Eur. Phys. J. C (2023) 83:932
https://doi.org/10.1140/epjc/s10052-023-12106-5

THE EUROPEAN PHYSICAL JOURNAL C



Regular Article - Experimental Physics

# The gamma/hadron discriminator $LCm$ in realistic air shower array experiments

**R. Conceição**[1,2,a], **P. J. Costa**[1,2], **L. Gibilisco**[1,2], **M. Pimenta**[1,2], **B. Tomé**[1,2]

[1] Laboratório de Instrumentação e Física Experimental de Partículas (LIP), Lisbon, Portugal
[2] Instituto Superior Técnico (IST), Universidade de Lisboa, Lisbon, Portugal

Received: 24 April 2023 / Accepted: 6 October 2023 / Published online: 15 October 2023


**Abstract** In this article, it is shown that the $C_k$ and $LCm$ variables, recently introduced as an effective way to discriminate gamma and proton-induced showers in large wide-field gamma-ray observatories, can be generalised to be used in arrays of different detectors and variable fill factors. In particular, the $C_k$ profile discrimination capabilities are evaluated for scintillator and water Cherenkov detector arrays.

## 1 Introduction

At energies surpassing approximately 100 GeV, gamma-rays originating from astrophysical sources cannot be directly detected. Instead, their detection is inferred indirectly by reconstructing the extensive air showers (EAS) they generate upon interacting with Earth's atmosphere [1]. A careful evaluation of the shower characteristics is essential to differentiate EAS induced by gamma rays from those produced by the dominant cosmic ray background. A noteworthy parameter in this discrimination process is the assessment of a shower's muon content, which is expected to be higher for showers triggered by hadronic processes than gamma rays.

Although muon counting offers a highly effective method for discriminating between gamma and hadron-induced showers (see for instance [2], its practical implementation necessitates some form of shielding - such as a layer of soil or water above the detector - to absorb the electromagnetic component of the shower. This requirement renders such experiments financially demanding and largely unfeasible in ecologically sensitive regions.

A recent study, documented in [3], demonstrated that crucial information for discriminating between gamma and hadron-induced showers can be derived from the azimuthal asymmetry of the ground-level shower footprint. Through comprehensive simulation investigations, the discriminatory potential of the newly introduced observables, denoted as $LCm$, is comparable to that achieved through muon counting. More importantly, these observables offer significantly improved experimental accessibility, addressing the challenges posed by the need for shielding in muon-based methods.

According to [3], the assessment of the fluctuations is done through the quantity $C_k$, defined in circular rings $k$ centred around the shower core position, with a width of 10 m and a mean radius $r_k$ as:

$$C_k = \frac{2}{n_k(n_k-1)} \frac{1}{\langle S_k \rangle} \sum_{i=1}^{n_k-1} \sum_{j=i+1}^{n_k} (S_{ik} - S_{jk})^2, \tag{1}$$

where $n_k$ is the number of stations in the ring $k$, $\langle S_k \rangle$ is the mean signal in the stations of the ring $k$, and $S_{ik}$ and $S_{jk}$ are the collected signals in the stations $i$ and $j$ of the ring $k$, respectively.

The shower azimuthal asymmetry level is stated by the quantity $LCm$ defined as the value of a parametrisation of the $\log(C_k)$ distribution at a given value of $r_k = r_m$, and $r_m$ was fixed to $r_m = 360$ m.

The behaviour of $LCm$ has been studied in [3] as a function of a scaling factor defined as such:

$$K = E^{\beta} \times FF, \tag{2}$$

where $E$ is the primary energy (in TeV), $\beta$ is the index of the power dependence of the mean number of muons at the ground and $FF$ is the fill factor, the fraction of instrumented array area. The parameter $\beta$ was fixed to 0.925, a typical mean value used in hadronic shower simulations. It has been shown that, for different energies and fill factors, but identical $K$ factors, the $C_k$ distributions are essentially the same.

In [3], a uniform array of single-layer water Cherenkov detectors (WCDs) stations with a constant fill factor and each

[a] e-mail: ruben@lip.pt (corresponding author)

with four photomultiplier tubes (PMTs) at its bottom was considered; an overview of these results will be given in Sect. 2.1.

Motivated by studies that try to use $C_k$ and $LCm$ in realistic experiments, such as its application to KASCADE data [4], this study has been hereby extended to setups with similar single-layer WCDs with the same radius and height, but with different numbers of photo-sensors at their bottom (Sect. 3). The applicability of the $C_k$ variable in scintillator arrays has been assessed as well in Sects. 2.2 and 2.3. Finally, a simple way to handle arrays with variable $FF$ is discussed in Sect. 4.

The results presented in this work were obtained using air shower simulations whose output was subsequently processed to emulate the behaviour of a detector array. CORSIKA (version 7.5600) [5] was used to simulate gamma-ray and proton-induced vertical showers assuming an observatory altitude of 5200 m a.s.l., and FLUKA [6,7] and QGSJET-II.04 [8] were used as hadronic interaction models for low and high energy interactions, respectively. The current investigations were conducted within a specific gamma-ray energy range, encompassing a single energy bin of width $\log(E) = 0.2$, originating at 100 TeV. The energies of the generated proton samples were chosen so that the total electromagnetic signal at the ground would be similar to the gamma-ray-generated showers.

A 2D-histogram emulated the ground detector array with cells with an area of $\sim 12\,\mathrm{m}^2$ covering the available ground surface. Each cell represents one station. The signal in each station was estimated as the sum of the expected signals due to the particles hitting the station, using calibration curves for each station type as a function of the particle energy for protons, muons and electrons/gammas. Fluctuations induced by the detector were mimicked by applying Gaussian distributions centred on the values given by the calibration curves and with sigmas equal to the respective RMS. The fill factor of the array was set in the interval $\in ]0, 1]$, by masking the 2D-histogram with the appropriate regular pattern. Following reference [3], it was chosen for the $\sim$ 100 TeV simulation a sparse array with a fill factor of 12%.

The same methodology as previously outlined will be employed to analyze the scintillator arrays, both with and without the inclusion of a lead converter.

## 2 $LCm$ and the detector technology

### 2.1 Water Cherenkov detectors

Due to the presence of hadronic sub-showers in hadron-induced extensive air showers, the $C_k$ variable, as defined in Eq. (1), is expected to be larger for proton-induced showers compared to gamma-induced showers with equivalent energies at the ground. This has been verified in [3], where the $C_k$ variable has been computed for proton and gamma showers using a uniform array of WCDs, each equipped with 4 PMTs at the bottom. The same results have been hereby replicated with an array of Mercedes (3 PMTs at the bottom) WCDs [9] of the same shape and size and a uniform $FF$ of 12% for gamma and proton showers with primary energy around 100 TeV. The detailed analysis of the dependence of $C_k$ from the number of PMTs in the WCDs is presented in Sect. 3. In Fig. 1, it is shown the distributions of the mean values of $\log(C_k)$ are represented as a function of the radius $r_k$ for both primaries. The error bars depict the standard deviation of the distributions. The $LCm$ has then been computed, and it is shown in Fig. 2. At an efficiency close to 100%, next to no background events are left within the current limits of the statistics.[1]

### 2.2 Scintillators

Scintillator arrays are widely used in cosmic-ray observatories. Without entering a detailed discussion of their advantages or disadvantages as compared to WCDs arrays, hereafter, their performance in gamma/hadron discrimination is explored regarding the gamma/hadron discriminating variables, $C_k$ and $LCm$.

Scintillators are very good for tagging charged particles but not for measuring their energies and/or identifying muons without using shielding to the other charged particles. Therefore gamma/hadron discriminating algorithms based on muon counting are not efficient for unshielded scintillator arrays. Furthermore, scintillators are mostly insensitive to the shower secondary photons.

However, as discussed in reference [3], the strong correlation between $LCm$ and the total number of muons hitting the detectors is still present without considering the contribution to the signal from the muons. Therefore, in this section, we investigate the possibility of using $LCm$, measured in scintillator arrays, as a gamma/hadron discriminator.

In the present simulation framework (see Sect. 1), the emulation of scintillator arrays can be easily done by introducing new calibration curves corresponding to their response to protons, muons, electrons and gammas as a function of the particle energy.

To be as realistic as possible, these calibration curves were built from the mean expected signal of a simulation of the plastic scintillator detectors using the Geant4 toolkit [11–13], in particular its capabilities to simulate the optical processes

[1] While the present study was done with $\mathcal{O}(10^4)$ events for proton showers, in a recent study, just submitted for publication [10], it is demonstrated using a simulation set with $\mathcal{O}(10^6)$ event that $LCm$ has a discrimination capability slightly higher that the measurement of the EAS muon content.

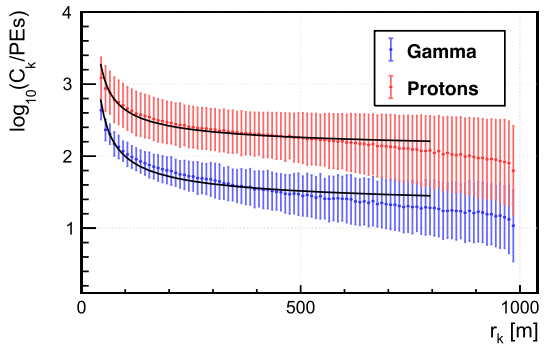


**Fig. 1** $C_k$ profile as a function of the distance to the shower core computed for showers with energies of $\sim$ 100 TeV, using an array of Mercedes WCD stations with $FF = 12\%$

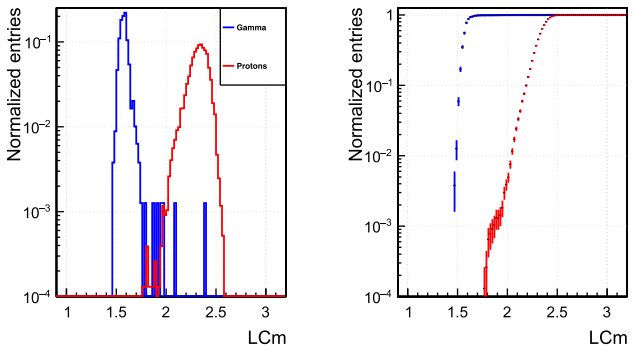


**Fig. 2** $LCm$ distribution (left) and cumulative (right) for showers with energies of $\sim$ 100 TeV, using an array of Mercedes WCD stations with $FF = 12\%$

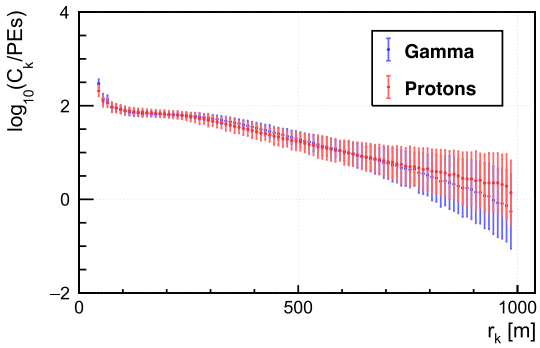


**Fig. 3** $C_k$ distributions for showers with energies of $\sim$ 100 TeV, using an array of scintillators with $FF = 12\%$

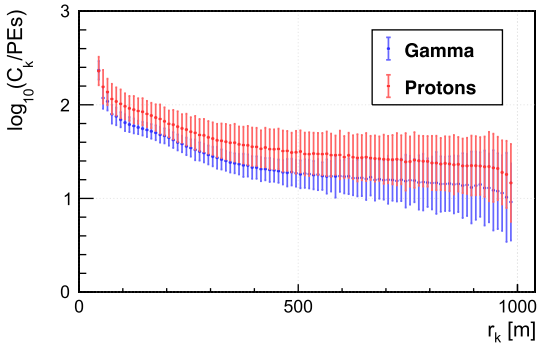


**Fig. 4** $C_k$ distributions for showers with energies of $\sim$ 100 TeV, using an array of lead-shielded scintillators with $FF = 12\%$

and describe the optical properties of the materials. The scintillator is 1 cm thick and 50 cm long. The light is readout at both ends by a photodetector with the same sensitive area as the Hamamatsu R9420, with a 38 mm bialkali photocathode. All relevant optical parameters [14] were implemented in this simulation, namely the scintillator's light yield, the emission spectra and the quantum efficiency of the PMT's photocathode. The optical properties of a white diffuser, wrapping the scintillator, were also included, using the unified model [15] implemented in the Geant4 toolkit.

Figure 3 shows the $\log(C_k)$ distributions as a function of the radius $r_k$ for $\sim$ 100 TeV gamma and proton-induced showers with similar energy at the ground, considering a scintillator array with a $FF = 12\%$. From this figure, it can be seen that the $C_k$ profile does not depend on the nature of the primary particle and, therefore, cannot be used as a gamma/hadron discriminator.

### 2.3 Scintillators coupled to lead converters

As a further exercise, to enhance the scintillator response to shower photons, a thin, 1 $X_0$ thick lead layer was placed on the top of the scintillators. The situation improves compared to the unshielded scintillators. In this case, the number of electrons produced during the ionization losses in the lead plate will scale with energy, making the apparatus relatively sensitive to the shower calorimetry. Consequently, $C_k$ will differ for gamma and hadron-induced showers, as seen in Fig. 4, thus enabling gamma/hadron discrimination. However, it should be noted that the WCD has a stronger discrimination power, as can be seen by comparing the separation between primaries in Figs. 1 and 4.

## 3 Impact on the number of photosensors

Having established the WCDs as a better choice than scintillators in terms of gamma/hadron discrimination capabilities, the study of the $C_k$ variable is now extended to WCDs with different numbers of photosensors. The PMTs are all placed at the bottom of the tank. The dimensions of the station are the same for all tested PMT configurations: a radius of 2m and a height of 1.7 m. In particular, arrays of 4-PMTs and Mercedes stations and arrays of stations with a single central PMT at their bottom (hereafter designated as Mercedes-1) are studied.

To simplify the comparison of results obtained using setups with stations with a different number of PMTs, $C_k$ and $LCm$ variables are normalised, in each setup, by the cor-

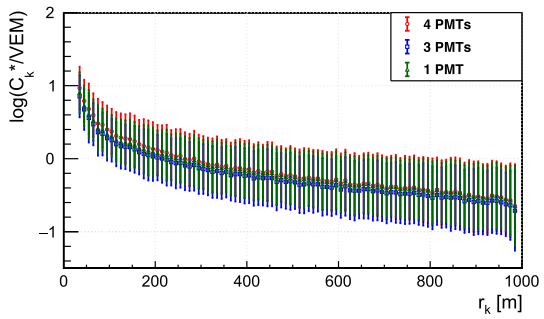


**Fig. 5** $C_k^*$ distributions for proton showers with energies of $\sim$ 100 TeV using a $FF = 12\%$ array of WCD stations equipped with four PMTs (red circles), three PMTs (blue squares) and one PMT (green triangles)

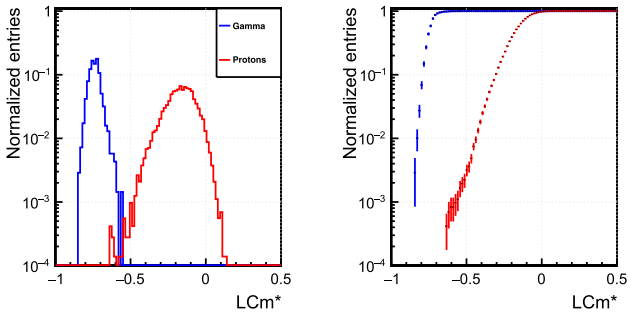


**Fig. 6** $LCm$ distribution (left) and cumulative (right) for showers with energies of $\sim$ 100TeV, using an array of Mercedes-1 WCD stations with $FF = 12\%$

responding mean signal produced by one relativistic vertical muon crossing one station at its centre (VEM), $Q_{\mathrm{VEM}}$, and renamed as $C_k^*$ and $LCm^*$:

$$C_k^* = \frac{C_k}{Q_{\mathrm{VEM}}}, \tag{3}$$

$$LCm^* = \frac{LCm}{Q_{\mathrm{VEM}}}. \tag{4}$$

It should be noted that each station (with a different number of PMTs) will have a specific $Q_{VEM}$. This value can be obtained using dedicated measurement [16] or the analysis of the omnidirectional atmospheric muons [17]. According to its definition, the mean of $C_k^*$ should not depend too much on the number of the PMTs placed at the bottom of each station as long as the expected mean signal of the station is high enough not to introduce significant statistical fluctuations. Indeed, this behaviour is confirmed in Fig. 5, where it is shown, for proton showers, the $\log(C_k^*)$ distributions for identical WCD stations but different numbers of PMTs. After the VEM normalisation, the differences become quite small. From the above considerations, an array of Mercedes-1 WCDs with a water height of 1.7m would guarantee the required level of gamma/hadron discrimination (rejection factors of the order or higher than $10^{-4}$) for energies and $FF$ ensuring a scaling factor $K$ (Eq. 2) of about 5–10. Such a premise is verified in Fig. 6, where the $LCm^*$ distributions, as well as their cumulative distributions, are shown for gamma (blue points) and proton-induced showers (red points). The primary energies of gamma showers are $\sim$ 100TeV, and the proton showers have been selected to have a similar energy at the ground, while the array of Mercedes-1 stations has $FF = 12\%$.

## 4 $LCm$ computation in inhomogeneous arrays

To cover a wide energy range, the layout of many present and future gamma-ray Observatories has a high $FF$ in the inner regions (the so-called compact arrays), primarily intended to cover the lower energy region, and a low $FF$ in the outer regions (usually designated as sparse arrays), conceived mainly to reach the higher energies. Ideally, the transition between these two regions should be smooth to optimise the observatory's sensitivity to intermediate energies.

In all these layout designs, the fill factor will not be constant throughout the array, inducing discontinuities in the $C_k$ distributions, not present beforehand. This effect is particularly evident in Fig. 7, where the shower core was placed at a distance of 300 meters from the centre of the array. This particular array is composed of two fill factors: a dense array ($FF = 100\%$) with a radius of 160 meters, encircled by a sparse array with a radius of 560 meters and a $FF$ of 12%. It is important to note that in addition to the observed discontinuities, the error bars in this figure are notably larger compared to those presented in Fig. 1. This discrepancy arises due to the increased number of stations involved in the computation of $C_k$ in the latter case.

To handle these discontinuities, an effective fill factor in each ring is introduced. This factor is defined as:

$$FF_k = \frac{n_k}{n_{k_1}}, \tag{5}$$

where $n_k$ is, as before, the number of stations in ring $k$, while $n_{k_1}$ is the number of stations in the ring $k$ if the $FF$ is 100%.

Consequently, Eq. (2) gets redefined for each ring as:

$$K_k = E^{\beta} \times FF_k. \tag{6}$$

According to equation 4.3 of [3], $LCm$ is a function of $K$ and may be parameterised for the proton sample as:

$$f_p(K) = LCm_p(K) \sim A_p + \frac{B_p}{\sqrt{K}} \tag{7}$$

where $A_p$ and $B_p$ are constants defined for the primary proton sample.

The function $f_p(K)$ can now be used as a correction factor:

$$C_{kcor} = C_k 10^{(f_p(K_{ref}) - f_p(K_k))}, \tag{8}$$

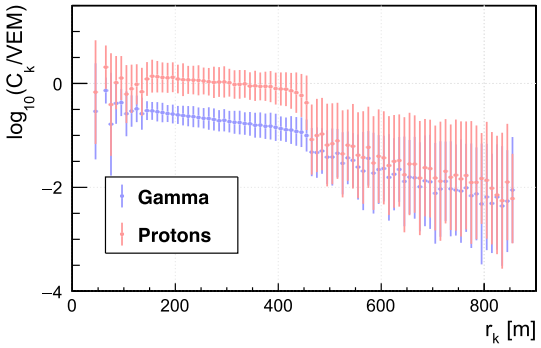


**Fig. 7** $C_k^*$ distributions for showers with energies of $\sim$ 100 TeV using an array of Mercedes WCD stations centred 300m away from the shower core and composed of a dense ($FF = 100\%$) 160 m-wide central region surrounded by a 560 m-wide ring with $FF = 12\%$

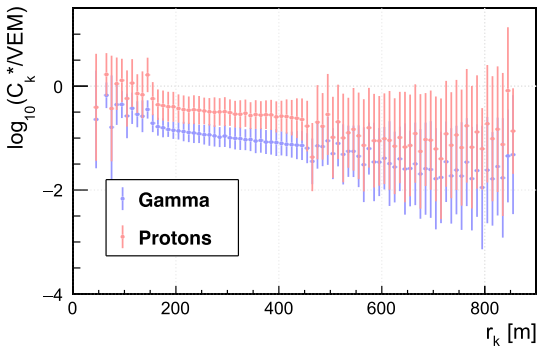


**Fig. 8** Same as Fig. 7, but with $C_k^*$ generalised following the definition in Eq. (8)

where $K_{\text{ref}}$ and $K_k$ are computed using Eq. (6), but using for the $FF_k$, respectively, a reference value (typically the mean $FF$ of the array) and the effective $FF_k$ of that specific ring $k$ (Eq. 5). Such a correction factor has to be applied whenever the effective $FF_k$ is not the reference $FF$, as in the case where two or more regions with different $FF$ are covered in the same ring. The same applies to rings partially covering regions outside the experiment instrumented region.

The correction factor computed for proton-induced showers was also applied when considering gamma primaries, even if it introduces a small systematic error. Such error, which may be minimised by fine-tuning the correction factor, will most likely induce a marginal inclusion of a few gamma-induced showers in the upper tail of the corresponding gamma $LCm$ distribution. This will slightly increase the efficiency for gamma showers, which is irrelevant for the purposes of this article. In fact, for $K$ greater than a few units, the lines describing the evolution of $LCm$ as a function of $K$ of protons and gammas are essentially parallel (see Figure 5 from reference [3]).

Such corrections bring the mean values computed in each ring to the level of the values expected in the case of an equivalent ring embedded in a uniform array with the reference $FF$. This effect is verified by comparing Fig. 7 with Fig. 8, produced in the same conditions but applying the correction factor.

In Fig. 8, it can also be seen that while there is no relevant discontinuity in the mean values after applying the correction factor, the error bars are considerably smaller in the high $FF$ array region, as expected.

Lastly, it is noteworthy that while the aforementioned investigations were centered on simulations at approximately 100 TeV, it was verified that the findings remain consistent regardless of the energy. This verification was carried out by analyzing two additional energy bins, each containing 10% of the simulation data used for the 100 TeV bin. These energy bins specifically encompassed 10 TeV and 1 PeV.

## 5 Conclusions

In this article, the applicability of the $C_k$ and $LCm$ gamma/hadron discriminator quantities to different realistic experimental scenarios has been addressed, namely:

- It has been shown that azimuthal fluctuations of the shower footprint are better measured with water Cherenkov detectors. It can also be measured using scintillator arrays coupled to a converter but with less discrimination power than with WCDs. It was also verified that scintillator arrays alone have no gamma/hadron discrimination power. This is an indication that quantities like $C_k$ and $LCm$ are exploring the shower calorimetry information and not the number of particles at the ground.
- The station signal converted into the Vertical Equivalent Muon (VEM) units makes the computation of $C_k$ and $LCm$ essentially insensitive to the number of photosensors in the station.
- The realistic scenario in which the array presents a higher $FF$ close to its centre and is sparser in the external regions has also been examined. In this case, the appropriate generalisation necessary to correctly handle the $C_k$ variable has been derived.

The above statements allow us to conclude that shallow WCDs equipped with as few as one PMT should be considered a valid option to deal with the high energies in the design of future observatories such as Southern Wide-field Gamma-ray Observatory (SWGO) [18].

**Acknowledgements** The work here presented has been funded by OE - Portugal, FCT, I. P., under project PTDC/FIS-PAR/4300/2020. R. C. is grateful for the financial support by OE - Portugal, FCT, I. P., under DL57/2016/cP1330/cT0002. L. G. is grateful for the financial support by FCT PhD grant PRT/BD/154192/2022 under the IDPASC program. P. J. C. wants to acknowledge the financial support by FCT PhD under grant UI/BD/153576/2022.

**Data Availability Statement** This manuscript has no associated data or the data will not be deposited. [Authors' comment: This work is solely based on simulations. Comprehensive information regarding the simulations and the tools employed, including their versions, is meticulously documented in the manuscript. Furthermore, all the simulation codes used in this study are openly accessible to the public. This ensures that anyone interested in reproducing the work will have all the necessary resources at their disposal.]



Funded by SCOAP[3]. SCOAP[3] supports the goals of the International Year of Basic Sciences for Sustainable Development.

## References

1. G. Di Sciascio, J. Phys. Conf. Ser. **1263**(1), 012003 (2019). https://doi.org/10.1088/1742-6596/1263/1/012003
2. F. Aharonian et al., Chin. Phys. C **45**(2), 025002 (2021). https://doi.org/10.1088/1674-1137/abd01b
3. R. Conceição, L. Gibilisco, M. Pimenta, B. Tomé, JCAP **10**, 086 (2022). https://doi.org/10.1088/1475-7516/2022/10/086
4. N. Arsene, (2023), ArXiv:2303.09889
5. D. Heck, J.N. Capdevielle, G. Schatz, T. Thouw, CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers, Report FZKA 6019, Forschungszentrum Karlsruhe (1998)
6. A. Ferrari, et al., CERN-2005-010, SLAC-R-773, INFN-TC-05-11 (2005)
7. T.T. Böhlen et al., Nuclear Data Sheets **120**, 211 (2014)
8. S. Ostapchenko, Phys. Rev. D **83**, 014018 (2011). https://doi.org/10.1103/PhysRevD.83.014018
9. P. Assis et al., Eur. Phys. J. C **82**(10), 899 (2022). https://doi.org/10.1140/epjc/s10052-022-10857-1
10. A. Bakalová, R. Conceição, L. Gibilisco, V. Novotný, M. Pimenta, B. Tomé, J. Vícha, Azimuthal fluctuations and number of muons at the ground in muon-depleted proton air showers at PeV energies (2023). arXiv:2304.02988 [hep-ph]
11. S. Agostinelli et al., Nucl. Instrum. Meth. A **506**, 250 (2003). https://doi.org/10.1016/S0168-9002(03)01368-8
12. J. Allison et al., IEEE Trans. Nucl. Sci. **53**, 270 (2006). https://doi.org/10.1109/TNS.2006.869826
13. J. Allison et al., Nucl. Instrum. Meth. A **835**, 186 (2016). https://doi.org/10.1016/j.nima.2016.06.125
14. A. Aab, et al., (2016). ArXiv:1604.03637
15. A. Levin, C. Moisan, A more physical approach to model the surface treatment of scintillation counters and its implementation into DETECT. in 1996 IEEE Nuclear Science Symposium. Conference Record, vol. 2, pp. 702–706 (1996). https://doi.org/10.1109/NSSMIC.1996.591410
16. A. Aab et al., JINST **15**(09), P09002 (2020). https://doi.org/10.1088/1748-0221/15/09/P09002
17. X. Bertou et al., Nucl. Instrum. Meth. A **568**, 839 (2006). https://doi.org/10.1016/j.nima.2006.07.066
18. R. Conceição, The Southern Wide-field Gamma-ray Observatory. PoS ICRC2023, 963 (2023). https://doi.org/10.22323/1.444.0963. arXiv:2309.04577