PROCEEDINGS OF SCIENCE

Mass Composition of Primary Cosmic Rays with GRAPES-3 Using Machine Learning Techniques

Subhalaxmi Rout[a,*], **Pravata Kumar Mohanty**[b], **Aruna Kumar Nayak**[a], and **Fahim Varsi**[c]

[a] *Institute of Physics,*
Sachivalaya Marg, Bhubaneswar 751005, India
[a] *Homi Bhabha National Institute,*
Training School Complex, Anushakti Nagar, Mumbai 400094, India
[b] *Tata Institute of Fundamental Research,*
Dr Homi Bhabha Road, Mumbai 400005, India
[c] *Karlsruhe Institute of Technology,*
Institute of Experimental Particle Physics , D-76021 Karlsruhe, Germany

E-mail: subhalaxmi.r@iopb.res.in

Precise measurements of the nuclear composition and energy spectrum of primary cosmic rays around the knee are essential to understand their origin, acceleration, and propagation. The GRAPES-3 experiment in Ooty, India, recently reported a spectral hardening in the proton spectrum at ~166 TeV using Gold's unfolding method based on muon multiplicity distributions. To enhance composition sensitivity by incorporating additional observables, we implement a Deep Neural Network (DNN) using both muon multiplicity and shower age, along with other high-level reconstructed shower parameters. This work presents the strategy, performance, and reliability of the DNN-based approach for mass composition studies at GRAPES-3.

39th International Cosmic Ray Conference (ICRC2025)
15–24 July 2025
Geneva, Switzerland

*Speaker

PoS(ICRC2025)374

1. Introduction

The GRAPES-3 (Gamma Ray Astronomy at PeV Energies – Phase-3) experiment [1] is located in Ooty, India (11.4°N, 76.7°E, 2200 m a.s.l.), and is designed to investigate extensive air showers (EAS) generated by cosmic rays (CRs) in the energy range from a few TeV to about 10 PeV. It consists of a dense array of 400 plastic scintillator detectors (SD), each with an area of 1 m^2, laid out in a hexagonal pattern with 8 m spacing. This compact layout ensures uniform lateral sampling of shower particles and improves the accuracy of core and direction reconstruction. Complementing the surface array is a 560 m^2 tracking muon detector composed of 3712 proportional counters (PRCs), arranged across four stations. Each station includes four modules, and each module consists of four orthogonal layers with 58 PRCs per layer. These detectors enable precise measurement of muon trajectories, and their energy threshold varies with the zenith angle, starting from ∼1 GeV for vertical muons. The GRAPES-3 array employs a two-level triggering system. The Level-0 trigger requires a 3-line coincidence within 100 ns, and Level-1 requires at least 10 detectors to fire within a 1 μs window. The fiducial area defined for high-quality event selection is about 14,560 m^2. Located at high altitude with an atmospheric overburden of ∼800 g cm^{-2} and equipped with a dense array and muon tracking capabilities, GRAPES-3 is well-suited to study cosmic-ray energy spectra and primary composition. In particular, the muon multiplicity distribution provides a composition-sensitive handle to probe nuclear masses of incoming cosmic-ray primaries [2].

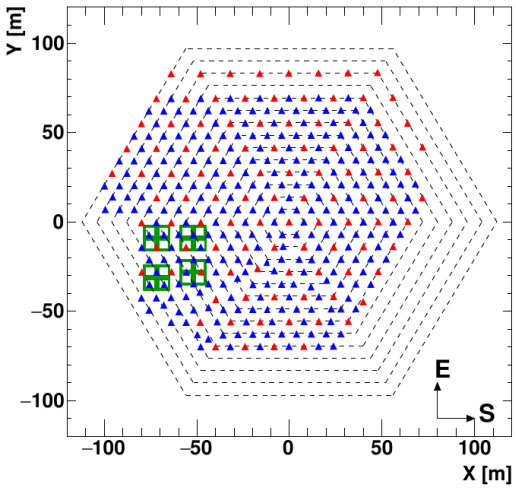

Figure 1: Schematic of GRAPES-3 EAS array, illustrating plastic scintillator detectors (Single PMT ▲, Double PMT ▲), and tracking muon detector modules (□).

Recent analyses from GRAPES-3 have revealed spectral features in the proton component near 166 TeV [3]. In this study, we extend the composition analysis by employing deep neural networks (DNNs) to classify primary cosmic rays (PCRs) nuclei into proton, helium, and heavier groups using key observables such as muon multiplicity distributions (MMDs), shower age along with other high level reconstructed shower parameters. The model is trained on Monte Carlo simulated events generated using the CORSIKA [4] air-shower simulation code with QGSJET-II-04 and FLUKA as the high- and low-energy hadronic interaction models, respectively, using a transition energy of 80 GeV. Simulated primary species include proton (H), helium (He), nitrogen (N), aluminium (Al), and iron (Fe), covering an energy range from sub-TeV to beyond 10 PeV. To ensure the accuracy of shower reconstruction and minimize edge effects, a series of quality cuts are applied. Only events passing a reconstruction flag criterion are selected. The reconstructed shower core is required to lie within the 50 m fiducial radius from the array center and at least 60 m from the center of the 16 muon detector modules to avoid contamination from poorly reconstructed events at the edges. The

reconstructed shower age parameter is restricted to the physical range $0.02 < s < 1.98$. Zenith angle selection is performed in the range $1.0 \leq \sec\theta < 1.1$, corresponding to approximately $\theta < 24.6°$.

This study employs two datasets derived from a common pre-simulated dataset with a spectral index of -2.5: (i) the `modelGST` dataset, and (ii) a dataset with spectral index of -2.7 for training purposes. To maintain dataset independence, the -2.7 dataset is partitioned into training, testing, and validation subsets, while the GST dataset is divided into two non-overlapping segments.

For the reliability assessment of the DNN model, the network trained on the -2.7 dataset is applied to the first segment of the GST dataset to construct the template, and the second segment is treated as pseudo-data for composition extraction. For the independent composition analysis, the training and testing subsets of the -2.7 dataset are used to generate the template, while the entire GST dataset serves as pseudo-data. This structure ensures a robust validation of the model's predictive power across both dependent and independent data scenarios. A detailed description of this methodology is provided in Section 3. All analyses are performed within shower size bins of width 0.2 in $\log_{10} N_e$, covering the range from 4.0 to 6.0, to study the composition evolution as a function of shower size.

2. Composition-Sensitive Parameters

A detailed understanding of the mass composition and energy spectrum of CRs around the knee region is essential to uncover the mechanisms responsible for their origin and acceleration. The GRAPES-3 experiment, operating in this energy regime, allows such investigations by measuring both the electromagnetic and muonic components of EAS. The key observables exploited for classification include MMDs and shower age, both sensitive to the mass of the PCRs. For showers of the same primary energy, heavier nuclei generate more muons because of their higher hadronic activity, while lighter primaries develop deeper in the atmosphere and appear younger at observational level as observed from the Figure. 2. Additional reconstructed parameters include the location of the shower core (NKGX, NKGY), the arrival direction (θ, ϕ), and the size of the shower (NKGSsize). These variables encode spatial and directional dependencies in the development of the shower.

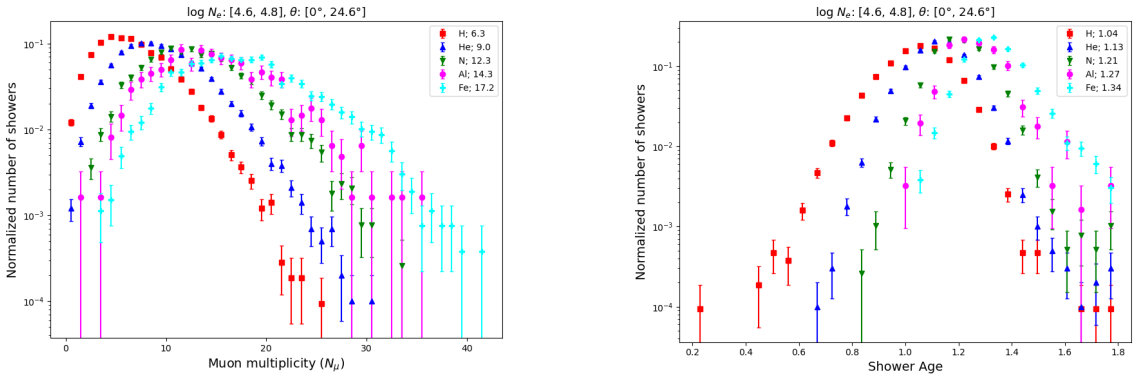

Figure 2: Normalized distributions of shower age and muon multiplicity (MMDs) for all simulated primaries in the range $4.6 \leq \log(N_e) < 4.8$ and $\theta < 24.6°$. Mean values for each primary from the GST dataset are indicated.

A feature correlation matrix shown in Figures 3 reveals nontrivial interdependencies among these parameters, justifying the use of a multivariate method such as a Deep Neural Network to exploit these correlations for effective classification.

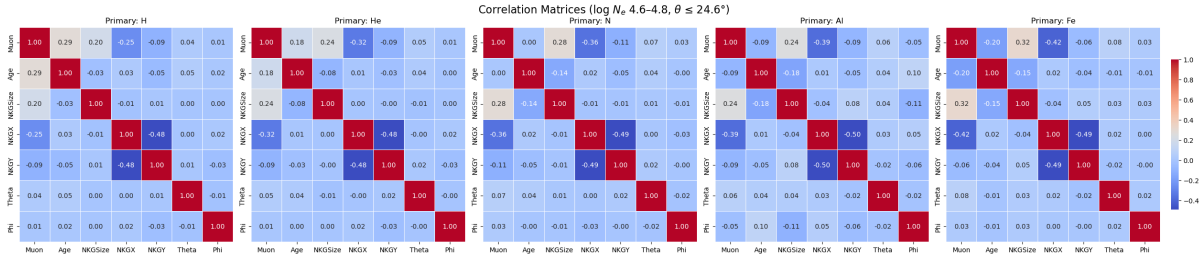

Figure 3: Feature-wise Pearson correlation matrix for input features.

Hence, the non-diagonal structure of the feature correlation matrix indicates interdependent variables, motivating a multivariate approach. Incorporating two composition-sensitive parameters—muon multiplicity (MMDs) and shower age—enhances the DNN's ability to distinguish mass groups.

3. Cosmic-Ray Composition Analysis Using Deep Neural Networks

DNNs are a class of machine learning models composed of multiple interconnected layers of artificial neurons capable of learning complex non-linear relationships from high-dimensional data. In the context of CR physics, DNNs provide a powerful tool for multivariate classification tasks, such as identifying the primary mass of cosmic rays based on composition-sensitive shower observables.

The DNN is trained in a supervised fashion using cross-entropy loss to classify events into three mass groups: proton (H), helium (He), and heavier nuclei (N + Al + Fe). Let $\mathbf{x} \in \mathbb{R}^d$ be the input feature vector, where $d = 7$ in our case (that is, MMD, Shower Age, Core X/Y, θ, ϕ, and Shower Size). The forward pass through a fully connected hidden layer is given by:

$$\mathbf{h}^{(l)} = \sigma^{(l)}\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right), \tag{1}$$

where:

- $\mathbf{h}^{(l-1)}$ is the input to layer l (with $\mathbf{h}^{(0)} = \mathbf{x}$),
- $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are the weight matrix and bias vector for layer l,
- $\sigma^{(l)}(\cdot)$ is the activation function (e.g., ReLU or LeakyReLU),
- $\mathbf{h}^{(l)}$ is the output of the l^{th} layer passed to the next layer.

This operation is repeated across all hidden layers. The output of the final hidden layer is then passed to the output layer followed by a softmax transformation [5]:

$$\mathbf{p} = \text{Softmax}\left(\mathbf{W}^{(\text{out})}\mathbf{h}^{(L)} + \mathbf{b}^{(\text{out})}\right), \tag{2}$$

where $\mathbf{p} = (p_0, p_1, p_2)$ is the probability vector over the three mass classes.

The final output of the network is a probability vector $\mathbf{p} = (p_0, p_1, p_2)$ obtained via the softmax function:

$$p_i = \frac{e^{z_i}}{\sum_{j=0}^{2} e^{z_j}}, \quad \text{for } i \in \{0, 1, 2\}, \tag{3}$$

where z_i is the logit corresponding to class i. Here, p_0 represents the predicted probability for class H, which is denoted as `score_class_0` and used for composition extraction.

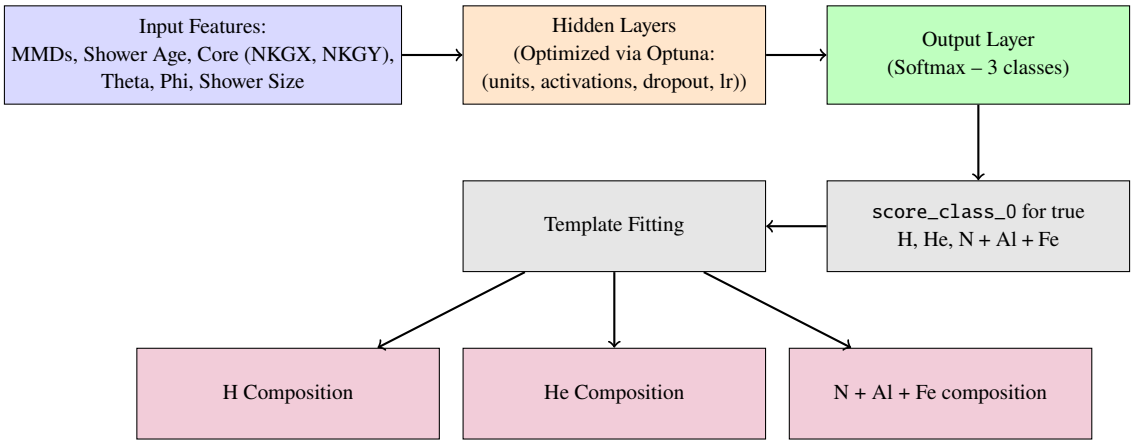

Figure 4: Schematic representation of DNNs pipeline used for primary mass classification in the GRAPES-3 experiment. The network takes 7 high-level reconstructed features as input, passes through hidden layers optimized using Optuna, and outputs class probabilities. The `score_class_0` values for true H, He, and N + Al + Fe events are used to construct templates. Template fitting is then used to extract the final mass composition.

To ensure optimal model performance, we employ the `Optuna` [6] framework for hyperparameter optimization. `Optuna` is an efficient, automated optimization library based on Bayesian sampling strategies, which searches for the best combination of hyperparameters such as learning rate, number of layers, nodes per layer, and other training parameters. This approach helps to reduce overfitting and improve generalization, ultimately improving the accuracy of primary mass classification.

The performance of the trained DNN is evaluated using ROC curves and a confusion matrix (Figure 5). High AUC values indicate strong separability, especially for heavier elements. Most confusion arises between H and He due to similar shower profiles. In general, the model achieves reliable classification among mass groups.

Beyond classification, we utilize the DNN output to estimate the mass composition of CRs through a template-fitting method. The softmax output provides a class probability vector for each event, from which the value corresponding to class 0, `score_class_0` is extracted for all simulated events with known labels (H, He, N + Al + Fe) to construct class-wise templates. The excellent agreement between the `score_class_0` distributions for the train and test sets (Figure 6, left) demonstrates the reliability of the trained model. Applying the model to an independent test set yields the pseudo-data distribution, which is then fit using a binned maximum likelihood method with the three templates to extract the mass composition. The result of this fitting procedure applied

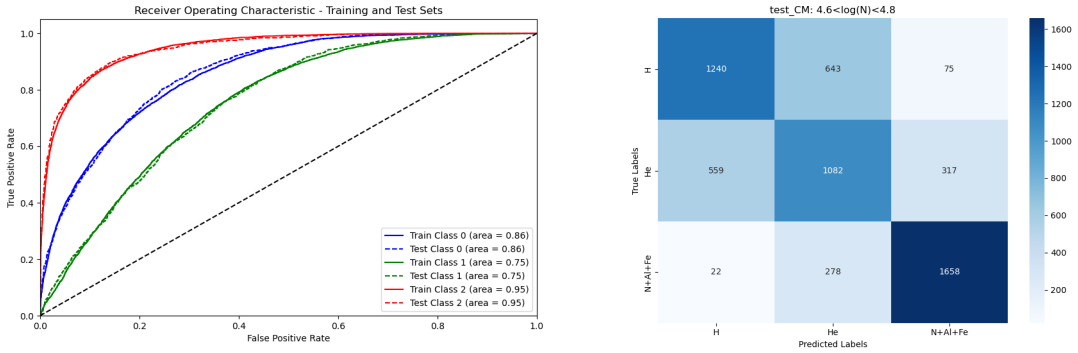

Figure 5: Receiver Operating Characteristic (ROC) curves and confusion matrix(On thr test dataset) illustrating the DNN classification performance across primary mass groups.

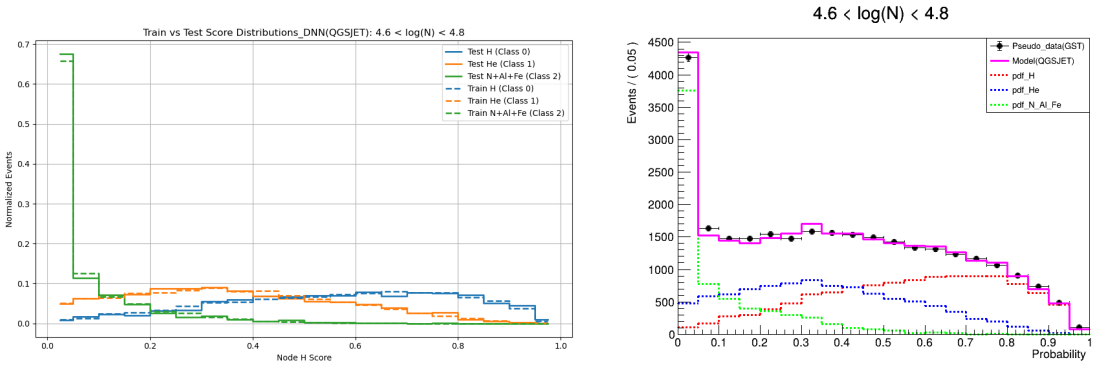

Figure 6: Comparison of `score_class_0` (Node H output) distributions for train and test sets, showing good agreement for true H, He, and N+Al+Fe classes, thereby confirming the reliability of the trained DNN (left). Mass composition fitting using templates from QGSJET-II-04 with spectral index −2.7 applied to the full GST dataset as pseudo-data (right).

to GST pseudo-data using QGSJET-II-04 templates is shown in Figure 6 (right), with the extracted fractions closely matching the injected composition.

The composition is extracted by fitting the distribution of `score_class_0` from Pseudo-data ($D(x)$) with a linear combination of MC templates ($T_i(x)$) for each class i:

$$D(x) = f_{\mathrm{H}} \cdot T_0(x) + f_{\mathrm{He}} \cdot T_1(x) + f_{\mathrm{Fe}} \cdot T_2(x), \tag{4}$$

where T_0, T_1, and T_2 are the normalized templates for H, He, and N + Al + Fe respectively, and f_{H}, f_{He}, and f_{Fe} are the fitted composition fractions subject to:

$$f_{\mathrm{H}} + f_{\mathrm{He}} + f_{\mathrm{Fe}} = 1. \tag{5}$$

A binned maximum likelihood method is used to determine the best-fit values of f_i. Let n_k be the observed number of events in bin k, and let the expected count be:

$$\mu_k = N_{\mathrm{tot}} \sum_{i=0}^{2} f_i T_{ik}, \tag{6}$$

where:

- f_i are the composition fractions for H, He, and N + Al + Fe,
- T_{ik} is the normalized value of the template for class i in bin k,
- N_{tot} is the total number of events in the pseudo-data.

Assuming Poisson statistics, the likelihood is:

$$\mathcal{L}(\{f_i\}) = \prod_k \frac{\mu_k^{n_k} e^{-\mu_k}}{n_k!}. \tag{7}$$

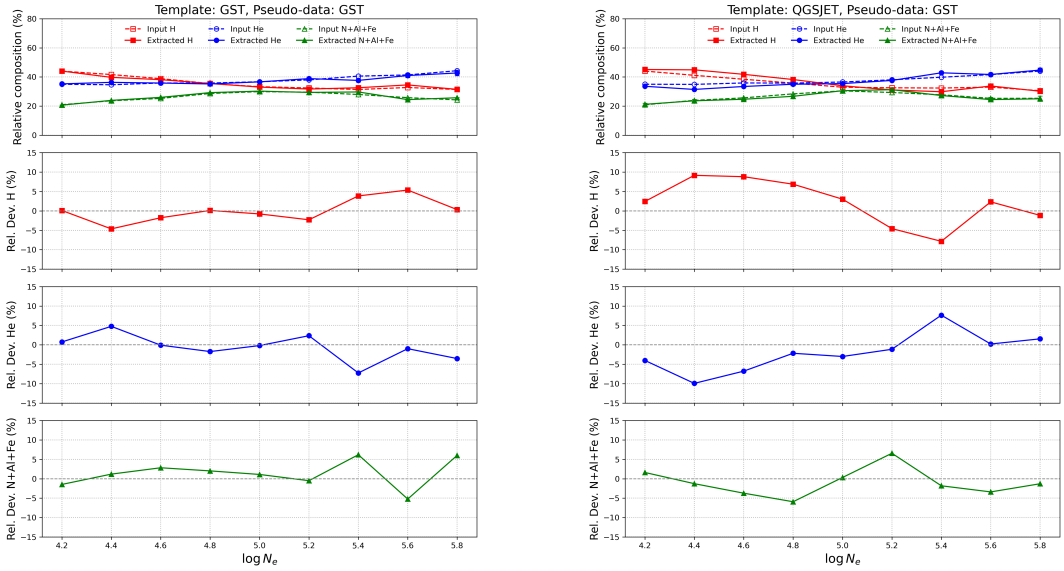

Figure 7: Comparison of extracted and input compositions for (left) model reliability test using GST-split template and pseudo-data, and (right) independent extraction using QGSJET-II-04 spectral index −2.7 template on full GST pseudo-data.

3.1 Model Reliability Assessment via Composition Fitting

In Figure 7 (left), DNN trained on QGSJET-II-04 spectral index -2.7 Monte Carlo data is validated for reliability. The first segment of the GST dataset is used to generate the composition template, while the second segment is treated as pseudo-data for composition extraction. The extracted mass fractions of H, He, and N+Al+Fe agree well with the input fractions, with relative deviations typically within ±5%. Specifically, the deviation for H remains within ±5%, while He and the heavy group fluctuate within ∼ 7%, demonstrating the DNN's capability to generalize effectively within the same physics framework. In Figure 7 (right), a more realistic test is performed where the full GST dataset serves as pseudo-data, and the template is constructed using the full

QGSJET-II-04 training and test sets. This provides smoother score distributions and better train-test consistency, particularly for the class score associated with protons (score_class_0). Despite the difference in spectral profile between template and pseudo-data, the extracted compositions still align closely with the injected fractions. Relative deviations for H vary within approximately ±10%, He remains within ±7%, and the heavy group exhibits deviations ranging from −6% to +6%. These results highlight the DNN model's stability and applicability for data-driven cosmic ray composition analysis, even in the presence of moderate model mismatches.

4. Conclusion and Outlook

We have developed a robust machine learning framework for mass composition analysis of cosmic rays in the knee region using GRAPES-3. A deep neural network trained on Monte Carlo simulations classifies events into H, He, and N+Al+Fe groups using composition-sensitive features. Score-based templates derived from the model outputs are fit to pseudo-data using a maximum likelihood method. Validation studies show agreement within 10% for known compositions. Future work includes application to GRAPES-3 observational data to extract energy-dependent mass composition and spectra from sub-TeV to 10 PeV. Systematic studies using alternative hadronic interaction models beyond QGSJET-II-04 are also planned to assess model-dependent uncertainties.

5. Acknowledgement

We thank Dr. D. B. Arjunan, Mr. Manjunath, Mr. K. Ramesh, Mr. Himanshu, Mr. Rupesh, Mr. B. Rajesh, Mr. S. Murugapandian, Mr. R. Suresh Kumar, Mr. V. Santhosh Kumar, Mr. S. Kingston, Mr. S. Sheriff, Mrs. C. Shobana, Mr. Pandurangan, Mr. Vigneswaran, Mr. B. Vivek, Mr. Sreeganth, Mr. V. Jeyakumar, Mr. Arun, Mr. Dinesh, Mr. Vinoth, Mr. A. Charles, Mrs. B. Ganga, and Mrs. Rani., and other colleagues for their support in the operation and maintenance of the GRAPES-3 experiment. I also acknowledge the GRAPES-3 collaboration for their support in this analysis. We acknowledge the support of Department of Atomic Energy, Government of India, under Project Identification No. RTI4002.

References

[1] S.K. Gupta, Y. Hayashi, S.H. Tokuda, S. Kawakami, H. Kojima, S. Sakai et al., *GRAPES-3 – A high-density air shower array for studies on the structure in the cosmic-ray energy spectrum near the knee*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **540** (2005) 311.

[2] P. Mohanty, S. Dugad, U. Goswami, S. Gupta, Y. Hayashi, A. Iyer et al., *Measurement of some eas properties using new scintillator detectors developed for the grapes-3 experiment*, *Astroparticle Physics* **31** (2009) 24.

[3] F. Varsi, S. Ahmad, M. Chakraborty, A. Chandra, S.R. Dugad, U.D. Goswami et al., GRAPES-3 Collaboration collaboration, *Evidence of a hardening in the cosmic ray proton spectrum at around 166 tev observed by the grapes-3 experiment*, *Phys. Rev. Lett.* **132** (2024) 051002.

[4] D. Heck, J. Knapp, J. Capdevielle, G. Schatz and T. Thouw, *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*. Karlsruhe Institute of Technology (KIT), 2021.

[5] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press (2016).

[6] T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama, *Optuna: A next-generation hyperparameter optimization framework*, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2623–2631, ACM, 2019, DOI.