PROCEEDINGS OF SCIENCE

Inference of the Mass Composition of Cosmic Rays with energies between 3 and 100 EeV using the data of the Pierre Auger Observatory and Deep Learning

Berenika Čermáková[a,*] **for the Pierre Auger Collaboration**[b]

[a] *Karlsruhe Institute of Technology, Institute for Astroparticle Physics, Campus North, Bldg. 401, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen*

[b] *Observatorio Pierre Auger, Av. San Martín Norte 304, 5613 Malargüe, Argentina*

Full author list: https://www.auger.org/archive/authors_2025_07.html

E-mail: berenika.cermakova@kit.edu

One of the open questions of astrophysics is the mass composition of ultra-high-energy cosmic rays (UHECRs). The flux of UHECRs is extremely low, demanding large observatories for indirect measurements of cosmic-ray air showers, cascades of secondary particles created by interactions of the cosmic ray with the atmosphere.

Located in Argentina, the Pierre Auger Observatory is the largest cosmic-ray observatory on Earth. The Observatory is a hybrid detector employing different detection principles to observe multiple components of air showers. The core part of the detector is the Surface Detector (SD), which comprises 1600 water-Cherenkov detectors with 1.5 km spacing in an area of 3000 km^2. The highly sensitive Fluorescence Detector (FD) overlooks the area above the SD. Since the FD can only operate on nights with good atmospheric conditions and low moon fraction, its duty cycle is limited to approximately 15%.

The indirect nature of measurements of the Pierre Auger Observatory poses several challenges. For example, estimating the mass of a primary cosmic ray. The atmospheric depth of the shower maximum X_{max} is a mass-sensitive observable. The FD measures the X_{max} directly, but the statistic is limited by the duty cycle.

On the contrary, the SD of the Pierre Auger Observatory, operating almost at 100% duty cycle, allows for a significant increase in the data. In this contribution, we present the X_{max} reconstruction based on deep neural networks that extends the energy range and statistics. We probe the energy evolution of the mean and standard deviation of the reconstructed X_{max}, which reflects the changes in the mass composition. The features found in the average X_{max} rate suggest a heavier and purer mass composition with increasing energy.

The European Physical Society Conference on High Energy Physics (EPS-HEP2025)
7-11 July 2025
Marseille, France

*Speaker

 ISSN 1824-8039 . Published by SISSA Medialab.

PoS(EPS-HEP2025)095

1. Introduction

Ultra-High-Energy Cosmic Rays (UHECRs) are charged particles with energies exceeding 1 EeV that originate in extraterrestrial environments. When an UHECR reaches Earth's atmosphere, it induces a particle cascade, denoted as an Extensive Air Shower (EAS). The particle cascade consists mostly of muons produced by charged pion decays. Neutral pions in the shower are responsible for the electromagnetic component. The development of the EAS components depends on the mass of the particle inducing the EAS. Therefore, information about the energy and mass of the primary particle is stochastically encoded in the EAS development. An important observable for discriminating the mass of the primary particle is the shower depth of the shower maximum X_{max}, where the number of secondary particles is the highest. On average, lighter UHECRs penetrate the atmosphere deeper than heavier UHECRs and therefore have a larger X_{max}.

Due to their scarcity, UHECRs can only be measured indirectly. A high exposure of the detector is required to obtain a significant statistic; therefore, it is necessary to compensate for the low flux by the size of the detector. The Pierre Auger Observatory [1], located in Malargüe in Argentina, is the largest operating hybrid detector specifically designed to measure different parts of EASs simultaneously. The Pierre Auger Observatory is mainly composed of two main detectors, the Surface Detector (SD) and the Fluorescence Detector (FD). The SD, consisting of a 3000 km^2 array of stations arranged in a triangular grid with 1.5 km spacing, has a duty cycle of almost 100%, and measures the signal of the secondary particles on the ground, the so-called EAS footprint. Each station consists of multiple subdetectors, the main detector being a water-Cherenkov detector (WCD). When a secondary particle crosses the detector, it emits Cherenkov light. The Cherenkov light is measured by three photomultipliers (PMTs). The shape of the time signal depends on the electromagnetic and the muonic EAS components [2, 3]. The particles in a shower front hit the WCD stations in a time sequence, depending on the inclination of the shower axis. From the signal timing at each station, the arrival direction of the primary cosmic ray can be reconstructed. The energy can be estimated using the lateral distribution of the measured signal from each WCD triggered by the EAS. The FD, consisting of 27 telescopes located at four sites, and overlooking the SD, measures the longitudinal profile of the EAS, the number of particles in the EAS during its propagation in the atmosphere, and has a duty cycle of around 15% as it can only operate during dark, moonless nights.

The FD measures the shower maximum depth X_{max} directly; however, due to its low duty cycle, the number of measurements is limited. The vast majority of events are detected with only the SD; therefore, it is essential to develop methods to reconstruct X_{max} from the shower footprint observations alone.

2. Estimation of X_{max} using deep neural networks

Deep Neural Networks (DNNs) excel at identifying non-trivial mappings between high-dimensional input data and carefully selected output data. They provide a straightforward way to test the potential of using most information contained in the shower footprint for estimating high-level, inaccessible shower observables. The DNN presented [4] employs the signal traces and SD station-level information predicting X_{max}. The triangular grid of the SD is transformed into

a Cartesian grid [5] of the size 13×13 stations, each station with 3 PMT traces. The size of the shower footprint grid guarantees a coverage of 99.9% of triggered WCDs in measurements. The first part of the DNN consists of recurrent long short-term memory (LSTM) layers, which extract features from the signal traces in each triggered station. The weights from the feature extraction are shared for all traces in the event because similar particles lead to similar detector responses. In the densely connected convolution layers [6], these trace features are combined with additional station information, e.g., about signal saturation, and arrival times of particles to each station at the footprint level. The output of the third part of residual convolutions is a single value of the $X_{\max}$. An in-depth explanation of the architecture is given in Ref. [4].

The DNN was trained with 400 000 simulated air shower events [7, 8] using Offline [9] for detector simulations. The air showers have been simulated with CORSIKA [10] air showers using EPOS-LHC [11] for modeling hadronic interactions and using the FLUKA model for modeling low-energy interactions [12]. The simulation library contains an equal mix of showers induced by protons, helium, oxygen, and iron. The events have energies ranging from 1 to 160 EeV with the spectral index $\gamma = -1$, full azimuth range, and zenith angles $\theta < 60°$.

To assess the performance of the DNN, we test a subset of events from the simulation library that were not used for training. Due to reconstruction biases, it is necessary to introduce a zenith angle- and energy-dependent fiducial event selection. At very low energies and high zenith angles, where the distance between $X_{\max}$ and the detector plane is large, the density of particles decreases. As a consequence, fewer WCD stations trigger, and therefore less signal for the DNN is to be analyzed, which decreases the precision and accuracy of the $X_{\max}$ reconstruction. The fiducial event selection is derived from the simulations of proton- and iron-induced showers, determining lower and upper limits on the zenith angle for different energies. Only the events that have zenith angles between the derived proton and iron bounds are kept for the analysis. The zenith angle dependence is similar across multiple hadronic interaction models, such as EPOS-LHC, QGSJetII-04 [13], and Sibyll2.3c [14]. Henceforth, the most conservative quality selection from Sibyll 2.3c, shown in Fig. 1, is used for this analysis.

3. Performance on simulations

After applying the quality selection, the mean and the standard deviation, $\langle X_{\max} \rangle$ and $\sigma(X_{\max})$ respectively, are compared with three composition estimates on simulated data generated from EPOS-LHC, QGSJetII-04, and Sibyll2.3c hadronic interaction models. The reconstruction shown in Fig. 2 is done for pure proton (red) and pure iron (blue) composition, as well as the Auger mix (yellow). Here, the Auger mix is the composition fractions obtained from the FD $X_{\max}$ distribution [15]. The true moments are shown as white boxes for reference. Since EPOS-LHC was used for training, no systematic bias of $\langle X_{\max} \rangle$ is observed. In contrast, both QGSJetII-04 and Sibyll2.3c reveal systematic shifts of $-5\,\mathrm{g\,cm^{-2}}$ and $-12\,\mathrm{g\,cm^{-2}}$ for all compositions. The shifts are caused by the differences between the hadronic interaction models, e.g., the different amount of muons produced during the shower cascade. The largest reconstruction bias is observed at low energies, whereas above 10 EeV all models perform similarly.

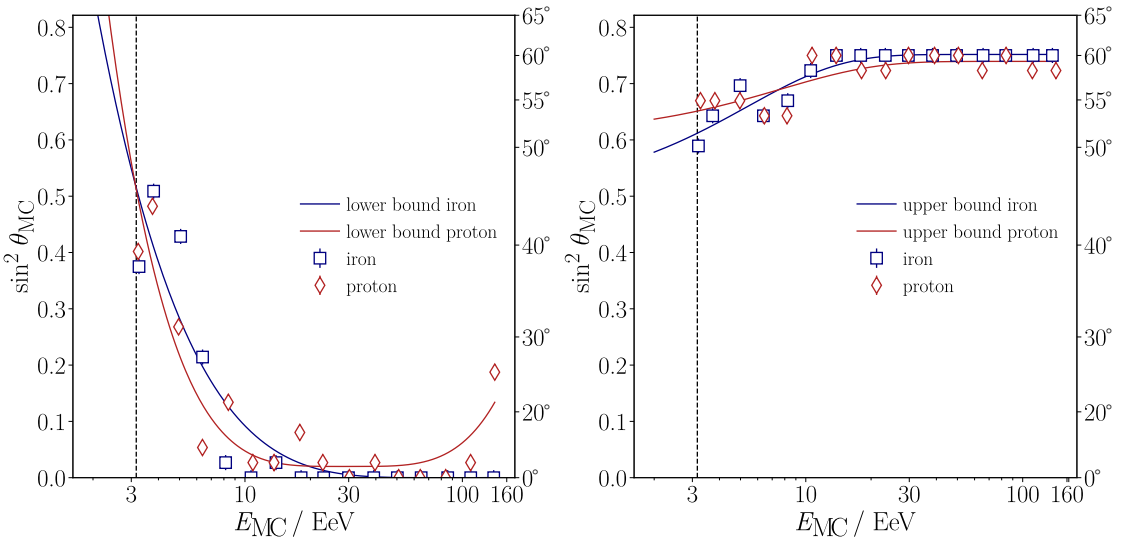


Figure 1: The fiducial event selection determined with simulations of proton- and iron-induced EAS with Sibyll2.3c. The lower (left) and upper (right) selection bounds for zenith angles are derived from pure proton and iron compositions. The markers show the minimum (left) and maximum (right) zenith angle bins for which the absolute reconstructed bias is lower than $10\,\mathrm{g\,cm^{-2}}$. The solid lines are parameterizations of this bias.

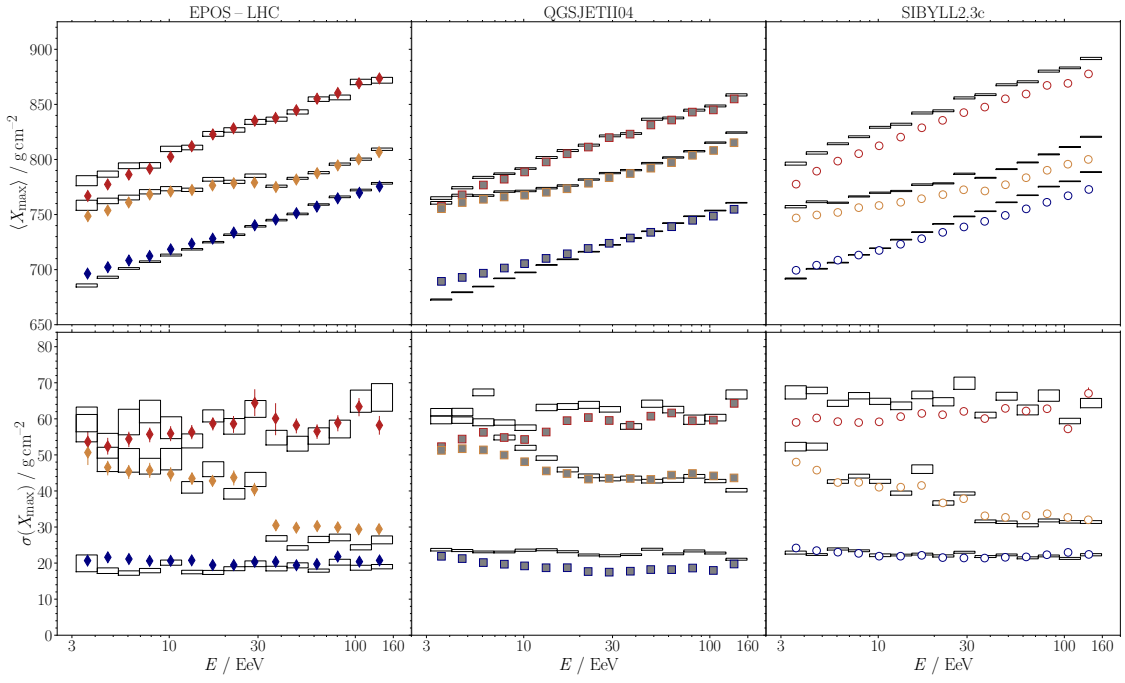
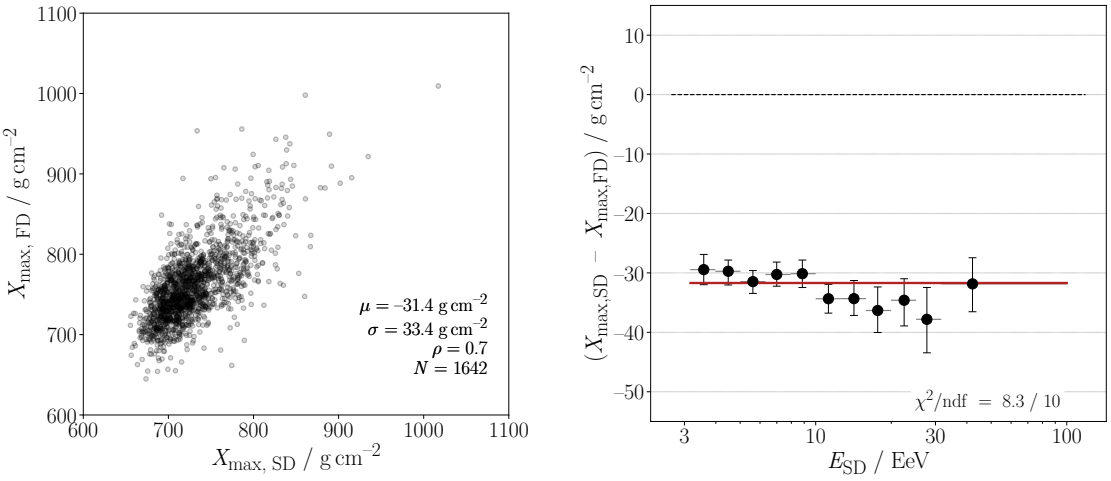


Figure 2: DNN reconstruction of the mean (top) and standard deviation (bottom) of X_{max} shown for the three hadronic interaction models: EPOS-LHC (left), QGSJetII-04 (middle), and Sibyll2.3c (right). The moments of the DNN predictions are shown for pure proton (red), pure iron (blue), and Auger mix (yellow) compositions. The white boxes show the true X_{max} moments from the Monte Carlo values.



(a) **(b)**

Figure 3: (a) Reconstruction of the hybrid data with DNN. Comparison of the reconstructed X_{max} and the FD measurements of the hybrid events on the left panel, with given bias μ, resolution σ, Pearson correlation coefficient ρ, and the number of events N. (b) The absolute bias of the reconstructed X_{max} (on the right panel) compared to the FD measurements in the hybrid data as a function of energy. The red line fits the constant negative shift originating from the difference between the data and the simulations.

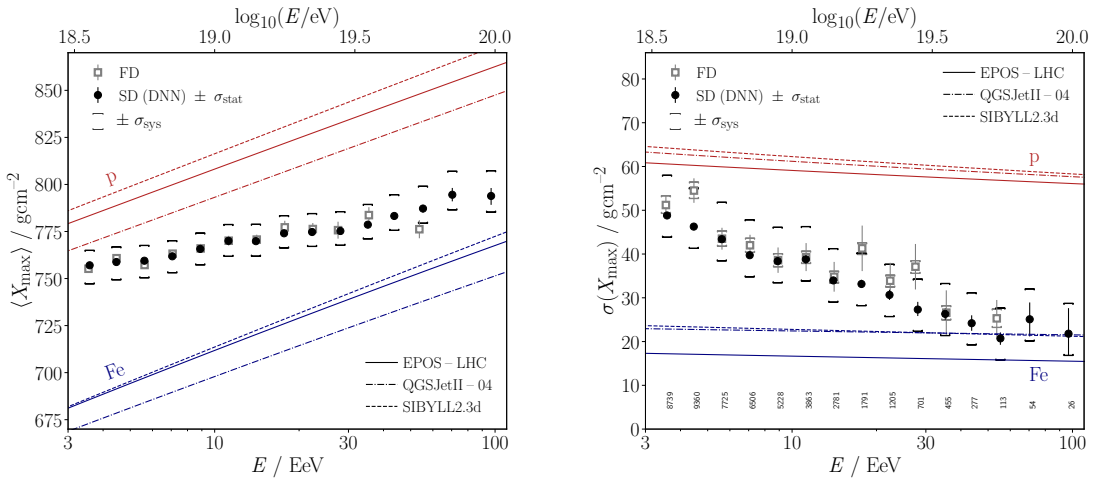
4. Calibration and performance on hybrid data

In order to use the DNN to reconstruct X_{max} from the SD measurements, a pre-selection of data is made to ensure a successful DNN reconstruction. The selection criteria are listed in Ref. [16].

The X_{max} reconstructions from the data must be corrected for effects that are not considered in simulations. The corrections are derived from non-physical dependencies on measurement data. Each correction is parameterized on the average X_{max} predictions. After parameterizing atmospheric conditions, station aging effects, altitude of the WCDs, and angular dependencies, these influences are effectively removed, resulting in a flat corrected X_{max} prediction. Details can be found in [16].

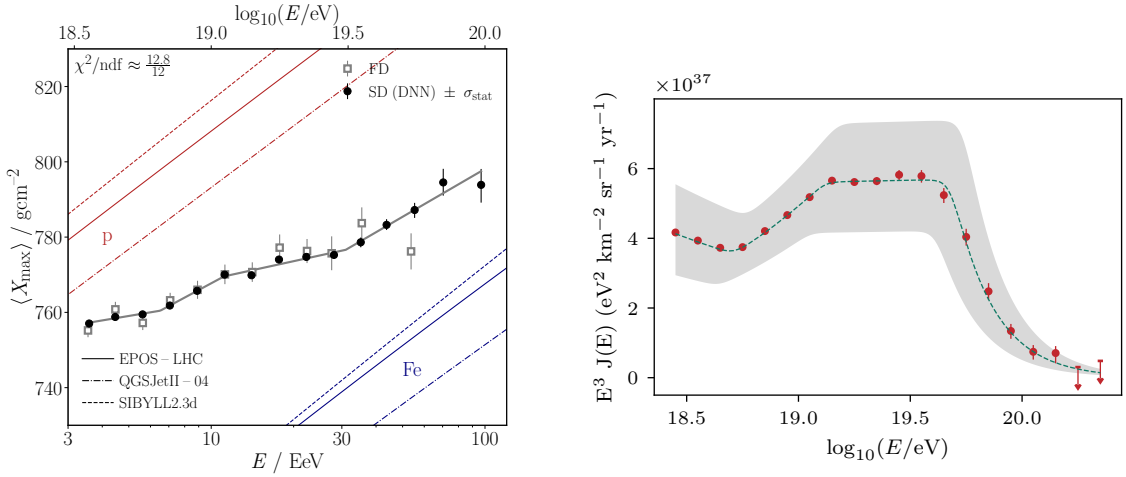
Events, simultaneously measured by SD and FD, so-called hybrid events, allow for a direct comparison between the true X_{max} and the DNN reconstruction on the event-by-event level. The Pearson correlation coefficient of the $X_{\mathrm{max}}^{\mathrm{DNN}}$ and $X_{\mathrm{max}}^{\mathrm{FD}}$ is $\rho = 0.70 \pm 0.03$ (see Fig. 3a), which is in agreement with the correlation of simulated events ($\rho_{\mathrm{MC}} = 0.73$).

Still, when comparing the FD measurements to the predictions of the NN on a event-by-event basis, they reveal an absolute bias of -31.4 ± 0.8 g cm^{-2}, which is larger than expected -15 g cm^{-2} from simulations. This bias is almost independent of the energy and can be attributed to the insufficiencies in the current generation of hadronic interaction models, the muon puzzle [17, 18]. The absolute bias is constant as a function of energy, as shown in Fig. 3b. It is calibrated out by a constant shift of each reconstructed X_{max}.



(a) **(b)**

Figure 4: Energy evolution of the first (a) and second (b) moment of $X_{\max}$. Red and blue lines denote the parametrized expectations for pure proton and iron compositions for the three hadronic interaction models.



(a) **(b)**

Figure 5: (a) The elongation rate model with three breaks fitting both the SD (black circles) and the FD (grey empty squares) data. The χ^2 represents the fit of the elongation rate model to the FD data. (b) The energy spectrum measured at the Pierre Auger Observatory with zenith angles up to 80°. The grey band shows the systematic flux uncertainty [19].

5. Results

After applying the calibration and data-driven corrections, the evolution of the mean and the standard deviation of $X_{\max}$ of the full SD dataset can be investigated. We study 48 824 events recorded by the SD which exceed a reconstructed energy of $10^{18.5}$ eV in energy bins of $\Delta \log_{10}(E/\text{EeV}) = 0.1$ with an integral bin above $10^{19.9}$ eV. The full FD dataset contains almost five times larger number of events compared to the hybrid dataset due to the different efficiencies

of the two detectors. Therefore, the SD reconstructed $X_{\max}$ values are adjusted by an additional constant shift of $-1.7\,\mathrm{g\,cm^{-2}}$ for a comparison to the full FD dataset. The mean $X_{\max}$ values of the FD data are listed in Ref. [20].

Fig. 5 depicts the energy evolution of $\langle X_{\max}\rangle$ and $\sigma(X_{\max})$ measured by FD (grey open squares) and reconstructed by DNN (black circles). The systematic uncertainties for the DNN reconstruction are derived in Ref. [16]. The agreement between the SD and FD data is exceptionally good, with only small deviations. The $\langle X_{\max}\rangle$ evolution of SD data indicates a transition from a light to a heavy composition with increasing energy.

The increase in statistics allows for a more precise fit of the elongation rate D_{10} defined as $\langle X_{\max}\rangle$ change per energy decade. Testing a fit with up to three breaks shows that the elongation rate with three breaks describes the data with the lowest $\chi^2/\mathrm{ndf} \approx 1.5$. It rejects a constant elongation rate with a significance of 4.6σ. The three-break fit is shown in Fig. 5a, the χ^2 in this figure refers to the FD data.

The breaks in the Xmax elongation rate at $E_0 = 6.5\,\mathrm{EeV}$, $E_1 = 11\,\mathrm{EeV}$, and $E_2 \approx 30\,\mathrm{EeV}$, are close to the breaks in the energy spectrum (see Fig. 5b), particularly the ankle at 5.1 EeV, the instep at 13 EeV and the suppression at 48 EeV [19].

The consequences of the three-break fit for the mass composition are left for further interpretation.

6. Conclusions

The presented DNN enabled improving mass composition studies of UHECRs, reconstructing $X_{\max}$ from the EAS footprint measured by the SD of the Pierre Auger Observatory. The DNN is trained with a simulation library with the EPOS-LHC hadronic interaction model as described in section 3. After training, the measured data are reconstructed. We find an excellent agreement of the reconstruction with the direct $X_{\max}$ measurements by the FD. The energy evolution of $\langle X_{\max}\rangle$ is described with a three-break model of the elongation rate. The $\langle X_{\max}\rangle$ evolution obtained from the SD data confirms the transition from lighter to heavier composition with increasing energy, consistent with the results obtained with FD data.

References

[1] P. A. Collaboration, et al., The Pierre Auger cosmic ray observatory, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 798 (2015) 172–213.

[2] A. Aab, P. Abreu, M. Aglietta, J. M. Albury, I. Allekotte, A. Almela, J. Alvarez-Muñiz, R. A. Batista, G. A. Anastasi, L. Anchordoqui, et al., Extraction of the muon signals recorded with the surface detector of the Pierre Auger Observatory using recurrent neural networks, Journal of instrumentation 16 (07) (2021) P07016.

[3] M. Ave, M. Roth, A. Schulz, A generalized description of the time dependent signals in extensive air shower detectors and its applications, Astroparticle Physics 88 (2017) 46–59.

[4] A. Aab, P. Abreu, M. Aglietta, J. M. Albury, I. Allekotte, A. Almela, J. Alvarez-Muñiz, R. A. Batista, G. A. Anastasi, L. Anchordoqui, et al., Deep-learning based reconstruction of the shower maximum $X_{\max}$ using the water-Cherenkov detectors of the Pierre Auger Observatory, Journal of instrumentation 16 (07) (2021) P07019.

[5] M. Erdmann, J. Glombitza, D. Walz, A deep learning-based reconstruction of cosmic ray-induced air showers, Astroparticle Physics 97 (2018) 46–53.

[6] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436–444.

[7] E. Santos on behalf of the Pierre Auger Observatory, Update on the Offline Framework for AugerPrime and production of reference simulation libraries using the VO Auger grid resources, PoS (ICRC2023) 248 (2023).

[8] E. Santos on behalf of the Pierre Auger Observatory, Monte Carlo simulations for the Pierre Auger Observatory using the VO auger grid resources, PoS (ICRC2021) 232 (2021).

[9] S. Argiro, S. Barroso, J. Gonzalez, L. Nellen, T. Paul, T. Porter, L. Prado Jr, M. Roth, R. Ulrich, D. Veberič, The offline software framework of the Pierre Auger Observatory, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 580 (3) (2007) 1485–1496.

[10] D. Heck, J. Knapp, J. Capdevielle, G. Schatz, T. Thouw, et al., CORSIKA: A Monte Carlo code to simulate extensive air showers, Report fzka 6019 (11) (1998).

[11] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, K. Werner, EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider, Physical Review C 92 (3) (2015) 034906.

[12] A. Fasso, A. Ferrari, S. Roesler, P. Sala, F. Ballarini, A. Ottolenghi, G. Battistoni, F. Cerutti, E. Gadioli, M. Garzelli, et al., The physics models of FLUKA: status and recent development, arXiv preprint hep-ph/0306267 (2003).

[13] S. Ostapchenko, QGSJET-II: towards reliable description of very high energy hadronic interactions, Nuclear Physics B-Proceedings Supplements 151 (1) (2006) 143–146.

[14] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, T. Stanev, Hadronic interaction model Sibyll 2.3 d and extensive air showers, Physical Review D 102 (6) (2020) 063002.

[15] Bellido J. on behalf of the Pierre Auger Observatory, Depth of maximum of air-shower profiles at the Pierre Auger Observatory: Measurements above $10^{17.2}$ eV and Composition Implications, PoS (ICRC2017) 506 (2017).

[16] A. Abdul Halim, P. Abreu, M. Aglietta, I. Allekotte, K. Almeida Cheminant, A. Almela, R. Aloisio, J. Alvarez-Muñiz, J. Ammerman Yebra, G. Anastasi, et al., Measurement of the depth of maximum of air-shower profiles with energies between $10^{18.5}$ and 10^{20} eV using the surface detector of the Pierre Auger Observatory and deep learning, Physical Review D 111 (2) (2025) 022003.

[17] A. Aab, P. Abreu, M. Aglietta, E.-J. Ahn, I. Al Samarai, I. Albuquerque, I. Allekotte, J. Allen, P. Allison, A. Almela, et al., Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events, Physical Review D 91 (3) (2015) 032003.

[18] A. Aab, P. Abreu, M. Aglietta, I. Al Samarai, I. Albuquerque, I. Allekotte, A. Almela, J. Alvarez Castillo, J. Alvarez-Muñiz, G. A. Anastasi, et al., Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory, Physical Review D 96 (12) (2017) 122003.

[19] D. Ravignani, Measurement of the cosmic-ray energy spectrum above 2.5 EeV using 19 years of operation of the Pierre Auger Observatory, arXiv preprint arXiv:2507.08573 (2025).

[20] A. Yushkov, P. A. Collaboration, et al., Mass Composition of Cosmic Rays with Energies above $10^{17.2}$ eV from the Hybrid Data of the Pierre Auger Observatory, in: 36th International Cosmic Ray Conference, Vol. 358, Sissa Medialab, 2021, p. 482.