PROCEEDINGS OF SCIENCE

Generation of Instrument Response Functions and Event-lists of data-level 3 for the ASTRI Mini-Array

F. Pintore,[a,*] **S. Lombardi,**[b,c] **S. Crestan,**[d] **A. Giuliani,**[d] **M. Mastropietro,**[b] **M. Cardillo,**[e] **A. D'Aì,**[a] **F. Lucarelli,**[b] **S. Vercellone,**[f] **P. Romano**[f] **for the ASTRI Project**[g]

[a] *INAF/IASF Palermo, Via Ugo la Malfa 153, I-90146 Palermo, Italy*
[b] *INAF/OAR, Via Frascati 33, I-00078 Monte Porzio Catone (Roma), Italy*
[c] *ASI, Space Science Data Center, Via del Politecnico s.n.c., I-00133 Roma, Italy*
[d] *INAF/IASF Milano, Via A. Corti 12, I-20133 Milano, Italy*
[e] *INAF/IAPS Roma, Via del Fosso del Cavaliere 100, 00133, Roma, Italy*
[f] *INAF/OA-Brera, via E. Bianchi 46, 23807, Merate, Italy*
[g] http://www.astri.inaf.it/en/library/

E-mail: fabio.pintore@inaf.it, saverio.lombardi@inaf.it

The ASTRI Mini-Array is an international project to build and operate a facility to study astronomical sources emitting very high-energy γ-rays in the 1–200 TeV energy band. It will be an array of nine small-sized (4-m diameter) and large field of view ($\sim$10°) imaging atmospheric Cherenkov telescopes (IACTs) currently under deployment at the *Observatorio del Teide* (Tenerife, Spain). Thanks to its expected performance, the ASTRI Mini-Array will be important to perform deep Galactic and extra-galactic γ-ray observations. In order to achieve the final science products, the high-level scientific analysis needs the Instrument Response Functions (IRFs) and event-lists of data level 3 (DL3) as primary inputs. IRFs are derived from the reduction of suitable Monte Carlo simulations, while the event-lists are achieved from the proper reconstruction and selection of the raw data collected during observations. The IRFs include all needed quantities to properly assess the true spectral and spatial properties of the γ-ray emitters. The IRFs generation mainly depends on hardware settings or configurations, telescope pointing direction, data-taking conditions (e.g. level of the Night Sky Background, atmospheric transparency, etc.) and analysis methods (e.g. reconstruction methods, cut optimisation criteria, etc.). In this contribution, we describe the tools and methods that are currently implemented in the data-reduction software package of the ASTRI Project to generate DL3 IRFs and event-lists.

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

*Speaker

PoS(ICRC2023)722

1. Introduction

The "*Astrofisica con Specchi a Tecnologia Replicante Italiana*" (ASTRI) Mini-Array is a project led by the Italian "*Istituto Nazionale di Astrofisica*" (INAF) aimed at observing the γ-ray sky at TeV and multi-TeV energies [1, 2]. The array will consist of nine small-sized (4-m diameter) imaging atmospheric Cherenkov telescopes (IACTs) with a Schwarzschild-Couder dual-mirror optical configuration, each telescope having a large field of view (~10°). The Mini-Array is currently under deployment at the *Observatorio del Teide* (Tenerife, Spain) and it will represent an optimal facility to perform deep Galactic and extragalactic γ-ray observations, in synergy with other ground-based γ-ray facilities in the Northern Hemisphere, such as MAGIC, CTAO, HAWC and LHAASO. A prototype of the ASTRI telescopes (ASTRI-Horn) is operative on Mount Etna at the INAF "M.C. Fracastoro" observing station in Serra La Nave (at an altitude of 1725 m a.s.l.). It provided the first detection of the Crab Nebula at TeV energies with a Cherenkov telescope in a dual-mirror Schwarzschild-Couder configuration [3]. Key features of the ASTRI Mini-Array are: 1) its angular resolution of a few arcmin, that will allow us to unveil the morphological properties of extended sources in unprecedented detail above ~10 TeV; 2) its expected sensitivity in the range ~1–200 TeV, of the order of tens of mCrab for 50 hr observing time; 3) its wide field-of-view (FoV) of ~10° of diameter that will permit to monitor multiple sources simultaneously during a single observing run [see 4, for further details]. Such a large FoV will also facilitate possible serendipitous discovery of γ-ray sources.

High-level scientific products of the ASTRI Mini-Array can be obtained once data level 3 (DL3) of each observing run are generated. These consist of instrument response functions (IRFs) and γ-ray event lists. IRFs codify the global array response to the detection of an event at time t with true energy $E_{\rm true}$ and true arrival direction (RA and Dec: α_{true}, δ_{true}). Generally, the observed distribution of events $C(E, \alpha, \delta, t)$ (at reconstructed energy E and direction <α, δ>) is given by the integral:

$$C(E,\alpha,\delta,t) = \int R(E_{\rm true},\alpha_{\rm true},\delta_{\rm true},t) f(E_{\rm true},\alpha_{\rm true},\delta_{\rm true},t)\, dt\, dE\, sin\delta\, d\alpha\, d\delta + B(E,\alpha,\delta,t) \quad (1)$$

where $f(E_{\rm true}, \alpha_{\rm true}, \delta_{\rm true}, t)$ is the true event distribution and $B(E, \alpha, \delta, t)$ defines the irreducible background (BKG) distribution as a function of reconstructed energy E and coordinates $[\alpha, \delta]$. The latter is due to hadron and high-energy electron events that could not be rejected in the gamma-hadron separation procedure. Finally, $R(E_{\rm true}, \alpha_{\rm true}, \delta_{\rm true}, t)$ is a convolution of the telescope effective area (AEFF), point spread function (PSF) and energy dispersion (EDISP). AEFF represents the effective collecting area of the facility; the PSF and EDISP are the probability distributions to reconstruct the direction of arrival and the energy of an event given their true values.

An IRF provides the main quantities (i.e. AEFF, PSF, EDISP and BKG) to determine the proper true spectral and spatial properties of γ-ray sources. Usually, the response properties are not uniform along the FoV, therefore the IRF quantities are also expressed in terms of offset angles from the FoV center. The (global) event lists are tables of reconstructed information (as energy, direction and nature) and time of the candidate γ-ray events. However, since they include both real γ-ray events and background events, it is crucial to reduce as much as possible the background contamination by discriminating between real γ-ray events and background during the DL3 generation.

PoS(ICRC2023)722

For the ASTRI Mini-Array, a dedicated pipeline was designed for the whole data-reduction chain from raw data to DL3 data [5]. The latter are the data that will be used by science tools (for example Gammapy [6]) to obtain scientific products (i.e. Data Level 4 and 5, DL4/5). To asses the correct properties of the analysed sources, the DL3 IRFs and event lists need to be generated following specific prescriptions. In These Proceedings, we focus on the final part of the pipeline, detailing the main workflow of the data-selection chain.

2. ASTRI Mini-Array tools for DL3 generation and IRF production

The ASTRI Team developed the `A-SciSoft` package for the reduction and analysis of the ASTRI data [5]. The primary inputs of the software are the raw Cherenkov data produced by triggering events. These are stored in data files and processed from raw data (level 0, DL0) up to high-level science-ready data (DL3). The pipeline is intended to serve for the whole operational period of the ASTRI Mini-Array, therefore it should be always updated and easily maintainable.

The response of the ASTRI Mini-Array to a photon of true energy $E_{\rm true}$ and true arrival direction $[\alpha_{\rm true}, \delta_{\rm true}]$ is estimated by performing Monte Carlo (MC) simulations [8] (trough the use of the CORSIKA [7] and *Sim_telarray* software packages) of extensive air showers produced by a large number of γ and background events. All events are generated according to a power-law model with spectral index $\Gamma = -1.5$. Then, these simulated events undergo the full array-wise MC data reconstruction [5] and are stored in a particular data format, which, in the ASTRI context, is dubbed as Monte Carlo Data Level 2b (MC2b). MC2b data are listed in tables, providing information for each event about its true and reconstructed energy and direction, *gammaness* parameter (i.e. a quantity between 0 and 1 that defines the likelihood to be a γ-ray event), detector coordinates, and the number of telescopes that triggered the event (i.e. the *multiplicity* parameter). The whole MC simulation is split into runs and, for each of them, a table with the survived events (and their reconstructed parameters), coupled to another one with all the simulated event properties, are created. MC2b are then given in input to the module called `astriirf` that produces IRFs of data level 2 (IRF2). At the same data-level, we also have the EVT2b that are the fully reconstructed (real data) events but not yet selected: these are the main input for the `astriana` module, responsible for the DL3 event list generation (see Section 2.2).

2.1 `astriirf`

`astriirf` is a `python` module that elaborates MC2b files and returns global IRF2. The module benefits of the `Numpy`, `Astropy`, `Scipy`, and `Pandas` libraries. We developed a set of general functions used by `astriirf` and contained into the `astripy.irf` library of the `A-SciSoft` software package.

The main workflow of `astriirf` can be summarised in the following steps:

- The module reads the survived and simulated MC2b γ-ray, proton and electron populations of a given MC production. The survived MC2b events (γ-ray, protons and electrons) are then grouped into a 4D hyper-bin according to a set of suitable azimuth, zenith, offset angles (with respect to the center of the FoV) and reconstructed energy ranges. To take into account the $\Gamma = -1.5$ power law spectrum used for the MC simulations, the number of events in each hyper-bin is reweighted for the measured spectrum (HEGRA [9], ATIC [10], and HESS [11] for the gammas, protons, and electrons, respectively).

- `astriirf` then scans a given range of multiplicity, gammaness and θ^2 [1] cuts and, for each tuple of cuts, determines the number of survived events in each hyper-bin. We note that the θ^2 cuts are applied only for γ-ray events and are calculated with respect to the nominal position of the source at each offset angle. For proton/electron events, instead, `astriirf` considers all the events that fall into a given offset bin: this number is then rescaled for the θ^2 region.
- For each hyper-bin, a rate of γ-rays and background (protons + electrons) is calculated. From these rates, a Li&Ma [12] significance is estimated for a fixed exposure time (e.g. 0.5, 5 and 50 hr) and assuming a background region 5 times larger than the source extraction region. If the significance is above (or below) a given threshold, the flux is rescaled to have a significance exactly at the threshold. In `astriirf`, the default threshold is set at a 5σ level.
- Once that the sensitivities in each hyper-bin and for each combination of cuts are estimated, `astriirf` computes the tuples of cuts that provide the best sensitivity in each hyper-bin and stores them. The selected cuts are labelled as *optimized class-cuts* and they maximise the discrimination between γ-ray and hadronic events and therefore reduce the level of background. Hence, they are applied to the survived MC2b events, which result in a new sample of MC2b survived events.
- The AEFF is calculated by dividing the number of survived MC2b γ-ray events after optimised cuts in *gammaness* and *multiplicity* and the MC simulated γ-ray events in each bin of azimuth, zenith, offset and *true* energies. For the simulated events, the maximum simulated impact parameter is considered (typically, 2400 m). No weighting to the γ-ray events is applied, i.e. we assume the simulated spectrum when estimating the AEFF. AEFF is in units of m^2.
- The PSF radius corresponds to the 68% containment radius of the survived MC2b γ-ray events after optimised cuts in *gammaness* and *multiplicity*, with respect to the nominal source position in each bin of azimuth, zenith, offset and *true* energy. We assume that the shape of the PSF is a 3D Gaussian and its widths (σ) are stored in the IRF. The PSF is in units of *deg*.
- The EDISP is given as the *dispersion* matrix (i.e. the ratio between the *reconstructed* and *true* energy of the γ-ray events as a function of true energy and offset) after optimised cuts in *gammaness* and *multiplicity* for each azimuth, zenith, offset and *true* energy bin.
- The BKG rate is finally estimated as the rate of survived proton + electron events after optimised cuts in *gammaness* and *multiplicity* in a given bin of azimuth, zenith, offset and *reconstructed* energies. The same weight used in the cut optimisation procedure is applied to the proton/electron events to take into account realistic proton/electron spectral models. The BKG in the IRF is provided in terms of reconstructed energies and detector coordinates (i.e. the instrumental X and Y coordinates, DETX–DETY, with respect to the center of the FoV) and then radially distributed over the FoV. The BKG is in units of ph s^{-1} MeV^{-1} sr^{-1}.

At the end, this process produces a global IRF2 in FITS format that is saved in the data archive. Both the main IRF extensions and the optimised cuts, from which the IRF2 was generated, are stored in this format. We finally note that IRF2s are "*global*", i.e. expressed in terms of azimuth

[1] The θ^2 parameter is defined as the squared angular distance between the reconstructed and nominal event source position.

and zenith angle bins. This allows us to have a rich pool of IRFs used as "bricks" for building the average DL3 IRFs of each observing run.

2.2 `astriana`

The final step of the data reduction chain is the production of reconstructed DL3 event lists and IRF3 for a given observing run. These are the basic input of each scientific analysis and their data formats can be distinct for different facilities. However, the ASTRI Mini-Array pipeline mostly follows the prescriptions of the *Gamma-ray Data Format* (GADF; [13]).

The module responsible for the production of DL3 data is called `astriana`. It is a `python` module that uses the `Numpy, Astropy, Scipy`, and `Pandas` libraries. It also exploits a set of functions contained in the `astripy.dl3_events` library of the `A-SciSoft` software. `astriana` takes as input the global IRF2 generated with `astriirf` and an EVT2b event list of a real observing run on a given target. This module determines the target sky trajectory during a Mini-Array observing run and estimates the amount of time spent by the source over a pre-defined set of alt-azimuth bins. Subsequently, `astriana` generates a DL3 IRF (IRF3) by collapsing, for each IRF2 extensions, the alt-azimuth dimensions into an average, weighted for the time spent in each alt-azimuth bin by the source. `astriana` also evaluates the AEFF, PSF and EDISP into a smaller number of energy bins than those in the IRF2, weighting for a realistic spectrum of the target. The final output is an IRF3 having AEFF, PSF, EDISP stored in terms of energy bins and offset angles, while the BKG rate is defined through detector coordinates (DETX–DETY) and energy bins.

The last step is the screening of the observing run EVT2b with the *optimised class-cuts* calculated by `astriirf` and used to create the global IRF2. The application of the *optimised class-cuts* builds a table of survived DL3 events with their time of arrival, reconstructed energies, RA and Dec coordinates, DETX–DETY coordinates and an identity number to tag the event.

The final product of `astriirf + astriana` is a reduced DL3 event list and the associated IRF3, ready for use with science tools.

3. A case study: analysis of a realistic Crab Nebula observation

We tested the `astriirf + astriana` chain with a realistic observation of the Crab Nebula with the ASTRI Mini-Array. We simulated an observing run with starting time at 2025-12-20 22:00:00 UTC and a realistic sky trajectory (at the Teide Observatory site) with zenith range [5°–35°], before and after source culmination. The elapsed time of the observation resulted in 5.08 hr. We created the global realistic EVT2b of the observation by first simulating a DL3 event list assuming the source at 0.5° in offset from the center of the FoV and with a HEGRA Crab spectrum. We simulated the data with *Gammapy* 1.0 using an IRF3[2] without any analysis cuts. The simulated background and γ events were binned in terms of energy, offset, azimuth and zenith and, to each of them, we associated a gammaness value sampled from the distribution of the gammaness of background and γ MC2b events in each hyper-bin. This DL3 event list was then converted into an EVT2b data format.

[2]This IRF3 originated from the *ASTRI Mini-Array Prod2-Teide* produced with an independent tool available for the ASTRI Mini-Array data reduction [14].

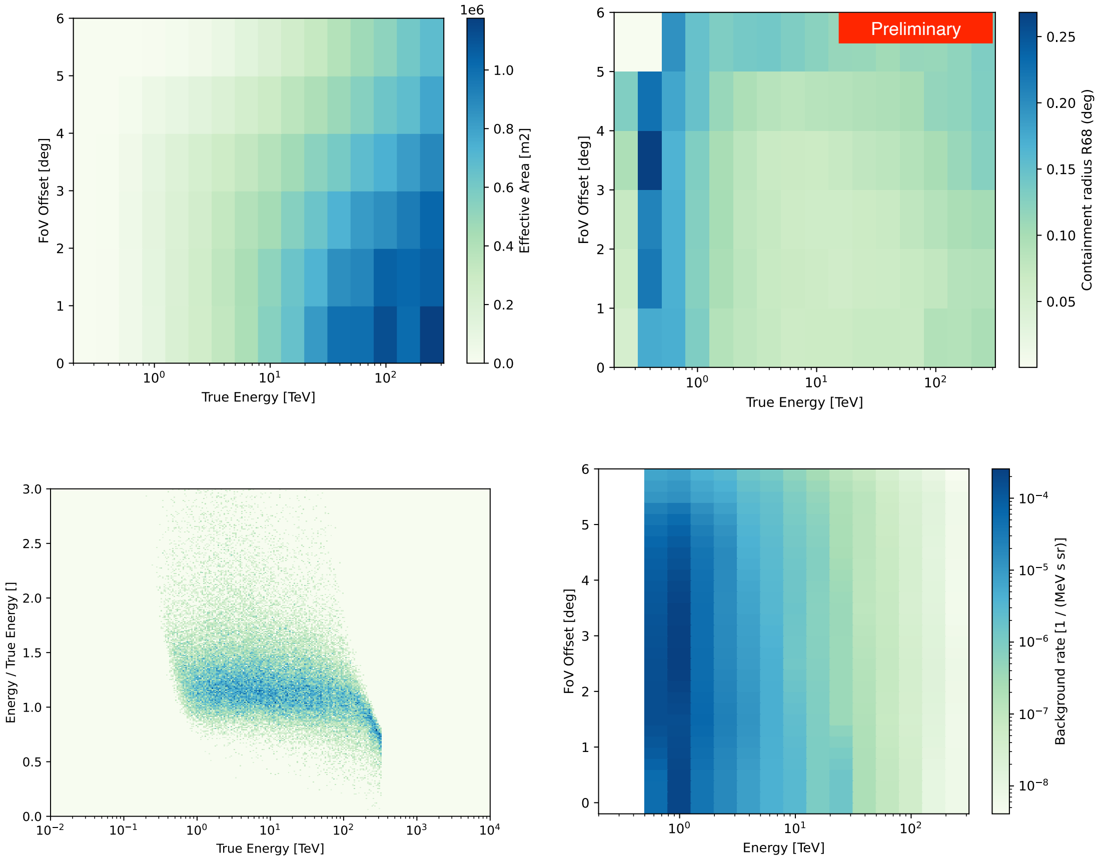

Figure 1: *Top-left*: effective area in terms of true energy and offset from the center of the FoV. *Top-right*: Containment radius at 68% of the PSF as a function of true energy and offset. *Bottom-left*: dispersion matrix (stacked over the whole offsets). *Bottom-right*: background rate as a function of energy and offset. We note that these images are preliminary.

As a first step, we ran `astriirf` by using as primary inputs MC2b data of the γ-ray, proton and electron populations. The MC2b data were taken from the ASTRI Mini-Array Prod2-Teide production [14] and *ad hoc* modified so to achieve continuous Azimuth and Zenith telescope pointing values in the range [0° – 360°] and [5° – 35°], respectively. We set: i) an energy range between 0.199 TeV – 316.23 TeV, split in 48 logarithmic bins, ii) four azimuth bins between 0°– 360° (i.e. 90° each bin), iii) four bins of cosine of zenith angles between 1 – 0.8, iv) six offset bins in the range 0°–6°. `astriirf` searched for the optimised cuts (assuming an exposure time of 5 hr) by scanning the parameter space of multiplicity, gammaness and θ^2 cuts between 2–9 (with steps of 1), 0–1 (with steps of 0.01), and $(0° - 0.1°)^2$ (with steps of $(0.002°)^2$), respectively. We then generated the global IRF2 by applying the optimized cuts to the data, as described in Section 2.1. Subsequently, the IRF2 and the EVT2b data of the Crab Nebula observation were passed as input to `astriana`, which produced as output a DL3 event list by applying the optimised cuts to the EVT2b events (see Section 2.2). Moreover, `astriana` generated an IRF3 by averaging the global IRF2 extensions over the time spent by the source in the two bins of azimuth (90°–180° and 180°–270°) and the four bins of

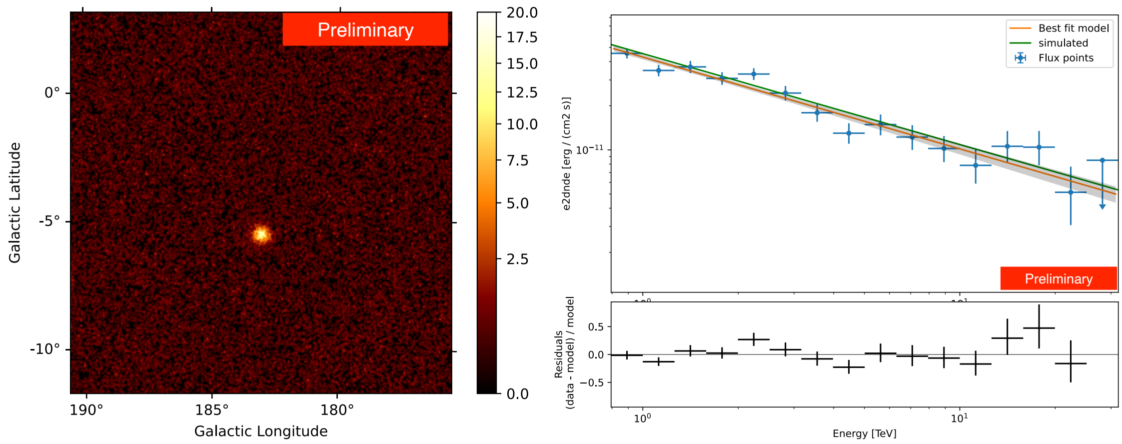

Figure 2: *Left*: Counts map, in Galactic coordinates, of the DL3 event list of the Crab Nebula realistic observing run. Colorbar is in counts per pixel. *Right*: Spectral points and best-fit model compared with the HEGRA Crab Nebula model. Downward pointing arrows indicate 3σ upper limits.

cosine of zenith crossed by the Crab Nebula during the observing run. In Fig. 1, we show the AEFF, PSF containment radius at 68%, EDISP and BKG, as a function of the energy and the offset angle, of the resulting IRF3. As a final check, we compared the IRF3 with the IRF released by the ASTRI Project [15] on Zenodo[3] and produced with an independent tool (foreseen to be superseded soon by the `astriirf + astriana` modules). A description of this independent software is beyond the scope of "These Proceedings" but, although the approaches can be slightly different from the one used in `astriif`, the results are quite compatible. Assuming the same optimized class-cuts, we found that the IRF extensions produced by the two independent approaches are generally compatible within less than 10%. Also, the two independent cut optimisation routines provide optimised cuts which are compatible within ~10%.

3.1 Gammapy data-analysis

We performed a complete scientific analysis with *Gammapy* v1.0 of the DL3 data of a realistic Crab Nebula observation. We selected only DL3 source + background events in the energy range 0.79–35 TeV and within a square region of 6° by side. This range was chosen considering the expected sensitivity of the ASTRI Mini-Array for exposure time of 5 hr [14]. A counts sky-map (not corrected for the exposure) is shown in Fig. 2-left: the Crab Nebula is clearly present in the FoV. We estimated the background with the reflected region method [16], using a circular source region of 0.2° of radius and 6 circular regions of the same size for the background.

We fitted the data with a power-law model, leaving all the power-law parameters free to vary. The fit converged to these best-fit parameters: $K = (2.70 \pm 0.10) \times 10^{-11}$ cm^{-2} s^{-1} TeV^{-1} and $\Gamma = -2.61 \pm 0.04$. They are well within 2σ uncertainty with respect to the expected HEGRA model (i.e. $K = (2.83 \pm 0.04_{\rm stat} \pm 0.6_{\rm sys}) \times 10^{-11}$ cm^{-2} s^{-1} TeV^{-1} and $\Gamma = -2.62 \pm 0.02_{\rm stat} \pm 0.05_{\rm sys}$). In Fig. 2-(right), we show the estimated flux points, the best-fit model with the 1 σ uncertainty butterfly and the simulated HEGRA model. The agreement between the best-fit from the scientific analysis and the input spectral model is rather high and proves the effectiveness of the `astriirf`

[3]https://doi.org/10.5281/zenodo.6827882

PoS(ICRC2023)722

+ `astriana` chain in generating correct (selected) event-list (EVT3) and instrument response functions (IRF3) from global event-lists (EVT2b) and γ-ray and background fully-reconstructed simulated events (MC2b). We also note that, from the scientific analysis, the source is robustly detected in ~ 5 hr, with a significance of about 3σ at ~20 TeV. Such a significance is compliant with the expected sensitivity of the ASTRI Mini-Array for an exposure time of 5 hr [14].

4. Conclusions

In this contribution, we presented a detailed description of the ASTRI Mini-Array data-reduction chain for generation of DL3 products (i.e. event lists and IRFs) needed for any scientific analysis. We showed the main steps related to the two modules of the standard analysis chain, `astriirf` and `astriana`, which are responsible for the creation of global IRF2 and DL3 IRF/event lists, respectively. The analysis of a realistic Crab Nebula observation proved that the current `astriirf + astriana` chain provides scientific results in line with expectations. The two modules are foreseen to be further tested with real data from the ASTRI-Horn telescope and, in particular, from the Mini-Array, as soon as the first observations will take place at the Teide Observatory site.

Acknowledgments

This work was conducted in the context of the ASTRI Project. We gratefully acknowledge support from the people, agencies, and organisations listed here: http://www.astri.inaf.it/en/library/. This paper went through the internal ASTRI review process.

References

[1] Scuderi S., et al.; 2022, JHEAp, 35, 52

[2] Giuliani A., et al.; 2023, Proceedings 38th ICRC, Nagoya, Japan, PoS(ICRC2023)892

[3] Lombardi S., Catalano O., Scuderi S., Antonelli L. A., Pareschi G., et al.; 2020, A&A, 634, A22

[4] Vercellone S., et al.; 2022, Journal of High Energy Astrophysics 35, 1–42

[5] Lombardi S., Lucarelli F., Bigongiari C. et al.; 2022, Proc. of SPIE, 121890P

[6] Deil C., Zanin R., Lefaucheur J., Boisson C., Khelifi B., Terrier R., et al.; 2017, Proc. 35th ICRC, 35, 766

[7] Heck D., Knapp J., Capdevielle J. N., Schatz G., Thouw T.; TIB Hannover, 998, 90

[8] Saturni F. G., et al.; 2023, Proceedings 38th ICRC, Nagoya, Japan, PoS(ICRC2023)717.

[9] F. Aharonian, A. Akhperjanian, M. Beilicke, K. Bernlöhr, et al.; 2004, ApJ, 614, 897

[10] H.S. Ahn, E.S. Seo, J.H. Adams, G. Bashindzhagyan, K.E. Batkov, et al.; 2006, Adv. in Space Research, 37, 10

[11] F. Aharonian, A. Akhperjanian, U Barres de Almeida, et al.; 2008, PRL, vol. 101, Issue 26, id. 261104

[12] Li T. P., Ma Y.Q.; 1983, ApJ, 272, 317.

[13] Nigro C., Hassan T., Olivera-Nieto L.; 2021, Universe, 7, 374.

[14] Lombardi S., et al.; 2021, ICRC21, 884

[15] ASTRI Project, ASTRI Mini-Array Instrument Response Functions (Prod2, v1.0), 2022, Zenodo

[16] Berge D., Funk S., Hinton J.; 2007, A&A, 466, 1219