Data Analysis Package

*P. Chandra, B. Ghosal and A. Tolamatti

The MACE gamma-ray telescope has the capability to observe Galactic and extra-galactic sources in the sky at very high energy (VHE; E > 20 GeV). The detection principle of MACE telescope is based on the Imaging Atmospheric Cherenkov Technique (IACT)[1]. Whenever an ultra-relativistic cosmic ray particle or VHE gamma-ray from an astrophysical source enters into the Earth's atmosphere, a cascade of relativistic particles and photons is initiated. Such a cascade is known as extensive air shower (EAS). The secondary relativistic charged particles, produced in the EAS, eventually generate a large number of Cherenkov photons which are detected by the telescope to form an image of the air shower at the camera. Since the shapes of the cascades are different for the different types of primary particles, the images of the air showers initiated by gamma-rays and cosmic-ray particles (mainly proton) show some fundamental differences at the camera of the telescope. These differences are used to segregate cosmic-ray background to detect VHE gamma-rays coming from an astrophysical source.

Due to very low photon flux in VHE range from the astrophysical sources, in comparison to huge cosmic-ray background, the detection of gamma-rays itself is a great challenge in gamma-ray astronomy. To accomplish this task for the MACE telescope, a software called the MACE data Analysis Package (MAP) has been developed. During an observation run, MACE records the raw data with a rate nearly 50 GB/hour. MAP processes the raw data recorded by MACE telescope and provides an end-to-end analysis of the data to extract gamma-ray signal from various potential astrophysical gamma-ray sources.

Availability and Installation

At present most of the ground-based gamma-ray telescopes keep the data and the analysis software private for various system-specific technical reasons. The MACE data and the current version of MAP (v. 2.1) both are currently available to only users from the Astrophysical Sciences Division. The package is developed in C++ utilizing ROOT data analysis framework and various graphical user interfaces of the software are developed in TCL/Tk on Linux platform. MAP can be installed at any Linux platform having g++ compiler and ROOT package.

The MAP software is currently available in two modes–terminal and desktop. The terminal mode is designed to analyze the MACE data using a single command which incorporates various stages of data analysis. This mode is used primarily to analyze data remotely on the Hanle server from Mumbai. Whereas, the desktop mode provides a full flash of graphical interface and control over the various aspects of data analysis to an individual user. The raw data of MACE is regularly transferred to 'Ambar' server in BARC, Mumbai. This data can be accessed and analyzed, using desktop mode at the server itself.

Analysis Work Flow

MAP is a collection of C++ programs based on ROOT data analysis framework which offers a high-level analysis of the data from the MACE gamma-ray telescope. The package is structured into several sub-packages to implement different steps of the data analysis chain. The very first input to MAP is the raw data (consisting of binary files) recorded by the MACE telescope. Broadly the raw data are divided into two categories: (i) telemetry data and (ii) event data. The raw data contains various information about each PMT of the camera (also called 'pixel') in the form of digitized counts for every triggered event. The current version of MAP (v. 2.1) uses both the telemetry and event data recorded by the MACE gamma-ray telescope for the detailed analysis.

Data Format

The first utility of MAP converts the raw binary data files into ROOT format for subsequent processing. Throughout the analysis process, both the telemetry and event data are organized in ROOT tree data structures for fast columnar access of the relevant parameters. It also reduces the size of the raw binary data files by many folds. The telemetry data contains various system related parameters like anode current, single channel rate (SCR), prompt coincidence rate (PCR), voltages of each channel, temperature in each module, etc., to monitor health of the telescope. On the other hand, the event data is used to find signal from the observed source. MAP uses all these information in the subsequent analysis by further processing these ROOT data files as input.

Telemetry Data Analysis

Telemetry parameters are extremely important to find the good time intervals (GTI) in the recorded event data to assure the best data quality for further analysis. There are two types of telemetry data available for

MACE - routine telemetry data and critical telemetry data.

The critical telemetry data such as anode current, single channel rate (SCR), prompt coincidence rate (PCR), chance coincidence rate (CCR) etc. are recorded in every 5 seconds, whereas the routine telemetry data like discriminator threshold, voltages, temperature etc. are stored in the system each minute. At this stage, MAP takes the telemetry data files in ROOT format as an input and processes those system/hardware parameters to assess the health of the system during the observation. It monitors the behavior of all the 1088 pixels of MACE camera and generates a list of bad channels (pixels) following a given criteria.

This program provides a utility to display all the telemetry parameters for each pixel, and also displays the average values (for the entire camera) of various telemetry parameters over time which are used to find Good Time Interval (GTI) based on some predefined algorithm. The GTI information is stored as output file.

Event Data Analysis

After telemetry data analysis, MAP executes the analysis of event data. The event data of MACE consists of three types of events-Cherenkov, Sky and LED. MACE records LED and sky events on the basis of predefined frequency and interval. Since the gain of each channel (PMT) of the MACE camera is different, at first, the code calibrates Cherenkov data using LED data by estimating relative gain of the each PMT with respect to some reference pixel. The package also provides the feature to check the response of each channel during observation. The code lists all the bad channels in a file which are essentially not considered in the further analysis. The Cherenkov photons from a typical EAS are expanded over only few pixels in the whole camera. However, the signals in the most of the camera pixels is dominated by the light of night sky background photons and hence the actual recorded images are observed to be noisy. At this stage the main task of MAP is to clean the images by removing the pixels which most likely do not contain Cherenkov photons from the EAS.

The decision to accept a particular pixel to be a part of the image depends on the strength of the signal in that pixel and also at the neighboring pixels. For this purpose, the code takes the Cherenkov and sky event files as input. It processes the sky event data to estimate the mean counts and standard deviation of the night sky background for each pixel. Following a standard algorithm based on the mean and standard deviation of the sky counts for each pixel, the 'image cleaning' of the Cherenkov events are executed. The MAP also provides an important feature to display both the raw and cleaned images. This utility of the MAP also shows profile of each channel. This information is used to check whether peak of the profile is falling around same location. If profile peaks of the trigger channels are not falling within the band of 32 nano-second, then those events are rejected for the further analysis. Another point to note that the Cherenkov data are also used to estimate PCR at various zenith angle. Normally PCR follows cosine dependence with the zenith angle. This behavior is also investigated to assure good quality of data at the beginning of the data analysis, along with telemetry data as discussed in the last section.

The clean shower images contain the information about the shower development, and thus can be used to find the type of primary particle (gamma-ray, proton or muon), its energy and arrival direction. Therefore, at first, it is important to characterize the cleaned images and the image parametrization module of MAP is employed to do this. The module takes the clean Cherenkov images as input and derives a set of quantities that characterize each image based on a technique called Hillas parameterization [2]. The technique is motivated by the roughly elliptical shape of the gamma-ray shower images on the camera and it is based on the estimation of moments of the two-dimensional distribution of the Cherenkov photons, up to few orders. Different parameters related to the Cherenkov images are estimated by calculating moments like length, width, asymmetry etc. These parameters are known as Hillas parameters, which describe the properties like orientation, shape of the images in a concise manner. These parameters are stored in an output ROOT file.

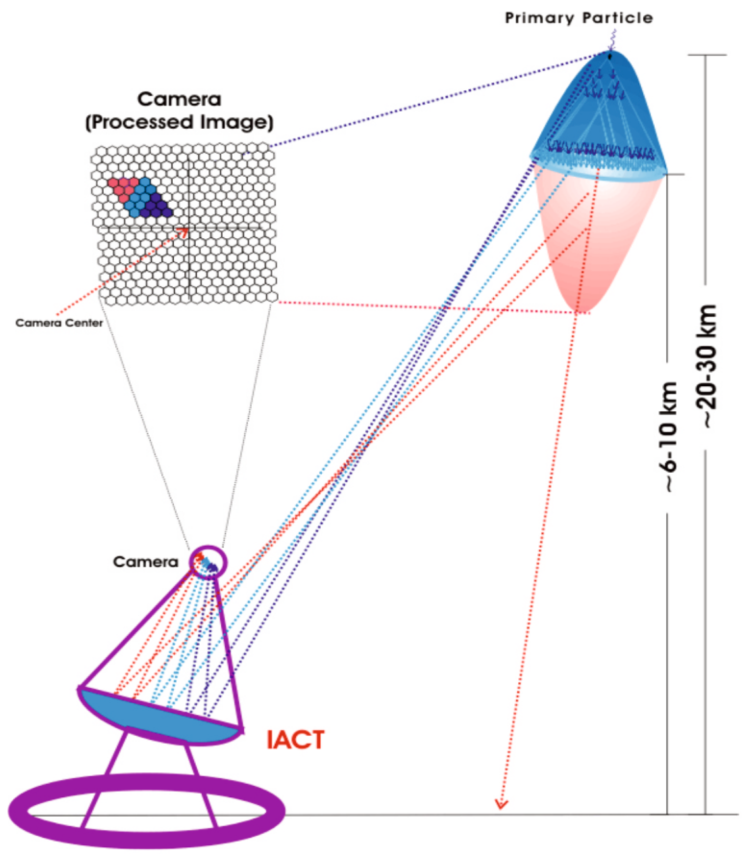
Signal Extraction

The standard procedure in MAP for extraction of gamma-ray signal by suppressing the cosmic-ray background is based on super-cut analysis. As mentioned earlier, the Cherenkov images of the EAS formed by gamma-ray and cosmic-ray exhibit some characteristics differences. A gamma-ray initiated shower image is more regular and pointed towards the camera center. On the other hand, in a proton-initiated shower image, it is more irregular in comparison to the gamma-ray initiated shower image.

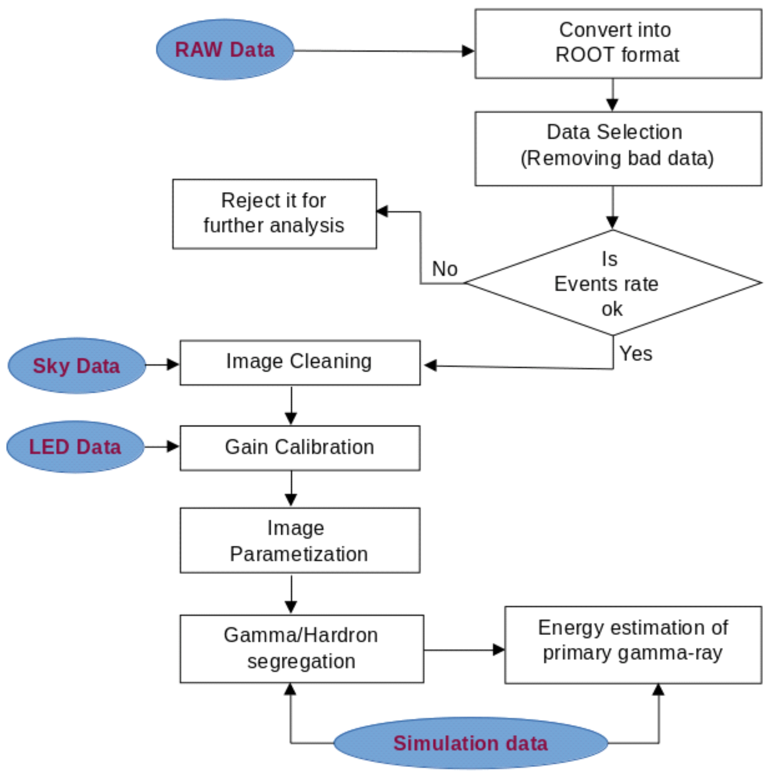
Because of such characteristic differences, the image parameters estimated for gamma-ray and proton-initiated showers show different distributions which are well studied through simulation studies of MACE telescope[3]. At this stage of data analysis, MAP processes the input file containing various image parameters of each Cherenkov image to select gamma-like shower images on the basis of static cuts/dynamic cuts on those parameters. These cuts are optimized from simulation studies. In this process, MAP segregates gamma-like events from the background Cosmic-ray events.

Since Cosmic rays come isotropically on the camera plane, their orientations are random on the camera plane. Whereas gamma-rays come only from certain source direction. Therefore, finally the orientation parameter, called alpha, is used to estimate the gamma-ray signal strength from a source. Alpha parameter basically depicts the angle between shower axis and the optical axis of the telescope. Therefore, events with small alpha point towards the source position. After applying the cuts on various image parameters, at the end a cut on alpha is applied to extract the gamma-ray signal from background events. The distribution of alpha parameter is expected to be nearly flat between 0^{o} and 90^{o}

> **THE** gamma-rays come only from a certain source direction. The orientation parameter, called alpha, is used to estimate the gamma-ray signal strength from a source.



▲ SCHEMATIC DIAGRAM for detection technique of IACTs.



▲ FLOW CHART for MACE data analysis.

for the background cosmic-ray events, whereas it peaks at lower alpha values for gamma ray events. The significance of the signal can be estimated using the formula proposed by Li and Ma[4]. Light curve of a source is the number of excess gamma-like events per unit time from the source direction at different time. The light curve is important to understand the variation in source flux over time.

Development of GUI

In the process of implementing complete data analysis chain the MAP provides a Graphic User Interface (GUI) at every stage. The interface is used for selecting data and taking various input parameters for detail analysis of data from the source. The utility also provides the facility for selection of channels based on the location of profile peak. The interface also provides various input parameters required by image cleaning and gamma-hadron segregation routines. Once the analysis of whole set of the data runs is over, the software generates various reports like source signal strength, telemetry and analyzed events information, calibration analysis information, hit channels information, logbook of analysis and GTI generated during telemetry analysis apart from various cuts values on image parameters used for gamma-hadron segregation.

Future Scope

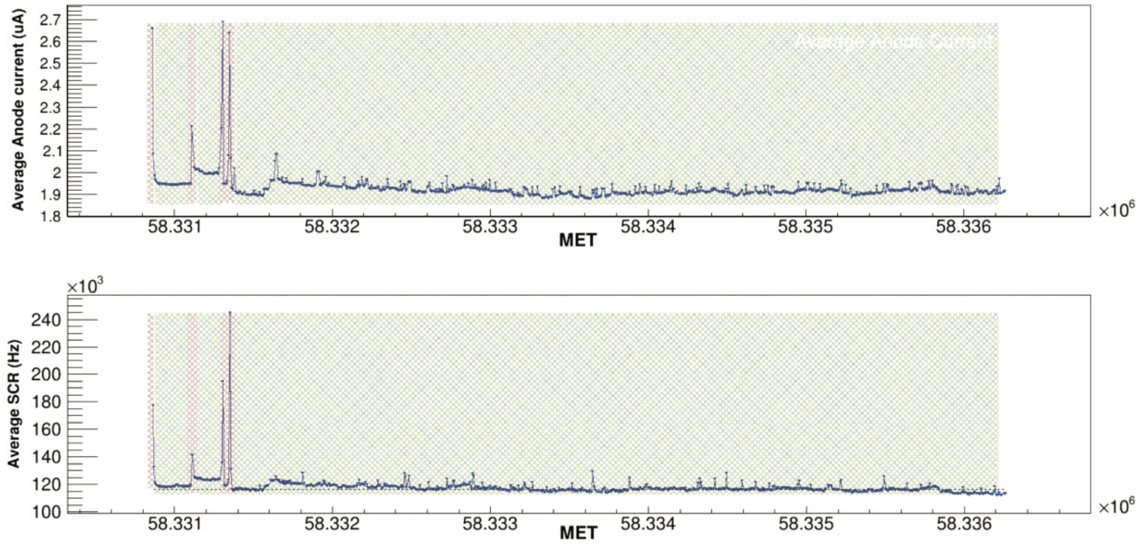
The current version of MAP is ready to extract signal from the observed data of MACE telescope. The future releases of this software will focus on the following areas:

Integration of DISP and Theta square module with MAP. This module will help in reconstructing the arrival direction of the incoming gamma-rays. Hence, the source position in the sky (sky map) can be determined. The next version of the MAP is expected to include the energy estimation module. This module is based on modern machine learning techniques like artificial neural network (ANN) or random forest (RF). This is essential to obtain the gamma-ray spectra of the observed sources.

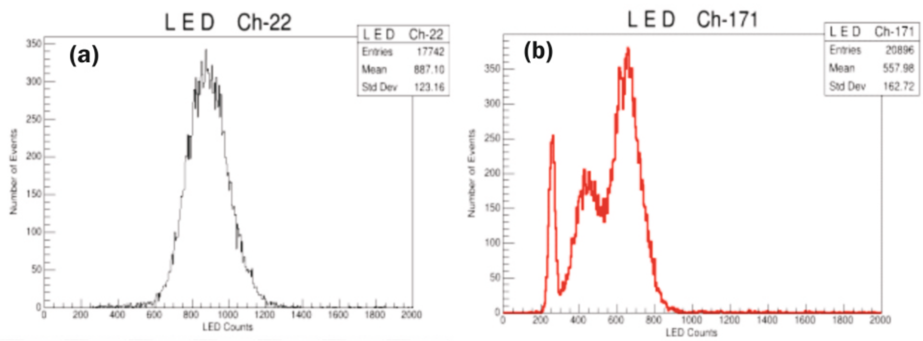
Apart from classical method of gamma-hadron segregation, a module using RF method will be included in the MAP for the same task. A python-based MACE data analysis program will be developed using the various standard scientific packages and machine learning libraries. This program will also be integrated into MAP.

Conclusions

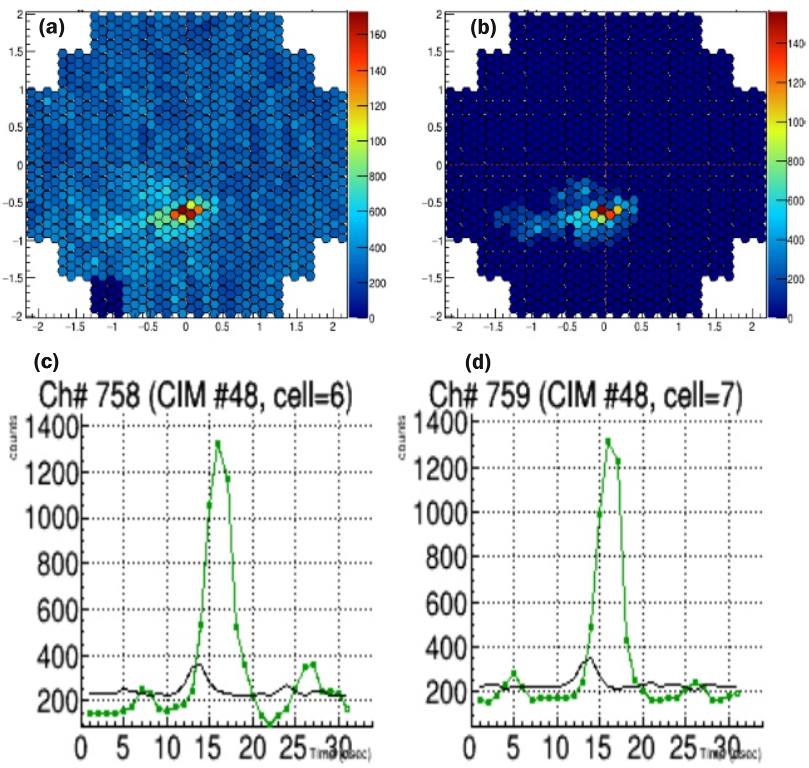
The MAP software provides a single platform for end-to-end analysis of the data collected by MACE telescope to study various astrophysical objects in gamma-ray energies. The package is developed in C++ using ROOT data analysis framework. MAP also provides various user interfaces, developed in TCL/Tk, to make the analysis user-friendly. The MAP contains various modules to perform different task of the data analysis chain like telemetry analysis, calibration, image



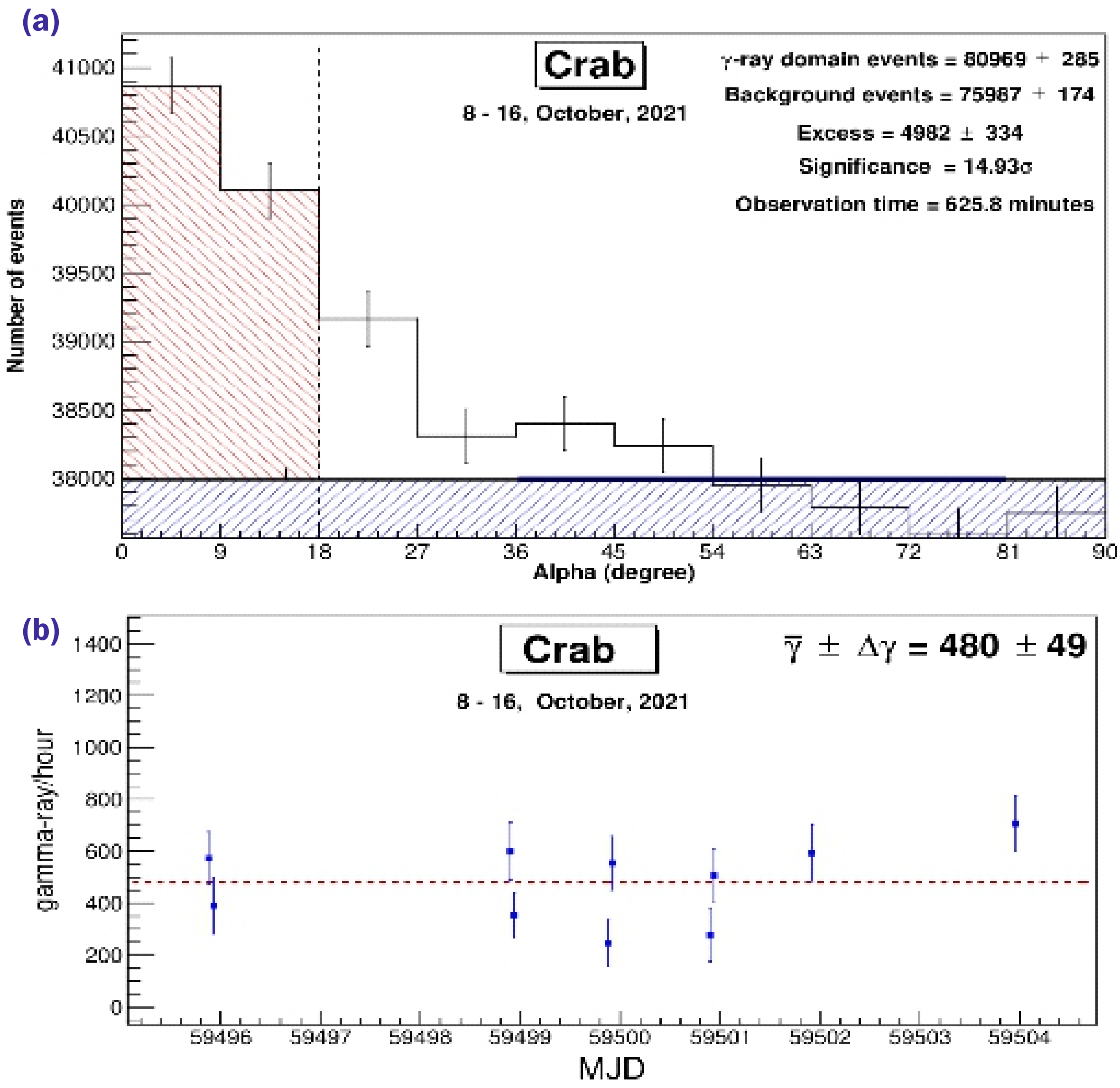
▲ VARIATION of to telemetry parameters anode current and SCR with time (in the unit Mission Elapse Time). Good and bad quality data are marked by green and red color respectively.



▲ LED COUNT DISTRIBUTION of one (a) good channel, (b) bad channel.



▲ (a) RAW (b) Cleaned Cherenkov image recorded by the MACE telescope; (c,d) Profile of two channels.



▲ (a) THE ALPHA parameter distribution of gamma like events obtained from the Crab Nebula observation by the MACE telescope during October 8-16, 2021.

(b) Corresponding gamma ray light curve generated by the MAP. Here the time is given in the unit of modified Julian Date (MJD).

processing, signal extraction. At every stage this software provides an extremely useful feature to display important parameters and images related to data analysis. Although, this package is now ready up to signal extraction, the development of MAP will remain active to introduce the advanced techniques of gamma-ray astronomy.

Acknowledgement

We would like to thank the CERN collaboration for developing and maintaining the ROOT data analysis package. We would also like to thank the observation team at Hanle, Ladakh for collecting data from the MACE telescope in extreme climatic conditions. We thank Anita Behere and her team from Electronics Division for providing various inputs during development of the software. We also thank ApSD colleagues for providing inputs during modification exercise and testing of the software.

Bibliography

[1] T. C. Weekes et al., Astrophysical Journal, 1989, 342, 379.
[2] A. M. Hillas, 19th International Cosmic Ray Conference, 1985, 3, 445.
[3] C. Borwankar et al., Nuclear Instruments and Methods in Physics Research A, 2020, 953, 163182.
[4] T. P. Li, Y. Q. Ma, Astrophysical Journal, 1983, 272, 317.

The authors are from Astrophysical Sciences Division of Bhabha Atomic Research Centre Trombay, Mumbai-400085.
*chandrap@barc.gov.in**