**THE EUROPEAN
PHYSICAL JOURNAL C**

Special Article - Tools for Experiment and Theory

The ARTI framework: cosmic rays atmospheric background simulations

Christian Sarmiento-Cano[1,a] , **Mauricio Suárez-Durán**[2,3,b] , **Rolando Calderón-Ardila**[4,5,c] ,
Adriana Vásquez-Ramírez[1,d] , **Andrei Jaimes-Motta**[1,e] , **Luis A. Núñez**[1,6,f] , **Sergio Dasso**[7,8,g] ,
Iván Sidelnik[9,10,h] , **Hernán Asorey**[4,11,12,i] , **For the LAGO Collaboration**

[1] Escuela de Física, Universidad Industrial de Santander, Carrera 27 calle 9 Ciudad Universitaria, Bucaramanga, Colombia
[2] Université Libre de Bruxelles, Boulevard du Triomphe, ACC.2 1050 Bruxelles, Belgium
[3] Departamento de Fisica y Geología, Universidad de Pamplona, Km 1 Vía Bucaramanga Ciudad Universitaria Pamplona - Norte de Santander, Pamplona, Colombia
[4] Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), Comisión Nacional de Energía Atómica, Centro Atómico Constituyentes, Av. Gral Paz 1499, (1650) Villa Maipú, Argentina
[5] Instituto Sábato, Comisión Nacional de Energía Atómica, Centro Atómico Constituyentes, Av. Gral Paz 1499, (1650) Villa Maipú, Argentina
[6] Departamento de Física, Universidad de Los Andes, Núcleo Universitario Pedro Rincón Gutiérrez, La Hechicera, 5101 Mérida, Venezuela
[7] Departamento de Ciencias de la Atmósfera y los Océanos - Grupo LAMP, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Intendente Güiraldes 2160, Pabellón cero-infinito, Ciudad Universitaria (C1428EGA), Buenos Aires, Argentina
[8] Instituto de Astronomía y Física del Espacio - LAMP Group, Universidad de Buenos Aires Consejo Nacionall de Investigaciones Científicas y Técnicas, Intendente Güiraldes 2160, Pabellón Cero-Infinito, Ciudad Universitaria (C1428EGA), Buenos Aires, Argentina
[9] Departamento de Física de Neutrones, Comisión Nacional de Energía Atómica, Centro Atómico Bariloche, Av. E Bustillo 9500, (8400) San Carlos de Bariloche, Argentina
[10] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina
[11] Departamento de Física Médica, Comisión Nacional de Energía Atómica, Centro Atómico Bariloche, Av. E Bustillo 9500, (8400) San Carlos de Bariloche, Argentina
[12] Unidad de Informática Científica (temporal stay), CIEMAT, Av Complutense 40, (28040) Madrid, Spain

Abstract ARTI is a complete framework designed to simulate the signals produced by the secondary particles emerging from the interaction of single, multiple, and even from the complete flux of primary cosmic rays with the atmosphere. These signals are simulated for any particle detector located at any place (latitude, longitude and altitude), including the real-time atmospheric, geomagnetic and detector conditions. Formulated through a sequence of codes written in C++, Fortran, Bash and Perl, it provides an easy-to-use integration of three different simulation environments: MagnetoCosmics, CORSIKA and Geant4. These tools evaluate the geomagnetic field effects on the primary flux and simulate atmospheric showers of cosmic rays and the detectors' response to the secondary flux of particles. In this work, we exhibit the usage of the ARTI framework by calculating the total expected signal flux at eight selected sites of the Latin American Giant Observatory: a cosmic ray Observatory all over Latin America covering a wide range of altitudes, latitudes and geomagnetic rigidities. ARTI will also calculate the signal flux expected during the sudden occurrence of a gamma-ray burst or the flux of energetic photons originating from steady gamma sources. It also compares these fluxes with the expected background when they are detected in a single water Cherenkov detector deployed in a high-altitude site. Furthermore, by using ARTI, it is possible to calculate in a very precise way the expected flux of high-energetic muons and other secondaries at the ground level and to inject them through geological structures for muography applications.

The LAGO Collaboration, see the complete list of authors and institutions at http://lagoproject.net/collab.html.

[a] e-mail: christian.sarmiento@correo.uis.edu.co
[b] e-mail: mauricio.suarezduran@ulb.be
[c] e-mail: rolando.calderon@iteda.cnea.gov.ar
[d] e-mail: adriana2168921@correo.uis.edu.co
[e] e-mail: mottandrei2@gmail.com
[f] e-mail: lnunez@uis.edu.co
[g] e-mail: sdasso@at.fcen.uba.ar
[h] e-mail: sidelnik@cnea.gov.ar
[i] e-mail: hernan.asorey@iteda.cnea.gov.ar (corresponding author)

1 Introduction

Except for very particular geographical locations, such as, for example, those active volcanic regions where degassing of ^{222}Rn into the atmosphere could increase natural radioactivity levels [1], the main contribution of the background radiation is due to the continuous arrival of a shower of particles produced during the interaction of primary cosmic rays (CR) with the atmosphere [2]. When an energetic cosmic ray impinges on the Earth's atmosphere, many secondary particles are produced via radiative and decay processes and known as Extensive Air Showers (EAS). At its maximum development – and depending on the energy E_p of the primary cosmic ray – the EAS could produce up to 10^{10} particles. The most common technique used in astroparticle physics to study these primary CRs, is the detection of secondary particles at ground level. We can distinguish two important cases: those belonging to a single EAS at the highest primary energy [3] and the ones corresponding to the integrated secondaries due to lower energies [4]. In this latter case, the precise measurement in the temporal variations of the flux is an important tool in the study of transient astrophysical events, such as the sudden occurrence of gamma-ray bursts [5,6], space weather phenomena [7,8] and even geophysical applications like muography [9–11]. Thus, knowing how the primary CRs are modulated by the Earth's magnetic field (EMF) and then reach a given geographical position is critical. The detailed response of the detector to the secondary particle flux is also essential for a better understanding of the different astroparticle scenarios.

The ARTI toolkit, a set of C++ and Fortran codes, and Bash, Perl and Python scripts, is the result of the efforts of the Latin American Giant Observatory (LAGO) to quantify the effect of the integrated galactic cosmic rays (GCR) flux and many transient astrophysics events at ground level [8,18]. This toolkit is publicly available at the LAGO GitHub repository [16]. Currently, several computational tools, such as MagnetoCosmics (MAGCOS) [12] enable the computation of charged particles' trajectories by using advanced and detailed EMF models. CORSIKA [13] or FLUKA [14] can be used to obtain detailed simulations of extensive air showers in the atmosphere. Finally, the detector response is simulated by Geant4 [15], a widespread tool used to simulate the interaction of radiation with matter developed by the high energy physics researchers. However, while these Monte Carlo tools are being extensively developed and tested within the Astroparticle Physics community, they cannot integrate secondary background flux produced by the total flux of primary cosmic rays, or to concatenate all the intermediate results. All these processes have to be done "by hand" every time a detection location needs to be characterized or evaluate the impact of the changing atmospheric or geomagnetic conditions. In the end, we are performing all these operations

manually, which is a hard and tedious jobs. Thus, we developed ARTI, to carry out these tasks in a semi-autonomous way [16]. The framework calculates and analyses effortlessly the total background flux of secondaries and the corresponding detector signals produced by the atmospheric response to the primary flux of GCR. This analysis can be implemented for single or small arrays of Water Cherenkov Detector (WCD) located at any geographic site, and under realistic and time-evolving atmospheric, geomagnetic and detector conditions.

As an example of the ARTI's functionality, we conduct a comparative study of the response of a LAGO WCD to the total flux of secondary particles at eight LAGO sites in Latin America. We select these sites – shown in Fig. 1 – to emphasize the differences in their altitudes, local atmospheric conditions and geographic and geomagnetic coordinates.

The outline of this document is as follows: in Sect. 2 we describe the computational architecture of ARTI. Next, Sect. 3 introduces the method implemented for estimating the nominal background flux of secondaries at the detector
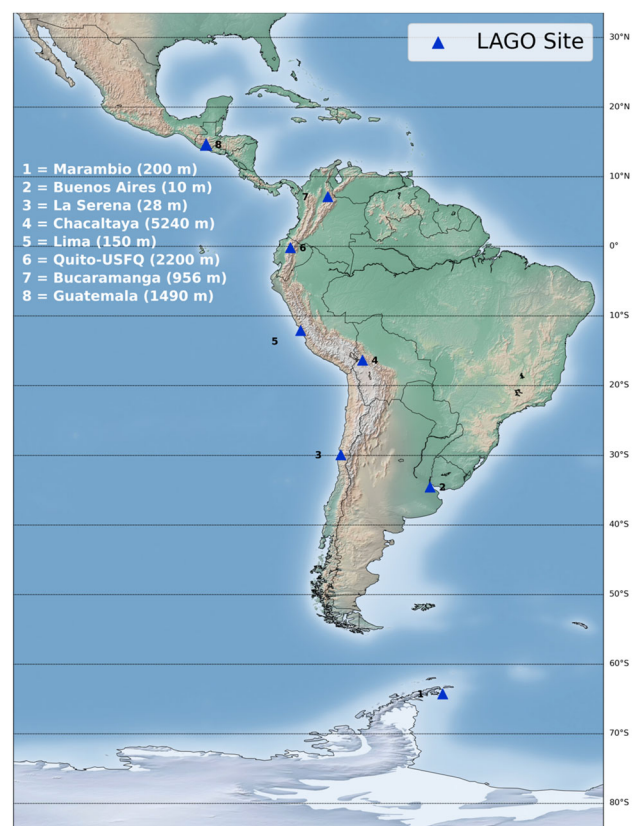

Fig. 1 In this work we studied these eight representative sites of the Latin American Giant Observatory, covering a wide range of altitudes and local geomagnetic rigidity cut-offs, from south to north: Marambio, Antarctica; Buenos Aires, Argentina; La Serena, Chile; Chacaltaya, Bolivia; Lima, Perú; Quito, Ecuador; Bucaramanga, Colombia; and Guatemala, Guatemala. Detailed information about these sites can be seen in Table 1

level. Then, in Sect. 4, we explain the methodology used to implement the correction due to the EMF condition in the selected LAGO sites [18]. The Geant4 models and the signals expected in a standard LAGO WCD are shown in Sect. 5, where we use the same type of detector to simplify the direct comparison of the results in different sites. We discuss the final remarks and future perspectives of this work in Sect. 6. Finally, in Appendices A and B we sketch the main parts of the framework and include a brief description of some main options to be used in different types of cosmic ray simulations.

2 ARTI framework

With ARTI it is possible to calculate the expected signal flux at any WCD and, depending on the user's initial configuration, ARTI calls the corresponding modules to interact with CORSIKA or Geant4.

The ARTI approach is in three stages:

1. site characteristics, primary spectrum calculations and EAS developments;
2. analysis of the secondary particles at the ground and for geomagnetic corrections;
3. and for detector simulation.

These three stages are sequentially accessed: the physical results from one stage are used as the input for the next stage. The files resulting from each stage are preserved and curated as LAGO datasets [21,22]. In this section we present the overall simulation sequence, while some main steps and the intermediate physical results are detailed in the following sections.

The ARTI framework consists of C++ and Fortran codes; Perl and various bash scripts which provides the interaction between the user and the different software (for details, please see Appendix A). This allows a straightforward approach for the user and a smooth interaction between ARTI and CORSIKA, MAGCOS and Geant4, as shown in Fig. 2.

Different tracks depend on the type of calculations the user would need. We will detail the necessary steps in calculating the expected signal flux at a particular WCD. We refer the reader to the ARTI website for further explanations and examples [16]. The main script for this calculation is do_sims.sh. All ARTI scripts can be configured using options and modifiers from the command line (cli). ARTI versions could change, but all the ARTI scripts and codes have an internal help modifier, (-?). The currently available set of options (do_sims.sh) is available in Appendix B. Here we describe the main simulation sequence with their corresponding options and modifiers for the initial configuration, i.e.:

-t , defines the time that will be used for the integration of the primary flux. Typical values range from 3600 to 86,400 s per run.

-s , specifies the geographic from a list of predefined sites around the World. Each site has an altitude above sea level, atmospheric parameters and local values of the EMF, as described in the LAGO Data Management Plan (DMP) [21,22]. Otherwise, the corresponding atmosphere model (-c), altitude (-k) or the B_x (north) and B_z (vertical) components of the magnetic field could be defined (-o and -q respectively). ARTI requires them to calculate the local EMF corrections at a later stage (see Sect. 4) and by CORSIKA to take into account the geomagnetic effects on the primary and secondary charged particles propagation in the atmosphere.

-c defines which type of atmospheric model will be used from these possibilities: (a) atmospheres from the MODTRAN atmospheric model [27]; (b) atmospheres characterised by the 20 parameters following the well known Linsley's atmospheric model; or (c) a realtime extraction of the instantaneous atmospheric profiles as compiled from the Global Data Assimilation System (GDAS) [19].

-u is the ORCID identifier of the user, for further reference of the results dataset [22].

-v represent the CORSIKA version to be used by ARTI. The Current default is version CORSIKA 7.7402.

-y is required to calculate the flux by considering a volumetric detector (such as a WCD), instead of a flat detector (such as a scintillator panel or a resistive plate chamber).

-x is the default for the non-selected options, just use the internal ARTI predefined values (energy range, arrival directions, etc.)

A typical call for the initial script could be:

```
$ do_sims.sh -w path_to/corsika/run/
   -p fluxBGA
   -s bga -v 76500 -t 3600 -u user -y -x
```

ARTI will calculate the one-hour (3600 s) integrated flux of secondaries for a volumetric detector (-y) located at ground level in the LAGO site of bga (Bucaramanga, Colombia, 965 m a.s.l.),[1] following the calculations described in Sect. 3, selecting the v7.6500 version of CORSIKA and the ARTI defaults (-x) for the non-selected options (see the complete list in Appendix B). All the files will be located in the subdirectory fluxBGA created in the CORSIKA executable's main path.

By default, ARTI divides the total flux of particles into 60 processes and produce 12 bash scripts to distribute the simulation load in a computer cluster. As described in Sect. 3,

---

[1] When possible, LAGO sites are internally known by the corresponding three or four letters IATA code of the local airport.
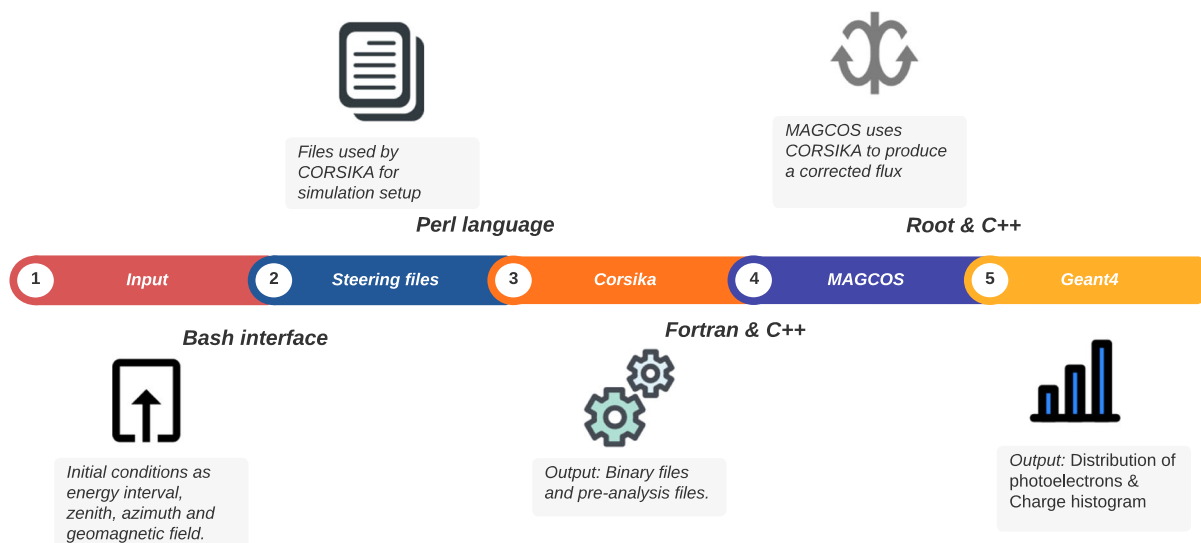
Fig. 2 As the simulation evolves, the different ARTI modules call CORSIKA, MAGCOS, or Geant4 to accurately calculate the signal flux at the detector expected by the integrated cosmic ray flux. The simulation moves forward as a sequence: the results from one steps are used as inputs for the next one

each process is determined by considering the flux of an individual particle that constitutes the primary background, by injected primaries from proton to iron. The default number of processes can also be modified by the user (-j option). Depending on the computing environment, it will produce an additional bash script to perform the automatic analysis of the results. It will automatically launch the CORSIKA stage simulations in, e.g., the corresponding user-accessible queues in high performance computing (HPC) clusters operating with the SLURM workload manager, or in cloud-based environments with docker containers [21,30]. At the end of this first stage, if the default number of processes was used, ARTI will produce 60 CORSIKA steering files (*.input), 60 bzip2-compressed unformatted binary fortran files containing all the simulated EAS information (DAT*.bz2), and 60 bzip2-compressed files containing the output produced by CORSIKA during the simulation run (*.lst.bz2). According to the LAGO DMP, all the files resulting from this first simulation stage are considered a data catalogue and are collectively known as the S0 datasets.

The S0 produces catalogues that can be read and analyzed with the ARTI set of tools which extracts the secondary particle flux information, i.e. the secondary particle ID, the particle momentum, the particle position relative to the primary core, and additional information related with its primary progenitor. The analysis stage can be performed automatically during the runtime or manually by calling the do_showers.sh script. The typical result of this script is several bzip2-compressed ASCII files containing the information of the injected primaries (*.pri.bz2), the resulting secondaries (*.sec.bz2), and four additional files:

- .shw.bz2 file containing the complete list of secondary particles at ground level and information about their corresponding progenitors;
- .hst file containing the energy distribution of the secondary particles by particle type;
- .dse and .dst files containing, respectively, the lateral distribution of the deposited energy and lateral distribution of the number of particles at the ground, respectively.

Depending on the computing environment, there are different types of parallelized analyses available that can be selected from cli. The recommended number of parallel processes are also automatically selected from the local node's total number of available processors. If the files to be analyzed are stored on cloud servers, they are first transferred to the local cluster before processing. For example, a standard manual call for running a parallelized analysis of a S0 catalogue stored in a cloud storage server, corresponding to the flux during one day at a site with altitude of 2400 m a.s.l., could be:

```
$ do_showers.sh -o path_to/cloud_server
/fluxBGA
-k 2400 -t 86400 -g site.geo 3 -l
```

In this case, the default ARTI option implements a parallel analysis using half of the available processors at each node. During the analysis the EMF is taken into account, and the local directional rigidities (see below) are read from the third column of the site.geo file. While this can be called by the user at any moment for, e.g., to reprocess the same S0 catalogue for different EMF conditions, this stage

is usually performed automatically by ARTI using the pre-defined default options without need of the user interaction. Please refer to the ARTI web site [16] and Appendix B for the detailed information about these scripts.

The predefined values for the local EMF are calculated and regularly updated[2] following the latest available version of the International Geomagnetic Field Reference (IGRF).[3] Notice that, at this stage, only the secular – very long term variations – of the EMF were taken into account. However, at this phase, we have not considered the effects due to the penumbra of the magnetic field. They are usually described as a system of forbidden and allowed trajectories within an intermediate range of local geomagnetic cut-off rigidities [24]. The Earth's magnetic field is continuously modulated and disturbed by the Solar activity. Therefore, as explained in Sect. 4, we provide an additional script to include both effects automatically. To perform this analysis in ARTI one must type:

```
$ STATIC_MAGNETOCOSMICS base.g4mac
```

creating a file with the corresponding values for the directional upper and lower limits of the local rigidities, at the site on a given date. This also is helpful for filtering and eliminating those secondaries in a forbidden primary within a `.shw` file. This method allows the re-use of the same `S0` catalogue for different instantaneous disturbances of the EMF. All the resulting files from the analysis stage are also considered a new LAGO data catalogue and are collectively identified as a `S1` dataset.

Finally, a Geant4-based code in ARTI calculates the WCD response to the secondary particles arriving at the detector. As explained in Sect. 5, ARTI injects every secondary particle within the detector volume with its initial momentum but with a random position above the WCD. The Geant4 simulation includes different parts that make up the detector, such as:

- the water container material, thickness and geometry;
- the reflectivity and diffusivity of the container's inner coating made of Tyvek®;
- the water absorption curves in the optical and UV wavelength range; the geometry and wavelength sensitivity of the photomultiplier tube (PMT);
- and the emulated response of the onboard LAGO electronic system.

The user can vary all these characteristics so as to incorporate the ones corresponding to the WCD to be simulated.

---

[2] The values for the magnetic field components can be automatically updated during runtime by the using the option -g, or forcing the values of B_x and B_z at launch time by using the corresponding options.

[3] Currently IGRF13-2019 [23].

Depending on the environment, it could run automatically during execution or will be called by the user once the previous stages are completed, by typing:

```
$ wcd -m input.in
```

Once this stage ends, we have a `.root` file containing the expected charge histograms at the detector, i.e., the histograms of the time integral of the individual pulses produced by each secondary particle impinging in the detector. As in the first two stages, ARTI identifies the resulting files from this last stage as the `S2` LAGO data catalogue.

The ARTI framework relies on the previous installation and compilation of ROOT [25], CORSIKA and Geant4. The current version of ARTI (v1r9) works with CORSIKA v7.7402, ROOT v6.20.08, Geant4.10.03 and Magnetocosmics v2.0 versions. Anyway, backward compatibility within previous minor versions and patches of the external dependencies is assured.

Once these tools are correctly installed, ARTI can be easily cloned from the LAGO GitHub repository [16] and automatically installed by running the provided installation script. As usual, ARTI performance will depend on the system and the installation options. For a typical installation in a modern HPC cluster, i.e. a cloud virtual cluster based on a Slurm Workload Manager job scheduler with one virtual master (v-master) and 10 v-nodes with 8 Intel Xeon Core E7 processors and 250 GB of shared memory [30]. All the computations are completed almost in real time. To compute the signals expected in a WCD deployed in a low-latitude site during 1 day of total primary flux ($\lesssim 10^9$ primaries) requires $\simeq 28$ h ($\simeq 2.2$ kCPU·h) and produce ~ 30 GB of compressed (`S0+S1+S2`) datasets.

The following sections describe the physical process and the results obtained for each of the three stages.

3 Cosmic background radiation at ground level

We model the GCR primary flux (Φ) as an isotropic flux impinging the Earth's atmosphere at an altitude of 112 km a.s.l. [33], where Φ is defined as

$$\Phi(E_p, Z, A, \Omega) \simeq j_0(Z, A) \left(\frac{E_p}{E_0}\right)^{\alpha(E_p, Z, A)}, \qquad (1)$$

with E_p the energy of the primary particle, $\alpha(E_p, Z, A)$ the spectral index for each type of injected particle and which can be considered constant, i.e. $\alpha \equiv \alpha(Z, A)$, in the energy range of interest in this work, from a few GeV to 10^6 GeV (1 PeV). Each kind of GCR considered is individualized by its mass number (A) and atomic number (Z) and $j_0(Z, A)$ is the measured flux in the top of the atmosphere at the reference energy $E_0 = 10^3$ GeV. See [18] and references therein for a detailed description of this method.

To calculate the expected number of primaries, we integrate Φ for each primary type from proton to iron (i.e., $1 \leq Z \leq 26$). Please notice that the integrated flux is simulated with the corresponding distribution for each type of particle [8]. We use time integration values typically ranging from 1 h to 1 day of flux impinging over an area of 1 m^2 and isotropically distributed in $0 \leq \theta \leq \pi/2$ and $-\pi \leq \phi \leq \pi$ where θ and ϕ are the zenith and azimuth angles respectively.

If the rigidity cutoff option is considered (option -b), then the minimum energy $E_{min}(Z)$ is selected for each type of primary as $E_{min}(Z) = R_0 \times Z$, where R_0 is lower than the minimum value of the local directional rigidity cut-off R_{min}, i.e., $R_0 < R_{min}$. Otherwise, a minimum value of 0.1 GeV for the kinetic energy is used for all the injected primaries, and so, in natural units, $E_{min} = m(Z, A) + 0.1$ GeV, where $m(A, Z)$ is the mass of the injected primary. In all the cases, the upper limit of the energy range is fixed to $E_{max} = 1$ PeV, as for $E_p > E_{max}$ a decrease in the GCR flux is observed for all species (the so-called knee of the cosmic ray spectrum, see e.g. [36]), and at this energy the flux can be considered so low that it does not affect the background calculations.[4]

Given the Poissonian behavior of the primary flux, all the computations can be easily done in parallel as EAS do not interact with each other. As for energies below the knee, protons then helium dominate the GCR flux. To rest of the nuclei corresponds to a small fraction of the total flux. Once we calculate the total number of particles to be injected for each primary (A, Z), they are grouped looking forward uniforming all computing aspects. For the physical aspects to be considered, the computing time not only depends on the total number of primaries but also on the energy ranges and the spectral index $\alpha(A, Z)$: low energy particles are much more abundant, but take much less time to compute. As described below, the site altitude effect is not trivial: at high altitudes sites, the cascades reach the ground during early stages of their development, but at the same time, the total number of secondaries that need to be tracked is much larger than at later stages of the EAS development. When the site altitude is close to the EAS X_{max} (the atmospheric depth where its reaches its maximum development), the total number of air shower increases.

As mentioned in Sect. 2, ARTI currently supports different types of cluster architectures and distributed computing solutions, such as those based on grid and federated or public clouds implementations [21]. So, from the computing point of view, it is relatively common than the number of available CPUs per nodes at HPC infrastructures was a power of 2. Additionally, if the binary files result too large, they can quickly fill the available local storage before process-

ing, with the corresponding loss of all the calculations on this node. Thus, after performing careful studies, the grouping or splitting of primary species results as follow: (a) the total flux of protons is splitted into 32 separated processes; (b) the total flux of Helium is splitted in 8 separated processes; and (c) the non-dominant components are grouped following this schema: (^{12}C-^{16}O-^{7}Li-^{24}Mg); (^{11}B-^{28}Si-^{14}N-^{20}Ne); (^{56}Fe-^{9}Be-^{32}S-^{27}Al); (^{23}Na-^{40}Ca-^{19}F-^{52}Cr); (^{40}Ar-^{48}Ti-^{55}Mn-^{39}K); and (^{51}V-^{31}P-^{35}Cl-^{45}Sc). Future releases of ARTI will be optimized by mixing protons and helium nuclei with the non-dominant species.

The atmospheric profile is another critical factor in producing secondary particles at ground level. Thus, we need to set up the model for each site's geographical location with at least the season profiles of MODTRAN [27], or the exact time instead of the atmosphere profile extracted from GDAS [19]. Depending on the atmospheric model selected by the user at launch time, ARTI will automatically handle all the internal CORSIKA steering keywords needed to run the EAS simulation.

Once the user sets all the primary simulation parameters (say, integration time and area; energy and solid angle ranges; site altitude, atmospheric model and EMF coordinates; atmospheric model, etc), ARTI calculates the number of particles of each kind that will be injected by integrating equation (1), split or group at the primaries following the above described schema; creating the corresponding CORSIKA .input data files,[5] and the scripts that will run the EAS simulation processes. These scripts are automatically adjusted depending on the computing environment where ARTI is running, and they are eventually launched or they are queued in the cluster's partition. After starting the first stage of ARTI calls the CORSIKA executable to simulate the cascades produced during the interaction of the primaries with the atmosphere.

For this example we used CORSIKA version 7.6500 compiled with the following options and modules: QGSJET-II-04 [34]; GHEISHA-2002; EGS4; curved and external atmospheric modules and flux calculations for a volumetric detector. For the selected eight sites of this example, we used the standard MODTRAN profiles: a tropical profile for Bucaramanga (BGA, Colombia), Ciudad de Guatemala (GUA, Guatemala) and Quito (UIO, Ecuador); a subtropical summer profile for Buenos Aires (EZE, Argentina), Lima (LIM, Perú), La Serena (LSC, Chile) and Chacaltaya (CHA, Bolivia), and the summer Antarctic profile for the Marambio Base (SAWB, Antarctica) [26,37]. The local EMF components B_x and B_z at the site of this example were obtained from the IGRF-12 model [35].

---

[4] This is ARTI's default behavior, but it can be modified at launch time by using the -r, -l, and -b options. See Appendix B and [16] for further information.

[5] A specially formatted ASCII file containing all the parameters needed to run the corresponding CORSIKA simulation.

As the EAS simulation evolves, each secondary particle is tracked down to ground level (the observation level), unless the particle decays or its energy E_s goes below the lowest secondary energy threshold E_{cut} allowed by CORSIKA.[6]

Once the EAS simulations end, the second stage of ARTI is called to obtain the total secondary flux of particles at the ground, Ξ. Again, after the trivial parallelization method we performed in the first stage, we take advantage of the Poissonian nature of the cascades and the uniformity of the primary flux to superpose all the resulting secondaries (contained in the .sec files) to generate a single file (the so-called "shower" .shw file) contained all the resulting particles at ground level. The stochastic nature of the primary flux and the EAS development, allows the possibility to normalize the flux by considering that all the secondaries produced by the primaries impinging in an area ΔA (typically, $\Delta A = 1\,\text{m}^2$) at the upper atmosphere during the time t, will reach the ground level in the same surface ΔA and in the same time period t. While the assumption on the later is easy to understand, as the time evolution of the cascade is very short (μs) when compared with the typical of values considered for t, the former is not obvious at all, as it is in the basis of the definition of an EAS: the almost simultaneous arrival of secondary particles distributed over a large area at ground. This distribution is originated by the transverse momentum $p_\perp$ acquired by the air shower particles at production and the subsequent scattering in the atmosphere. As an order of magnitude, the electromagnetic Moliere radius is of $X_m = 9.3$ g cm^{-2}, which in air and at sea level corresponds to $R_m = 80$ m [38]. However, it is important to recall on the uniformity and isotropy of the flux at this energy ranges. In average, a secondary particle that miss the target area at ground by, say, 50 m to the north, will be compensated by, a quite similar secondary particle produced during the cascade development of an equivalent primary impinging the upper atmosphere at 50 m to the south. This kind of sib-similarity properties of the flux of secondaries at ground is the basis of the ARTI spatial normalization of the flux Ξ of secondaries at ground, and of the flux Ξ^{D} measured at the detector (see Sect. 5).

As an example, in Fig. 3 the results obtained for the expected secondary spectra at four of the eight representative sites are shown: Chacaltaya (Bolivia) at 5240 m a.s.l., La Serena (Chile) at sea level (28 m a.s.l.), Lima (Perú) at 150 m a.s.l. and Marambio (Antarctica) at 200 m a.s.l. There are several features that can be seen in the secondary spectra. Typically, secondary particles are grouped in three main groups: the electromagnetic component, composed of $e^\pm$ and γs; the muonic $\mu^\pm$ component; and the hadronic component composed by all the barions and mesons that are present in the cascade. However, in this work we separate neutrons and protons from other hadrons as they can be used as tracers for some cascade mechanisms (protons) or space weather activity (neutrons) [39]. As can be seen, while at low secondary momentum p_s the flux is dominated by the electromagnetic component, the high energy flux is totally dominated by muons, and even for this short integration time it is possible to observe some muons that could reach up to several tens of TeV/c. These muons posses enough energy to traverse hundreds and up to thousands of meters of rock and could be the main source for signals in muography studies [40] or background noise at underground laboratories [41]. In general, atmospheric absorption produces the well known decrease in the total flux of secondaries at low altitudes in all components, as can be seen comparing, the two left side panels of Fig. 3, where a difference of up to an order of magnitude exists between the integrated flux at Chacaltaya (CHA) that is at more than 5000 m a.s.l., and Lima at sea level. Even more, at this stage, the simulation results are so detailed that it is even possible to see a significant increase in the photon flux at the 510–520 keV energy bin, corresponding to the 511 keV photons produced during pair annihilation in the atmosphere, as it is also reflected in the differences in the flux of electrons and positrons at low p_s.

Additionally, the cascade evolution through the atmosphere can be inferred from the relative fraction of components. For example, at Chacaltaya altitudes the charged-pions-to-muons fraction is larger than at sea level, as most of the muon component is originated from the charged pions decay, a process that typically occurs below 5000 m a.s.l. At $\sim 3.5 \times 10^{-1}$ GeV/c a comparison between the flux of neutrons at Chacaltaya and La Serena shows that at the highest altitudes the flux in this p_s range is dominated by neutrons instead of muons, while at lower altitudes the effect is opposite. Since the LAGO detector calibrations are based on muons, it is important to notice also that the prediction for the muon component is larger than the charged electromagnetic component at La Serena, at sea level, while, at the altitude of Chacaltaya, the $e^\pm$ dominates with respect to $\mu^\pm$ due to atmospheric development of the cascades. This kind of studies are frequently used in LAGO and in other astroparticle observatories to characterise the planned sites for different types of astrophysical studies.

4 Geomagnetic field corrections

The Earth's magnetic field deflects low energy GCRs ($E_p < 100$ GeV) trajectories which are usually parametrized by the magnetic rigidity term (R_m) [42–44]. During these events,

---

[6] These thresholds depends on the CORSIKA version and the type of secondary. For the current versions of CORSIKA, these thresholds are $E_s \geq 5$ MeV for $\mu^\pm$s and hadrons (excluding π^0s); and $E_s \geq 5$ keV for $e^\pm$s, π^0s and γs. While these limits can't be decreased, they can be increased in ARTI by using the option -a, as this can be very useful for high energy applications such as muography, where the secondaries needs then to go through, e.g., 500 m of rock [31].
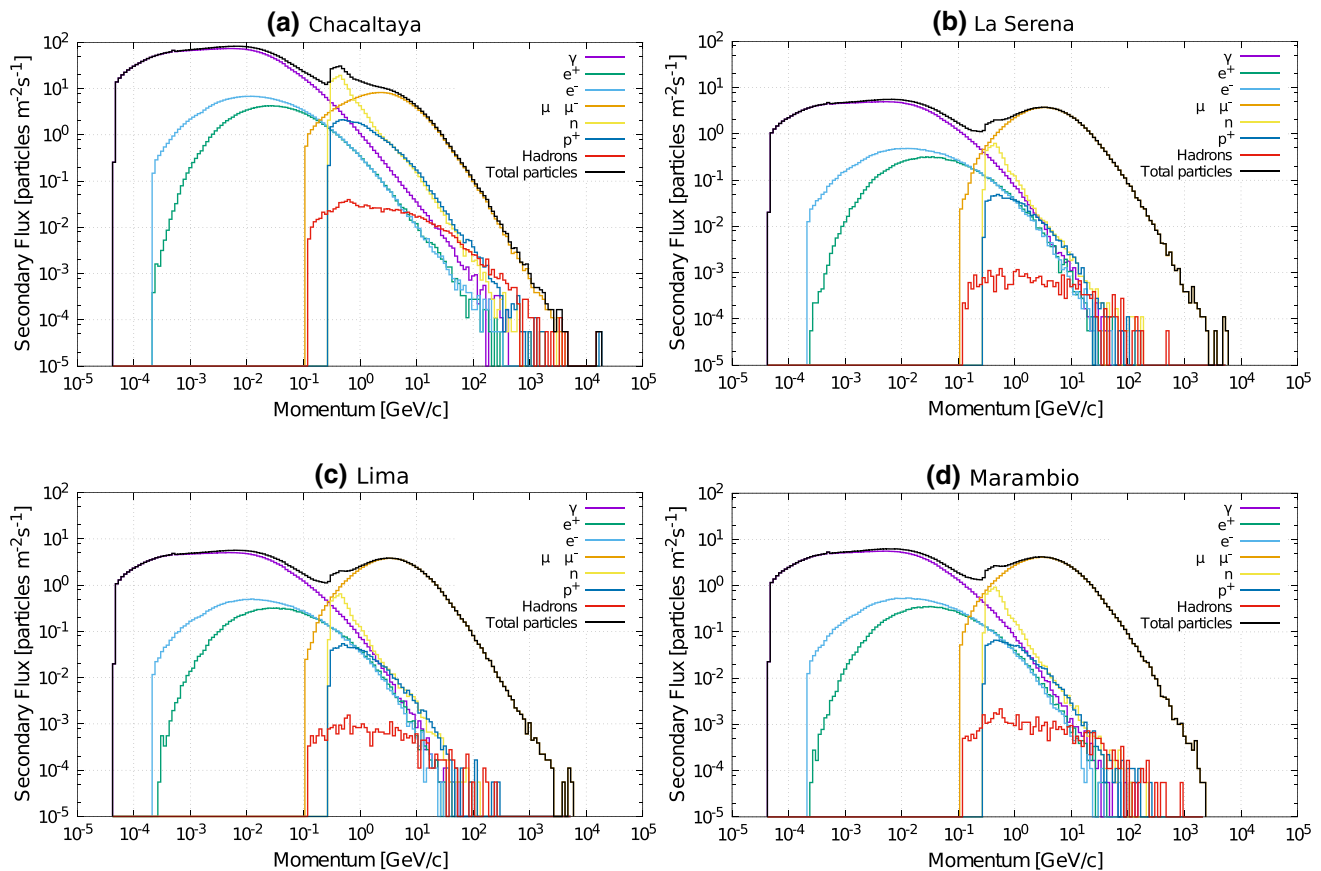
Fig. 3 Spectra of the secondary particles discriminated by particle type at four LAGO sites: **a** Chacaltaya (CHA), Bolivia (5240 m a.s.l.); **b** La Serena (LSC), Chile (28 m a.s.l.); **c** Lima (LIM), Perú (150 m a.s.l.) and **d** Marambio Base (SAWB), Antarctica (200 m a.s.l.). The comparison of these four plots evidences the atmospheric absorption, reaching up to one order of magnitude (at $\sim 10^{-2}$ GeV/c) in the total flux of secondaries (black line). See a detailed description of this figure in the text

the geomagnetic field can also be disturbed due to its interaction with the magnetized solar wind plasma, consequently changing the geomagnetic shielding on energetic particles and modifying their trajectories. Thus, Solar activity could change the primary flux and so, also change the observed flux of secondaries at ground level [7,45,46], as e.g., the well known Forbush decreases (FD) are manifestation of this phenomena.

Several FDs have been registered by different cosmic ray observatories using WCDs [8,46–49]. In 2013 the LAGO Collaboration developed its LAGO Space Weather (LAGO-SW) program [8,17], to study the variations in the flux of secondary particles at ground level and their relation to the heliospheric modulation of GCRs.

The EMF effect on the flux Ξ has been included in this work for each of the eight selected LAGO sites, following our filtering method during the second stage of the ARTI framework. In this method, the local magnetic rigidity cut-off (R_c) defined as,

$$R_c = R_c \left(\text{Lat, Lon, Alt, T}, \theta, \phi, P \left(R_m \left(\theta, T \right) \right) \right), \quad (2)$$

is a function of the geographical latitude (Lat), longitude (Lon), altitude above sea level (Alt), primary arrival direction (θ, ϕ) and epoch time T, is used to determine whether a secondary particle should reach the ground or not depending on the progenitor primary rigidity, and including a cumulative probability distribution function for the penumbra region, $P \left(R_m \left(\theta, T \right) \right)$, as explained in [18]. These methods allow us to determine the expected background flux baseline and to evaluate the impact of the changing geomagnetic conditions during, for example, a geomagnetic storm.

Table 1 displays some of the main characteristics of the selected LAGO sties and the results for the estimated flux of particles at ground level, including the EMF correction. It is clear that, as mentioned in Sect. 3, for sites with similar secular values of R_c the site's altitude is the predominant variable.

This effect is also noticeable in Fig. 4, which shows both effects for different EAS components: the atmospheric absorption changing with the altitude, and the EMF shielding depending on the latitude. For example, the geomag-

Table 1 The estimated flux of cosmic background radiation at ground level for eight LAGO sites: Chacaltaya, Quito, Guatemala, Bucaramanga, Marambio Base, Lima, La Serena, and Buenos Aires, characterized by their geographical coordinates, atmospheric profiles and altitude above sea level. The total flux for all secondaries ($\varXi^{\mathrm{All}}$) is presented along with the relative geomagnetic effect over the total flux (GE$^{\mathrm{All}}$[%]), estimated as the percent difference with respect to the flux $\varXi$ without the EMF corrections

LAGO site	Site name	Country or region	Coordinates (Lat, Lon)	MODTRAN profile	Alt m a.s.l.	R_c GV/c	$\varXi^{\mathrm{All}}$ m^{-2} s^{-1}	GE$^{\mathrm{All}}$ %
CHA	Chacaltaya	Bolivia	(16.35 S, 68.13 W)	Subtropical summer	5240	11.6	4450	−18
UIO	Quito	Ecuador	(0.20 S, 78.5 W)	Tropical	2800	12.2	1260	−12
GUA	Guatemala	Guatemala	(14.63 N, 90.59 W)	Tropical	1490	9.1	730	−4
BGA	Colombia	Colombia	(7.14 N, 73.12 W)	Tropical	956	11.6	540	−6
SAWB	Marambio	Antarctica	(64.24 S, 56.62 W)	Antarctic summer	200	2.2	430	−1
LIM	Lima	Perú	(12.1 S, 77.02 W)	Subtropical summer	150	12.0	390	−5
LSC	La Serena	Chile	(29.90 S, 71.25 W)	Subtropical summer	28	9.3	380	−3
EZE	Buenos Aires	Argentina	(34.54 S, 58.44 W)	Subtropical summer	10	8.2	390	−2

netic effect influences only the low energy primary particles, and so, the impact on the total secondary particle flux at ground level is more important on high altitude sites, where the impact of very low energy primaries is more significant. However, the geomagnetic effect on the flux of secondary neutrons is dramatic, as can be seen in the lower panel of Fig. 4; and this is one of the main reasons for using the neutron flux as an indicator to study the Solar activity from ground level observatories. In this Figure we show a comparison in the flux of electromagnetic, muon and neutron components, at four sites with similar altitude but very different geomagnetic rigidities: Marambio (SAWB), Buenos Aires (EZE), La Serena (LSC), and Lima (LIM). This considerable latitudinal dependence of the flux of secondary neutrons, combined with the capacity of WCDs to indirectly observe neutrons [50], has a major impact on space weather studies, and strongly supports the installation of WCDs at sites with very low rigidity cut-off, such as the ones deployed by LAGO in the Antarctic continent.

5 Water Cherenkov detector simulation

The third main stage of ARTI corresponds to a complete and detailed Geant4-based-simulation on a water Cherenkov detector [15]. As previously, the main result of the previous stage – the resulting shower .shw file after considering the local geomagnetic field effects – is used as the main input for this last stage.

A WCD consists of a closed, light-tight water container with optical detectors, typically PMTs or solid state photon counters, which are in close contact with the water volume so as to register the Cherenkov radiation. The corresponding Cherenkov photons propagate inside the water volume and may be absorbed in the water or in the inner coating of the container or, instead, reach the optical detector producing a
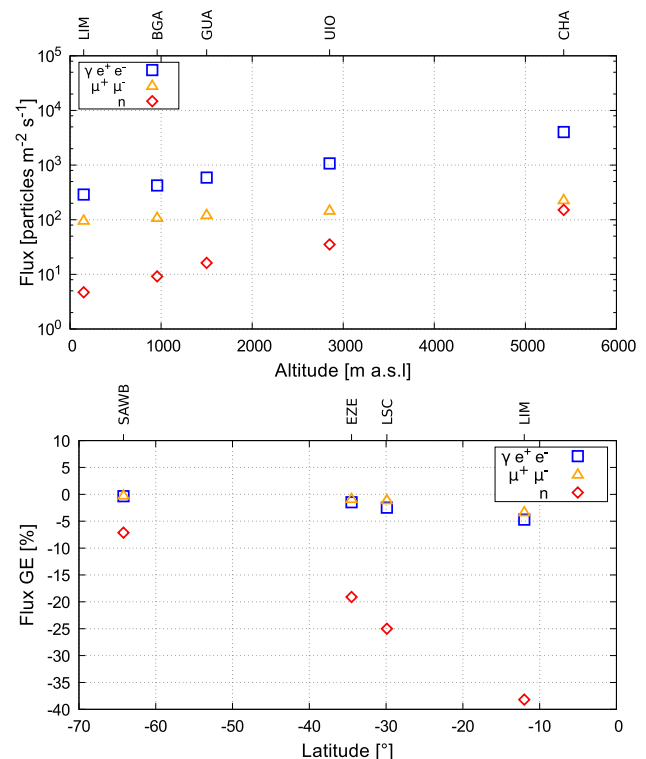

Fig. 4 The flux ($\varXi$) of cosmic background radiation at the ground as a function of the altitude at five LAGO sites with very different altitudes but similar geomagnetic rigidities are shown in the upper panel for different components of the cascades: the electromagnetic component in blue squares (γ, e^-, e^+), the muonic component in yellow triangles (μ^-, μ^+) and neutrons as part of the hadronic component are shown in red diamonds (n). The altitude effect on the flux is visible. Instead, in the lower panel, the same flux is shown as a function of the geomagnetic rigidity for four LAGO sites with similar altitudes but very different values of the local rigidity cut-offs R_c. While the effect on the total flux is not as significant as the altitude effect, the impact of the latitude on the neutron flux is quite significant

variable number of photo-electrons (pe). These pe generate a current that is multiplied in the corresponding stages of the optical detector producing a pulse. This small signal will be amplified and shaped by the detector electronics, to produce the so called pulse, i.e., the time evolution of the total charge deposited in the optical detector that it is measured at specific sampling rates. Finally, if this pulse complies with a pre-established trigger conditions, it is registered, transmitted and stored for further processing, analysis and curation.

The duration and shape of the pulse registered will depend on several factors:

- the type and energy of the particle impinging the detector and its corresponding response;
- the water quality,
- the detector geometry and age;
- the characteristics of the optical device;
- the sampling rate and even the components of the detector electronics

are some of the main factors needed to be taken into account by the simulation to represent the detector behaviour in the most possible accurate way.

The ARTI third stage incorporates different Geant4 macros and C++ codes so as to configure the detector characteristics and the type of processing that should be done with the corresponding simulated signals. It also includes an injector macro that reads the secondaries .shw file and injects the particles in the water volume, and different Geant4 physics lists depending on the type and energy of the injected particle. All the physical processes that need to be taken into account are considered: Cherenkov production in water, neutron scattering and neutron–proton reaction, muon and electronic captures, pair production, Compton scattering, photo-electric effect, Cherenkov propagation and absorption in the water, Cherenkov absorption and diffusive reflection in the inner coating, are some of the main processes included. The simulation starts by placing all the secondary particles on a normalized area at the detector top located just above the WCD, and they are propagated and injected into the detector with their corresponding momenta p_s. Depending on the particle type, the corresponding physical processes can take place and the propagation of relativistic charged particles also produce Cherenkov photons. As soon as a Cherenkov photon of wavelength λ_{Ch} is created, a Metropolis–Hastings algorithm is used to determine if this photon will either produce an electronic signal at the optical detector or not. The detection probability is calculated from the optical detector response using quantum efficiency reported by manufacturer, typically QE(λCh) $\lesssim 25-30\%$. This largely improves the computing efficiency, as only those Cherenkov photons that will eventually produce electronic signals are propagated inside the detector volume ($\lesssim 5-10\%$).

As an example of the ARTI capabilities, in this work we have evaluated the expected performance of the same generic LAGO WCD installed at the previously described sites. LAGO WCDs are typically built with commercial plastic or stainless-steel cylindrical tanks filled with $1\,m^3$ to $4\,m^3$ of purified water and with an inner coating made of Tyvek®, a highly reflective and diffusive material at the ultraviolet spectrum region [51,52]. A single 8"-9" PMT is located at the top centre of the tank, pointing downwards and in close contact with the water surface [54]. In this trial run, the simulated detector is a cylindrical tank ($h = 0.90\,m$ and $r = 1.05\,m$), with the Tyvek coated inner surface, containing $3.12\,m^3$ of pure water, and an 8" Hamamatsu R5912 PMT. After fixing the geometry and characteristics of the detector and its constituents, we use the same calibration procedure that as used at most astroparticle observatories using WCD [20,56,57]: signals are measured in VEMs, i.e., the total signal produced by the passage of vertical and central muons trough the water volume, as sketched in the upper panel of Fig. 5. Muons are typically used for the calibration of the deposited energy in WCD, and at the typical energy range of atmospheric muons ($2-6\,GeV$), their stopping power is nearly constant and close to the so-called MIP (minimum ionization particle), with an energy loss of $\sim 2\,MeV\,cm^2\,g^{-1}$, i.e., $2\,MeV\,cm^{-1}$ in water. To estimate the signal value corresponding to 1 VEM, we inject in the detector vertical axis 10^5 monoenergetic muons of $3\,GeV$ moving in the $-z$ direction and obtaining the distribution shown in the lower panel of Fig. 5. The number of pe produced in the PMT is $60 \lesssim VEM_{pe} \lesssim 250$, which peaked at $1VEM_{pe} \simeq 100\,pe$, and thus $1\,VEM \simeq 100\,pe \simeq 180\,MeV$ of deposited energy, E_d.

Once the VEM was obtained, the standard WCD calibration relies on identifying the main features of the so-called charge histogram, i.e., the histogram of the total charge by each of the secondary particles present in the flux Ξ at the detector level [56]. While the exact characteristics of these features are essentially determined by the detector geometry and its response to the different EAS components, all the charge histograms share the same characteristics: (a) a peak at low E_d values corresponding to the detector response to the EM component and the trigger effect; (b) a second peak at larger E_d values, the so called muon hump, product of the detector response to muons and corresponding to the typical charge produced by vertical muons, i.e., from the position of this hump in the measured histogram the value of the VEM can be obtained [58]; and (c) a change in the histogram slope at higher values of E_d corresponding to the transition from single to multiple particles impinging in the detector [8].

This features can be observed in the simulated charge histograms obtained by ARTI and shown in Fig. 6. These histograms were obtained by simulating the response of the same generic LAGO detector exposed to the total secondary flux Ξ at the sites of Chacaltaya, La Serena, Lima
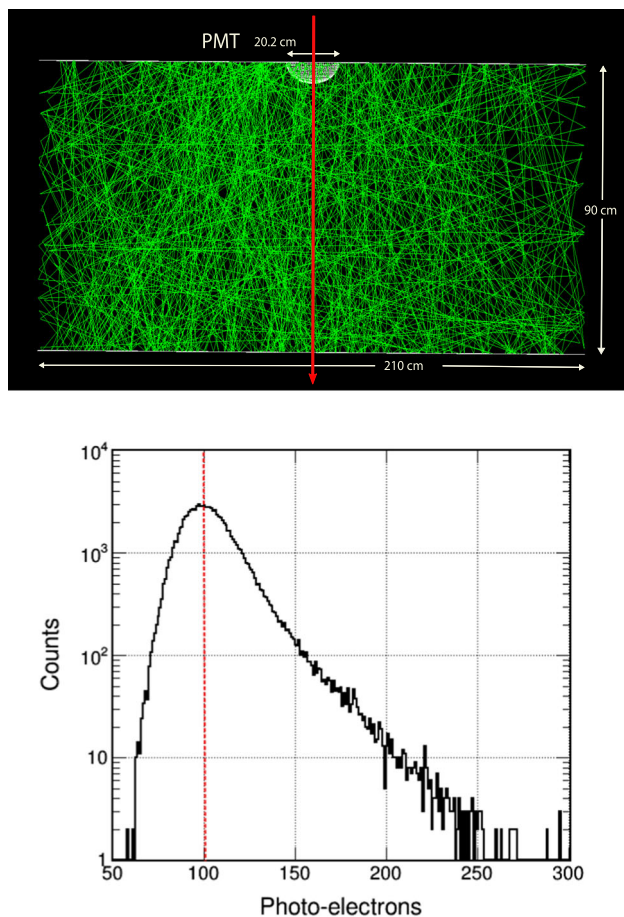
Fig. 5 Top: A vertical and central muon of 3 GeV (red arrow) moves through the detector producing thousands of Cherenkov photons (green lines) that are propagated in the water volume and diffusely reflected in the Tyvek® coating. Eventually, these photons can be either absorbed or finally reach the PMT, producing a certain number of photo-electrons (pe). Bottom: distribution of the expected number of pe obtained due to the passage of 10^5 central and vertical muons of $E_s = 3$ GeV, peaked at a value of 100 pe per VEM, that represents the main unit of calibration for this particular simulated detector. By considering the muon stopping power in water at these energies, then 1 VEM $\simeq$ 100 pe $\simeq$ 180 MeV

and Marambio. For all detectors, the deposited energy for the muon hump, ranges from $150 \text{ MeV} \lesssim E_d \lesssim 300 \text{ MeV}$ with slight differences between the sites, is peaked for all cases at a value consistent with the VEM estimations, i.e., 1 VEM$\simeq$ 180 MeV. The contribution of the different components of Ξ to the charge histogram also depends on altitude; since at higher energy the fraction of muons is not as dominant as near sea level, where the EAS are totally developed and the EM component is more absorbed than the muon component in the atmosphere.

Of course, most of the low energy particles present in the total flux Ξ are not be detected by a WCD and so the total signal flux Ξ^D will be lower that Ξ, as shown in Table 2, where Ξ^D is presented along with the percent reduction of about $55-60\%$ when compared with Ξ. The table also shown

the different contributions of each EAS component to the total deposited energy in the detector.

Another parameter that is related with the EAS evolution is the relative fractions of the different components to the total muon content, since as the cascade evolves charged pions and kaons start to decay in to muons. When the shower $X_{\max}$ is behind, the electromagnetic component is being absorbed by the atmosphere. Thus, it is to be expected that the muon component become more dominant as the altitude decreases, and so it's impact on the total deposited energy. This behaviour can be seen in Fig. 7, where the relative ratios in the deposited energy at the detector by muons and other components, i.e. E_d^i/E_d^μ; where i represents the rest of the cascade component (photons, electrons, neutrons and other hadrons), are shown as a function of altitude, and as a possible observational trigger in our detectors for the ratio Ξ^i/Ξ^μ. This can be used to determine, e.g., the optimal altitude for detecting photon-initiated showers, as those produced during the sudden occurrence of a GRB [6]; or for space weather studies, as in this case we want to measure the Solar activity effects on the measured flux Ξ^D.

6 Conclusions

In the first stage, based on user selections, ARTI performs the calculation of the total number of primaries to be injected into the chosen type of atmosphere and calls on the corresponding CORSIKA routines to simulate the atmospheric response to the primary flux. Then, in the second stage, ARTI analyzes the first stage results to produce the expected flux of secondaries at ground level, Ξ, corrected by the real-time geomagnetic conditions. Finally, during the third stage, the ARTI Geant4 macros are properly configured and used to simulate the expected response of a simulated water Cherenkov detector to the flux of secondaries.

In this work, we present a detailed description of the ARTI simulation toolkit, a publicly available and highly configurable framework designed to obtain in a semi-autonomous way, the detailed flux of the secondary particles and the corresponding signals at any type of water Cherenkov detector, produced by the interaction of the total flux of galactic cosmic rays with the atmosphere; as well as, for example, the expected flux at ground produced by GRBs or other transient astrophysical phenomena. Even more, this can also be done
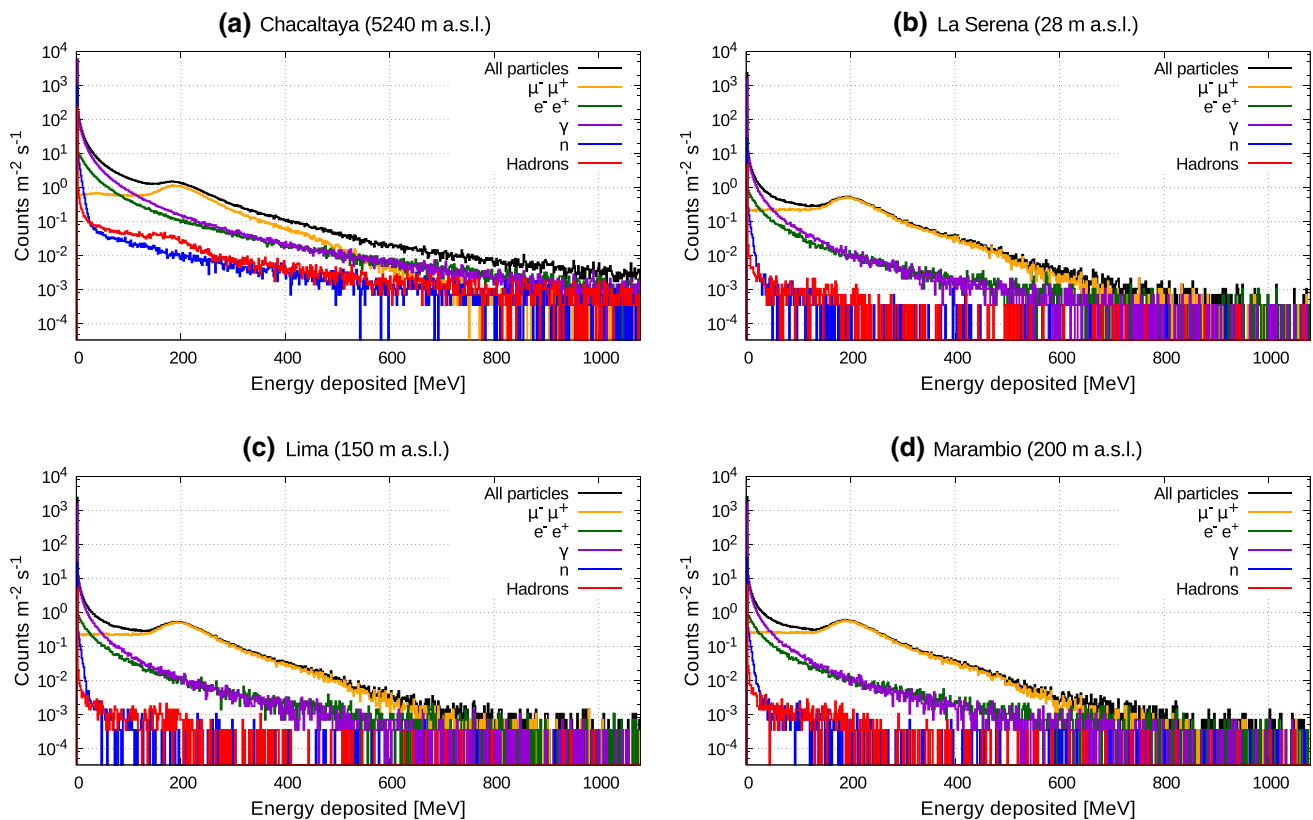
Fig. 6 Charge histograms for four LAGO sites of the total signal (black line) and single component contributions (colour lines): Chacaltaya (**a**), La Serena (**b**), Lima (**c**) and Marambio (**d**). The histograms have been normalized to areas of $1\,m^2$ and exposure time of 1 s

Table 2 Expected signal flux, Ξ^D, in a single LAGO like WCD deployed at the ground level of the eight selected sites used as examples of the ARTI capabilities in this work. The percent relative difference of the expected signals to the expected flux of particles, i.e., $\Delta\Xi/\Xi = (\Xi^D - \Xi)/\Xi\%$, and the deposited energy for the different EAS components of the are also shown: electromagnetic component E_d^{EM}, muon component E_d^μ, neutrons E_d^n, and the total deposited energy E_d^{All}. For the hadronic component, we only show the neutron contribution. The remaining hadronic contribution can be obtained by subtraction

Site	Ξ^D	$\Delta\Xi/\Xi$	E_d^{EM}	E_d^μ	E_d^n	E_d^{All}
CHA	1800	−60	1.49	0.22	0.49	1.77
UIO	520	−59	0.40	0.14	0.11	0.55
GUA	310	−57	0.22	0.11	0.05	0.34
BGA	230	−57	0.16	0.10	0.03	0.26
SAWB	180	−58	0.11	0.10	0.02	0.21
LIM	170	−56	0.11	0.09	0.02	0.20
LSC	170	−55	0.11	0.09	0.01	0.20
EZE	170	−56	0.11	0.09	0.01	0.20

for time-evolving atmospheric and geomagnetic conditions at any place in the World, and including, e.g., the fast disturbances produced in the Earth's magnetic field by the Solar activity.

As an example of the ARTI capabilities, we calculated the signal flux expected at a simulated WCD virtually deployed at eight astroparticle observation sites in Latin America. These sites were selected due to their distinctive characteristics in altitude (different atmospheric depths), and in latitude (different geomagnetic responses). Throughout this work, the three main stages of the ARTI simulation were presented and as well as the description of the physical basis for all the calculations and assumptions.

ARTI can be easily obtained from the LAGO github repository [16], and can be prepared for running automatically at different types of computing facilities: from personal notebooks or desktops up to high performance computing clusters. More recently at cloud-based environments trough `OneDataSim` [30], the ARTI implementation for the European Open Scientific Cloud (EOSC) and other federated and public clouds, which is publicly available at the EOSC Marketplace [59].
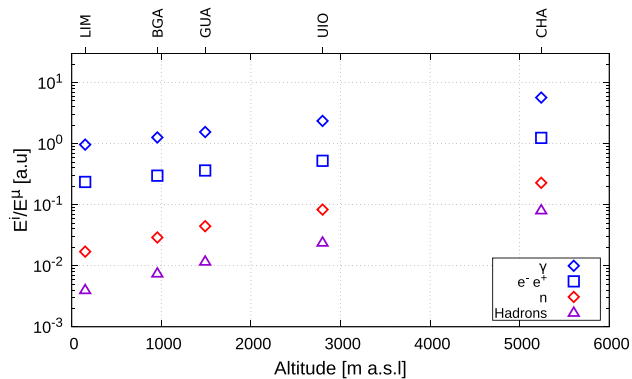
With the help of ARTI, we were able

Fig. 7 The component-to-muon ratio, i.e., the ratio between the deposited energy within the detector volume by a particular shower component (photons, blue diamonds; electrons, blue squares; neutrons, red diamonds; and hadrons, violet triangles) divided by the energy deposited by muons, is shown for five LAGO sites with similar geomagnetic rigidity but increasing altitude at Lima, Bucaramanga, Guatemala, Quito and Chacaltaya. The observed trend for this ratio reflects the expected evolution of the cascades with altitude, as the impact of the atmospheric absorption is much lower for muons than for other shower components. High altitude sites are optimal for studying photon-initiated showers, such as those produced by GRBs

- to characterize the expected response and sensitivity of LAGO or any other astroparticle observatory to the flux of cosmic rays in the galactic energy range [4,29];
- to determine the sensitivity of LAGO at high altitude sites for the observation of steady gamma sources or astrophysical transients, such as the sudden occurrence of a gamma ray burst within the LAGO field of view [6,53];
- to study the impact of space weather phenomena from ground level by using water Cherenkov detectors [8,28, 30,39];
- to calculate the most statistically significant flux of high energy muons at underground laboratories, equivalent to one year of the expected primary flux at the site [30,41];
- to help in the assessment of active volcanoes risks in Latin America [31,32,55,60];
- to design new safeguard radiation detectors for detecting the traffic of fissile materials [50,61];
- to contribute to the detection of improvised explosive devise at warfare fields in Colombia [62];
- to estimate the expected radiation dose received by the crew during commercial flights;
- and even to determine the radiation exposure of equipment and people at Mars' surface during the incoming exploration missions.

Appendix A: ARTI pseudocode

The following algorithm represents the three main stages that make up ARTI for flux simulations and EAS developments via CORSIKA, magnetic field correction via MagnetoCosmics and detector simulation via Geant4.

Simulation flux;

Input : E = [1, 10^6] GeV; energy range
Injected primary nuclei, Z = [1, 26]
θ = [0, 90]; zenith angle range
ϕ = [-180, 180]; azimuth angle range
Altitude in m a.s.l.
B_x, B_z; site's magnetic field
atmospheric model of the site
time;

Output: $\varXi$, particle flux at ground

begin
| $\Phi(E_p, Z, A, \Omega)$; Integrate cosmic-ray spectra
| Z, #part(E) → built steering Corsika files
| run block, via Corsika software
| Analysis block; read and uncompress binary files
end
return $\varXi$;
Magnetic field correction;

Input : $\varXi$, particle flux at ground
Rm, magnetic rigidity
IGRF, magnetic model
Output: $\varXi_{corr}$, particle flux at ground corrected
begin
| $R_C(\phi, \theta)$, magnetic rigidity cutoff
end
return $\varXi_{corr}$;

Water Cherenkov detector simulation;

Input : $\varXi_{corr}$, geomagnetic corrected particle flux at ground
D(r,h), detector's dimensions and characteristics
W(λ), water properties
PMT, photo-multiplier's features
Output: E_D, energy deposited

for *particles* **do**
| Interacting with water
| E_i^D, energy deposited by i-particle
end
return E_D & charge histograms;

Algorithm 1: ARTI is divided in three consecutive stages: flux calculations, atmospheric model and geomagnetic conditions to simulate EAS developments via Corsika; geomagnetic secular variations and disturbances via Magneto-Cosmics and their impact on the particle flux at ground; and the detector simulation via Geant4.

Appendix B: ARTI command line options example

As described in the text, the simulation can be totally configured by selecting the corresponding options from the command line at launch time. These include the flux time, the CORSIKA version, the observatory site, the number of process to use, the activation of the Cherenkov mode in the shower, etc. All the user options override the default values included in ARTI. As an example, here we show the currently available options for the main scripts of the first two ARTI stages: do_sims.sh and do_showers.sh. All the ARTI scripts have their internal help that can be seen by using the -? modifier.

```
$ do_sims.sh -?

USAGE do_sims.sh:

Simulation parameters
  -w <working dir>                  : Working directory, where bin (run) files are located
  -p <project name>                 : Project name (suggested format: NAMEXX)
  -v <CORSIKA version>              : CORSIKA version
  -h <HE Int Model (EPOS|QGSII)>    : Define the high interaction model to be used
  -u <user name>                    : User Name.
  -j <procs>                        : Number of processors to use
Physical parameters
  -t <flux time>                    : Flux time (in seconds) for simulations
  -m <Low edge zenith angle>        : Low edge of zenith angle.
  -n <High edge zenith angle>       : High edge of zenith angle.
  -r <Low primary particle energy>  : Lower limit of the primary particle energy.
  -i <Upper primary particle energy> : Upper limit of the primary particle energy.
  -a <high energy ecuts>            : High energy cuts for ECUTS
  -y                                : Select volumetric detector mode (default=flat array)
Site parameters
  -s <site>                         : Location (several options)
  -k <altitude, in cm>              : Fix altitude, even for predefined sites
  -c <atm_model>                    : Fix Atmospheric Model even for predefined sites.
  -o <BX>                           : Horizontal comp. of the Earth's mag. field.
  -q <BZ>                           : Vertical comp. of the Earth's mag. field.
  -b <rigidity cutoff>              : Rigidity cutoff; (if set value in GV = enabled).
  -g <Lat, Lon>                     : Obtain the current values of BX and BZ for a
                                      site located at (Lat,Lon,Alt). Unless the -s option
                                      is used, -k option is mandatory.

Modifiers
  -l                                : Enables SLURM cluster compatibility (with sbatch).
  -e                                : Enable CHERENKOV mode
  -d                                : Enable DEBUG mode
  -x                                : Enable other defaults (It doesn't prompt user
                                      for unset parameters)
  -?                                : Shows this help and exit.

$ do_showers.sh -?

USAGE do_showers.sh:

  -o <origin directory>   : Origin dir, where the DAT files are located
  -r <ARTI directory>     : ARTI installation directory, generally pointed by $LAGO_ARTI
                            environment variable (default)
  -w <working directory>  : Working dir, where the analysis will be done (default is current
                            directory)
  -e <energy bins>        : Number of energy secondary bins (default: 20)
  -d <distance bins>      : Number of distance secondary bins (default: 20)
  -p <project base name>  : Base name for identification of S1 files (don't use spaces).
                            Default: odir basename
  -k <site altitude, in m> : For curved mode (default), site altitude in m a.s.l. (mandatory)
  -s <type>               : Filter secondaries by type: 1: EM, 2: MU, 3: HD
  -t <time>               : Normalize energy distribution in particles/(m2 s bin), S=1 m2;
                            <t> = flux time (s).
  -m <bins per decade>    : Produce files with the energy distribution of the primary flux
                            per nuclei.
  -g <file> <col>         : Include geomagentic effects. Read the local rigidities from
                            column <col> of <file>.
  -j                      : Produce a batch file for parallel processing. Not compatible
                            with local (-l)
  -l                      : Enable parallel execution locally (N procs). Not compatible
                            with parallel (-j)
  -v                      : Enable verbosity. Will write .log processing files.
  -?                      : Shows this help and exit.
```

References

1. L. Terray et al., Radon activity in volcanic gases of Mt. Etna by passive dosimetry. JGR Solid Earth **125**(9), e2019JB019149 (2020)
2. K.H. Kampert, A. Watson, Extensive air showers and ultra high-energy cosmic rays: a historical review. EPJ H **37**(3), 359–412 (2012)
3. The Pierre Auger Collaboration, The Pierre Auger Observatory and its upgrade. Sci. Rev. from the end of the world **1**(4), 8–33 (2020)
4. I. Sildenik, H. Asorey for the LAGO Collaboration, LAGO: The Latin American Giant Observatory, NIM A **876**, 173–175 (2017)
5. K. Hurley et al., Detection of a γ-ray burst of very long duration and very high energy. Nature **372**, 652–654 (1994)
6. C. Sarmiento-Cano et al., The Latin American Giant Observatory (LAGO) capabilities for detecting Gamma Ray Bursts, in *Proceedings of Science*, Berlin, Germany, vol. PoS(ICRC2021), p. 929 (2021)
7. I.G. Usoskin et al., Forbush decreases of cosmic rays: energy dependence of the recovery phase. JGR Space Phys. **113** (A07102), p. 1–7 (2008)
8. H. Asorey et al., The LAGO Space Weather Program: directional geomagnetic effects, background fluence calculations and multi-spectral data analysis, in *Proceedings of Science*, The Hague, The Netherlands, vol. PoS(ICRC2015), p. 142 (2015)
9. K. Jourde et al., Monitoring temporal opacity fluctuations of large structures with muon radiography: a calibration experiment using a water tower. Sci. Rep. **6**(1), 1–11 (2016)
10. K. Morishima et al., Discovery of a big void in Khufu's Pyramid by observation of cosmic-ray muons. Nature **552**, 386–390 (2017)
11. J. Peña-Rodríguez et al., Design and construction of MuTe: a hybrid Muon Telescope to study Colombian volcanoes. JINST **15**, P09006 (2020)
12. L. Desorgher, MAGNETOSCOSMICS, Geant4 application for simulating the propagation of cosmic rays through the Earth magnetosphere. Technical report, Physikalisches Institut, University of Bern, Bern, Germany (2003)
13. D. Heck et al., CORSIKA: a Monte Carlo code to simulate extensive air showers, FZKA **6019**, p. 1–90 (1998)
14. A. Ferrari et al., Fluka: a multi-particle transport code. CERN 2005-010, 1–405 (2005)
15. S. Agostinelli, J. Allison, K. Amako et al., Geant4—a simulation toolkit. NIM A **506**, 250–303 (2003)
16. H. Asorey et al., The ARTI framework: cosmic rays atmospheric background simulations (2015). https://github.com/lagoproject/arti. Accessed 2022
17. H. Asorey for the LAGO Collaboration, The LAGO solar project, in *Proceedings of the 33rd ICRC*, Brazilian Journal of Physics, Rio de Janeiro, Brazil, vol. IV (2013), pp. 1–4
18. H. Asorey, L.A. Núñez, M. Suárez-Durán, Preliminary results from the Latin American Giant Observatory space weather simulation chain. Space Weather **16**, 461–475 (2018)
19. J. Grisales-Casadiegos, C. Sarmiento-Cano, L.A. Núñez, Impact of Global Data Assimilation System atmospheric models on astroparticle showers. Can. J. Phys. **100**(3), 152–157 (2022)
20. H. Asorey, S. Dasso, for the LAGO Collaboration, LAGO: the Latin American Giant Observatory, in Proceedings of the 34th ICRC, The Hague, The Netherlands, Proceedings of Science PoS(ICRC2015) (2015), p. 247 arXiv:1605.02151
21. A.J. Rubio-Montero et al., A novel cloud-based framework for standardized simulations in the Latin American Giant Observatory (LAGO), in *IEEE Proceedings of the WSC2021* , IEEE, Phoenix, USA (2021), p. 9715360
22. A.J. Rubio-Montero et al., For the LAGO Collaboration, LAGO data management plan (2021). https://lagoproject.github.io/DMP/. Accessed 2022
23. P. Alken et al., International geomagnetic reference field: the thirteenth generation. Earth Planets Space **73**, 49–56 (2021)
24. P. Bobik, K. Kudela, I. Usoskin, Geomagnetic cutoff Penumbra structure: approach by transmissivity function, in *Proceedings of the 27th ICRC*, Copernicus Gesellschaft, Hamburg, Germany vol. 4056 (2001)
25. R. Brun, F. Rademakers, ROOT—an object oriented data analysis framework. NIM A **389**, 81–86 (1997)
26. S. Dasso et al., for the LAGO Collaboration, A project to install water-Cherenkov detectors in the Antarctic peninsula as part of the LAGO detection network, in *Proceedings of Science*, The Hague, The Netherlands, (2015), p. 105
27. F.X. Kneizys et al., *The MODTRAN 2/3 Report and LOWTRAN 7 Model* (Phillips Laboratory, Hanscom AFB, 1996)
28. C. Sarmiento-Cano et al., Modeling the LAGO's detectors response to secondary particles at ground level from the Antarctic to Mexico, in *Proceedings of the 36th ICRC*, vol. PoS(ICRC2019) (2019), p. 412
29. The Pierre Auger Collaboration, Studies on the response of a water-Cherenkov detector of the Pierre Auger Observatory to atmospheric muons using an RPC hodoscope. JINST **15**(09), P09002 (2020)
30. J.A. Rubio-Montero et al., The EOSC-Synergy cloud services implementation for the Latin American Giant Observatory (LAGO), in *Proceedings of Science*, Berlin, Germany, vol. PoS(ICRC2021), p. 261 (2021). arXiv:2111.11190
31. A. Taboada, C. Sarmiento-Cano, A. Sedoski, H. Asorey, Meiga, a dedicated framework used for muography applications. JAIS **2022**(1), p. 1–4 (2022)
32. J. Peña-Rodríguez, A. Vesga-Ramírez, A. Vásquez-Ramírez et al., Muography in Colombia: simulation framework, instrumentation and data analysis. JAIS **2022**(6), 1–8 (2022). arXiv:2201.11160
33. M. Aguilar et al., Relative composition and energy spectra of light nuclei in cosmic rays: results from AMS-01. ApJ **724**, 329 (2010)
34. S. Ostapchenko, Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model. PR D **83**, 014018 (2011)
35. E. Thébault et al., International geomagnetic reference field: the 12th generation. Earth Planets Space **67**, 79 (2015)
36. A. Letessier-Selvon, T. Stanev, Ultrahigh energy cosmic rays. Rev. Mod. Phys. **83**, 907–916 (2011)
37. J.J. Masías-Meza, S. Dasso, Geomagnetic effects on cosmic ray propagation under different conditions for Buenos Aires and Marambio, Argentina. SunGe **9**, 41–47 (2014)
38. P.A. Zyla et al. (Particle Data Group), Review of particle physics. PTEP **2020**(8), 083C01 (2020)
39. I. Sidelnik et al., Simulation of 500 MeV neutrons by using NaCl doped water Cherenkov detector. Adv. Space Res. **65**(9), 2216–2222 (2020)
40. L. Bonechi, R. D'Alessandro, A. Giammanco, Atmospheric muons as an imaging tool. Rev. Phys. **5**, 100038 (2020)
41. C. Perez Bertolli, C. Sarmiento-Cano, H. Asorey, Estimation of the muon flux expected at the ANDES underground laboratory. ANALES AFA **32**(4), 106–111 (2022)
42. D. Smart, M. Shea, Fifty years of progress in geomagnetic cutoff rigidity determinations. Adv. Space Res. **44**, 1107–1123 (2009)
43. R. Modzelewska, M. Alania, The 27-day cosmic ray intensity variations during solar minimum 23/24. Sol. Phys. **286**, 593–607 (2013)
44. J. Masías-Meza et al., Superposed epoch study of ICME substructures near earth and their effects on galactic cosmic rays. A&A **592**, A118 (2016)
45. H. Cane, Coronal mass ejections and Forbush decreases. Space Sci. Rev. **93**, 55–77 (2000)
46. H. Asorey for the Pierre Auger Collaboration, Measurement of low energy cosmic radiation with the water Cherenkov detector array of the Pierre Auger Observatory, in *Proceedings of the*

33rd ICRC (2011), pp. 41–44. http://icrc2011.ihep.ac.cn/paper/, Beijing,China

47. I. Angelov, E. Malamova, J. Stamenov, The Forbush Decrease after the GLE on 13 December 2006 detected by the Muon Telescope at BEO–Moussala. Adv. Space Res. **43**, 504–508 (2009)

48. S. Dasso, H. Asorey for the Pierre Auger Collaboration, The scaler mode in the Pierre Auger Observatory to study heliospheric modulation of cosmic rays. Adv. Space Res. **49**, 1563–1569 (2012)

49. M. Mostafá, For the HAWC Collaboration, The high-altitude water Cherenkov Observatory. Braz. J. Phys. **44**, 571–580 (2014)

50. I. Sidelnik et al., Enhancing neutron detection capabilities of a water Cherenkov detector. NIM A **955**, 163172 (2020)

51. A. Filevich et al., Spectral-directional reflectivity of Tyvek immersed in water. NIM A **423**, 108–118 (1999)

52. R. Calderón, H. Asorey, L.A. Núñez for the LAGO Collaboration, Geant4 based simulation of the Water Cherenkov Detectors of the LAGO Project. Nucl. Part. Phys. **267**, 424–426 (2015)

53. D. Allard et al., Use of water-Cherenkov detectors to detect gamma-ray bursts at the large aperture GRB observatory (LAGO). NIM A **595**, 70–72 (2008)

54. I. Sidelnik for the LAGO Collaboration, The sites of the Latin American Giant Observatory, in Proceedings of Science, The Hague, The Netherlands, vol. PoS(ICRC2015) (2015), p. 665

55. A. Vásquez-Ramírez et al., Simulated response of MuTe, a hybrid Muon Telescope. JINST **15**(08), P08004 (2020)

56. X. Bertou et al., Calibration of the surface array of the Pierre Auger Observatory. NIM A **568**, 839–846 (2006)

57. A. Galindo for The LAGO Collaboration, Calibration and sensitivity of large water-Cherenkov detectors at the Sierra Negra site of LAGO, in Proceedings of Science, The Hague, The Netherlands, vol. PoS(ICRC2015) (2016), p. 673

58. The Pierre Auger Collaboration, Studies on the response of a water-Cherenkov detector of the Pierre Auger Observatory to atmospheric muons using an RPC hodoscope. JINST **15**(09), P09002 (2020). arXiv:2007.04139

59. A.J. Rubio-Montero et al., EOSC marketplace—service One-DataSim (2021). https://marketplace.eosc-portal.eu/services/onedatasim. Accessed 2022

60. A. Vesga-Ramírez et al., Simulated annealing for volcano muography. JSAES **109**, 103248 (2021)

61. I. Sidelnik et al., Neutron detection capabilities of water Cherenkov detectors. NIM A **952**, 161962 (2020)

62. A. Vásquez Ramírez et al., Improvised explosive devices and cosmic rays, in Proceedings of Science, Berlin, Germany, vol. PoS(ICRC2021) (2021), p. 480