PROCEEDINGS OF SCIENCE

A combined study of HiSCORE and TAIGA scintillation detector array for the identification of EASs

A. Vaidyanathan* [1,3] **and E. Kravchenko**[1,2]

[1] *Novosibirsk State University, NSU, Novosibirsk, Russia*
[2] *Budker Institute of Nuclear Physics SB RAS, Novosibirsk, Russia*
[3] *St. Thomas College (Autonomous), Thrissur, India*

E-mail: arunneelakandaiyer@hotmail.com

The TAIGA experiment in Tunka Valley is expanding the present scintillation detector array with new TAIGA-Muon detector stations. A simulation model was developed to optimise the layout of the new stations and study the identification performance of the array. The extensive air showers (EASs) were simulated with the CORSIKA simulation tool, and the detector response was simulated with the GEANT4 package. The EASs induced by gamma quanta or proton, Helium, Carbon, Nitrogen, Oxygen, Iron in the energy range of 1 PeV and the zenith angle range from 0°to 45°, are used for these studies. For the identification of high-energy extensive air showers, a method based on a neural network was suggested. With this method, the proton identification efficiency is more than 99.9%, while the gamma identification efficiency is not less than 50%.

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

*Speaker

PoS(ICRC2023)260

1. Introduction

The TAIGA experiment at Tunka Valley has three different detector setups: Imaging Air Cherenkov Telescopes (IACT), non-imaging wide-angle optical detectors (HiSCORE detectors), and scintillation detector array (Tunka-Grande and TAIGA-Muon) [1]. Up to now three IACTs, 120 HiSCORE stations, and 19 Tunka-Grande and three TAIGA-Muon stations are taking data. In the nearest future, 10 TAIGA-Muon stations are planned to install at the observation site.

In this simulation study, an identification method based on machine learning is put forward. For this study, simulation models of the Tunka-Grande and HiSCORE stations are utilized. In other words, it is a combined study of charged particles and Cherenkov photons detected in the extensive air shower (EAS).

2. Simulation model

The simulation model can be divided into two parts: the EAS model and the model of scintillation detector response. The simulation models are connected with two subsidiary programs based on the libraries: Eventio [3] and COAST packages [3]. The EAS model in the CORSIKA package (version-77410) also includes the simulation of the HiSCORE detector response [4]. The information about Cherenkov photons and other secondary particles present in the EAS is stored in two different files. The Eventio library is used to extract information about Cherenkov photons. The COAST library is a bridge program between the EAS model and the model of the scintillation array. The detector response of the scintillation array is simulated by using the secondary particle list made with the COAST library. The Tunka-Grande scintillation detector stations are modelled using the GEANT4 toolkit (10.06.p02) [5].

The EAS model is developed in the CORSIKA package by choosing the libraries for the high-energy hadronic interaction library QGST04-II, the low-energy hadronic interaction library GHEISHA, and the electromagnetic interaction library EGS4. The atmospheric absorption and the quantum efficiency of PMTs are included in the model. The PMTs which are used in the HiSCORE station were tested and the quantum efficiency profile of the PMT is added to the model. Based on the test result the wavelength is fixed region between 280 nm to 640 nm. The details of PMT testing are explained in the thesis [6].

The emission angle of the Cherenkov photon is considered wavelength independent. The Cherenkov bunch size is fixed as 30 after several tests. Most of the Cherenkov bunches in an event are about 30. In addition to that we have to consider the time consumption of the simulation. The EAS core of each primary event is randomised 200 times in 200 m radial circumference. The observation level was fixed at 675 m above sea level based on the location of the experiment. The minimum energy cut of secondary particles has been used: 50 MeV for hadrons, 10 MeV for muons, and 0.5 MeV for electrons (positrons) and gammas. The 'thinning' option is excluded from the model. The zenith angle of the primary particle is selected between 0°to 45°. The simulation study is conducted in the energy range of 1 PeV. For this study, seven different elements induced EAS is simulated: proton, gamma-quanta, He, C, N, O, and Fe.

The HiSCORE station is considered as a small telescope. The detection area of one HiSCORE station is 0.55 m^2. The station is tilted 25°towards the south. According to the detection area of the station, the radius of the sphere is selected in the telescope option of CORSIKA. The HiSCORE simulation model is explained in the referred article [7].

The HiSCORE stations which are close to the Tunka-Grande stations are only selected for this simulation study. The position of these 62 stations is given as in the observation site.

The Eventio library is used to extract the simulation result of the CORSIKA simulation. The EAS core and the total number of photons in each HiSCORE station by each event are saved for further study. The Tunka-Grande scintillation detector experimental setup is used for this study. The detailed explanation of the Tunka-Grande model and the selection procedure of secondary charged particles are explained in the referred article [8]. The secondary particles at ground level in the EAS are extracted from CORSIKA simulation result using the COAST library package. The random shower cores which are generated in the CORSIKA simulation are used to recreate the EASs.

The extracted details of the secondary particles at the ground level have been used as input variables for the simulation of the Tunka-Grande and TAIGA-Muon models in the GEANT4 toolkit. The detector response as the total energy loss per event (MeV) in each scintillation counter is saved for further analysis. The Tunka-Grande and TAIGA-Muon Geant4 model is explained in the article [8].

3. Procedure

One of the goals of the TAIGA experiment is the identification of mass composition and the search of high energy gamma-quanta at the energy beyond 1 PeV. In the case of the HiSCORE optical array, the Cherenkov amplitude at the distance between 150-200 m from the shower core is used for EAS energy determination [9]. It is known that proton-induced EASs show a resemblance with other element-induced EASs having different energy. This overlapping of the amplitude of various elements having different energy is one of the challenges facing the EAS identification study. The simulation study is categorised into three parts.

- Preliminary study: calculate the energy of the primary cosmic particle having the same Cherenkov photon density at 150-200m from the shower core.
- Study with fixed energy: identification of the mass composition and gamma-proton separation at overlapping condition.
- Study with spread energy: identification of mass composition and gamma-proton separation in a certain energy range.

The aim of the preliminary study was to calculate the primary CR energy. For this, a set of EAS simulations was conducted with fixed energy and three sets of the zenith-angle regions. A total of 10000 EAS events are simulated for each primary particle with fixed energy and in a certain range of zenith angles.

In general, the energy reconstruction using the HiSCORE system was conducted using Cherenkov photon densities of 150-200 m. The total number of photons registered in a station by each event was calculated. According to the position of the station, the distribution of mean amplitudes along with the shower core is calculated. From this data set, the Cherenkov amplitudes between 150 m and 200 m from the shower core are only selected. The mean amplitude in each event was calculated and fitted with a simple line function. Using the linear fitting function the primary energy of all elements was calculated with respect to the energy of proton events. The proton-induced EAS with 1 PeV energy and the overlapping energy of other elements was only taken for this study. The calculated energies of elements having the same Cherenkov photon density at 150-200m proton-induced EAS are shown in Table:1.

Table 1: overlapping point of energy values in GeV

angle	p	γ	He	C	N	O	Fe
0-15	1000	717	1102	1216	1227	1235	1431
15-30	1000	718	1107	1228	1239	1261	1465
30-45	1000	726	1119	1245	1269	1290	1505

The amplitude distribution in the HiSCORE station was plotted along with the distance from the shower core (Figure:1).

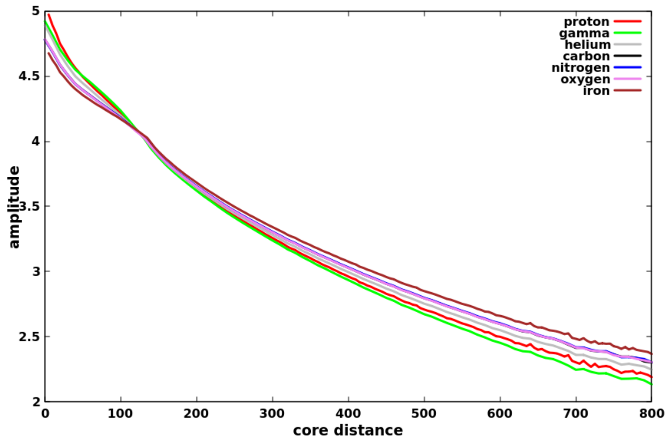

Figure 1: The Cherenkov amplitude distribution at overlapping points of energies.

After the verification of the amplitude distribution of Cherenkov photons along with the distance from the shower core, another set of EAS was simulated with the selected energies (Table 1) in the 1 PeV range. In the third step, for further verification, we spread the energy ±20% for all elements. It was done to consider the energy resolution of the HiSCORE system which is about 20%. For this part of the study, we have simulated an additional set of EASs.The detector response of the scintillation array and the optical detector system was analysed using Machine learning algorithm. The identification of mass composition and γ/p separation were conducted separately.

4. Analysis method

The simulation output was analyzed by using two different machine-learning algorithms. The hadron-proton separation was studied by using the Random forest method and the gamma-proton separation was conducted with the Binary cross-entropy method. In the previous article on gamma-proton separation, we got a second-order suppression factor [10]. This value is not enough for reliable identification of gamma-quanta induced EAS. This study was conducted by using only the scintillation detector array and the output was analysed by the Binary cross-entropy method.

In this article, we are following the same method for gamma-proton separation using additional information from HiSCORE system.

Before implementing the Random forest method for hadron-proton separation various algorithms were considered. We have considered four different classification methods: Random forest, Gradient Boosting Classifier, SV Classifier, and Naive Bayes Classier. When comparing the precision values Random forest method has shown better value. The hyperparameters were optimised with several tests.

The simulation results of optical stations and scintillation detector stations are listed separately. So there are 62 Cherenkov amplitude parameters belonging to HiSCORE stations. In the case of scintillation detector stations, the mean amplitude in surface and underground counters is measured separately. So, there are 38 amplitude parameters belonging to Tunka-Grande stations. A total of 100 amplitude parameters were utilised for the analysis.

5. Simulation results

At the beginning of the investigation using fixed overlapping points of energy, the Identification efficiency of different detector systems was studied. The simulation result was verified with the random forest method only for the 0-15 angular range. The identification efficiency was calculated for 62 HiSCORE stations, 19 Tunka-Grande stations, HiSCORE and Tunka-Grande stations, and 29 scintillation detector stations (Tunka-Grande and TAIGA-Muon) (Table 2). In this study, the network was trained with four individual elements: gamma, proton, nitrogen, and iron. The HiSCORE stations are mainly contributing to the identification efficiency in the case of the hadron-proton separation. In the case of gamma-proton separation scintillation array is also providing information for the calculation of identification efficiency. There is an increase in identification efficiency because of the additional 10 TAIGA-Muon stations. This result highlights the importance of increasing the number of scintillation detector stations.

Table 2: The calculated identification efficiency for different detector systems.

	HiSCORE				HiSCORE + Tunka-Grande			
	gamma	proton	nitrogen	iron	gamma	proton	nitrogen	iron
gamma	**81.3**	15.2	3.5	0.0	**87.8**	10.4	1.8	0.0
proton	32.5	**44.6**	21.9	1.0	17.0	**61.3**	20.7	1.0
helium	19.3	30.2	45.9	4.5	7.2	43.9	44.6	4.4
carbon	3.4	11.8	**64.4**	20.4	1.0	15.0	**63.2**	20.8
nitrogen	2.4	9.2	**65.1**	23.2	0.8	11.4	**64.3**	23.4
oxygen	3.1	10.0	**69.3**	17.6	0.9	13.5	**67.9**	17.7
iron	0.0	0.1	16.1	**83.8**	0.0	0.1	15.5	**84.4**

The data set including HiSCORE and scintillation array was used to study the variation of identification efficiency. For this, the network trained with individual elements (proton, nitrogen, and iron) and combination of elements (Proton-Helium (7:3), CNO (1:1:1), and iron). When comparing these two sets of results there is no significant variation in identification efficiency (Table:3). So we planned to use combination of elements for training the network.

	Tunka-Grande				Tunka-Grande + TAIGA-Muon			
	gamma	proton	nitrogen	iron	gamma	proton	nitrogen	iron
gamma	**81.2**	11.0	5.2	2.6	**89.0**	10.2	0.7	0.1
proton	24.3	**31.7**	21.1	22.9	9.3	**50.7**	22.5	17.4
helium	18.9	27.3	24.0	29.8	5.3	41.7	27.1	25.9
carbon	15.8	21.6	**25.6**	37.9	2.6	30.1	**29.6**	37.6
nitrogen	15.3	21.2	**25.6**	37.9	2.9	28.7	**29.3**	39.1
oxygen	14.2	21.0	**26.3**	38.4	2.3	28.2	**29.2**	40.2
iron	9.6	15.0	26.0	**49.4**	0.9	15.6	26.8	**56.7**

Table 3: The calculated identification efficiency with training individual elements and combination elements.

	0-15			0-15		
	p	N	Fe	p-He	CNO	Fe
proton	80.3	18.9	0.8	80.9	18.2	0.9
helium	51.6	44.2	4.2	51.7	43.8	4.6
carbon	16.4	63.4	20.2	22.9	58.1	19.0
nitrogen	12.3	65.4	22.3	16.4	62.1	21.5
oxygen	14.4	68.2	17.3	19.1	62.7	18.2
iron	0.1	15.6	84.3	0.1	12.4	87.5

The hadron-proton separation was studied by using the Random forest method for different zenith angles (Table:4). The network was trained with combination of elements (Proton-Helium (7:3), CNO (1:1:1), and iron). The study shows better identification efficiency for the combined detector systems than the study using an individual detector system.

Table 4: The calculated identification efficiency for different zenith angles.

angle	0-15			15-30			30-45		
	p-He	CNO	Fe	p-He	CNO	Fe	p-He	CNO	Fe
proton	79.5	18.0	2.4	74.0	19.5	6.5	68.4	21.4	10.1
helium	67.3	26.4	6.3	59.7	28.4	11.9	55.2	27.8	16.9
carbon	29.8	43.2	27.0	29.8	36.6	33.5	30.6	33.3	36.1
nitrogen	27.4	45.7	26.8	28.5	36.1	35.4	31.3	32.7	36.0
oxygen	27.3	45.7	26.9	30.1	39.6	30.3	32.4	32.6	34.9
iron	1.5	20.1	78.4	2.0	20.7	77.2	6.3	25.8	67.9

The study was repeated for overlapping points of energy with a 20% spread (Table:5). In this case, the data set is more realistic and this result can be used for further discussion. The identification efficiency of proton-helium was pushed to a maximum while maintaining the identification efficiency of other elements at least 30%. The random forest method to accept the last hypothesis compares the values of the variables. To change (increase) the probability of proton-helium identification we could require strict values for CNO or pHe identification.

Table 5: The calculated identification efficiency with overlapping point of energy having 20% of spread.

angle	thr$_{ph}$ = 0.3, thr_{cno} = 0.3			thr$_{ph}$ = 0.26, thr_{cno} = 0.29			thr$_{ph}$ = 0.28, thr_{cno} = 0.32		
	p-He	CNO	Fe	p-He	CNO	Fe	p-He	CNO	Fe
proton	91.6	8.0	0.4	91.4	7.7	0.9	88.5	9.0	2.4
helium	83.4	15.5	1.0	83.7	13.8	2.5	76.8	18.9	4.3
carbon	50.2	41.8	8.0	58.6	33.2	8.2	56.1	29.9	13.9
nitrogen	48.4	46.1	5.5	55.5	35.9	8.6	55.1	31.7	1.2
oxygen	47.4	46.5	6.1	60.1	32.3	7.6	56.7	30.6	12.7
iron	5.7	53.5	40.8	13.7	53.5	32.8	21.3	45.8	32.9

The binary cross-entropy method has been utilised for this gamma-proton separation. In this method, the network was trained with a mixture of elements: (proton-helium - 85%, CNO - 10%, iron – 5%) and gamma. The resulting identification efficiency of background air shower events was suppressed by maintaining the gamma identification efficiency at least 50%. The result gives a suppression in the factor of $\sim 10^3$ (Table:6).

Table 6: The calculated identification efficiency with fixed overlapping point of energy and overlapping point of energy having 20% of spread.

angle	fixed overlapping point of energy			overlapping energy with 20% spread		
	threshold	gamma	mixture	threshold	gamma	mixture
0-15	0.005	50.5	99.9	0.003	50.0	99.9
15-30	0.005	49.4	99.9	0.010	50.9	99.8
30-45	0.002	51.2	99.9	0.040	49.7	99.8

6. Conclusion

There is an increase in identification efficiency because of the 10 TAIGA-Muon stations. The combined study gives better results than the identification study using individual detector systems. The combined study of gamma-proton separation gives a third-order suppression factor while having 50% gamma identification efficiency.

Acknowledgments

The work was performed at the UNU "Astrophysical Complex of MSU-ISU" (agreements 13.UNU.21.0007, EB-075-15-2021-675). The work is supported by the Russian Science Foundation, 23-72-00054 (Sections 4,5)), the Russian Federation Ministry of Science and High Education (projects FZZE-2020-0024, FZZE-2023-0004, FSUS- 2022-0015 and FSUS-2020-0039).

References

[1] N. Budnev et al, *TAIGA - an advanced hybrid detector complex for astroparticle physics and high energy gamma-ray astronomy in the Tunka valley* JINST **15** C09031-C09031.

[2] https://pypi.org/project/eventio/

[3] https : //web.ikp.kit.edu/rulrich/coastreleases.html

[4] D.Heck, T.Pierog, *Extensive Air Shower Simulation with CORSIKA:A Users Guide* **Ver-7.6300** (Sept-2017).

[5] GEANT4 Collaboration, *GEANT4 Users Guide for Toolkit Developers Release 10.4* **Rev1.0** (2017).

[6] Epimakhov S *Exploring cosmic ray origins with ground-based EAS arrays Tunka and HiS-CORE* PhD thesis, Hamburg University (2015)

[7] M Ternovoy et al, *Simulation of the Tunka-Grande, TAIGA-Muon and TAIGA-HiSCORE arrays for a search of astrophysical gamma quanta with energy above 100 TeV* Journal of Physics Conference Series **1847(1)** (2021)

[8] I.Astapov et al, *Optimisation studies of the TAIGA-Muon scintillation detector array* Journal of Instrumentation **17(06)** (2022)

[9] V.Prosin et al, *Primary Cosmic Rays Energy Spectrum and Mean Mass Composition by the Data of the TAIGA Astrophysical Complex* https://doi.org/10.48550/arXiv.2208.01689 (2022)

[10] I.Astapov et al, *Identification of electromagnetic and hadronic EASs using neural network for TAIGA scintillation detector array* Journal of Instrumentation **17(05)** (2022)

Full Authors List: TAIGA Collaboration

I. Astapov[6], P. Bezyazeekov[2], E. Bonvech[1], A. Borodin[7], N. Budnev[2], A. Bulan[1], D. Chernov[1], A. Chiavassa[4], A. Dyachok[2], A. Gafarov[2], A. Garmash[8,9], V. Grebenyuk[7,11], E. Gress[2], O. Gress[2], T. Gress[2], A. Grinyuk[7], O. Grishin[2], M. Ilyushin[2], A.D. Ivanova[2], A.L. Ivanova[8,2], N. Kalmykov[1], V. Kindin[6], S. Kiryuhin[2], R. Kokoulin[6], N. Kolosov[12], K. Kompaniets[6], E. Korosteleva[1], V. Kozhin[1], E. Kravchenko[8,9], A. Kryukov[1], L. Kuzmichev[1,2], A. Lagutin[1]0, M. Lavrova[7], Y. Lemeshev[2], B. Lubsandorzhiev[1,3], N. Lubsandorzhiev[1], S. Malakhov[2], R. Mirgazov[2], R. Monkhoev[2], E. Okuneva[1], E. Osipova[1], A. Pakhorukov[2], A. Pan[7], L. Pankov[2], L. Panov[1], A. Petrukhin[6], D. Podgrudkov[1], E. Popova[1], E. Postnikov[1], V. Prosin[1], V. Ptuskin[5], A. Pushnin[2], R. Raikin[10], A. Razumov[1], G. Rubtsov[3], E. Ryabov[2], V. Samoliga[1], I. Satyshev[7], A. Silaev[1], A. Silaev(junior)[1], A. Sidorenkov[3], A. Skurikhin[1], A. Sokolov[8,9], L. Sveshnikova[1], V. Tabolenko[2], A. Tanaev[2], M. Ternovoy[2], L. Tkachev[7,11], N. Ushakov[3], A. Vaidyanathan[8], P. Volchugov[1], N. Volkov[10], D. Voronin[3], A. Zagorodnikov[2], D. Zhurov[2,12], I. Yashin[6].

[1]Skobeltsyn Institute of Nuclear Physics MSU, Moscow, Russia. [2]Institute of Applied Physics ISU, Irkutsk, Russia. [3]Institute for Nuclear Research of RAS, Moscow, Russia. [4]Dipartimento di Fisica Generale Universiteta di Torino and INFN, Torino, Italy. [5]IZMIRAN, Moscow, Russia. [6]NRNU MEPhI (Moscow Engineering Physics Institute), Moscow, Russia. [7]JINR, Dubna, Russia. [8]Novosibirsk State University, NSU, Novosibirsk, Russia. [9]Budker Institute of Nuclear Physics SB RAS, Novosibirsk, Russia. [10]Altai State University, Barnaul. [11]Dubna State University, Dubna, Russia. [12]Irkutsk National Research Technical University, Irkutsk, Russia.