Experimental Astronomy (2025) 59:23
https://doi.org/10.1007/s10686-025-09993-9

RESEARCH

Machine learning applications to energy reconstruction of gamma-ray showers for the Tibet ASγ experiment

Y. Meng[1,2] · J. Huang[1] · D. Chen[3] · K Y. Hu[1,2] · Y. Zhang[1] · L M. Zhai[3] · Y H. Zou[1,2] · Y L. Yu[1,2] · Y Y. Li[1,2]

Received: 11 September 2024 / Accepted: 14 February 2025

Abstract

In order to improve the energy reconstruction accuracy of gamma-ray events observed by ground-based array experiments, this work propose a new energy estimator based on machine learning (ML) algorithm to determine the energies of gamma ray induced air showers in the energy range between 1 TeV and 10 PeV. We carry out a full Monte Carlo (MC) simulation using the Tibet air shower array and underground muon detector array, located at an altitude of 4,300 m above sea level. The MC simulated gamma-ray data are used to extract characteristic parameters depicting the air shower information, which are then fed into the ML model for training on both high-energy data sets ($E >\sim 10$ TeV) and low-energy data sets ($E < 10$ TeV). In our simulation data tests, we found that the ML method showed significant advantages over traditional energy estimators (S50, N_e, and $\sum \rho$), with improved energy resolution for both low and high energy datasets. Compared to the traditional estimator, the energy resolution improves by approximately 30% for the inner array events and 55% for the outer array events at $E < 10$ TeV. At around 100 TeV, the energy resolution for large zenith angle events in the outer array improves by approximately 20%. This work also found that while the energy resolution of events falling the inside array can only be slightly improved, however, events outside array and at large zenith shower clear improvements. Moreover, it is particularly noteworthy that the ML method has little difference in the energy resolution of the inner and outer array events. The enhanced energy resolution achieved through the machine learning method for outer array events reduces the limitations imposed by the observation area, resulting in an approximately 30% improvement in statistical events. This method is suitable for ground-based array experiments in gamma-ray astronomy, and provides some technical support for further study of the primary gamma-ray energy reconstruction.

Keywords High energy · Gamma ray · Cosmic ray · Extensive air shower array · Machine learning

J. Huang, D. Chen, K Y. Hu, Y. Zhang, L M. Zhai, Y H. Zou, Y L. Yu, and Y Y. Li contributed equally to this work.

Extended author information available on the last page of the article

Published online: 31 March 2025

1 Introduction

Experimental particle physics typically utilizes accelerators and detectors to study the behavior of high-energy particles, with cosmic ray observations also serving as a crucial means of investigating fundamental particles. Based on the observation methods of cosmic rays, observations can be categorized into space-based direct observations and ground-based indirect observations. Space-based experiments directly observe the properties of primary cosmic rays, such as Fermi-LAT [1], but due to payload constraints, the highest observable energy range is limited to a few hundred GeV [2, 3]. Cosmic ray flux at energies above the TeV level is low, necessitating observations to rely on ground-based experiments.

Through large-area detector arrays, it is possible to record secondary particles generated by the interactions of high-energy particles with the Earth's atmosphere. These detectors may include scintillators, Cherenkov telescopes, and water Cherenkov detectors, among others. By analyzing the energy distribution of the secondary, researchers can infer the energy of the primary cosmic rays and the direction of their arrival [4]. In recent years, significant progress has been made in gamma ray observations in the TeV to PeV energy range through ground-based experiments. For example, the Tibet ASγ experiment, using an air shower (AS) array ($\sim$ 65,700 m^2) and an underground Muon Detector (MD) array ($\sim$3,400 m^2), observed gamma rays emitted in the direction of the Crab Nebula in excess of 100 TeV [5]. This pioneering work opens a new high-energy window for exploring the extreme universe. Subsequently, HAWC [6] and LHAASO [7] have detected dozens of gamma-ray sources with $E > 100$ TeV. The detection of gamma rays above 100 TeV is a key to understanding the origin of cosmic rays in the "knee" energy region [8–11]. The origin of cosmic rays in the "knee" region has remained a mystery since the discovery in 1958 [12]. With further observations using this new window, we expect to find "PeV cosmic-ray sources" in the Milky Way that accelerate cosmic rays to PeV energies, known as "Pevatron" [8]. Therefore, precisely measuring the spectral shape of gamma rays and then distinguishing whether it is electron origin or cosmic-ray origin is the key to solving the problem. Therefore, it is important to accurately measure the energy of gamma rays.

In recent years, machine learning has revolutionized particle physics, with machine learning methods being widely applied to numerous issues in high-energy physics [13, 14]. Its applications include particle identification, energy estimation, anomaly detection, and more [15, 16]. Machine learning techniques have been applied and have made significant progress in ground-based gamma-ray and cosmic-ray experiments, as evidenced by reports from experiments such as MAGIC [17], HAWC [18], and Auger [19]. Firstly, machine learning algorithms can efficiently process and analyze vast amounts of data, enhancing the speed and efficiency of data processing. Secondly, automated data analysis and processing reduce human errors, improving the reliability of results. Furthermore, machine learning excels at identifying patterns and features within complex data, aiding in more accurate estimation of cosmic ray energy and other properties [20]. Among these methods, LightGBM gained widespread attention and application in the machine learning and data science community after its official open-source release in 2017 [21]. LightGBM is a gradient boosting framework based on decision trees, designed to provide fast and accurate model training capabilities.

LightGBM excels in handling large-scale data and complex models, widely applied across multiple domains to offer efficient and accurate solutions, making it a preferred tool in many data analysis and machine learning tasks.

In this paper, we will study the energy determination of gamma-ray showers in the 1 TeV - 10 PeV energy region based on LightGBM by Monte Carlo (MC) simulation assuming the present Tibet ASγ experiment. The MC simulation data sets are divided into two sets, namely low-energy data set and high-energy data set, which are used to reconstruct the low energy events and high energy events respectively when used for ML training. Compared with the traditional energy reconstruction method of Tibet ASγ experimental group [22], we found that the energy reconstruction of gamma events by using this machine learning method can greatly improve its energy resolution. This method shows applicability to ground-based array experiments in gamma astronomy, providing technical support for further study of the primary gamma-ray energy reconstruction.

2 Experiment

The Tibet ASγ experiment is an international joint experiment between China and Japan, which has been successfully operated at Yangbajing (90.522° E, 30.102° N; 4300 m above sea level) in Tibet, China since 1990 [23]. It has continuously made a wide field-of-view (approximately 2 steradian) observation of cosmic rays and gamma rays in the northern sky. After several upgrades, the current surface air shower array (AS) [5] consists of 597 plastic scintillation detectors with an area of 0.5 m^2, covering an area of 65,700 m^2 as shown in Fig. 1. The Tibet muon detector (MD) array (3,400 m^2 in total area) consists of 64 water-Cherenkov-type detectors located at 2.4 m underground of the AS array as shown in Fig. 1. Each MD detector is a waterproof cell filled up with water of 1.5 m in depth, 7.5 m × 7.5 m in area, equipped with a 20-inch-diameter downward-facing photomultiplier tube (PMT) on the ceiling (Fig. 1). A gamma-ray induced air shower has much less muons compared with a cosmic-

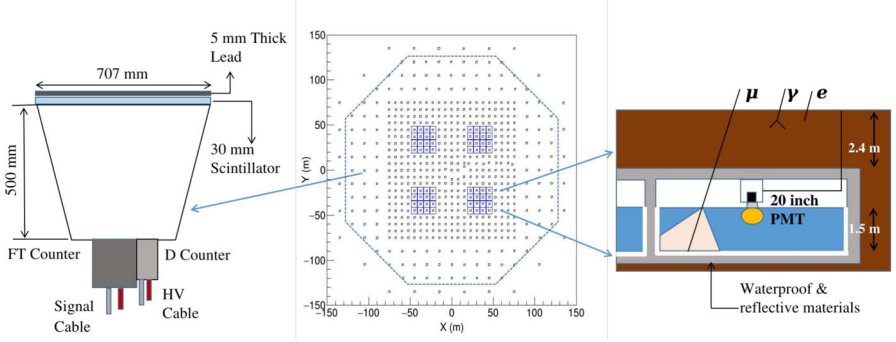

Fig. 1 Schematic diagram of the Tibet ASγ experiment array, in which the black square and blue square represent the AS and MD array respectively. In the middle figure above, the inside of the dashed octagon frame is defined as the inner array. Octagon dashed line frame outside, defined as the outer array

ray induced one. The MD array enables us to significantly discriminate a cosmic-ray background event from a gamma-ray signal by means of counting the number of muons in an air shower.

The surface air shower array (AS) is used for reconstructing the primary particle energy and direction, while the underground muon detectors are used for discriminating gamma-ray induced muon-poor air showers from cosmic-ray (proton, helium, ...) induced muon-rich air showers. For gamma rays, the Tibet ASγ experiment achieves angular resolutions of approximately 0.5° and 0.2° at 10 and 100 TeV, respectively [5].

In this work, under the detector configuration shown in Fig. 1, we analyze MC data to estimate the energy of gamma events using the machine learning method.

3 Simulation and analysis

We have carried out a full Monte Carlo simulation on the development of extensive air showers (EAS) in the atmosphere and the response in the Tibet hybrid experiment (AS + MD) array, located at an altitude of 4,300 m. The simulation code CORSIKA (ver. 76400) [24], which includes EPOS-LHC [25] and FLUKA [26] interaction model are used to generate air-shower events. All detector responses are simulated using Geant4 (version 10.06) [27]. In the CORSIKA simulations, we utilized the default atmospheric model while taking into account the altitude of the ASγ experiment (4,300 m) along with the local magnetic field conditions (34.55 μT directed north and 35.058 μT directed toward the Earth's center). Additionally, to ensure that the processing of MC events aligns with the experimental data analysis, we utilized Geant4 to input simulated air shower events into the detectors configured identically to the AS and MD array. We implemented selection criteria consistent with the experimental trigger conditions, specifically requiring that any four scintillator detectors record 1.25 or more particles to trigger an event. For gamma-ray events, the primary particle energy ranges from 300 GeV to 100 PeV with a spectral index of -2.0 and a zenith angle coverage range of 0 to 60 degrees. A total of 10^9 events were generated to cover a wide energy range.

The MC simulated data were reconstructed in the same manner as in the procedure for the experimental data analysis. Using these MC simulated data, the following characteristic parameters can be reconstructed, they can describe the cosmic-ray air shower, and subsequently used for machine learning to reconstruct the energy of the gamma-ray events. The following describes the individual feature parameters that will later be used for machine learning.

Nch - the number of detectors hit;

$\sum \rho$ - the sum of particle densities;

$(core_x, core_y)$ - the reconstructed core location of an air shower on the AS array, as shown in the following function:

$$(core_x, core_y) = \left(\frac{\sum_i \rho_i^2 x_i}{\sum_i \rho_i^2}, \frac{\sum_i \rho_i^2 y_i}{\sum_i \rho_i^2} \right), \tag{1}$$

where (x_i, y_i) and ρ_i are the coordinates of the i^{th} detector and the number density (m^{-2}) of detected particles;

θ - the zenith angle of the incoming air shower;

σ - the residual error in reconstructing the arrival direction of an event, which is indicative of the quality of direction reconstruction.

$< R >$ - the mean density weighted lateral spread, described as follows:

$$< R >= \frac{\sum_i \rho_i r_i}{n}, \tag{2}$$

where r_i and n are the distance from the i^{th} detector and the number of detectors involved in the calculation;

N_e, s - the air shower size and age [22, 28], estimated by fitting the lateral density distribution using the Nishimura-Kamata-Greisen (NKG) function [29, 30]:

$$\rho_{NKG}(r) = \frac{N_e}{r_m^2} \frac{\Gamma(4.5-s)}{2\pi\Gamma(s)\Gamma(4.5-2s)} \left(\frac{r}{r_m}\right)^{s-2} \left(1+\frac{r}{r_m}\right)^{s-4.5}, \tag{3}$$

where r represents the distance from air shower axis, r_m denotes the Moliere radius, set to be 130 m in this context, and N_e and s are the fitting parameters. The lateral distribution of a MC gamma-ray event fitted with the NKG function is shown in Fig. 2.

S50 - the particle density at 50 m from the air shower axis [22], which corresponds to the value $r = 50$ m in (3), thus S50 = $\rho_{NKG}(50\text{ m})$.

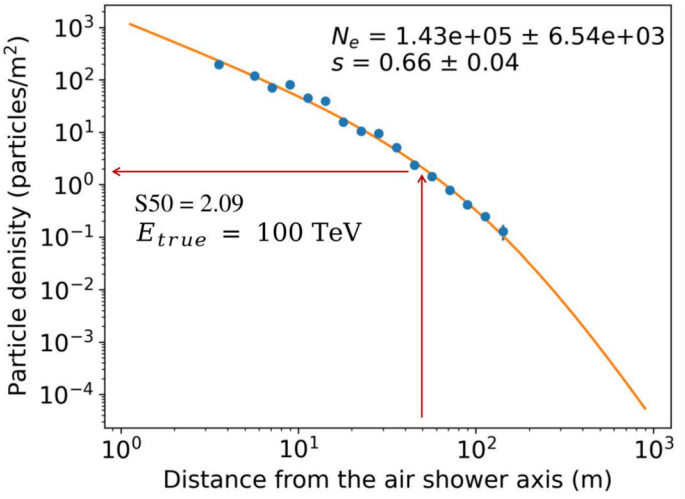

Fig. 2 The lateral distribution of a MC gamma-ray event, with an energy of 100 TeV and a zenith angle of 20°, is depicted. Solid circles represent data points, while the curves depict the fitting of the NKG function to the data, with N_e and s as the fitted parameters. S50 (ρ(50 m)) refers to the particle density at $r = 50$ m derived from completing the lateral distribution for this event

4 Method

4.1 Machine learning method

In this work, we explore the application of the machine learning method in the ASγ experiment. The Gradient Boosting Decision Tree (GBDT) [31], renowned for its effectiveness in the machine learning community, optimizes performance through iterative training of weak decision tree classifiers. LightGBM [21], an optimized implementation of GBDT, is particularly suited for large-scale data processing. It significantly enhances model training and prediction efficiency through histogram-based optimization and a leaf-wise split strategy. We incorporate this successful algorithm into the data reconstruction of the ASγ experiment, where it is employed to learn the characteristics of simulated gamma-ray events, thereby enabling the prediction of gamma-ray's energy.

To effectively implement and evaluate the ML model, we segment the dataset into training, validation, and testing sets. The training and validation sets are used to construct the model for predicting gamma-ray's energy, while the testing set is utilized to assess the model's performance. Details regarding model training can be found in Appendix A.

4.2 Features evaluation

Evaluating the importance of input features is crucial for understanding the contributions of different features to the accuracy of model predictions. To achieve this goal, we employ the Permutation Importance method, a model-agnostic technique that clearly reveals the impact of each feature within the dataset [32]. Permutation Importance is achieved by randomly shuffling individual variables in the validation dataset and observing changes in model accuracy. A significant decrease in model performance after a feature is shuffled indicates its high importance in the prediction process. The advantage of this method is that it does not rely on model-specific metrics and provides a direct, empirical measure of feature importance based on changes in model accuracy.

In Section 5, we will discuss in detail the effectiveness of machine learning in estimating gamma-ray energy, comparing it with traditional estimators.

5 Results and discussions

For the traditional method of gamma-ray energy reconstruction in Tibet ASγ experiment, if the energy of gamma ray is greater than 10 TeV, the energy of each air shower is reconstructed using the lateral distribution of ρ. As an energy estimator, we use S50 [22], which is defined as ρ at a distance of 50 m from the air shower axis in the best-fit Nishimura-Kamata-Greisen (NKG) function. On the other hand, the energy below 10 TeV was estimated directly from $\sum \rho$ corresponding to the sum of the particle density measured by each scintillation detector, as the number of hit detectors is too low to fit S50, N_e and s. In this work, to enhance the accuracy of energy estima-

tion, low-energy and high-energy data are trained separately. The selection criteria for low-energy data sets ($E < 10$ TeV) are: ($Nch \geq 8\ and\ Nch < 16$) or $\sum \rho < 100$; while the selection criteria for high-energy data sets ($E >\sim 10$ TeV) are: $Nch \geq 16$ and $\sum \rho \geq 100$. Subsequently, we present the specific performance of the trained machine learning model on the testing sets of high-energy and low-energy data. The model's performance is quantified using energy resolution as the metric.

5.1 Energy reconstruction of high-energy data sets

This work utilizes features reconstructed from MC simulated data by the AS array to train the machine learning model. For the high-energy data sets, all ten features from Section 3 are employed. And the importance scores of these features, as shown in Fig. 4a. The results indicate that S50, θ, and N_e are the most critical features for accurately predicting gamma-ray energy, with S50 also being one of the traditional energy estimators used in the ASγ experiment.

To demonstrate the performance of the model trained in this work for high-energy data energy estimation, we compare it with traditional energy estimators (S50, N_e, and $\sum \rho$) used in the ASγ experiment. Figure 3 illustrates the comparison between the energy estimates from the ML and traditional methods against simulated true primary energies. It can be observed that the distributions obtained from the ML method are narrower than those from the traditional methods, indicating that the energy reconstructed using this method is closer to the true primary gamma-ray energy. This clearly demonstrates the superiority of the machine learning approach in energy reconstruction. The result shows that the ML method, by utilizing information from multiple parameters, can predict primary gamma-ray energy more accurately. The input parameters for the ML include not only characteristics sensitive to primary air-shower energy, such as S50 and N_e, but also parameters that describe the direction and air-shower core position of the air shower, such as θ, $< R >$ and $(core_x, core_y)$. In contrast, traditional energy reconstruction methods usually rely on a single characteristic parameter [22], which puts traditional methods at a disadvantage, especially when reconstructing "far-core" location events that are relatively far from the center of the detector array, resulting in inaccurate energy reconstruction. As depicted in Fig. 3d, "bulge" data, where Fig. 4b shows the distribution of core ($\sqrt{core_x^2 + core_y^2}$), illustrates that "bulge"-data air-shower core predominantly lie far from the center of the AS array. Unlike the blue points representing data without the "bulge", these "bulge" events are generally "far-core" events, whose energy estimates are frequently underestimated by traditional estimators due to their inadequate consideration of the air-shower core's influence. Notably, this "bulge" is absent in Fig. 3a, which depicts results using the ML estimator.

Furthermore, we analyze the energy resolution of the four estimators across different energy ranges. The energy resolution is derived from Gaussian fits to the distribution of $\ln(E_{\rm rec}/E_{\rm true})$, where $E_{\rm rec}$ represents the reconstructed gamma-ray energy, and $E_{\rm true}$ represents the true primary gamma-ray energy in MC simulation. Figure 4c-h illustrate the energy resolution for both the inner and outer arrays, with the inner array delineated by the dashed lines in Fig. 1 and the outer array corresponds to the region outside these

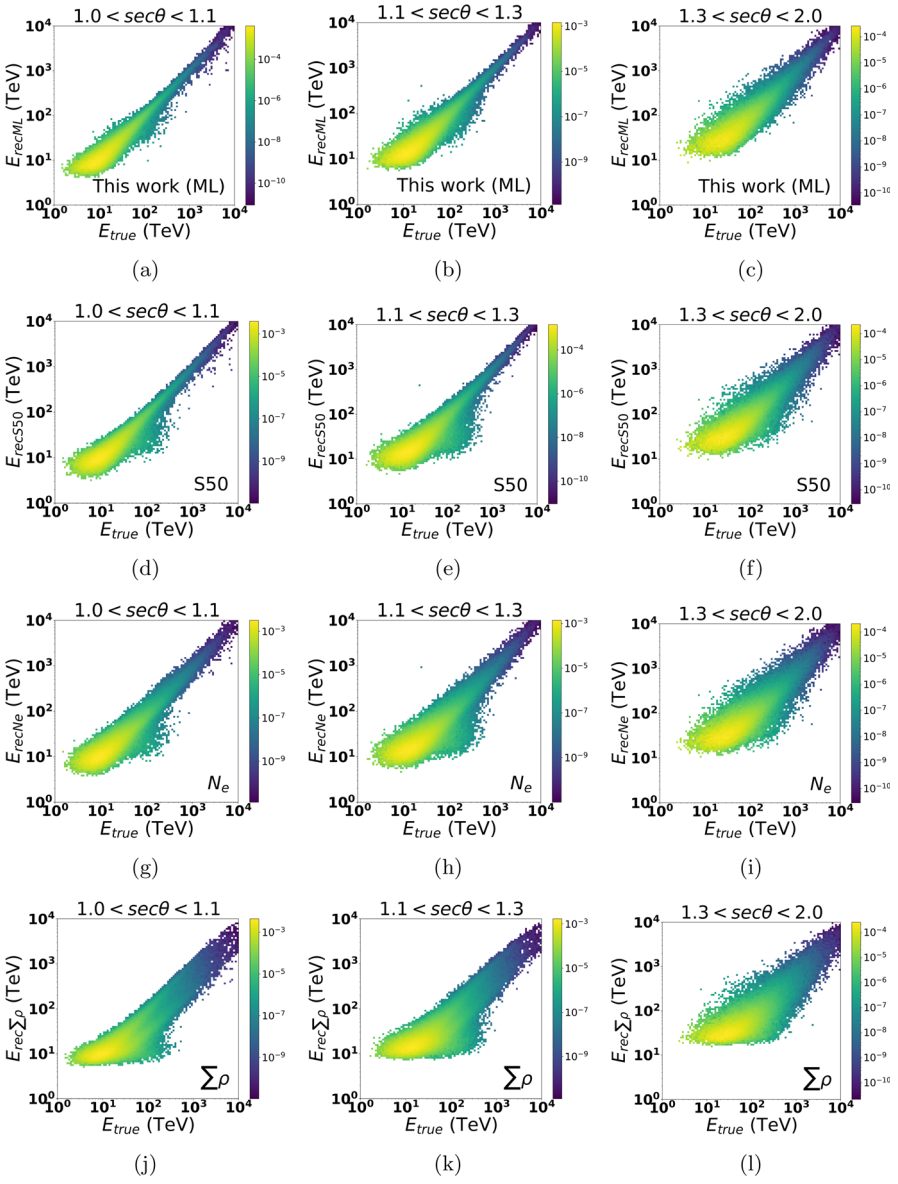

Fig. 3 For high-energy data, a comparison of the results from different energy estimators across various zenith angle intervals: (a)-(c) represent the results from this work, while (d)-(l) correspond to those from traditional methods

dashed lines. The results indicate that the energy resolution achieved through the ML method outperforms that of traditional methods. Among these traditional estimators, S50 demonstrates superior performance compared to N_e and $\sum \rho$. Consequently, S50 is selected as the representative estimator for comparisons with the ML method. At

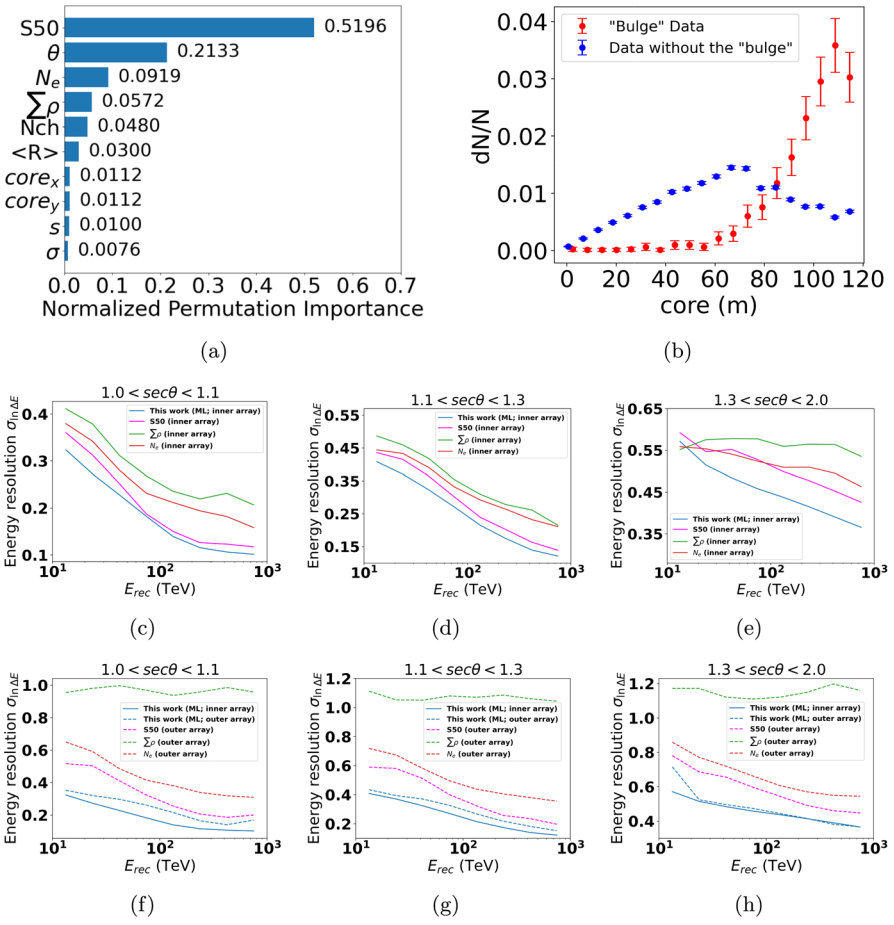

Fig. 4 For high-energy data: (a) Importance of input features. (b) Distribution of air-shower core ($\sqrt{core_x^2 + core_y^2}$) in Fig. 3d, with red points denoting "bulge" data in Fig. 3d and blue points representing data excluding the "bulge". (c)-(e) Energy resolution of different estimators for the inner array: ML (blue lines), S50 (magenta lines), N_e (red lines), and $\sum \rho$ (green lines) across zenith angle ranges of $1.0 < sec\theta < 1.1$, $1.1 < sec\theta < 1.3$, and $1.3 < sec\theta < 2.0$. (f)-(h) Energy resolution of different estimators for the outer array (solid lines for the inner array and dashed lines for the outer array) in the same zenith angle ranges

$E \sim 40$ TeV, the ML method improves the energy resolution by about 10% for inner array events and 25% for outer array events, respectively, compared to S50. In contrast, at $E \sim 100$ TeV, the enhancement in energy resolution with the ML method is minimal for small zenith angles and inner array events. However, for large zenith angle interval ($sec\theta > 1.3$), there is a notable improvement of roughly 20% for outer array events. In this work, it is found that ML method can achieve the best energy resolution for

both inner and outer array events. It is particularly worth emphasizing that ML method has little difference in energy resolution for inner and outer array events. Therefore, in this work, the statistical events can be improved by about 30% through this ML method (the events in the outer array account for about 30% of the total events). On the other hand, S50 performs nearly as well as the ML method for energies exceeding 100 TeV in the small zenith angle interval, suggesting that S50 remains a strong candidate for gamma energy estimation. Nevertheless, the ML method demonstrates more consistent performance, exhibiting advantages across different zenith angle and energy ranges. Additionally, aside from $\sum \rho$, the ML method, S50, and N_e display relatively small energy bias, generally within 0.05. A detailed discussion regarding energy bias can be found in Appendix B. We also explored other machine learning models for the energy reconstruction of gamma events, which achieved a consistent energy resolution differing within 3% in most cases. Further details can be found in Appendix C.

5.2 Energy reconstruction of low-energy data sets

For the Tibet ASγ experiment, due to the 7.5 m gap between detectors in the Tibet-III air shower array, fewer detectors were hit by cases with energies lower than 10 TeV, resulting in the inability to reconstruct the three characteristic parameters of N_e, s and S50. Therefore, in the low-energy data sets, we cannot give three characteristic parameters, N_e, s and S50, to participate in machine learning. Therefore, the energy reconstruction of low-energy data sets is different from that of high-energy data sets. In ML model training, when reconstructing low-energy cases, we exclude these three feature parameters and only train the remaining seven features. Figure 5a shows the importance scores of input feature parameters during machine learning. It can be seen that Nch, $< R >$ and θ are the most sensitive feature parameters for energy estimation in low energy data sets.

Here, due to the absence of S50 and N_e, we retain $\sum \rho$ as the traditional energy estimator for comparison with the ML estimator. Figure 5b-d present the energy resolution for gamma rays reconstructed by the ML method and $\sum \rho$ for both inner and outer array events across three zenith angle ranges for low-energy data. The energy resolution of low-energy gamma rays reconstructed by the ML method is significantly superior to that achieved by $\sum \rho$. Notably, at $E < 10$ TeV, the ML method enhances the energy resolution for inner and outer array events by approximately 30% and 55%, respectively, compared to the traditional estimator. This improvement can be attributed to the multi-input feature parameters learning utilized by the ML approach, which captures more detailed information about the air shower. As a result, the ML method is capable of predicting energy more accurately than $\sum \rho$. It is also important that the ML method exhibits slight variation in energy resolution between inner and outer array events. Therefore, this approach can increase the statistical events by around 30%, considering that outer array events account for roughly 30% of the total events.

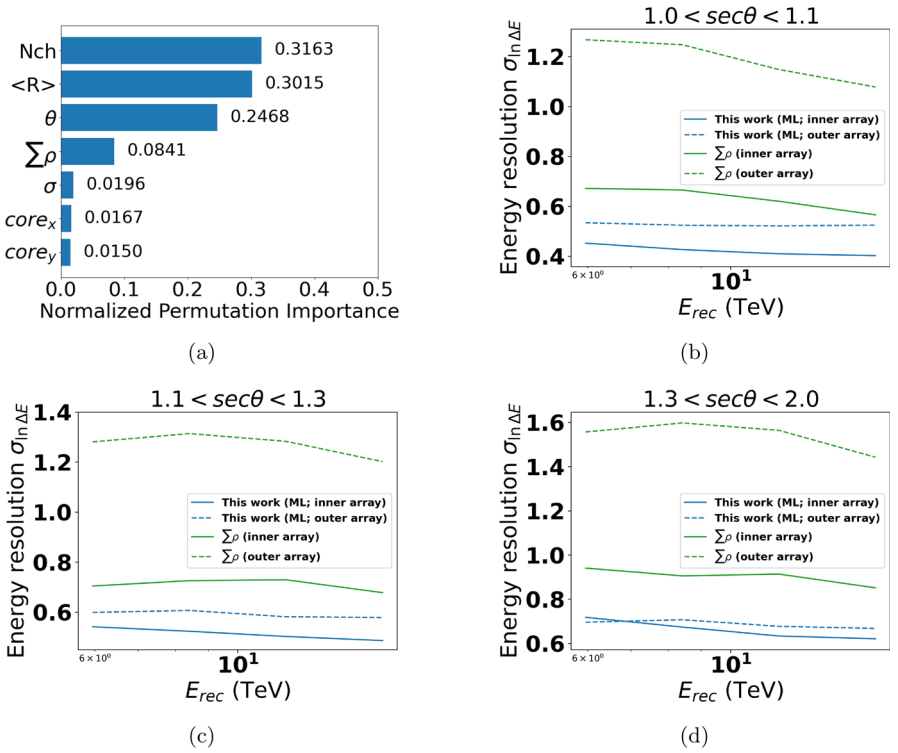

Fig. 5 For low-energy data: (a) Importance of input features. (b)-(d) Energy resolution of different estimators for the inner and outer arrays: ML (blue lines) and $\sum \rho$ (green lines) across zenith angle ranges of $1.0 < sec\theta < 1.1$, $1.1 < sec\theta < 1.3$, and $1.3 < sec\theta < 2.0$

6 Conclusion

This work employs the machine learning approach to predict the energy of gamma rays for both high-energy and low-energy data sets. Various feature parameters reconstructed from the AS array in the Tibet ASγ experiment are utilized as inputs for training the ML model. The results show that for different zenith angles, compared with traditional energy reconstruction methods such as S50, N_e and $\sum \rho$, the machine learning method consistently has better energy resolution than the traditional reconstruction of gamma-ray events. This improvement is particularly pronounced in large zenith angle ranges, the outer array region and low-energy data. For high-energy data, at approximately 40 TeV, the energy resolution for events in the outer array using the ML method exceeds that of the best-performing traditional estimator S50 by about 25%. At around 100 TeV, while the energy resolution obtained by the ML is nearly identical to that of S50 at small zenith angles, it still achieves an improvement of about 20% at large zenith angles in the outer array. On the other hand, for low-energy data ($E < 10$ TeV), the ML method enhances the energy resolution for inner and outer array events by approximately 30% and 55%, respectively, compared to the traditional method $\sum \rho$. Therefore, we find that although the energy resolution of events falling

inside the array is only slightly improved, however, events outside array and at large zenith shower clear improvements. Moreover, it is particularly noteworthy that the ML method has little difference in the energy resolution of the inner and outer array events. Therefore, the improved energy resolution achieved by ML method for outer array events allows for the retention of these events, resulting in an increase of approximately 30% in the statistical sample. In addition, we also tested different machine learning models and found that if the energy is reconstructed using different machine learning models under the same conditions, the energy resolution remains consistent within a 3% error range (see Appendix C). In the near future, this effective machine learning-based data reconstruction method will be applied to reconstruct additional features in the Tibet ASγ experiment, such as the direction and composition of primary particles, with the aim of further improving the overall accuracy and efficiency of the experiment. Of course, this ML method is also suitable for other ground-based array experimental groups to reconstruct the energy of air shower.

Appendix A: Machine learning model training

In this work, we utilized the LightGBM, a gradient boosting decision tree-based algorithm, to reconstruct the energy of gamma-ray events using machine learning models. Below, we list the main hyperparameter settings used in the trained models and their respective roles:

1. objective : regression

The task is framed as a regression problem to minimize the difference between predicted and actual values, thereby enhancing the reconstruction of gamma-ray energy.

2. num_leaves : 50

This parameter defines the maximum number of leaves in one tree. More leaves can capture more complex patterns but can also lead to overfitting.

3. min_data_in_leaf : 600

This parameter sets the minimum number of samples that must be present in a leaf node. It is a important parameter for preventing overfitting in the leaf nodes of the tree.

4. learning_rate : 0.1

This parameter controls how much to update the model with each iteration.

5. bagging_fraction : 0.7

This parameter controls the fraction of data to be used for each iteration of training. It helps to prevent overfitting by randomly selecting a subset of data for training.

6. feature_fraction : 0.7

This parameter determines the fraction of features to be randomly selected for each iteration. Similar to bagging_fraction, it helps to reduce overfitting by using only a subset of features.

7. bagging_freq : 7

This parameter controls how often to perform bagging. It can also prevent overfitting.

Appendix B: The bias of gamma-ray energy reconstruction

The performance evaluation of energy estimators relies not only on energy resolution but also on bias. Bias is defined as the average difference between the reconstructed energy and the true energy in the logarithmic space. Figure 6 illustrates the bias in gamma-ray energy reconstruction between the machine learning (ML) method and traditional energy estimators. The results indicate that, apart from the $\sum \rho$ estimator, which exhibits a large bias across all energy and zenith angle ranges, other estimators maintain a bias within 0.05 (with a bias of within 0.15 for energies near 10 TeV when $sec\theta > 1.3$). Furthermore, as shown in Fig. 6, the ML method consistently demonstrates minimal bias across all energy and zenith angle ranges in gamma-ray energy reconstruction. This further highlights the advantages of the ML approach in gamma-ray energy estimation.

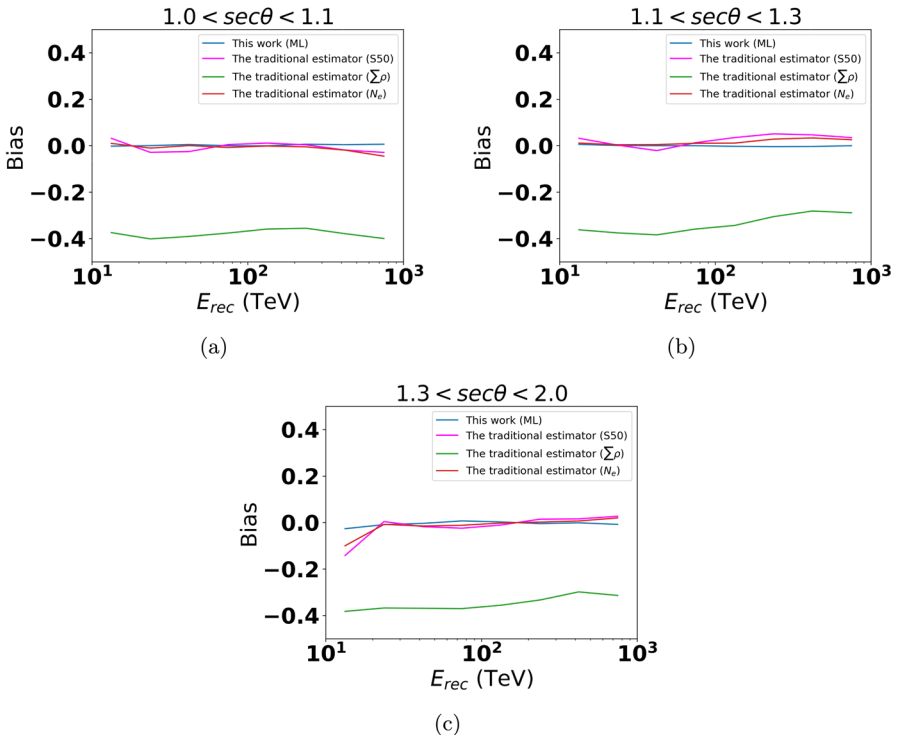

Fig. 6 The energy bias of different energy estimators within the zenith angle ranges of (b) $1.0 < sec\theta < 1.1$, (c) $1.1 < sec\theta < 1.3$, and (d) $1.3 < sec\theta < 2.0$, respectively

Appendix C: Test the dependence of reconstruction energy on different machine learning models

In our investigation of gamma-ray energy reconstruction, we explored other alternative machine learning techniques beyond the primary model discussed in the main text. Notably, we evaluated the performance of Multi-Layer Perceptrons (MLP) neural networks, and tree-based XGBoost. We examined the energy resolutions achieved by these different models across various zenith angle intervals and energy ranges. As illustrated in Fig. 7, all three methods exhibit a high degree of consistency, with energy resolutions differing within 3% error range. However, in the case of the large zenith angle range ($1.3 < sec\theta < 2.0$) and an energy of approximately 10 TeV, the observed difference increases to 9%. This variation is deemed acceptable given the large zenith angle involved.

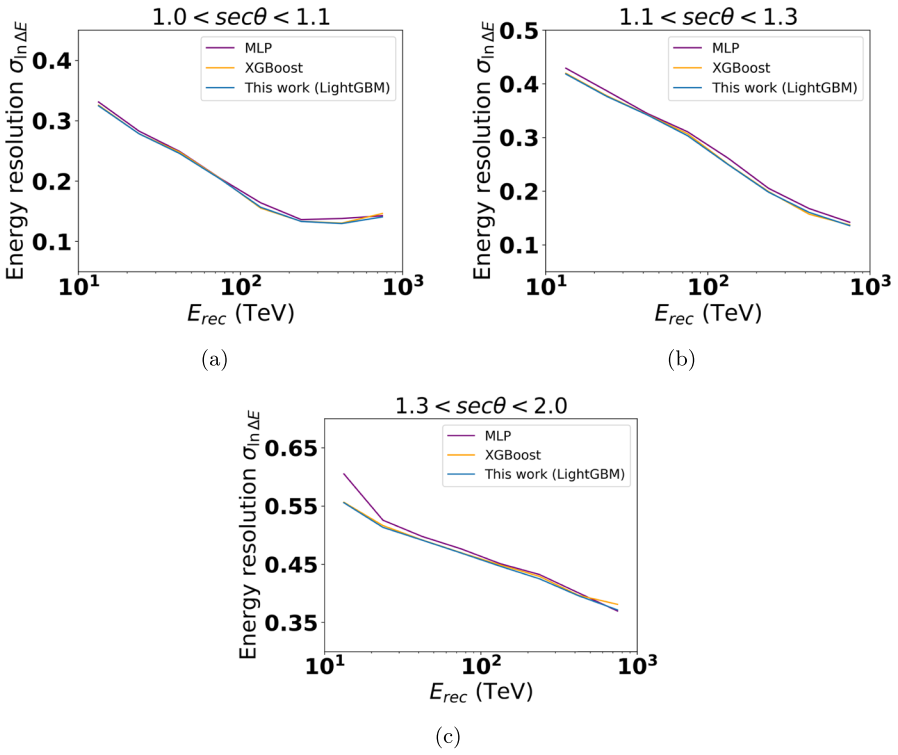

Fig. 7 The energy resolution of different machine learning models within the zenith angle ranges of (b) $1.0 < sec\theta < 1.1$, (c) $1.1 < sec\theta < 1.3$, and (d) $1.3 < sec\theta < 2.0$, respectively

Given these findings, we opted to streamline our discussion in the main text by focusing on the LightGBM method. We selected LightGBM as a representative example of machine learning applications in the context of gamma-ray energy reconstruction, as it demonstrated stable performance.

Acknowledgements The authors would like to express their thanks to the members of the Tibet ASγ collaboration for fruitful discussions. This work is supported by the National Natural Science Foundation of China (grant nos. 12227804,12275282,12073050) and by the Key Laboratory of Particle Astrophysics at the Institute of High Energy Physics of the Chinese Academy of Science.

Author Contributions All authors contributed to the study conception and design. The first draft of the manuscript was written by Y. Meng and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work is supported by the National Natural Science Foundation of China (grant nos. 12227804,12275282,12073050) and by the Key Laboratory of Particle Astrophysics at the Institute of High Energy Physics of the Chinese Academy of Science.

Data Availability Data sets generated during the current study are available from the corresponding author on reasonable request.

Code availability The codes utilized in this work are integrated within the analysis framework of the Tibet AS + MD array, making it impractical to extract them. As a result, the codes are not publicly accessible.

Declarations

Conflicts of interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval Not applicable

Consent for publication Not applicable

References

1. Atwood, W., et al.: The large area telescope on the fermi gamma-ray space telescope mission. Astrophys. J. **697**(2), 1071 (2009). https://doi.org/10.1088/0004-637X/697/2/1071
2. Schlickeiser, R.: Direct Observations of Cosmic Rays, pp. 25–71. Springer, Berlin, Heidelberg (2002). https://doi.org/10.1007/978-3-662-04814-6_3
3. Zombeck, M.V.: Handbook of Space Astronomy and Astrophysics. Cambridge University Press, England (2007)

4. Amenomori, M., et al.: The all-particle spectrum of primary cosmic rays in the wide energy range from 10^{14} to 10^{17} eV observed with the Tibet-III air-shower array. Astrophys. J. **678**(2), 1165–1179 (2008). https://doi.org/10.1086/529514
5. Amenomori, M., et al.: First detection of photons with energy beyond 100 TeV from an astrophysical source. Phys. Rev. Lett. **123**(5), 051101 (2019). https://doi.org/10.1103/PhysRevLett.123.051101. arXiv:1906.05521 [astro-ph.HE]
6. HAWC Collaboration: 3HWC: The third HAWC catalog of very-high-energy gamma-ray sources. Astrophys. J. **905**(1), 76 (2020). https://doi.org/10.3847/1538-4357/abc2d8. arXiv:2007.08582 [astro-ph.HE]
7. Cao, Z., et al.: The first LHAASO catalog of gamma-ray sources. Astrophys. J. Suppl. Ser. **271**(1), 25 (2024). https://doi.org/10.3847/1538-4365/acfd29. arXiv:2305.17030 [astro-ph.HE]
8. Amenomori, M., et al.: First detection of sub-PeV diffuse gamma rays from the galactic disk: evidence for ubiquitous galactic cosmic rays beyond PeV energies. Phys. Rev. Lett. **126**(14), 141101 (2021). https://doi.org/10.1103/PhysRevLett.126.141101. arXiv:2104.05181 [astro-ph.HE]
9. Amenomori, M., et al.: Potential PeVatron supernova remnant G106.3+2.7 seen in the highest-energy gamma rays. Nature Astron. **5**(5), 460–464 (2021). https://doi.org/10.1038/s41550-020-01294-9. arXiv:2109.02898 [astro-ph.HE]
10. Albert, A., et al.: HAWC J2227+610 and its association with G106.3+2.7, a new potential Galactic PeVatron. Astrophys. J. Lett. **896**, 29 (2020). https://doi.org/10.3847/2041-8213/ab96cc. arXiv:2005.13699 [astro-ph.HE]
11. Cao, Z., et al.: Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray Galactic sources. Nature **594**(7861), 33–36 (2021). https://doi.org/10.1038/s41586-021-03498-z
12. Kulikov, G., Khristiansen, G.: On the size spectrum of extensive air showers. Sov. Phys. JETP **35**(8), 441–444 (1959)
13. Icecube Collaboration: Observation of high-energy neutrinos from the Galactic plane. Science **380**(6652), 1338–1343 (2023). https://doi.org/10.1126/science.adc9818. arXiv:2307.04427 [astro-ph.HE]
14. The Pierre Auger Collaboration: inference of the mass composition of cosmic rays with energies from $\mathbf{10^{18.5}}$ to $\mathbf{10^{20}}$ eV using the pierre auger observatory and deep learning. arXiv e-prints, pp. 2406–06315 (2024). https://doi.org/10.48550/arXiv.2406.06315. arXiv:2406.06315 [astro-ph.HE]
15. Bourilkov, D.: Machine and deep learning applications in particle physics. Int. J. Mod. Phys. A **34**(35), 1930019 (2019)
16. Feickert, M., Nachman, B.: A living review of machine learning for particle physics. arXiv preprint (2021). https://doi.org/10.48550/arXiv.2102.02770. arXiv:2102.02770
17. Albert, J., et al.: Implementation of the random forest method for the imaging atmospheric cherenkov telescope magic. Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerators Spectrometers Detectors Assoc. Equip. **588**(3), 424–432 (2008). https://doi.org/10.1016/j.nima.2007.11.068
18. Alfaro, R., et al.: Gamma/hadron separation with the hawc observatory. Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerators Spectrometers Detectors Assoc. Equip. **1039**, 166984 (2022). https://doi.org/10.1016/j.nima.2022.166984
19. Pierre Auger Collaboration: Extraction of the muon signals recorded with the surface detector of the Pierre Auger Observatory using recurrent neural networks. J. Instrum. **16**(7), 07016 (2021). https://doi.org/10.1088/1748-0221/16/07/P07016
20. Erdmann, M., Glombitza, J., Walz, D.: A deep learning-based reconstruction of cosmic ray-induced air showers. Astropart. Phys. **97**, 46–53 (2018). https://doi.org/10.1016/j.astropartphys.2017.10.006
21. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y.: Lightgbm: a highly efficient gradient boosting decision tree. Advances in neural information processing systems, vol. 30 (2017)
22. Kawata, K., Sako, T., Ohnishi, M., Takita, M., Nakamura, Y., Munakata, K.: Energy determination of gamma-ray induced air showers observed by an extensive air shower array. Exp. Astron. **44**, 1–9 (2017). https://doi.org/10.1007/s10686-017-9530-9
23. Amenomori, M., et al.: Search for steady emission of 10-TeV gamma rays from the Crab Nebula, Cygnus X-3, and Hercules X-1 using the Tibet air shower array. Phys. Rev. Lett. **69**(17), 2468 (1992). https://doi.org/10.1103/PhysRevLett.69.2468
24. Heck, D., Knapp, J., Capdevielle, J.N., Schatz, G., Thouw, T.: CORSIKA: A Monte Carlo code to simulate extensive air showers (1998)

25. Pierog, T., Karpenko, I., Katzy, J.M., Yatsenko, E., Werner, K.: Epos lhc: test of collective hadronization with data measured at the cern large hadron collider. Phys. Rev. C **92**, 034906 (2015). https://doi.org/10.1103/PhysRevC.92.034906
26. Battistoni, G., et al.: Overview of the fluka code. Annals Nucl. Energy **82**, 10–18 (2015). https://doi.org/10.1016/j.anucene.2014.11.007
27. Agostinelli, S., et al.: GEANT4-a simulation toolkit. Nucl. Instrum. Meth. A **506**, 250–303 (2003). https://doi.org/10.1016/S0168-9002(03)01368-8
28. Lipari, P.: Concepts of "age" and "universality" in cosmic ray showers. Phys. Rev. D **79**(6) (2009). https://doi.org/10.1103/physrevd.79.063001
29. Kamata, K., Nishimura, J.: The lateral and the angular structure functions of electron showers. Prog. Theor. Phys. Suppl. **6**, 93–155 (1958). https://doi.org/10.1143/PTPS.6.93
30. Greisen, K.: Cosmic ray showers. Ann. Rev. Nuclear Sci. **10** (1960). https://doi.org/10.1146/annurev.ns.10.120160.000431
31. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232 (2001)
32. Altmann, A., Toloşi, L., Sander, O., Lengauer, T.: Permutation importance: a corrected feature importance measure. Bioinformatics **26**(10), 1340–1347 (2010). https://doi.org/10.1093/bioinformatics/btq134

Authors and Affiliations

Y. Meng[1,2] · J. Huang[1] · D. Chen[3] · K Y. Hu[1,2] · Y. Zhang[1] · L M. Zhai[3] · Y H. Zou[1,2] · Y L. Yu[1,2] · Y Y. Li[1,2]

✉ Y. Meng
mengyu@ihep.ac.cn

1 Key Laboratory of Particle Astrophysics, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, People's Republic of China

2 University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, People's Republic of China

3 National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, People's Republic of China