Cosmic-ray mass composition with the TA SD 12-year data

Yana Zhezher\textsuperscript{a,b,\ast} on behalf of the Telescope Array Collaboration
(a complete list of authors can be found at the end of the proceedings)

\textsuperscript{a}Institute for Cosmic Ray Research, University of Tokyo, Kashiwanoha 5-1-5, Kashiwa, Chiba, Japan
\textsuperscript{b}Institute for Nuclear Research, Russian Academy of Sciences, 60-letya Oktyabrya 7a, Moscow, Russia
E-mail: zhezher.yana@physics.msu.ru

Telescope Array (TA) is the largest ultra-high-energy cosmic-ray (UHECR) observatory in the Northern Hemisphere. It is dedicated to detect extensive air showers (EAS) in hybrid mode, both by measuring the shower’s longitudinal profile with fluorescence telescopes and their particle footprint on the ground from the surface detector (SD) array. While fluorescence telescopes can measure the most composition-sensitive characteristic of EAS, the depth of the shower maximum (X\text{max}), they also have the drawback of small duty cycle. This work aims to study the UHECR composition based solely on the surface detector data. For this task, a set of composition-sensitive observables obtained from the SD data is used in a machine-learning method – the Boosted Decision Trees. We will present the results of the UHECR mass composition based on the 12-year data from the TA SD using this technique, and we will discuss of the possible systematics imposed by the hadronic interaction models.

\textsuperscript{37th} International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany

\ast Presenter

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).
https://pos.sissa.it/
1. Introduction

Telescope Array is the largest ultra-high-energy cosmic-ray experiment in the Northern hemisphere, operating since 2008 and located in Utah, USA. It observes the cosmic rays in the hybrid mode, which means that extensive air showers are registered both by a grid of surface detectors on the ground level and fluorescence telescopes, which overlook the sky above the array and assess the longitudinal development of an EAS.

The surface detector array [1] is a square grid of scintillator detectors arranged to have a separation of $1.2 \text{ km}$ between nearest-neighbors. Each detector is composed of two layers of $1.2 \text{ cm}$ thick extruded scintillator each with $3 \text{ m}^2$ effective area. Altogether the TA SD array is comprised of 507 detectors covering an area of approximately $700 \text{ km}^2$.

Mass composition is one of the characteristics of cosmic rays related to the cosmic-ray acceleration mechanisms in the source and source population as well as to the propagation of the UHECR. It is directly measurable from the extensive air showers observed at the Earth. The most widely used technique employs the fluorescence observations to derive the depth of the shower maximum, $X_{\text{max}}$, as a composition-sensitive observable. $X_{\text{max}}$-based results were previously derived by HiRes [2], Pierre Auger Observatory [3] and Telescope Array [4]. HiRes and TA results are compatible with light primary scenario resembling mostly protons up to $10^{19.8} \text{ eV}$, while the Auger results show an indication towards heavier composition for energies higher than $10^{18.3} \text{ eV}$.

The current study is an update of the previous results complemented with three more years of data from the Telescope Array surface detector. Besides augmented data set, the analysis is also performed with another hadronic interaction model, QGSJET II-04 and compared with the results from QGSJET II-03.

2. Dataset and Monte-Carlo simulations

2.1 Data events

We employ the 12-year data set from the TA surface detector, covering from 11th May 2008 up to 10th May 2020. After applying the quality cuts, the final dataset used for the analysis consists of 23159 events in the energy range from $10^{18} \text{ eV}$ to $10^{20} \text{ eV}$.

2.2 Monte-Carlo simulations

For the Monte-Carlo simulations, the CORSIKA package [5] is used along with the QGSJET II-03 and QGSJET II-04 models for high-energy hadronic interactions [6, 7], FLUKA [8, 9] for low-energy hadronic interaction and EGS4 [10] for electromagnetic processes.

To reduce the computational cost, we implement a thinning procedure [11] during the EAS modelling. With thinning, all particles with energies greater than a certain fraction of the primary energy $\epsilon_f$ are followed in detail, but below the threshold only one particle out of the secondaries produced in a certain interaction is randomly selected. This effective particle is assigned a weight to ensure energy conservation.

To restore the shower properties, we implement a consequent detheining procedure [12]. The detector response is simulated by the GEANT4 package [13]. Real-time array status and detector calibration information for 12 years of observations are used for each simulated event [14].
both QGSJET II-03 and QGSJET II-04, four Monte-Carlo sets are created: initiated by primary protons, helium, nitrogen and iron nuclei, and stored in the same data format as the TA SD data.

3. Method

3.1 Boosted Decision Trees

To discriminate between primaries, we implement the Boosted Decision Trees (BDT) [15, 16] technique. This is a method usually used in tasks where one wants to discriminate between signal and background events in a multivariate data set.

BDTs are trained using a set of composition-sensitive observables which we derive for Monte-Carlo sets: “signal”, in our case a set of events initiated by iron nuclei, and “background” one, corresponding to a proton MC set. The result of the BDT classifier is a single value $b$ for each data and Monte-Carlo event. The value of $b$ resides in the range $b \in [-1; 1]$, where $b = 1$ for a pure signal event and $b = -1$ – for a pure background event. The variable $b$ is used in the following one-dimensional analysis.

We separate proton and iron MC sets into three parts with equal statistics: the first one to build and train the classifier, the second to estimate the $\langle \ln A \rangle$ and the third to perform a bias correction. A separate classifier is constructed for each energy bin with the width of 0.2 decade. The classifier is applied to the data set as well as to the remaining proton and iron Monte-Carlo sets and to the helium and nitrogen MC sets.

3.2 Composition-sensitive observables

For each event from data and Monte-Carlo sets, we reconstruct a set of 16 composition-sensitive observables:

1. Linsley shower front curvature parameter.
2. Area-over-peak (AoP) of the signal at 1200 m and AoP slope parameter [17].
3. Number of detectors hit.
4. Number of detectors excluded from the fit of the shower front by the reconstruction procedure [18].
5. $\chi^2/d.o.f.$ of the joint geometry and LDF fit.
6. $S_p$ parameter for $b = 3$ and $b = 4.5$ [19].
7. The sum of the signals of all the detectors of the event.
8. Asymmetry of the signal at the upper and lower layers of detectors.
9. Total number of peaks within all FADC (flash analog-to-digital converter) traces.
10. Number of peaks for the detector with the largest signal.
11. Number of peaks present in the upper layer and not in the lower.
12. Number of peaks present in the lower layer and not in the upper.

13. Zenith angle of an event.


3.3 $\langle \ln A \rangle$ estimation

After compiling the $\xi$ distributions for data and MC sets, we can directly transform them to the average logarithm of atomic mass, $\langle \ln A \rangle^{(1)}$. Following the two-component approximation, the distribution of $\xi$ in the data is compared to different mixtures of the $p$ and $Fe$ Monte-Carlo events with the Kolmogorov-Smirnov test. The second part of the Monte-Carlo is used in this step and the mixtures are made with the proton fraction $\epsilon_p$ step of 2.5%. The first estimate of an average atomic mass is then given by the following equation:

$$\langle \ln A \rangle^{(1)} = \epsilon_p \times \ln (M_p) + (1 - \epsilon_p) \times \ln (M_{Fe}) ,$$

where $\epsilon_p$ is the fraction of protons in the mixture with the smallest KS-distance. The same procedure is applied to the remaining proton and iron MC sets and to the helium and nitrogen sets.

The third part Monte-Carlo sets are used to correct possible bias of the estimator given by the Eq. 1. We estimate $\langle \ln A \rangle$ for this remaining MC set and linearly approximate the result for proton and iron sets with $y_p (x)$ and $y_{Fe} (x)$, where $x = \ln E$. These lines slightly differ from the constants $\ln A = \ln(M_p)$ and $\ln A = \ln(M_{Fe})$ due to limited statistics.

The result is then corrected with the following linear function:

$$\langle \ln A \rangle = \ln (M_p) + \frac{\langle \ln A \rangle^{(1)} - y_p (x)}{y_{Fe} (x) - y_p (x)} \times (\ln (M_{Fe}) - \ln (M_p)) ,$$

where $y_p (x)$ and $y_{Fe} (x)$ are linear approximations for the MC $\langle \ln A \rangle$ distributions.

3.4 Systematic uncertainties

![Figure 1](https://via.placeholder.com/150)

**Figure 1:** $\langle \ln A \rangle$ values for proton (red), helium (green), nitrogen (purple) and iron (blue) MC sets. QGSJET II-03 (left, solid line) and QGSJET II-04 (right, dashed line) hadronic interaction models.
To validate the method based on the two-component mixture hypothesis, we employ the helium and nitrogen Monte-Carlo sets and perform the whole $\langle \ln A \rangle$ analysis procedure with them, with the results shown in Figure 1. $He$ and $N$ Monte-Carlo sets don’t perfectly fit the straight lines $\ln A = 1.38$ and $\ln A = 2.63$ correspondingly, due to simplified two-component description of $\langle \ln A \rangle$. The “shift” between $He$ and $N$ MCs and $\ln A = 1.38$ and $\ln A = 2.63$ is an estimate of the method’s systematic errors.

4. Results

The average atomic mass of primary particles in the 12 year TA data shows no significant energy dependence and yields $\langle \ln A \rangle = 1.50 \pm 0.08 (\text{stat.}) \pm 0.50 (\text{syst.})$ using QGSJET II-03 and $\langle \ln A \rangle = 0.90 \pm 0.05 (\text{stat.}) \pm 0.30 (\text{syst.})$ using QGSJET II-04.

Comparison with the results of other experiments for QGSJET II-03 is shown in Figure 2. We compare our results with the TA MD hybrid composition results [20], with the Pierre Auger Observatory $X_{\text{MAX}}^\mu$ and risetime asymmetry results [21], with the HiRes stereo [2] results and with the Yakutsk $\rho_\mu$ results [22] based on the QGSJET II-03 model.

Comparison with the results of other experiments for QGSJET II-04 is shown in Figure 3. We compare our results with the TA MD hybrid results [4] results and with the Pierre Auger Observatory SD delta results [23], all derived with QGSJET II-04 hadronic interaction model.

Acknowledgments

The Telescope Array experiment is supported by the Japan Society for the Promotion of Science (JSPS) through Grants-in-Aid for Priority Area 431, for Specially Promoted Research JP21000002, for Scientific Research (S) JP19H04006, for Specially Promoted Research JP15H05693, for Scientific Research (S) JP15H05741 and JP19H05607, for Science Research (A) JP18H03705, for Young Scientists (A) JPH26707011, and for Fostering Joint International Research (B) JP19KK0074, by the joint research program of the Institute for Cosmic Ray Research (ICRR), The University of Tokyo; by the Pioneering Program of RIKEN for the Evolution of Matter in the Universe (r-EMU); by the U.S. National Science Foundation awards PHY-1404495, PHY-1404502, PHY-1607727, PHY-1712517, PHY-1806797 and PHY-2012934; by the National Research Foundation of Korea (2017K1A4A3015188, 2020R1A2C1008230, & 2020R1A2C2102800); IISN project No. 4.4501.18, and Belgian Science Policy under IUAP VII/37 (ULB). The development and application of the machine learning analysis method is supported by the Russian Science Foundation grant No. 17-72-20291 (INR). This work was partially supported by the grants of the joint research program of the Institute for Space-Earth Environmental Research, Nagoya University and Inter-University Research Program of the Institute for Cosmic Ray Research of University of Tokyo. The foundations of Dr. Ezekiel R. and Edna Watts Dumke, Willard L. Eccles, and George S. and Dolores Doré Eccles all helped with generous donations. The State of Utah supported the project through its Economic Development Board, and the University of Utah through the Office of the Vice President for Research. The experimental site became available through the cooperation of the Utah School and Institutional Trust Lands Administration (SITLA), U.S. Bureau of Land Management (BLM), and the U.S. Air Force. We appreciate the assistance of the State of Utah and Fillmore offices of the...
Figure 2: \( \langle \ln A \rangle \) values for 12-year dataset derived with QGSJET II-03 (black crosses) compared with TA MD (upper left, green circles) [20] and TA SD 9 year results (red circles) [24], with the Pierre Auger Observatory \( X_{\text{MAX}}^\mu \) and risetime asymmetry results (upper right, blue squares and red triangles) [21], with the HiRes stereo results (lower left, orange triangles) [2] and with the Yakutsk \( \rho_\mu \) results (lower right, green squares) [22].

BLM in crafting the Plan of Development for the site. Patrick A. Shea assisted the collaboration with valuable advice and supported the collaboration’s efforts. The people and the officials of Millard County, Utah have been a source of steadfast and warm support for our work which we greatly appreciate. We are indebted to the Millard County Road Department for their efforts to maintain and clear the roads which get us to our sites. We gratefully acknowledge the contribution from the technical staffs of our home institutions. An allocation of computer time from the Center for High Performance Computing at the University of Utah is gratefully acknowledged.

References


Figure 3: \langle \ln A \rangle values for 12-year dataset derived with QGSJET II-04 (orange pentagons) compared with the TA MD hybrid (left, green circles) [4] results and with the Pierre Auger Observatory SD delta results (blue squares and red triangles) [23].

Indications of Proton-Dominated Cosmic-Ray Composition above 1.6 EeV, \textbf{104} (Apr., 2010) 161101, [0910.4184].


Full Authors List: Telescope Array Collaboration


1 Department of Physics, Loyola University Chicago, Chicago, Illinois, USA
2 The Graduate School of Science and Engineering, Saitama University, Saitama, Saitama, Japan
3 High Energy Astrophysics Institute and Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah, USA
4 Graduate School of Science, Osaka City University, Osaka, Osaka, Japan
5 Department of Physics and The Research Institute of Natural Science, Hanyang University, Seongdong-gu, Seoul, Korea
6 Department of Physics, Tokyo University of Science, Noda, Chiba, Japan
7 Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba, Japan
8 The Hakubi Center for Advanced Research and Graduate School of Science, Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto, Japan
9 Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, Kashiwa, Chiba, Japan
10 Information Engineering Graduate School of Science and Technology, Shinshu University, Nagano, Nagano, Japan
11 Faculty of Engineering, Kanagawa University, Yokohama, Kanagawa, Japan
12 Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Yamanashi, Japan
13 Academic Assembly School of Science and Technology Institute of Engineering, Shinshu University, Nagano, Nagano, Japan
14 Astrophysical Big Bang Laboratory, RIKEN, Wako, Saitama, Japan
15 Department of Physics, Sungkyunkwan University, Jang-an-gu, Suwon, Korea
16 Department of Physics, Tokyo City University, Setagaya-ku, Tokyo, Japan
17 Institute for Nuclear Research of the Russian Academy of Sciences, Moscow, Russia
18 Faculty of Systems Engineering and Science, Shibaura Institute of Technology, Minato-ku, Tokyo, Japan
TA SD mass composition

Yana Zhezher

19 Department of Engineering Science, Faculty of Engineering, Osaka Electro-Communication University, Neyagawa-shi, Osaka, Japan
20 Department of Physics, Chiba University, Chiba, Chiba, Japan
21 Service de Physique Théorique, Université Libre de Bruxelles, Brussels, Belgium
22 Department of Physics, Yonsei University, Seodaemun-gu, Seoul, Korea
23 Center for Astrophysics and Cosmology, University of Nova Gorica, Nova Gorica, Slovenia
24 Faculty of Science, Kochi University, Kochi, Kochi, Japan
25 Nambu Yoichiro Institute of Theoretical and Experimental Physics, Osaka City University, Osaka, Osaka, Japan
26 Department of Physical Sciences, Ritsumeikan University, Kusatsu, Shiga, Japan
27 Quantum ICT Advanced Development Center, National Institute for Information and Communications Technology, Koganei, Tokyo, Japan
28 Sternberg Astronomical Institute, Moscow M.V. Lomonosov State University, Moscow, Russia
29 Department of Physics, School of Natural Sciences, Ulsan National Institute of Science and Technology, UNIST-gil, Ulsan, Korea
30 Earthquake Research Institute, University of Tokyo, Bunkyo-ku, Tokyo, Japan
31 Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Hiroshima, Japan
32 Institute of Particle and Nuclear Studies, KEK, Tsukuba, Ibaraki, Japan
33 Graduate School of Science and Engineering, Tokyo Institute of Technology, Meguro, Tokyo, Japan
34 Department of Research Planning and Promotion, Quantum Medical Science Directorate, National Institutes for Quantum and Radiological Science and Technology, Chiba, Chiba, Japan
35 CEICO, Institute of Physics, Czech Academy of Sciences, Prague, Czech Republic
36 Department of Physics and Institute for the Early Universe, Ewha Womans University, Seodaemun-gu, Seoul, Korea

** Deceased