The Cosmic Muon Images Demonstrator within the REINFORCE Project

Theodore Avgitas and Marteau Jacques on behalf of REINFORCE consortium
Institut de Physique des deux Infinis de Lyon (IP2I), Villeurbanne, France

Corresponding author: Theodore Avgitas
Email: avgitas@ipnl.in2p3.fr

Abstract

Cosmic Muon Images is a muon tomography citizen science framework developed within the EU-funded (GA-872859) REINFORCE project [1] (REsearch Infrastructures FOR Citizens in Europe). REINFORCE aspires to decrease the knowledge gap between research and society. Citizen scientists will engage in the process of scientific discovery by analyzing detector data from four major physics domains. Muon Tomography, Gravitational Waves, Neutrino Astronomy, and High Energy Physics join forces with people from all over the world to do science and discuss detection techniques, signal versus background discrimination, and environmental effects on measurement and broadcast the impact of fundamental research on everyday life and on societal progress. In this effort, we want to reach the broader audience possible, and this means also making our data accessible to as many people as possible. SonoUno [2] is a user centered software developed within REINFORCE that allows people with different sensory styles to explore scientific data, both visually and through sonorization. Cosmic Muon Images utilizes the Zooniverse [3] website to bring muon tomography to the general public and the citizen science community. Muon telescope data are visualized with 3D and 1D plots with the goal being the identification of patterns through a series of lines and points on these plots. The results of this pattern identification will be used to train machine learning (ML) algorithms to discriminate between signal and background events. We will then evaluate the performance of these ML algorithms to more traditional track reconstruction and event selection algorithms that are already in use. The key factor is how they compare with respect to speed and pattern recognition accuracy. Furthermore, a large dataset cataloged by the eye will act as a stepping-stone toward other studies like background identification and Monte Carlo simulation development. Citizen scientists will benefit through their exposure to calibration and detection techniques for muon telescopes to make their scientific contributions more robust. This will also provide a more fulfilling experience by learning new stuff which is for the most part reserved for higher education audiences. They are going to learn about the different applications of muon tomography in fields like volcanology, archeology, civil engineering, hazard monitoring, and others. A great effort is made toward the inclusion of school students through a series of schools and seminars coorganized together with other EU-funded projects (e.g., FRONTIERS Summer School 2021) since young people have much to gain from learning about interdisciplinarity between sciences and how scientists from different domains collaborate toward a common goal.

Keywords: muon tomography, cosmic rays, citizen science, machine learning, artificial intelligence, trajectographs, diaphane
DOI: 10.31526/JAIS.2022.252

1. INTRODUCTION

Citizen Science is the active participation of the public in scientific research that produces new knowledge or understanding [4]. It stems from human beings' natural desire for knowledge (Aristotle, Metaphysics) combined with the modern way of discovering this knowledge systematically through science. Citizen science projects can be found all over the world and in very diverse scientific domains, from zoology (e.g., wildlife, ornithology, lepidopterology) to geology (e.g., seismology, petrology/lithology) and from physics (e.g., particle physics, astronomy/astrophysics) to social sciences and arts. A citizen science project, depending on its nature and framework, gives the opportunity to the citizen scientists to participate in various stages of scientific discovery from the collection of data to their analysis and from the interpretation of the results to their publication. There are many studies that address the interplay between professionals and amateur scientists within the context of a citizen science project along with the ethical and practical aspects that arise from these kinds of collaborations [5]. In this publication, we document the aspirations and intents behind *Cosmic Muon Images* project, we present *REINFORCE* as the context within which our framework develops, and we focus our presentation mainly on our *Zooniverse* demonstrator that is the main channel of communication between our team and the public.

REINFORCE

“REINFORCE aspires to engage more than 100,000 citizens across Europe and beyond, to help enhance society’s science literacy and awareness, and to contribute in the production of scientific knowledge by citizens for society” [6]. This goal is realized by 11 partners involving leading Large Research Infrastructures and research institutions in the fields of frontier physics, citizen science, Science Education, Social Innovation, Policy, and Economic Competitiveness [1]. This partnership will focus on investigating the extent to which citizen scientists can contribute to the development of new knowledge in frontier science, whether citizens can apply this new knowledge in solving societal problems and ways to integrate the feedback that citizen scientists will provide. These questions will be answered through a series of steps that go beyond outreach and focus on the development and implementation of 4 cutting-edge citizen science projects in the fields of Gravitational Waves, Neutrino Astronomy, Particle Physics, and Cosmic Ray interplay with geoscience and archeology. Citizen scientists will analyze data and help enhance the discovery potential of the detectors behind these data. A very important aspect is the inclusion and empowerment of the public to contribute in the scientific mission of these Large Research Infrastructures, and these goals are addressed not only by the design of the projects themselves but also by efforts put toward the development and the training of various citizen communities. Great effort is also put toward the inclusion of sensitive citizen groups as well as of participants from secondary and tertiary education.

Zooniverse

“The Zooniverse is the world’s largest and most popular platform for people-powered research” [3]. Zooniverse provides a template for the development of citizen science projects along with a multitude of tools for analyzing data accordingly to the specific needs of the project. The nature of the projects is very diverse and covers multiple scientific disciplines like history, literature, physics, and many others. Many Zooniverse projects lead to discoveries and achieve new ways of understanding data, and this results in the publication of papers and studies. In the physics domain specifically, projects like Gravity Spy [7], Muon Hunters [8], and Higgs Hunters [9] have been very successful within the Zooniverse community, with large numbers of participants and noteworthy results that would be very difficult, if not impossible, to achieve without the public’s participation. This evident success was one of the reasons that REINFORCE bases the development of the four citizen science demonstrators on the Zooniverse platform.

The Demonstrators

The four demonstrators developed for the Zooniverse platform are GWitchHunters [10] (Gravitational Waves), Deep Sea Explorers [11] (Neutrino Astronomy), New Particle Search at CERN [12] (High Energy Physics), and Cosmic Muon Images [13] (Muon Tomography). Each one of these demonstrators uses real data from experimental instruments that are considered to be in the forefront of their respective fields.

GWitchHunters is utilizing data from the Virgo detector [14]. It focuses on the identification of a specific type of event registered by the Gravitational Wave detector known as a “glitch.” A wide tool set has been developed so that citizen scientists can analyze these data and in doing so “help scientists to study and reduce the noise in the gravitational detectors, with the goal of improving their sensitivity and catching more gravitational waves” (from [10]). Furthermore, “Deep Sea Explorers” uses data taken from the Cubic Kilometer Neutrino Telescope (KM3NeT) [15]. KM3NeT opens a new window to the Universe through the detection of neutrinos but at the same time provides a unique opportunity for oceanologists and other specialists to probe one of the deepest regions of the Mediterranean Sea. KM3NeT registers optical and acoustic data, and both are becoming available to the public through the Zooniverse demonstrator. The systematic categorization of these data is going to provide invaluable help for scientists “... to study bio-activity in the deep sea!” and to “... better understand marine sources of noise in the KM3NeT detector, making our search for neutrinos much easier”, as it is very eloquently written in the demonstrator’s page.

The goal of the “New Particle Search at CERN” demonstrator is to actively engage nonexpert citizen scientists in the searches for new elementary particles produced from high-energy proton-proton collisions at the LHC (CERN). This is achieved in a three-stage process with analysis tools developed for each of the stages [18]. This demonstrator not only uses tools from the Zooniverse website but also utilizes the HYPATIA [19] visualization tool in order to facilitate the citizen scientists’ understanding of events developed within the ATLAS particle detector [16].

The goal of the “Cosmic Muon Images” demonstrator is the use from various muon tomography experiments for cataloging by citizen scientists. The result of this cataloging will be used to further develop our experimental methods and to better understand the response of our detectors to cosmic particle atmospheric cascades.

Data Sonification

Scientific materials become widely available through visualization. Most of the people working with data are familiar with this method of realizing data through various kinds of plots (histograms, scatter plots, etc.). REINFORCE partners understand that this excludes some people from accessing the data and deprive them so much from the information included within this data but from the joy of discovery as well. An answer to the problem of sonorization comes from the team behind the sonoUno software [20]. SonoUno can be used as a standalone program installed by the user or through its recently developed web interface. It gives the users the opportunity to explore datasets in real time by also facilitating the different sensory styles of its users. The sonorization it provides is not specific to the REINFORCE datasets and this makes its development “futureproof” for use outside the scope of REINFORCE as well.

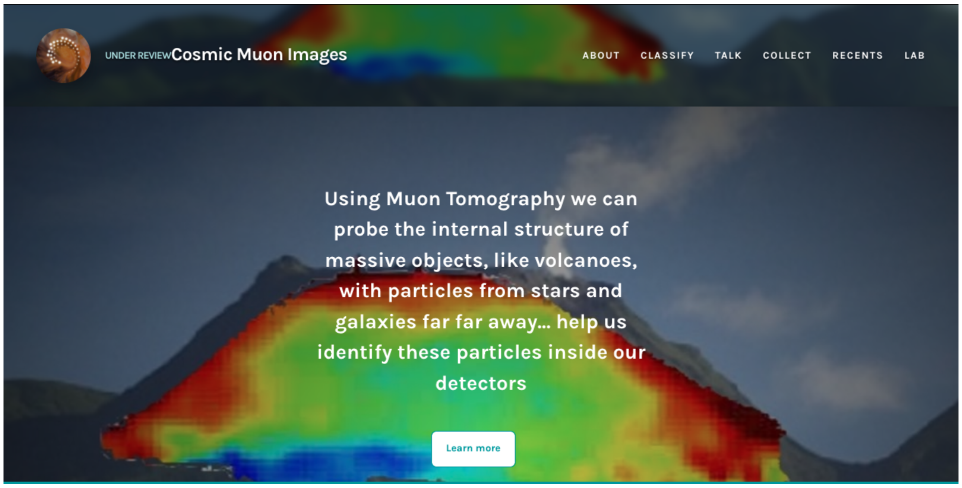

FIGURE 1: Cosmic Muon Images Welcome page. The Volcano is *La Grande Soufriere* and the colored histogram is the product of muon tomography measurement of the internal density distribution. In the top right part, there are several links to navigate the project.

2. COSMIC MUON IMAGES DEMONSTRATOR

The Cosmic Muon Images demonstrator is the online platform for the Cosmic Muon Images citizen science project developed in the context of REINFORCE. The goal is to use data from various muon tomography experiments for cataloging by citizen scientists. The results of this cataloging will be used to further develop our experimental methods and to better understand the response of our detectors to cosmic particle atmospheric cascades. The primary data catalog consists of data collected during the 2018 ARCHé (Archeology with Cosmic Rays, to probe Hellenic tumuli) muon tomography mission at the Tumulus site of Apollonia in Chalkidiki [21].

Every Zooniverse project follows a predefined structure for the citizen science website development. There are three basic parts, the introductory/welcome page, the main/about page, and the classification segment. Our demonstrator follows this same structure in order to lead the participants through the basics of muon tomography and the specifics of our project, to the event cataloging/classification.

2.1. Welcome Page

The welcome page on the Zooniverse site for Cosmic Muon Images illustrates the La Grande Soufriere volcano in Guadeloupe, with a muon tomography density histogram superimposed on it (Figure 1).

La Soufriere is of paramount importance, as much for volcanology as for the Diaphane team, since much of the R&D, methodology testing, and measurements have focussed around this very interesting active volcano. Just below this image, we find the “Get Started” section, which very briefly describes the two workflows that the citizen scientists are going to use in order to categorize the data provided to them. The first workflow uses reconstructed lines to depict the passage of muons through the detector, while the second workflow treats more complicated event topologies, in which the reconstruction algorithm has failed completely. This is a simple reconstruction algorithm based on the geometric characteristics of the triggered channels of the detector and the “true” muon events are found by applying empirical selection criteria. All events that do not comply with these criteria are flagged as background.

At the bottom of the page is a text very briefly describing the extent of muon tomography applications using the Diaphane team’s detectors and the purpose of the Cosmic Muon Images citizen science project. The goal is to create better images of the internal structure of massive objects by augmenting the signal-to-noise ratio through the work of the citizen scientists. Next to this text, there are seven links, pointing to the most important websites that have information related to several aspects of the Cosmic Muon Images project, such as the REINFORCE and DIAPHANE official sites.

2.2. About

The About page currently consists of three sections: *Research*, *The Team*, and *Frequently Asked Questions* (FAQ). There is also one other section that will be populated with materials as the project progresses. This is the *Education* section, which hosts materials with the purpose of helping citizen scientists to understand different aspects of muon tomography, from particle physics and the detection principles to the science behind the studied targets, like Geology, Volcanology, and Civil Engineering.

Research

The first goal of the research page is to expose all the scientific aspects of a muon tomography project. The citizen scientists do not need prior knowledge of particle physics or detection techniques. Furthermore, the text has been written in simple terms and is adapted for middle schoolers. A few keywords are placed to stimulate curiosity on the different concepts that are important to the muography methodology in general. The second goal of the research section is to guide the readers seamlessly and quickly toward the workflows, by providing sufficient knowledge for them to perform data processing in a fulfilling and creative manner. This is very important for us since a well-informed individual is more likely to perform better and more data examinations.

The research section text starts with some generalities on our team's projects and is followed by an image taken from one of our team's papers on the combined results of muography and gravimetry [22]. The purpose of the image is to familiarize the reader with the density histograms that result from a set of muon flux measurements and to illustrate the interdisciplinarity of the field. Additionally, it highlights the conjunction with different measurement techniques used by other scientific fields, such as geology, for example. The rest of the text is divided into paragraphs, each with a descriptive title.

An *Introduction* makes the popular analogy of muon tomography with the common X-ray scanning used in modern healthcare facilities. Afterward, it provides some interesting facts on cosmic particles and introduces in a sentence the two kinds of tomography, absorption/transmission and scattering. The *Cosmic-ray showers* paragraph takes the reader through the processes that create a footprint of particles at the ground level as a result of a high-energy particle entering the Earth's atmosphere. Here, we also mention a few interesting facts on the resulting muon flux in order to paint a more complete picture for the reader. Moving on to *The Method* paragraph, we utilize an image to demonstrate the way a muon tomography measurement is done. The text describes the way these measurements are used to determine the inner-density distribution of a massive object. This description is given in simple terms so that the reader is not discouraged from continuing reading further. At the same time, the information is accurate and communicates the principle behind muography.

The text continues with *The Detector* description of a typical muon telescope of the Diaphane type. An image provides information on the different segments of the detector. In this part, we also describe the data acquisition chain under which the detector operates. The rest of the text describes what is considered as background events for tomography studies. The goal of the proposed workflows is to identify patterns in these events. Some details are hence given in the *Background* and *Background Types* paragraphs as well as some concrete examples [23]. The next paragraph, *background rejection*, presents common techniques adopted to facilitate the minimization of this background finding its way into our data-analysis sample. It also presents pattern recognition as an additional technique to background rejection. The research page concludes with The task at hand paragraph that is our segue to the workflows. It describes what data the citizen scientists are going to work with and how the results of their work are going to be used in the context of event identification.

2.3. Classify

The Classify section is the core segment of any Zooniverse project. The philosophy is to present data to the citizen scientists and to ask them to classify/perform various tasks on each event. The results can be used to help scientific studies by creating catalogs, by training machine-learning algorithms, or even by discovering cases in the data that have never been found before. The purpose of our classification scheme is to identify patterns in the data recorded by our detector. The typical signal events in our detector are straight lines produced by muons crossing the detector. Sometimes, these lines are "contaminated" by other particles that cross parts of our detectors. Some other times, a straight line can be approximated by several particles that propagate on the same plane as a result of a local cascade. Some studies [24] have revealed that only 30% of the appropriate signal should be considered as such in the data recorded by a muon telescope. This leads to a different perspective for the muon tomography experiments where the response of the detector can be studied against extensive air shower products rather than against just the muon component. These two perspectives are complementary and strive to increase the accuracy of the muography methodology. The two proposed workflows will help define patterns that can be cataloged. While there are several ways of using the newly found patterns, we plan to use them to train a machine-learning algorithm. This will allow us to identify similar patterns in new datasets. Other potential uses of the classifications will be investigated and might lead to the creation of new workflows within the demonstrator.

2.3.1. The Workflows

The two proposed workflows are similar to each other. The first, the *Introductory* one, treats the simpler cases of background events for which the reconstruction algorithm can find a track passing through the three detector planes. This track is depicted through two red lines that are the projections of the reconstructed track on the xz and yz planes of the detector. The presence of reconstructed tracks makes the first workflow (Figure 2) a good first step on the road to becoming familiar with the demonstrator workflows. Additionally, by construction, it treats simpler topologies and thus a smaller percentage of the total background events end up in this workflow.

The classification interface provides two event representations (3D and 1D) accompanied with a question to be answered: "Are the drawn lines the ones that you would draw?". This provides three possible answers: "yes", "no", and "I do not see anything resembling a track". The third answer is there to take care of unforeseen cases, where the reconstruction algorithm provided some kind of line, but no realistic track is present.

If the participants select "yes" or the third option, they move to the next event. When the citizen scientists select "no", they are asked to draw the proper lines to substitute those provided by the algorithm. There is also the possibility of marking "orphan" points on the detector surfaces. The "orphan" refers to the fact that they do not participate in the creation of a track.

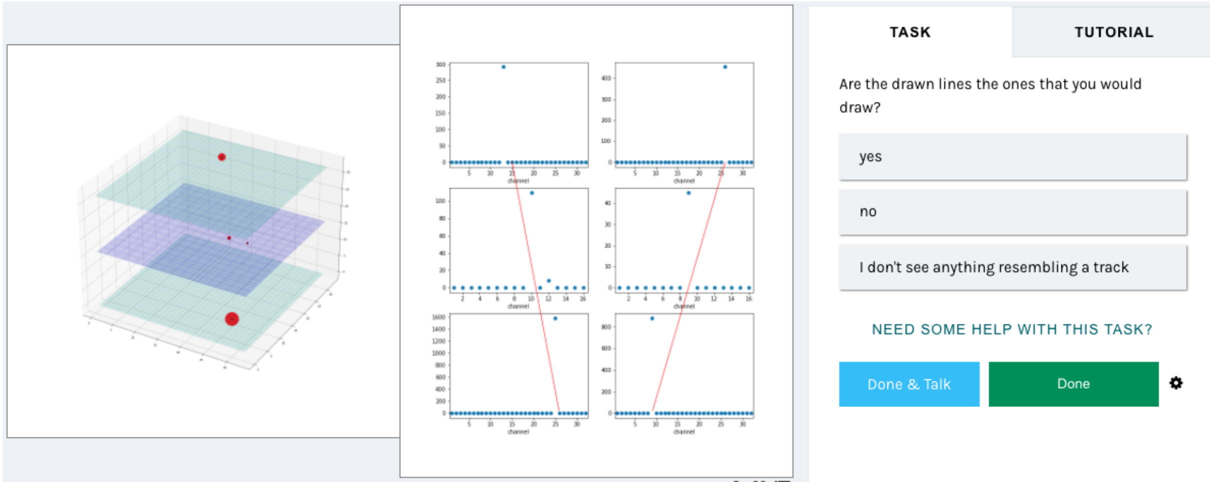

FIGURE 2: The introductory classification workflow, in which the citizen scientist is aided by the reconstruction algorithm to the identification of patterns.

FIGURE 3: The *FreeStyle* classification workflow treats the classification of a more complicated topology for which the reconstruction algorithm fails to provide a track.

The second workflow, *FreeStyle*, focuses on more complicated topologies (Figure 3). The reconstruction algorithm is not able to provide a track within acceptable confidence limits for these events. This means that there are no proposed lines, and the discovery of patterns for these events relies completely upon the citizen scientist. This also means that the second workflow gets the majority of background events, and it is expected to be the preferred workflow once the citizen scientists have acquired experience in classifying events. This workflow is built similarly to the first one but with no provided track reconstruction and in this case the question to be answered is very general: “Is there a pattern?” and gets two possible answers. If a pattern exists, the citizen scientists will identify it through a series of lines and “orphan” points, while if no pattern is found, according to their experience, the next event is provided.

Tutorials, Help, Field Guide, and Talk Boards. Both these workflows are accompanied by a *Tutorial*, which explains the required steps in each case so that the participants perform the classifications in the most fruitful way. Furthermore, a *Help* text is provided in a pop-up window with very precise instructions. The Tutorial and the Help text have been created in such a way that they can work as standalone explanations but work in a complementary manner as well. Furthermore, a *Field Guide* accompanies the user through every step of his classification experience. The field guide includes more detailed explanations regarding the detector operation, the data acquisition, and the relationship between these two and the event visualization we already presented. Finally, if these materials are not enough and the citizen scientists need more help with a specific event, they can highlight it and bring our attention to it. This discussion takes place in the *Talk Board* section where we are provided feedback and answer questions on topics related to the project and the science behind it.

3. CONCLUSIONS

The Cosmic Muon Images demonstrator consists of two main parts: one that describes the science and one that is dedicated to cataloging events. We aspire to bring together people with diverse interests, from particle physics to geology, and the content of the demonstrator was prepared accordingly. The classification of events based on their patterns is expected to be useful in different analyses. The primary analysis method is the training of a machine-learning algorithm to better identify signal events. The classified events can be compared with simulated events in order to investigate the main origin of the identified patterns. We look forward to working with the citizen scientists and to providing them with new materials in order to keep the project fresh and interesting. We really hope that Cosmic Muon Images is a fulfilling experience for all participants.

CONFLICTS OF INTEREST

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

ACKNOWLEDGMENTS

REINFORCE has received funding from the European Union's Horizon 2020 research and innovation program, under Grant Agreement no. 872859.

References

[1] https://www.reinforceeu.eu.
[2] http://sion.frm.utn.edu.ar/sonoUno/.
[3] https://www.zooniverse.org/.
[4] ECSA, Ten Principles of Citizen Science, Berlin (2015) (http://doi.org/10.17605/OSF.IO/XPR2N).
[5] H. Riesch and C. Potter. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. Public Understanding of Science. 2014; 23(1): 107–120. doi: 10.1177/0963662513497324.
[6] https://cordis.europa.eu/project/id/872859.
[7] S. B. Coughlin, S. Bahaadini, N. Rohani, M. Zevin, O. Patane, M. Harandi, C. Jackson, V. Noroozi, S. Allen, and J. Areeda et al. "Classifying the unknown: discovering novel gravitational-wave detector glitches using similarity learning," Phys. Rev. D **99**, no. 8, 082002 (2019), doi: 10.1103/PhysRevD.99.082002.
[8] Q. Feng and J. Jarvis. "A citizen-science approach to muon events in imaging atmospheric Cherenkov telescope data: the Muon Hunter," 2017, arXiv:1708.06393
[9] Alan James Barr, Charles William Kalderon, and Andrew C. Haas. "'That looks weird' - evaluating citizen scientists' ability to detect unusual features in ATLAS images of LHC collisions," 2017, arXiv:1610.02214.
[10] https://www.zooniverse.org/projects/reinforce/gwitchhunters.
[11] https://www.zooniverse.org/projects/reinforce/deep-sea-explorers.
[12] https://www.zooniverse.org/projects/reinforce/new-particle-search-at-cern/.
[13] https://www.zooniverse.org/projects/reinforce/cosmic-muon-images.
[14] T. Accadia et al. [The Virgo Collaboration]. "Virgo: a laser interferometer to detect gravitational waves," 2012, Journal of Instrumentation, doi: 10.1088/1748-0221/7/03/p03012.
[15] S. Adrian-Martinez et al. [The KM3Net Collaboration]. "Letter of intent for KM3NeT 2.0," J. Phys. G **43**, no. 8, 084001 (2016), doi: 10.1088/0954-3899/43/8/084001.
[16] G. Aad et al. [The ATLAS Collaboration]. "The ATLAS Experiment at the CERN Large Hadron Collider," Journal of Instrumentation, 2008, doi: 10.1088/1748-0221/3/08/s08003.
[17] N. Lesparre and J. Marteau et al. "Design and operation of a field telescope for cosmic ray geophysical tomography," Geoscientific Instrumentation, Methods and Data Systems, 2012, doi: 10.5194/gi-1-33-2012.
[18] S. Angelidakis on behalf of the REINFORCE consortium. The REINFORCE citizen-science project and the search for new long-lived particles at the LHC, 021 J. Phys.: Conf.Ser. 2105 012028, doi: 10.1088/1742-6596/2105/1/012028.
[19] C. Kourkoumelis and S. Vourakis. "HYPATIA - An online tool for ATLAS event visualization," Phys. Educ. 49 (2014) 21–32.
[20] Garcia Beatriz et al. "SonoUno: a User-Centred Approach to Sonification," 2021, Proceedings of the IAUS367-Cambridge University Press.
[21] J. Marteau et al. "Applied muography: from volcanology to archaelogy with a mobile muon detector (DIAPHANE/ARCHé)," 2018, adsurl=https://ui.adsabs.harvard.edu/abs/2018AGUFMNS23B0705M.
[22] M. Rosas-Carbajal et al. "Three-dimensional density structure of La Soufrière de Guadeloupe lava dome from simultaneous muon radiographies and gravity data," 2017, Geophys. Res. Lett., 44, 6743–6751, doi: 10.1002/2017GL074285.
[23] Bonechi Lorenzo et al. "Atmospheric muons as an imaging tool," 2020, Reviews in Physics, doi: 10.1016/j.revip.2020.100038.
[24] Peña Rodríguez et al. "Characterization of the muography background using the Muon Telescope (MuTe)," 2021, PoS(ICHEP2020), doi: 10.22323/1.390.0984.