PROCEEDINGS OF SCIENCE

PoS(ICRC2023)533

Energy dependence of the optimal distance used to determine the size of air showers: implications for the energy spectrum of ultra-high-energy cosmic rays

---

O. Deligny,[a] **I. Lhenry-Yvon,**[a] **Q. Luce,**[b] **M. Roth,**[c] **D. Schmidt**[c,*] **and A.A. Watson**[d]

[a] *CNRS/IN2P3, IJCLab, Université Paris-Saclay, Orsay, France*
[b] *Institute of Experimental Particle Physics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*
[c] *Institute for Astroparticle Physics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*
[d] *School of Physics and Astronomy, University of Leeds, LS2 9JT, UK*

E-mail: `david.schmidt@kit.edu`

For modern air-shower arrays, a singular distance from the shower axis is commonly used to estimate the energies of cosmic ray primaries. The optimal distance — i.e. the distance that minimizes statistical and systematic uncertainty — is largely defined by the spacing between detectors in an array. However, an energy dependence in this optimal distance may arise for some array configurations due to a complex interplay between array geometry, the type and dynamic range of the detectors, and the form of the function used to fit the lateral distribution of signals. In such cases, the use of a single reference distance for showers of all sizes can result in significant, energy-dependent systematic and statistical uncertainties in the estimation of primary energy. These uncertainties can translate into discrepancies in the reconstructed energy spectrum of ultra-high-energy cosmic rays on the order of those observed at the Pierre Auger and Telescope Array installations. We present the full chain of analysis demonstrating the possible emergence of such discrepancies in the energy spectrum from the array properties and reconstruction effects.

38th International Cosmic Ray Conference (ICRC2023)
26 July - 3 August, 2023
Nagoya, Japan

---

[*] Speaker

https://pos.sissa.it/

1. Introduction

Surface detector arrays sample the distribution of particles from extensive air showers (EASs) arriving at the ground. A fundamental step in estimating the energy of a given cosmic ray primary involves fitting this lateral distribution. The composite measurement of many air showers allows for a reconstruction of the energy spectrum whose features provide astrophysical insight.

To reconstruct the energy of a primary particle from measurements taken by the individual detectors of a surface array, a fit is performed to the lateral distribution (i.e. the logarithmic decrease of signals measured in detectors as a function of the distance from the shower axis). The spacing of more than a kilometer between stations in the sparse arrays designed for measuring ultra-high-energy cosmic rays is many times larger than the Molière radius of air (~90 m). The resulting core uncertainty, coupled with the limitations on the dynamic range of the particle detectors, makes measurement of the shape of the lateral distribution within a distance of even a few times the Molière radius exceptionally difficult.

Nevertheless, two experiments have performed this measurement for showers with energies of up to ~10 EeV using relatively dense arrays. In the last years of the Volcano Ranch experiment (i.e. after 1972), 80 scintillator detectors were arranged on an isometric triangular grid with a spacing of 147 m (~1.5 Molière units) between detectors, and the dynamic range of each 0.8 m^2 detector was configured to measure between 0.5 and $5 \cdot 10^5$ particles per square meter [1]. The Haverah Park group performed a similar measurement with 21 water-Cherenkov detectors with a depth of 120 cm arranged on a grid with 150 m spacing [2, 3]. Both experiments demonstrated that fluctuations in the slope parameter of the logarithmic functions used to fit the lateral distributions exceeded expectations from sampling fluctuations. Even for a fixed shower size or primary energy, it could be concluded that the slope of the lateral distribution notably differs from shower-to-shower due to fluctuations in initiation and development.

These considerations posed a problem for energy estimation of individual primaries with sparse arrays, where the shape of the function used to fit the lateral distribution must be fixed for most measurements due to the low multiplicity of triggering stations and the lack of reliable measurements close to the shower axis. Historically, a line integral in distance of the lateral distribution function (LDF) was performed to estimate the total number of particles reaching the ground, which in turn served as an estimator for the energy of the primary. Use of an average slope in place of the slope appropriate for the shower in question translated into significant systematic uncertainties in the energy estimation of individual events. Hillas [4] proposed using the signal at a singular distance from the shower axis, chosen as it it fluctuated from shower-to-shower by a relatively small amount, to define the shower. At this time, the Haverah Park group, recognizing the impossibility of measuring the total number of particles in a shower, was using the energy flow between 100 and 1000 m from the shower axis to estimate the primary energy. Using a power-law lateral distribution function, Hillas found that the spread of this quantity was ~70% for a range of slopes of ±0.3. By contrast, the fluctuations in the signal at 500 m from the shower axis was only 12% for the same range of indices. The efficacy of Hillas's proposal and the dependence of the estimates of the primary energy on mass and models is further discussed in [5].

This method of using the signal at a singular distance from the shower axis to estimate primary energy has been adopted by the AGASA, the Telescope Array, and the Pierre Auger Observatory,

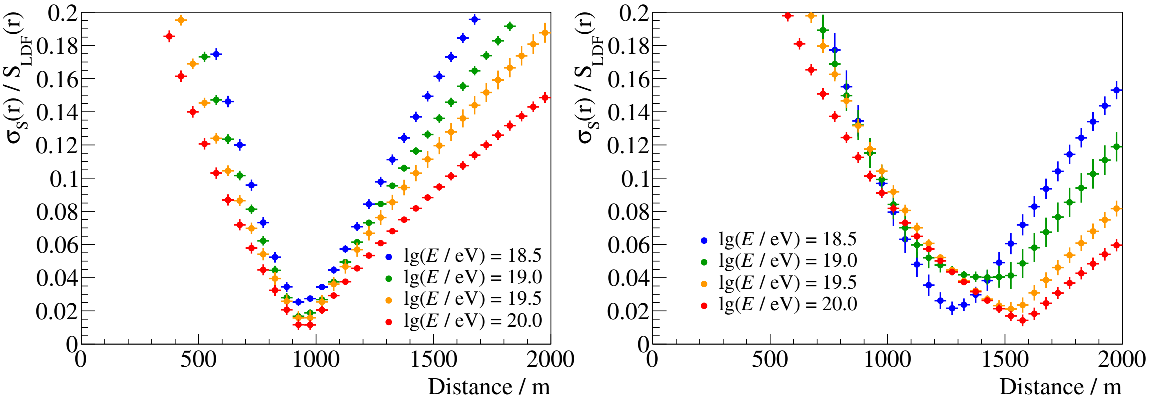

Figure 1: Relative fluctuations in the predicted signal, i.e. $\sigma_S(r)/S(r)$, as a function of distance from the shower axis. Results are shown for vertical air showers induced by proton primaries of four different energies incident on an Auger-like isometric triangular array of WCDs with 1500 m spacing. *Left*: Events without a saturated station. *Right*: Events with a saturated station. No events exhibited more than one saturated station.

all of which employ variants of an NKG-type LDF derived from [6, 7]. The logarithmic slope of the LDF used by Auger has been parameterized using events from within its own acquired data set that exhibit a high multiplicity of triggering stations [8]. The Telescope Array has adopted the functional form and slope of the lateral distribution as measured by AGASA [9], despite the difference of ~900 m in elevation and in atmospheric conditions. Investigations of which distance minimizes the impact of uncertainties on the exact shape of the LDF for a given shower have been performed for AGASA and the Auger surface detector array.

For AGASA, which consists of scintillator detectors distributed on a nominally square grid of 1000 m spacing, the optimal distance was determined to be ~600 m, with dependencies on the primary energy and the inclination of the shower[1] [10]. For an Auger-like isometric triangular grid of water-Cherenkov detectors (WCDs) with 1500 m spacing, the optimum distance was shown to be largely defined by the geometry of the array and to have a value of ~1000 m for a grid spacing of 1500 m with minimal dependence on shower inclination and energy, as long as the signal in all stations did not exceed their intended dynamic range [11]. In the case where a station saturates, the optimal distance is, on average, 300 to 500 m further from the axis and the mean optimal distance was shown to increase by ~200 m over two decades in logarithmic energy.

In the analysis documented in this note, we seek to demonstrate the intricacies of energy estimation using a singular reference distance for detectors with a limited dynamic range. We demonstrate that biases in energy may arise if care is not taken in the choice of reference distance and the slope of the LDF that is fixed in the reconstruction of individual events. We question the assumption that an energy calibration with an independent estimate of primary energy (e.g. fluorescence detectors) will fully compensate for biases that do arise, and we demonstrate that appreciable energy-dependent biases can arise in the surface detector energy estimation on an order that could explain a significant fraction of the magnitude of differences in the energy spectrum measured by the modern detectors of ultra-high-energy cosmic rays, namely the Telescope Array and the Pierre Auger Observatory.

---

[1] A second order dependence on primary mass was also observed.

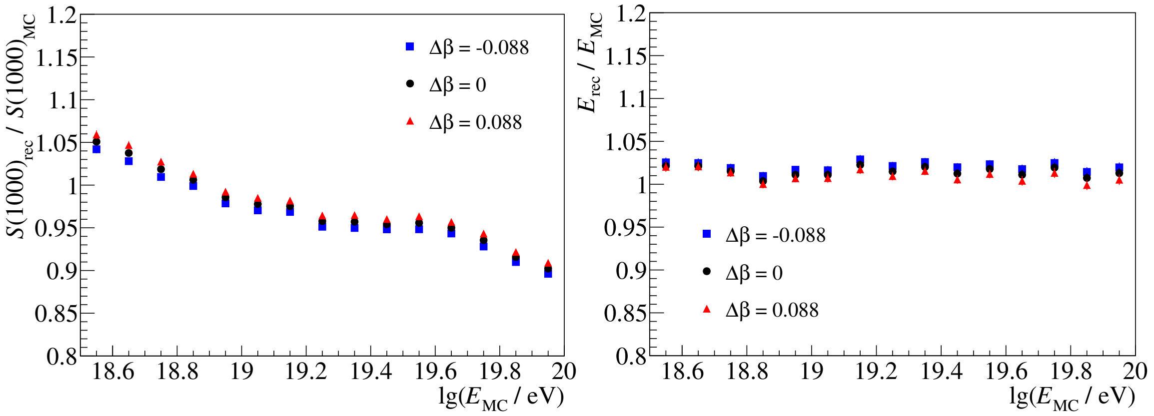

Figure 2: *Left*: Relative residuals in $S(r_{\mathrm{opt}})$ as a function of true primary energy. *Right*: Relative bias in reconstructed energy as a function of true primary energy. The case where the slope parameter β is fixed to the nominal value used by Auger is shown in addition to two cases where β is fixed to values above and below the nominal value.

2. Validation of the optimal distance

As a precursor to the studies that follow, we reproduce the results of [11] with simulations of an isometric triangular array of WCDs with 1500 m spacing. The dimensions of the active water volume match those of the WCDs of the Auger surface detector, as does the dynamic range[2]. Proton primaries at four different energies were simulated (120 per energy). The event geometry is uniquely reconstructed for each event as is the lateral distribution with the functional form

$$S_{\mathrm{Auger}}(r) = S(r_{\mathrm{opt}}) \left(\frac{r}{r_{\mathrm{opt}}}\right)^{-\beta(\theta,S(r_{\mathrm{opt}}))} \left(\frac{r + 700\ \mathrm{m}}{r_{\mathrm{opt}} + 700\ \mathrm{m}}\right)^{-\beta(\theta,S(r_{\mathrm{opt}}))}, \tag{1}$$

as described in [8]. Each event is reconstructed 100 times with a distribution of values for β drawn from a Gaussian distribution with its center and standard deviation set to the nominal values for of β measured and employed by Auger [8]. The results using the QGSJetII-04 hadronic interaction model are shown in Figure 1, where the lack of any appreciable energy dependence is clear for the case where no stations saturate. For the case where a station does saturate[3], an energy dependence is observed, and the optimal distance increases by ∼200 m over a decade of energy. The results are consistent for showers with zenith angles of 32° and 48° and for shower simulations performed with the EPOS-LHC hadronic interaction model. These results confirm the values for the optimal distance and its dependencies derived by Newton et al. in [11].

The energy dependence that exists in the average optimum distance for the given array and choice of lateral distribution is driven by the increasing fraction of events with a saturated station with energy. For the simulated detector configuration, slightly over 10% of events at 3 EeV saturate the electronics of a station compared to close to 75% at 100 EeV.

---

[2] The Offline software framework [12] was used to perform these simulations.

[3] Only in extremely rare cases for the highest energy showers do two stations saturate.

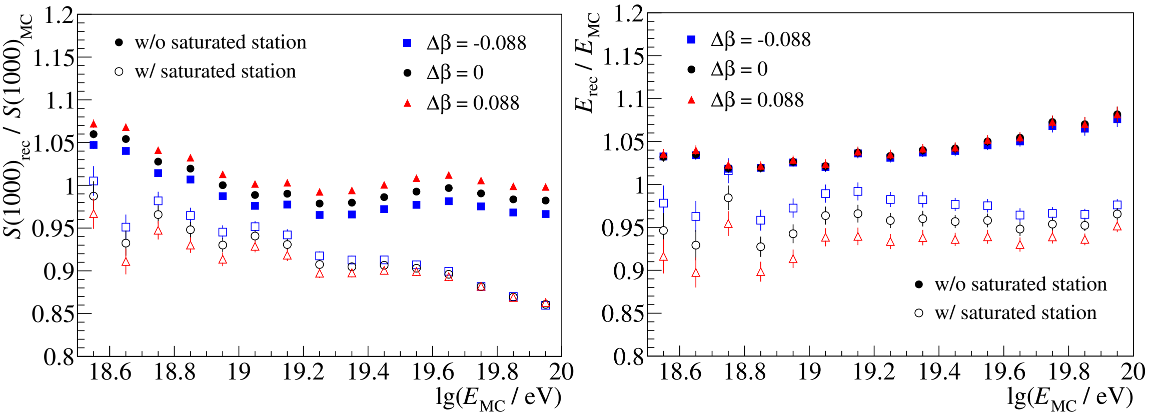

Figure 3: Relative residuals in $S(r_{\text{opt}})$ (*left*) and relative bias in reconstructed energy (*right*) as a function of true primary energy. Events with and without a saturated station are factorized. The case where the slope parameter β is fixed to the nominal value used by Auger is shown in addition to two cases where β is fixed to values above and below the nominal value.

3. Impact of a non-optimal shape

We go on to test the efficacy of using a singular, representative optimal distance to determine the shower size for events spanning the energy and zenith ranges covered by the largest observatories. The showers in these simulations libraries contain intrinsic fluctuations in the true slopes of their logarithmically falling lateral distributions as they are produced with full Monte-Carlo simulations using the QGSJetII.04 and EPOS-LHC hadronic interaction models.

Each event is fit with the function given in Equation 1. The slope is fixed to the nominal values of β parameterized by Auger as well as to values of β with positive and negative offsets of 0.088, which is approximately equal to the magnitude of the intrinsic fluctuations in β for a given shower size, as measured by [2, 3]. The idea behind this is to probe whether using an LDF slope that differs systematically from the mean slope induces significant biases. That such differences between the slope of the employed LDF and the true average LDF for a given shower size and shower geometry exist is likely given the inherent limitations in parameterizing β for such sparse arrays, as outlined in Section 1. The biases in the shower size estimator $S(r_{\text{opt}})$ (where $r_{\text{opt}} = 1000\,\text{m}$) for this exercise are shown in the left panel of Figure 2. An energy-dependent bias evolving at ~10% per decade in energy is observable for all cases.

We go on to derive and apply the remaining steps of the reconstruction to convert from the shower size estimator $S(r_{\text{opt}})$ to the energy of the shower. A measurement-driven normalization accounting for differences in the attenuation of showers at different zenith angles making use of the constant intensity cut method documented in [13] is parameterized and applied. Next, a calibration of the now zenith-independent estimates of the shower size with the true energy (as a proxy for an independent measurement of the energy that could be performed with fluorescence detectors) is performed with function linear in logarithmic energy. Determination of the parameters for this calibration closely follows the methods documented in [14]. The results are then applied to the zenith-independent shower size estimates to obtain the reconstructed energy for each event. These two steps are performed independently for each offset of β from the nominal value to emulate

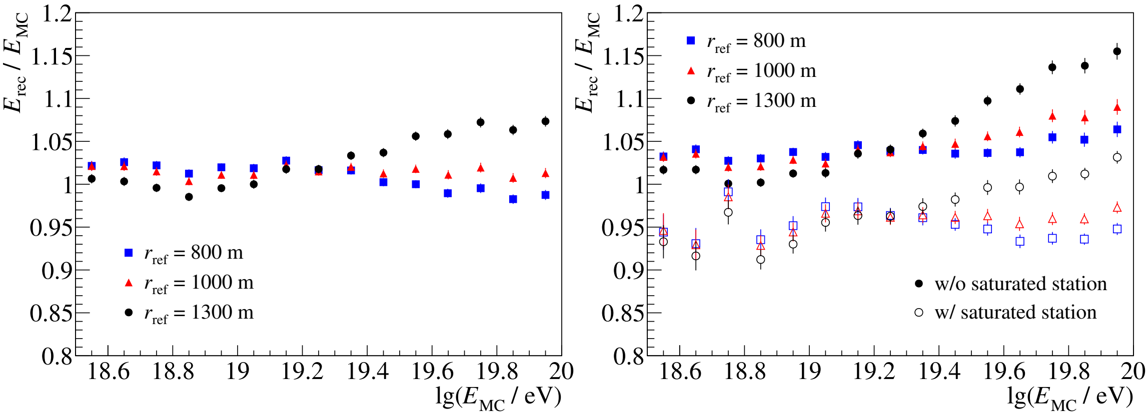

Figure 4: Relative bias in reconstructed energy as a function of the true primary energy for reconstructions of the same events performed using three different reference distances. β was fixed to the nominal value used by Auger in all cases.

the procedure applicable to measurements from end-to-end. The relative biases that remain after performing this procedure are shown in the right panel of Figure 2. The energy-dependent bias in the shower size does not translate to the reconstructed energies, where biases remain within 3% across all energies. The energy calibration successfully removes the dominant, linear component of the bias in logarithmic energy.

Inspecting events with and without a saturated station (see Figure 3), it can be seen that the underlying dynamics of the bias are more complex than those visible when simply viewing the average bias for all events. An energy-dependent bias in $S(r_{\mathrm{opt}})$ is visible at low energies for events without a saturated station but $S(r_{\mathrm{opt}})$ is essentially unbiased for such events above 10^{19} eV. For events with a saturated station, a bias distinct in magnitude and evolution with energy may be observed. That it is distinct is consistent with the understanding that the optimum distance differs both on average and in its evolution between events with and without a saturated station. By inspecting the respective biases in the reconstructed energy, the impact of calibrating away the mean bias for all events (keeping in mind that the fraction of events with a saturated station evolves with energy) can be observed for the different cases.

4. Impact of a non-optimal reference distance

Having confirmed the robust estimation of primary energy using the signal at the average optimal distance, we go on to examine what occurs when departing from this optimal distance. We use the reference distances of 800 m, 1000 m, and 1300 m. For each, the same procedure as described in Sections 2 and 3 is independently applied. The nominal value for β is used in all cases. The resulting energy biases considering all events are shown in the left panel of Figure 4. The results distinguishing between events with and without a saturated station are shown on the right. A bias in the reconstructed energy increasing by ~5% between primaries of 10^{19} eV and 10^{20} eV is visible for the reference distance of 1300 m. This demonstrates that, depending on the choice of reference distance, such energy-dependent biases may manifest whereas they do not with the choice of a more appropriate reference distance.

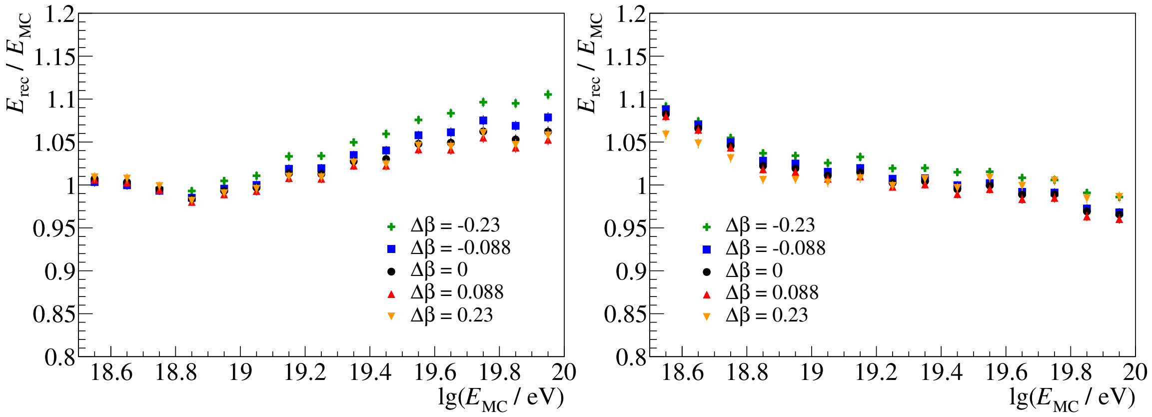

Figure 5: Relative bias in reconstructed energy as a function of the true primary energy. Reconstructions of the same events were performed fixing β to the nominal value used by Auger as well as to values above and below the nominal value. A reference distance of 1300 m was used in all cases. Results where the calibration of the shower size estimator with the true energy was performed using events with energies above $10^{18.5}$ eV and 10^{19} eV are shown on the *left* and *right*, respectively.

5. Impact of using a non-optimal shape and reference distance

We perform one further exercise to probe for sources of energy-dependent bias in reconstructed energy, namely the reconstruction of events with an LDF slope that significantly differs from the nominal value using a reference distance distant from that of the majority of events. A range of values for the slope β are tested including deviations from the nominal β parameterized by Auger of up to 0.23, which is approximately three times the magnitude of the intrinsic fluctuations measured in [2, 3]. The reference distance of 1300 m is used in all cases. After performing and applying the parameterizations to normalize for differences in attenuation independently for each β offset and performing and applying the energy calibration independently for each β offset, we calculate the average bias for all events, as shown in the left panel of Figure 5 as a function of the true primary energy. An LDF slope that differs systematically from the mean slope for a given shower size can clearly result in significant biases in the reconstructed energy. For a deviation of 0.23 from the nominal values of β, the bias in energy increases by more than 10% between primaries with energies of 10^{19} eV and 10^{20} eV. The energy range of events used to parameterize the energy calibration can not eliminate a bias that evolves non-linearly in logarithmic energy; however, it can influence where the bias manifests. The right panel of Figure 5 shows the relative bias in reconstructed energy using the same procedure as that applied to produce the previous results with the only difference being the events used to perform the energy calibration were restricted to those with energies of greater than 10^{19} eV. Whereas the relative bias in energy now remains within 5% for energies above 10^{19} eV, the bias instead manifests in the more abundant lower-energy events, which were not represented in the calibration data set. An evolution of the relative bias by ~5% within half a decade of energy can be observed for the largest offsets from the nominal value of β.

6. Conclusions

An energy-dependent bias in reconstructed energy of 10% per decade could account for a significant fraction of the discrepancies between the energy spectra measured by the Telescope Array and Pierre Auger collaborations. The studies described in this note demonstrate how such a bias can emerge when a singular reference distance that differs significantly from the optimal distance of the majority of events is employed. Coupled with the fixation of the LDF to a slope which differs significantly from the mean slope for a given shower size and inclination can result in the emergence of energy-dependent biases of this magnitude and larger.

Great caution should be exercised in adopting the LDF form and slope measured by other experiments, whose differing elevation and atmosphere could make such a choice inappropriate. Scintillators may be particularly sensitive to such a choice given their higher sensitivity to the electromagnetic component of EASs. Care should also be exercised in choosing the reference distance for energy determination. The dynamic range of detector stations plays a clear role in defining the optimal distance. If a significant fraction of events exhibit one or more saturated stations, the optimal reference distance will differ from the nominal reference distance based on the array geometry alone. It may also come with a possibly troublesome energy dependent bias.

References

[1] J. Linsley, in *Proceedings of the 15th International Conference on Cosmic Rays* **12** 89 (1977)

[2] M. Ave et al., *Astropart. Phys.* **19** 47-60 (2003), [astro-ph/0112253].

[3] R. N. Coy et a1., *Proceedings of the 17th International Conference on Cosmic Rays* **14** (1981)

[4] A.M. Hillas, Acta Phys. Acad. Sci. Hung. **29** Suppl. 3 355 (1970).

[5] A.M. Hillas, D.J. Marsden, J.D. Hollows and H.W. Hunter, *Proceedings of the 12th International Conference on Cosmic Rays*, **3** 1001 (1971).

[6] K. Kamata and J. Nishimura, *Progress of Theoretical Physics Supplement* **6** 93-155 (1958).

[7] K. Greisen, *Ann. Rev. Nucl. Part. Sci.* **10** 63-108 (1960).

[8] A. Aab et al. [Pierre Auger Coll.], *JINST* **15** P10021 (2020), [2007.09035].

[9] K. Shinozaki et al. [AGASA], Nucl. Phys. B Proc. Suppl. **136** (2004)

[10] H.Y. Dai, K. Kasahara, Y. Matsubara et al., *J. Phys. G: Nucl. Phys.* **17** 793 (1988).

[11] D. Newton, J. Knapp and A.A. Watson, *Astropart. Phys.* **26** 414-419 (2007), [astro-ph/0608118].

[12] S. Argiró, S.L.C Barroso, and J. Gonzalez et al., Nucl. Instrum. Meth. A **580** 1485 (2007), [0707.1652].

[13] A. Aab et al. [Pierre Auger Coll.], *Phys. Rev. D* **102** 6 (2020), [2008.06486].

[14] H. Dembinski et al., *Astropart. Phys.* **73** 44-51 (2016), [1503.09027].