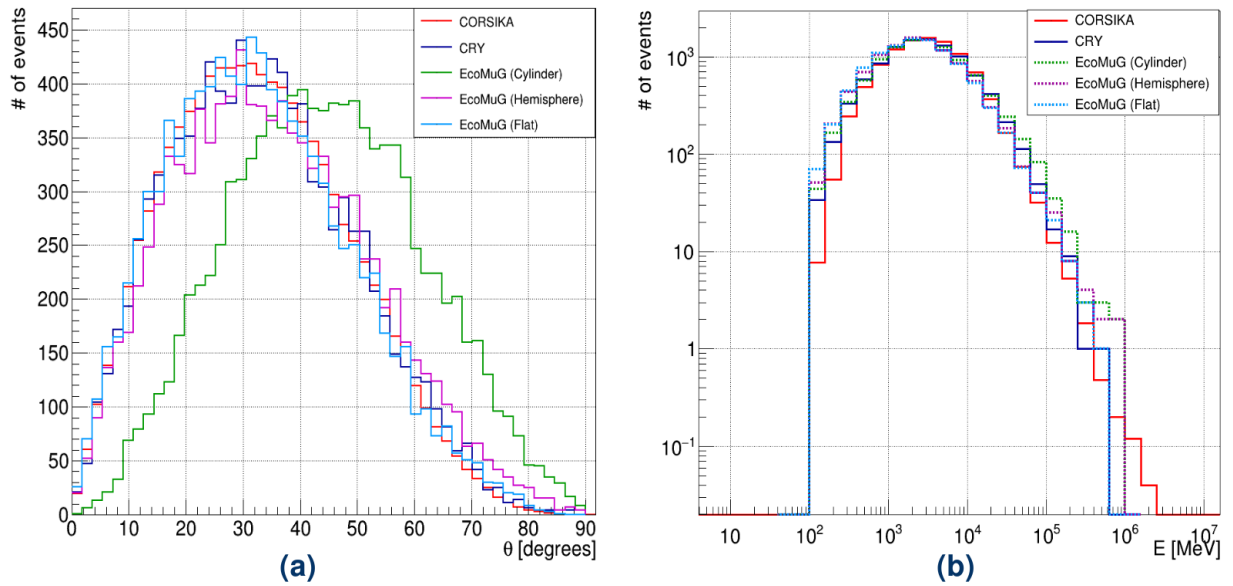

FIGURE 3: Distributions of zenith angle θ (a) and kinetic energy E (b) generated using CRY, CORSIKA, and three EcoMug modes (corresponding to flat, cylindrical, and hemispherical generation planes).

Table 1 compares the relative merits of the three generators examined, from the point of view of what is important for a muon radiography experiment similar to ours. For the purpose of MURAVES data analysis, we chose to use CRY as the main generator, while also maintaining CORSIKA and EcoMug (with its flat-surface option) for the purpose of estimating systematic uncertainties related to physics modeling. Both CRY and EcoMug can easily be interfaced with GEANT4 [9], meaning that by compiling the appropriate libraries it is possible to run muon generation and detector response simulation in the same job. Running GEANT4 on CORSIKA events requires, instead, saving the generator output in HepMC format [16], which complicates the storage logistics. Additional considerations, although less crucial, include the ability to simulate nonmuon backgrounds from cosmic showers (e.g., electrons, positrons, and hadrons) and multiple muons from the same shower (hence detected simultaneously in the same hodoscope, leading to rejection by the current tracking algorithm). Both features are naturally included in CORSIKA, as it develops the full showers. CRY contains parametrizations of all particles including backgrounds, based on MCNPX [11], and can handle multiple particles per event. EcoMug is not designed for multiple muons or for backgrounds, and examples are provided only for

muons; however, we implemented and tested a technical solution for simulating any kind of particle, after private communication with the authors, which might be of interest for future studies.

Generator	CORSIKA	CRY	EcoMug
Speed (10^5 muons)	O (hours)	O (min)	O (sec)
Modeling	Ab initio, several models available	Immutable, derived from MC [11]	User-defined, default from [10]
Accuracy	State of the art	Approximate	Approximate
Use with GEANT4	Complex	Easy	Easy
Systematics	Hadronic interaction models	Time dependence (solar cycle)	Alternative parameterizations

TABLE 1: Qualitative comparison of the features of CORSIKA [4], CRY [5], and EcoMug [6] which are of particular interest for muon radiography of a volcano.

3. PASSAGE OF MUONS THROUGH VESUVIUS

The probability of a muon surviving the crossing of a certain amount of rock is approximately determined by the opacity of the latter (defined as the integrated matter density ρ over the path length, i.e., $\omega \equiv \int_{\text{entry}}^{\text{exit}} \rho(x)dx$) and the initial energy of the muon. It also depends on the chemical composition of the rocks [17], but these effects are expected to be small, in the Vesuvius case, with respect to the precision expected to be achievable with MURAVES. The measurement of the transmitted flux of atmospheric muons through a target in different directions from a given point of view (the position of the muon detector) provides the two-dimensional projective measurement of the target, and the average density ($\bar{\rho}$) of the material of this target can be calculated by knowing the opacity using the equation $\omega = L\bar{\rho}$, where L is the total muon path length within the material. The path lengths corresponding to different zenith and azimuth angles of arrival of the muons, across the Vesuvius crater, are evaluated using a Digital Terrain Model (DTM) with 1 m resolution, derived using the data from [18].

The computation of the expected muon flux through Mt. Vesuvius is conducted using PUMAS [7, 8], a muon transport library based on a Backward Monte Carlo (BMC) technique. The concept of a BMC can be concisely described as "an MC run backward," i.e., from the detector to the sky. A BMC deals with the exact same problem as a forward MC, but it allows generating only the final states that are actually observable, thus reducing the computation time considerably. In our use case, it is important to simulate the scattering of the muons when they pass through the mountain. In locations that are close to the mountain, we can expect a considerable amount of low-energy muons suffering multiple scattering near the surface and being deflected into the acceptance of the detector in a fake direction.

The choice of the optimal observation point where to install the MURAVES detector was based on PUMAS (in collaboration with its authors), integrated with TURTLE [19] to extract the path lengths from the DTM. Various potential installation points, preselected based on logistical considerations, were studied with PUMAS, and the muon flux was simulated given the thickness and average density of the material. Among them, one was chosen based on its clear view of Mt. Vesuvius and on the signal purity expected over most directions. In this context, purity is defined as the ratio of the muon flux without scattering over the muon flux with scattering, and it is used as a figure of merit; low purity means a large contamination from low-energy muons whose entry and exit directions are different due to multiple scattering, hence blurring the muography resolution.

The accurate simulation of muon transport through the volcano is also important in order to translate the measured transmission map into an opacity map. The larger computation efficiency of PUMAS with respect to forward MC programs makes it a very appealing option, and this program has also additional appealing features such as the ability to easily simulate macroscopic mixtures of, e.g., rock and water; but the relative novelty of the BMC approach imposes the need to validate it against a well-known alternative. For that reason, we studied the energy loss and scattering of muons through homogeneous standard rock using PUMAS and GEANT4, and the results of the two simulations were compared. Muons with a flat energy distribution ranging from 0 GeV to 1000 GeV were simulated, and the muon path length was set to 500 m in standard rock. In GEANT4, standard rock was defined as a mixture of $CaCO_3$ and $MgCO_3$, i.e., with a density of 2.65 g/cm^3 and mass fractions of 52% O, 27% Ca, 12% C, and 9% Mg, while in PUMAS the predefined material definition for standard rock was selected with atomic number $Z = 11$, atomic mass $A = 22.0$, average excitation energy $I = 136.4$ eV, and density 2.65 g/cm^3. Given that in both cases the same material density was used and that also the radiation lengths appear similar,[1] we expect the physical difference between the two definitions to be insignificant. Figure 4 shows the distribution of energy loss (a) and scattering angle (b) evaluated using PUMAS (red line) and GEANT4 (blue line). We conclude that the two programs give reasonably close results.

4. DETECTOR SIMULATION

GEANT4 [9] is the most popular MC particle transport simulation in particle and nuclear physics, and it is also used in medical and space applications. For the MURAVES simulation studies, we built a detailed model of a MURAVES hodoscope in GEANT4,

---

[1]In the case of GEANT4, we can estimate the $1/X_0$ of the standard rock material as the mass fraction weighted average of the $1/X_0$ values of the individual elements, leading to $X_0 = 26.25$ g/cm^2, which matches closely the PDG value of $X_0 = 26.54$ g/cm^2 for standard rock as specified in [20].

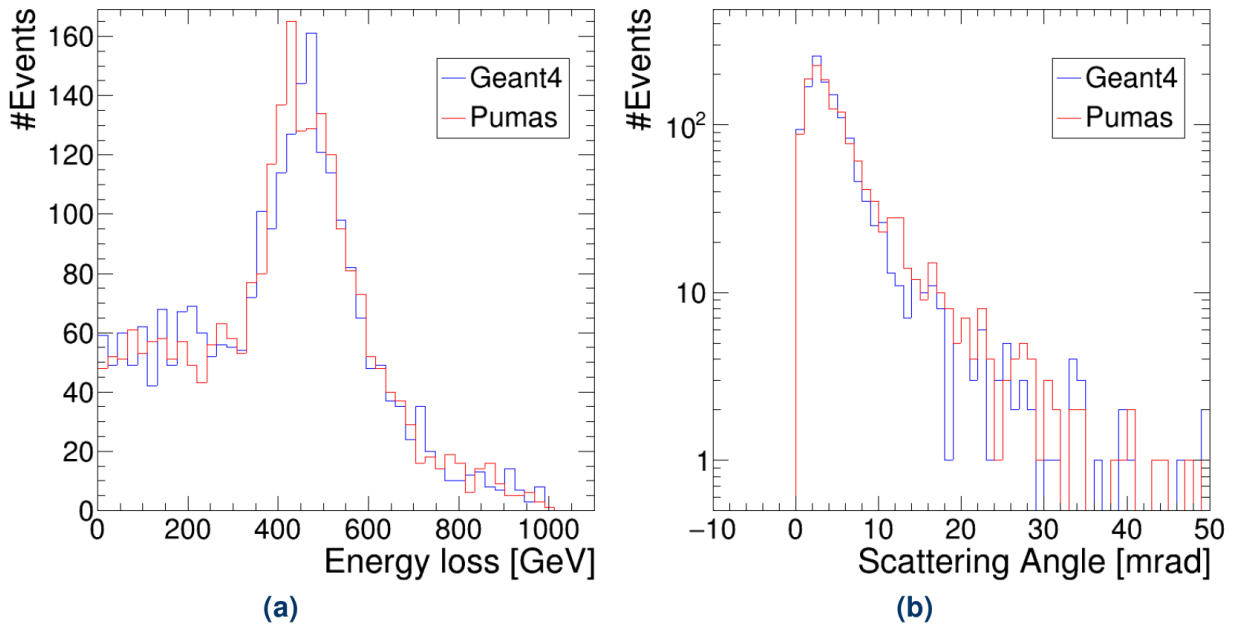**(a)**   **(b)**

FIGURE 4: Distributions of energy loss (a) and scattering angle (b) of 0–1000 GeV muons in 500 m of standard rock generated using GEANT4 and PUMAS.

interfaced with CRY as a source for all signal and background particles. For specific studies, single muons have been simulated at specific energies and incident positions using GEANT4 "particle gun" generator or General Particle Source class. We also produced samples where GEANT4 was interfaced with EcoMug, and where the input was taken from the HepMC files previously produced with CORSIKA.

The MURAVES setup consists of three identical and independent tracking hodoscopes, constituted of scintillator bars coupled to Silicon Photo-Multipliers (SiPMs). Two of them are positioned in "forward-looking mode" to take Mt. Vesuvius data, while one is installed in reverse orientation, to acquire a free-sky control sample. The MC simulation is applied to the three hodoscopes, considered strictly identical.

Each hodoscope includes four tracking stations. A thick lead wall of thickness 60 cm between the 3rd and the 4th stations acts as a "minimum muon energy" selector with a turn-on point of 0.9 GeV [12] and also reduces very effectively the background from electrons, positrons, or hadrons. Each station has two planes (X and Y) with 64 adjacent scintillator bars (divided into two modules) housed inside an aluminum container and of 1 m^2 sensitive area placed orthogonal to each other to extract the bidimensional information of the muon hits. The scintillator bars are of the triangular cross-section with a base width of 33 mm and height 17 mm and alternate upward and downward pointing tips.

We used GEANT4 to guide the choice of a possible analysis selection step based on the scattering induced by the lead wall between the 3rd and the 4th stations. The purpose is to further reduce the fraction of low-energy muons, after the passive filtering provided by the absorption in the lead wall, while keeping as many as possible of the high-energy ones, which carry the cleanest information about the depths of rock traversed. To perform this study, we generated 40 million muons with CRY, in free-sky conditions, and we selected the events containing a signal in all hodoscope stations. GEANT4 gives access to the scattering points within the lead wall; however, those are unobservable in real data; therefore, we approximate the actual scattering angle with the angle between the vector connecting the 1st- and 3rd-station hits, and the vector connecting the 3rd- and the 4th-station hits. This approximation mimics the practical requirement that can be applied on tracks in real data. Finally, we counted the frequency at which the events survive an uppercut on this angle. Figure 5 shows the outcome of this study, for muons belonging to various incoming energy categories. The horizontal axis indicates the upper cut on the scattering angle (in steps of 2.5°), while the vertical axis shows the frequency of survival of a muon, defined as the fraction of muons passing this cut, among those that give hits in all stations (hence, that were not absorbed in the lead wall). It can be seen that even a very restrictive cut of ≤5° preserves the vast majority of muons above 2 GeV (better than 90% between 2 and 3 GeV, and almost 100% above 3 GeV), while removing more than 40% of the muons between 1 and 2 GeV and about 70% of the sub-GeV muons that pass the lead wall, whose angular distribution is very broad.

5. DIGITIZATION, CLUSTERING, AND TRACKING

The output of GEANT4 contains "hits," corresponding to each interaction of a particle with the sensitive detector material (scintillator bars) for each simulation step. Hits are characterized by 3D position, time, energy deposition, and identity of the particle
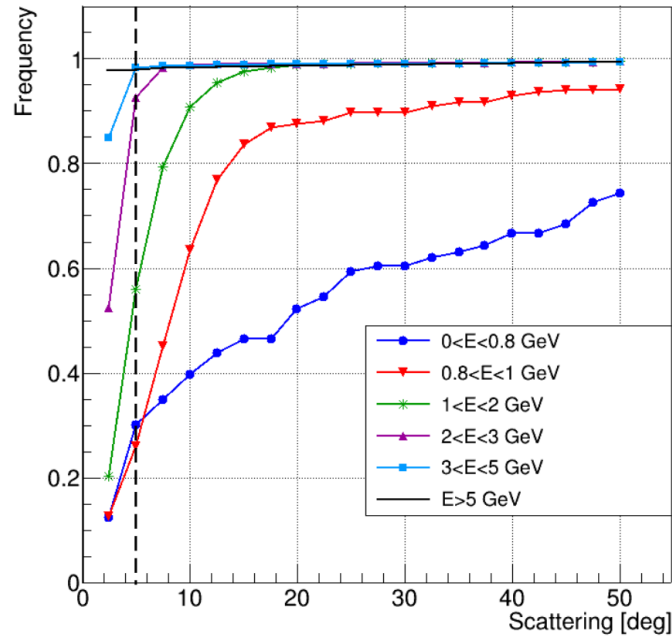
FIGURE 5: Efficiency of various possible upper cuts on the scattering angle for different ranges of incoming muons energy.

involved. In GEANT4, when a muon passes through a detector plane, it typically produces multiple hits in the scintillator bars, while we get only a single signal per bar in the real case.

Our next simulation step is the "digitization" of these GEANT4 raw hits, i.e., the quantization of both their positions and energy deposits. Energy deposits are first summed per scintillator bar and then converted to number of photoelectrons (nPE) using a conversion factor of 10 nPE every 1.4 MeV, based on the comparison of the peak deposited energies observed in real and simulated data, respectively. At this point, we have a single value of the energy deposit per bar, expressed in proper units, and we can apply the clustering algorithm. A cluster is a collection of adjacent bars whose signals pass a certain threshold; clusters may also be constituted of a single bar, in which case; however, the signal must pass a stricter threshold. When a cluster contains more than one bar, its position is calculated from their barycenter, i.e., the weighted average of the positions of the centers of the bars involved, where the weights are given by the corresponding signals, measured in a number of photoelectrons. The same clustering algorithm and thresholds are applied in MC as in real data.

Finally, the simulated clusters are fed into the tracking algorithm, which again is rigorously applied as in real data. The tracking is performed for X planes and Y planes independently. To validate the full simulation chain, the χ^2 values of the simulated tracks are compared with those of the real data from a free-sky calibration run. For this, 10^5 muon events were simulated, all crossing at least the first three planes of the detector, and the full reconstruction chain was applied. The result is shown in Figure 6, which also includes an alternative MC dataset where no thresholds are applied in the clustering algorithm. Although the trend is qualitatively similar between data and MC, the latter is significantly too optimistic; i.e., it has a lower tail than real data. This difference is found to be unrelated to the applied cluster thresholds, as demonstrated by the similarity of the MC distributions with and without thresholds. A possible insight into the source of this discrepancy is offered in Figure 7, which shows the comparison of the size of the clusters involved in the tracks selected with the standard cut $\chi^2 \leq 5$. As per the results, it is clear that MC is currently unable to simulate the long tail that we observe in the real data, and more investigations are needed in order to find out which unmodeled effects are the culprits; in fact, our simulation chain does not consider (yet) several nuisances, such as dark noise and various front-end electronics effects. The discrepancy in χ^2 distribution could be solved either by emulating such effects or by further cleaning the real data by more optimal selections at the cluster level.

6. CONCLUSIONS

We presented a complete Monte Carlo simulation chain for the MURAVES experiment. For the generation of cosmic muons, a comparison study between three different particle generators (CRY, CORSIKA, and EcoMug) was conducted, and it was decided to use CRY as the main generator for our analysis, with the other two being kept as options for the estimation of modeling systematics. The problem of efficient simulation of the muon passage through a large body, such as Vesuvius, is approached using PUMAS, whose accuracy we verified by comparing with GEANT4 in reference conditions. A detailed geometry of a MURAVES hodoscope was built using GEANT4 for the simulation of the detector response. The output of GEANT4 is digitized into a format that mimics raw data. In order to validate the performance of the CRY-GEANT4-digitization simulation chain, we compared the χ^2 distribution of real and simulated tracks, finding a difference that will deserve further investigation.
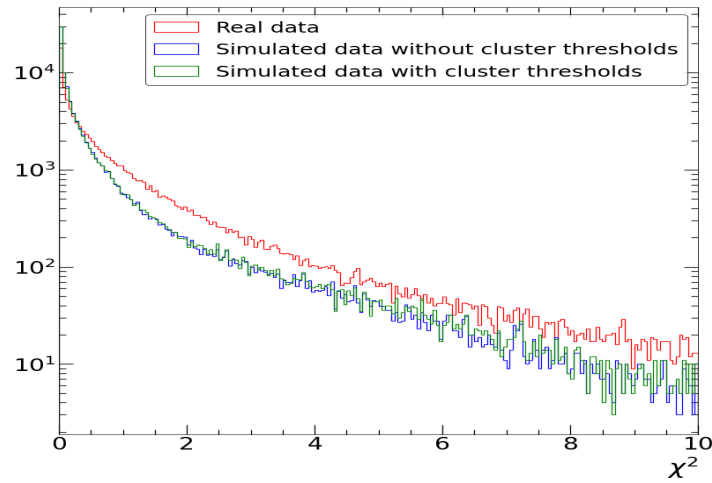
FIGURE 6: Comparison of the χ^2 of the tracks in real data and simulated data (the latter with and without clustering thresholds).
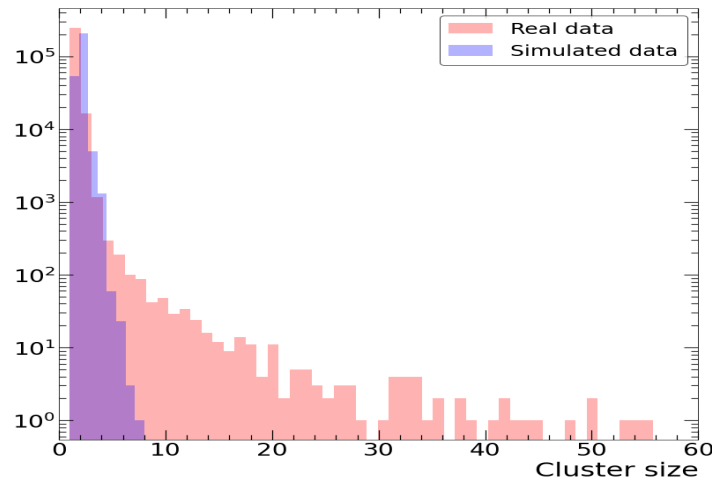

FIGURE 7: Comparison of the sizes of the clusters involved in the selected tracks ($\chi^2 < 5$) for real and simulated data.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

ACKNOWLEDGMENTS

References

[1] L. Bonechi, R. D'Alessandro, and A. Giammanco. Atmospheric muons as an imaging tool. *Rev. Phys.*, 5: 100038, 2020.

[2] A. Lechmann, D. Mair, A. Ariga, T. Ariga, A. Ereditato, R. Nishiyama, C. Pistillo, P. Scampoli, F. Schlunegger, and M. Vladymyrov. Muon tomography in geoscientific research—A guide to best practice. *Earth-Science Reviews*, 222: 103842, 2021.

[3] M. D'Errico et al. Muon radiography applied to volcanoes imaging: the MURAVES experiment at Mt. Vesuvius. *JINST*, 15(03): C03014, 2020.

[4] D. Heck et al. CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers. Technical Report FZKA 6019, Forschungszentrum Karlsruhe, 1998.

[5] C. Hagmann, D. Lange, and D. Wright. Cosmic-ray shower generator (CRY) for Monte Carlo transport codes. In *2007 IEEE Nuclear Science Symposium Conference Record*, volume 2, pages 1143–1146. IEEE, 2007.

[6] D. Pagano, G. Bonomi, A. Donzella, A. Zenoni, G. Zumerle, and N. Zurlo. EcoMug: An Efficient COsmic MUon Generator for cosmic-ray muon applications. *Nucl. Inst. Meth. A*, 1014: 165732, 2021.

[7] V. Niess, A. Barnoud, C. Cârloganu, and E. Le Ménédeu. Backward Monte-Carlo applied to muon transport. *Comput. Phys. Comm.*, 229: 54, 2018.

[8] V. Niess. The PUMAS library. *Comp. Phys. Comm*, 279: 108438, 2022.

[9] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Inst. Meth. A*, 506: 250, 2003.

[10] L. Bonechi, M. Bongi, D. Fedele, M. Grandi, S. Ricciarini, E. Vannuccini. Development of the ADAMO detector: test with cosmic rays at different zenith angles. *29th International Cosmic Ray Conference*, 9: 283, 2005.

[11] D. B. Pelowitz et al. MCNPX Users Manual Version 2.7.0. Technical Report LA-CP-11-00438, Los Alamos National Laboratory, 2011.

[12] A. Samalan, S. Basnet, L. Bonechi, L. Cimmino, R. D. Alessandro, M. D. Errico, A. Giammanco, R. Karnam, G. Macedonio, M. Moussawi, C. Rendon, G. Saracino, M. Tytgat. End-to-end simulations of the MUon RAdiography of VESuvius experiment. *JINST*, 17(01): C01015, 2022.

[13] J. Ranft. New features in DPMJET version II.5. Technical Report FZKA 5828, Forschungszentrum Karlsruhe, 1999.

[14] H. Fesefeldt. The Simulation of Hadronic Showers - Physics and Application. Technical Report PITHA-85/02, RWTH Aachen, 1985.

[15] D. Heck. Low energy hadronic interaction models. *Nucl. Phys. B Proc. Suppl.*, 151: 127–134, 2006.

[16] M. Dobbs and J. B. Hansen. The HepMC C++ Monte Carlo event record for High Energy Physics. *Comput. Phys. Commun.*, 134: 41–46, 2001.

[17] A. Lechmann, D. Mair, A. Ariga, T. Ariga, A. Ereditato, R. Nishiyama, C. Pistillo, P. Scampoli, F. Schlunegger, and M. Vladymyrov. The effect of rock composition on muon tomography measurements. *Solid Earth*, 9(6): 1517–1533, 2018.

[18] G. Vilardo, G. Ventura, E. Bellucci Sessa, and C. Terranova. Morphometry of the Campi Flegrei caldera (southern Italy). *Journal of maps*, 9(4): 635–640, 2013.

[19] V. Niess, A. Barnoud, C. Cârloganu, and O. Martineau-Huynh. TURTLE: A C library for an optimistic stepping through a topography. *Comp. Phys. Comm.*, 247: 106952, 2020.

[20] Particle Data Group: Atomic and nuclear properties of standard rock. `https://pdg.lbl.gov/2021/AtomicNuclearProperties/HTML/standard_rock.html`.

[21] N. Mori. GGS: a Generic Geant4 Simulation package for small- and medium-sized particle detection experiments. *Nucl. Instrum. Meth. A*, 1002: 165298, 2021.