PROCEEDINGS OF SCIENCE

STeVECat, the Spectral TeV Extragalactic Catalog

Lucas Gréaux,[a,*] **Jonathan Biteau,**[a] **Tarek Hassan,**[b] **Olivier Hervet,**[c] **Mireia Nievas Rosillo**[d,e] **and David A. Williams**[c]

[a] *Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France*
[b] *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, E-28040 Madrid, Spain*
[c] *Santa Cruz Institute for Particle Physics and Department of Physics, University of California, Santa Cruz, CA 95064, USA*
[d] *Instituto de Astrofísica de Canarias, E-38205 La Laguna, Spain*
[e] *Universidad de La Laguna, Dept. Astrofísica, E-38206 La Laguna, Tenerife, Spain*

E-mail: lucas.greaux@ijclab.in2p3.fr, jonathan.biteau@ijclab.in2p3.fr, tarek.hassan@ciemat.es, ohervet@ucsc.edu, mireia.nievas@iac.es, daw@ucsc.edu

The Whipple Observatory's discovery of extragalactic TeV photons from Markarian 421 in 1992 marked the beginning of extragalactic gamma-ray astronomy at very high-energy (VHE, $E >$ 100 GeV). Three decades later, both the number of detected extragalactic VHE sources and the variety of their observed spectral states have greatly increased. We present a catalog that summarizes the main spectral analysis results from extragalactic observations made by imaging atmospheric Cherenkov telescopes (IACTs). The Spectral TeV Extragalactic Catalog, STeVECat, is a collection of 403 very-high energy spectra with at least two data points published in journals between 1992 and 2021. In addition to spectral points and associated physical units, the catalog features observational metadata in the form of observation periods, livetime, excess counts over background and significance when available. STeVECat also includes the coordinates and types of the observed sources, as well as their redshifts whenever available from dedicated spectroscopic studies. This work enables studies using TeV photons that require a large corpus of extragalactic VHE observations, such as population studies of extragalactic VHE sources, studies of the GeV-TeV connection and studies of the optical to infrared content of the universe. As the most extensive collection of extragalactic IACT observations so far, STeVECat allows a comprehensive view of the currently known extragalactic sky before the advent of the new generation of IACT, the Cherenkov Telescope Array (CTA). The catalog will be made public and made accessible through the `catalog` subpackage of Gammapy, the Science Analysis Tool selected by the CTA Observatory.

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

*Speaker

PoS(ICRC2023)751

1. Introduction

In 1992, the Whipple Observatory detected TeV photons from Markarian 421, an extragalactic object located in the Ursa Major constellation [1]. This detection marked the dawn of extragalactic gamma-ray astronomy at very high-energies (VHE, $E > 100\,\mathrm{GeV}$). Three decades later, both the number of detected extragalactic VHE sources and the variety of their observed spectral states have greatly increased [2], thanks to observations carried out by ground based imaging atmospheric Cherenkov telescopes (IACTs). Three collaborations make up the current generation of IACTs: the H.E.S.S. Collaboration, observing in Namibia's Khomas Highlands, the MAGIC Collaboration, observing in La Palma, Spain, and the VERITAS Collaboration, observing in Arizona, USA.

Extragalactic sources detected at VHE are mostly active galactic nuclei (AGN), whose emission spans wavelengths from radio to gamma rays. The component generally assumed to come from inverse Compton radiation peaks in the gamma-ray band, and often corresponds to more than half of the bolometric emission of the source. Contemporaneous observations at GeV and TeV are thus key to understand AGN [2]. Observations in the GeV energy range are conducted using a single instrument, Fermi-LAT, facilitating the natural construction of coherent datasets, even with an order of magnitude more sources observed at GeV compared to TeV [3]. In contrast, the TeV observations carried out by ground-based instruments pose greater challenges of unification, making large population studies at VHE difficult without adequate spectral corpora. Moreover, the extragalactic VHE spectra observed from Earth are not the spectra emitted by the sources. The extragalactic background light (EBL, see [4]) induces absorption features in the observed spectra, which can be studied to probe the electromagnetic content of the Universe. Population studies of extragalactic VHE sources, studies of the GeV-TeV connection and EBL studies with TeV gamma-rays all require a large corpus of extragalactic observations [3, 5].

While each IACT collaboration publicly releases data associated to their published observations (e.g. VTSCat for VERITAS [6]), there is currently no complete unified catalog of extragalactic observation at VHE. Efforts have been made to create such a repository, but projects like GammaCat[1] are missing the most recent observations. We present a catalog of high-level results from extragalactic observations made by IACTs, the Spectral TeV Extragalactic Catalog, or STeVECat. STeVECat gathers IACT results from 1992 to 2021, in the form of observed spectra and their associated observational metadata.

The catalog can be accessed through the open science tool Zenodo.[2] Following GammaCat and VTSCat, we use the ECSV format developed by the Astropy project[3] to store astronomical data as well as their associated metadata. Two data tables are available in the main directory of STeVECat, containing information regarding all spectra available in the catalog (`table_spectra.csv`) as well as information regarding the sources associated with these spectra (`table_sources.csv`). The spectral data can be accessed with the file `stevecat.fits.gz` or with individual ECSV files. In Sec. 2, we describe the data tables and their contents. In Sec. 3, we describe the spectral content of the catalog and the associated collection process.

[1]GammaCat: source catalog for VHE γ-ray astronomy, see https://gamma-cat.readthedocs.io/index.html
[2]STeVECat: see https://doi.org/10.5281/zenodo.8152245
[3]Astropy: Python library for astronomy, see https://www.astropy.org/

PoS(ICRC2023)751

2. STeVECat data tables

2.1 Table of spectra

To compile data for STeVECat, we searched all journal publications up to December 31st 2021 referencing extragalactic γ-ray sources in the TeVCat[4] database. From those, we selected all observations with an associated TeV energy spectrum displayed in a figure with two spectral points or more (excluding upper limits). Datasets reported in multiple publications only appear once in STeVECat, such that one spectrum is always associated with one single publication. However, publications can report multiple datasets, and can therefore be associated with more than one spectrum. These spectra are not necessarily independent of one another, which we discuss in Sec. 3.1. In total, STeVECat gathers data from 403 observations, associated with 173 distinct journal publications.

All this information is synthesized in the table file `table_spectra.csv`, where each line corresponds to a single observation. Each entry is associated with three identification elements, which are the NASA ADS Bibcode of the publication (column named `reference`), the corresponding figure number (`reference_figure`) and the name of the observatory that produced the spectrum (`instrument`). For each spectrum, we report the source that was observed (`source_name` and `source_id`, corresponding to the name and id of the source following the GammaCat and VTSCat conventions). We extracted from the publication the associated observation period (`observation_period` for common calendar dates, `mjd_start` and `mjd_stop` for the observation start and stop in modified Julian date, MJD). When disclosed by the authors, we also report the total livetime of the observation (`livetime`, in hours), the excess count over background (`excess`) and the significance of the detection (`significance`, in number of standard deviations). When this information was not reported, these fields are filled with the `NA` value.

For each observation, we report a redshift value corresponding to the redshift of the observed source, as well as a quality flag for this value (`redshift` and `redshift_flag`, see Sec. 2.2). We also provide a flag to highlight observations that are independent of one another (`overlap_flag`, see Sec. 3.1). Each observation is associated to a single ECSV file, whose name is reported in the table (`file_id`, see Sec. 3.2).

2.2 Table of sources

The 403 spectra in STeVECat are associated to 73 distinct extragalactic objects. Information regarding these sources has been compiled in the table `table_sources.csv`, where each entry corresponds to one single object. For each source, we report one commonly used name, a list of alternative names (`source_name` and `alt_names`, the former matching the tag in `table_spectra.csv`) and its corresponding id following the GammaCat and VTSCat convention (`source_id`). Using the SIMBAD Astronomical Database, we also provide sky coordinates in the J2000 equatorial system, right ascension and declination (`coord_ra` and `coord_dec`).

We report the object class associated with each source in the 4FGL-DR3 catalog (`class`, see [7]). STeVECat mainly gathers observations from BL Lac objects (1 low-energy peaked source, `class:lbl`; 6 intermediate-energy peaked sources, `class:ibl`; 33 high-energy peaked source,

[4]TeVCat: online catalog for TeV astronomy; see http://tevcat2.uchicago.edu/

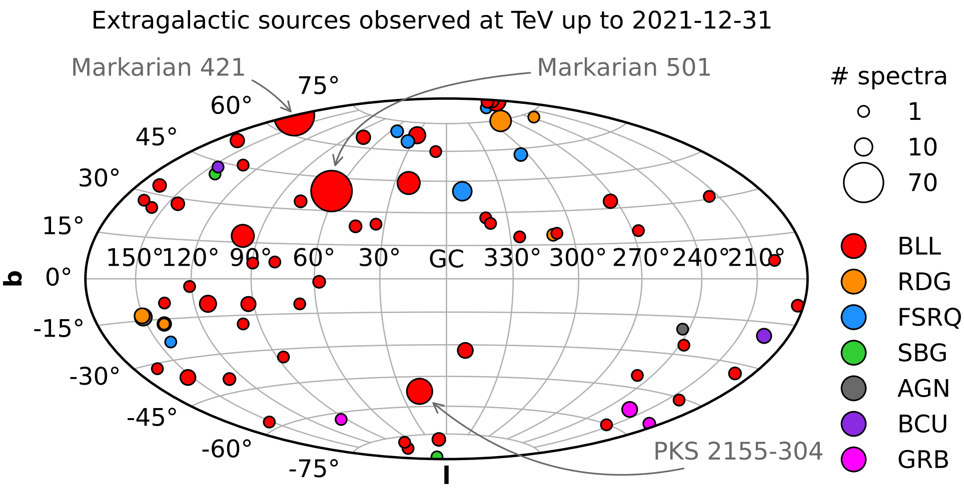

Figure 1: Sky-map of all sources referenced in STeVECat in Galactic coordinates. The area of each circle is a linear function of the number of associated spectra, and the color represents the source type: BL Lac (BLL), radio galaxy (RDG), flat spectrum radio quasar (FSRQ), starburst galaxy (SBG), AGN of unknown class (AGN), blazar candidate of uncertain class (BCU) or gamma-ray burst (GRB). The three most observed sources are highlighted: Markarian 421, Markarian 501, and PKS 2155-304.

`class:hbl`; and 13 extremely high-energy peaked sources, `class:ehbl`), flat spectrum radio quasars (`class:fsrq`, 6 sources), blazar candidates of uncertain class (`class:bcu`, 2 sources), radio galaxies (`class:fr-i`, 5 sources) and active galactic nuclei of uknown class (`class:agn`, 2 sources). Observations from 3 long gamma-ray bursts (`class:lgrb`) and 2 starburst galaxies (`class:sbg`) are also available in the catalog. The sky-map of all sources referenced in STeVECat is shown in Fig. 1, with each color corresponding to a different class of object.

In order to simplify the use of the catalog, we assign to each source a redshift measurement, corresponding to the column `redshift`. These values are extracted either from a literature review or from dedicated spectroscopic redshift measurements [8], which are reported in the `redshift_ref` column. Each value is assigned a quality flag `redshift_flag`, corresponding to the reliability of the measurement. This reliability can correspond to a lower limit (`redshift_flag:0`, 7 sources), a solid measurement (`redshift_flag:1`, 58 sources), an unknown redshift (`redshift_flag:2` with a reported redshift of `NaN`, 7 sources) or a tentative value with uncertain association or lack of spectral details (`redshift_flag:3`, 1 source).

3. STeVECat spectral data

3.1 Data collection and verification

In STeVECat, we report spectral energy distribution corresponding to each spectrum referenced in the table `table_spectra.csv` (see Sec. 2.1). These data were extracted, by order of priority: from the associated publications themselves, in the form of tables or supplementary material; from the public repositories of the collaboration that published the spectrum;[5] from the existing catalogs GammaCat and VTSCat; by contacting the corresponding authors associated with the publication; or by digitizing the published figures. The spectra obtained by digitizing figures are marked as such in their associated `comments` metadata field (see Sec. 3.2 for the metadata structure). Some published spectra were deabsorbed for the EBL-absorption effect by the publishing collaborations. In such cases, we incorporated back this absorption using the same EBL model and source redshift as the publishing collaboration. We report these corrections in the associated `comments` metadata field of the spectral data. With these corrections, all the data in STeVECat correspond to observed spectra of the sources, as opposed to intrinsic spectra.

To ensure that STeVECat was complete for spectra observed by H.E.S.S., MAGIC and VERITAS, we verified that we included all spectra from the corresponding public repositories, as well as all spectra from GammaCat. To crosscheck the validity of the collected data, we plotted each spectrum in the catalog using GammaPy [9] and matched it to the corresponding published figure. We paid close attention to the correctness of the reported observation period, which can be used to perform contemporaneous analysis of the data.

Some publications report observations from campaigns spanning several years, or covering highly variable states of the sources. Authors frequently report spectra corresponding to different subdivisions of the total observation period, as well as the time-averaged spectrum over the full campaign. These spectra are not necessarily independent of one another, but we decided to include them all in STeVECat. To enable studies requiring independent spectra, we considered all the sets of spectra with no observational overlap that provided the greatest temporal coverage. We selected the set with the largest number of spectra as a reference set and refer to it as the non-overlapping set. We marked all spectra with the `overlap_flag`. Spectra with `overlap_flag: 1` belong to the non-overlapping set, while spectra with `overlap_flag: 0` do not. Other set selection can be applied by the users of the catalog. The non-overlapping set contains 350 spectra, from 153 distinct journal publications.

3.2 Spectral data

The STweVECat main directory contains 173 sub-folders, each corresponding to a distinct journal publication and named after the corresponding NASA ADS Bibcode. Each folder contains one or more ECSV file, each corresponding to one entry of the `table_spectra.csv` table (see Sec. 2.1). Each line of this table corresponds to the file named `{reference}/{file_id}.ecsv`, each file id being of the form `tev-{id}-sed-{n}`, with `id` the id of the source and `n` a disambiguation number.

In each of the ECSV files, the spectral data are stored in a table-like structure which contains entries corresponding to the energy bin (`e_ref` for the energy bin center, and optionally `e_min` and `e_max` for the energy bin edges), the flux levels (`dnde` for the differential flux) and the errors on the flux (`dnde_errn` and `dnde_errp` for negative and positive asymmetrical errors). When the data

[5]For H.E.S.S., MAGIC and VERITAS, see the following repositories:
H.E.S.S.: see https://www.mpi-hd.mpg.de/HESS/pages/publications/
MAGIC: see http://vobs.magic.pic.es/
VERITAS: see VTSCat, https://github.com/VERITAS-Observatory/VERITAS-VTSCat

were available, we also report flux upper limits (dnde_ul). All energy columns are reported in TeV, and all differential fluxes are reported in $\mathrm{TeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

A set of metadata is also provided inside each ECSV file, in the form of a YAML-like structure. The metadata contains almost all information provided in the table_spectra.csv data table described in Sec. 2.1: the publication and the file name (reference and file_id), the id and name of the observed source (source_id, source_name), the name of the observing experiment (instrument) and the observation period (observation_period, mjd_start and mjd_stop). When this information was available, we also reported the total livetime (livetime in hours), the excess counts (excess) and the significance (significance in number of standard deviations). We finally report the overlap status of each spectrum as well as any additional information regarding the data (overlap_flag and comments, see Sec. 3.1).

As a complement to the ECSV file, we also provide in the main STeVECat directory a FITS file combining all the metadata from the data table table_spectra.csv and all the associated spectral data. The file stevecat.fits.gz contains entries for all the fields discussed in Sec. 2.2 (associated publication, observational parameters and redshift information), as well as entries for the energy bins and differential fluxes of the spectra. When information is lacking (for instance, livetime not reported in the associated publication or upper limits dnde_ul not provided), we report a value of NaN.

4. Conclusion

The VHE gamma-ray sky has expanded greatly since Markarian 421 was detected for the first time at TeV energies in 1992. To enable analysis of the entire extragalactic TeV sky, we created STeVECat, the largest database of extragalactic VHE spectra to date. STeVECat collects published extragalactic very high-energy spectra and associated observational metadata. The catalog is publicly accessible on Zenodo. The data are formatted in accordance with the standards used in publicly accessible repositories, and the entire catalog can be quickly loaded with the catalog subpackage of GammaPy, the Science Analysis Tool chosen by the Cherenkov Telescope Array Observatory.

STeVECat collects 403 spectra from 173 journal publications, making it the most comprehensive collection of VHE extragalactic spectra to date. Each of these spectra reports the observed spectrum rather than the intrinsic one, which means it can be used to infer EBL properties. We provide a subset of spectra called the non-overlapping set, which corresponds to observations that can be considered independent, to enable population studies of extragalactic gamma-ray sources, studies of the GeV-TeV connection, and studies of absorption on the EBL. Other contributions to this conference highlight the preliminary results obtained with STeVECat [10–13].

References

[1] M. Punch, C. Akerlof, M. Cawley, M. Chantell, D. Fegan, S. Fennell et al., *Detection of TeV photons from the active galaxy Markarian 421*, *Nature* **358** (1992) .

[2] E. Prandini and G. Ghisellini, *The Blazar Sequence and Its Physical Understanding*, *Galaxies* **10** (2022) 35.

[3] J. Biteau and M. Meyer, *Gamma-Ray Cosmology and Tests of Fundamental Physics*, *Galaxies* **10** (2022) 39.

[4] J. Biteau and D.A. Williams, *The Extragalactic Background Light, the Hubble Constant, and Anomalies: Conclusions from 20 Years of TeV Gamma-ray Observations*, *ApJ* **812** (2015) 60.

[5] M. Nievas Rosillo and T. Hassan, *A data-driven evaluation of Fermi-LAT extrapolation schemes to the VHE regime.*, *PoS* **ICRC2021** (2021) 722.

[6] W. Benbow, A. Brill, M. Capasso, J.L. Christiansen, A.J. Chromey, M.K. Daniel et al., *VTSCat: The VERITAS Catalog of Gamma-Ray Observations*, *Research Notes of the American Astronomical Society* **7** (2023) 6.

[7] S. Abdollahi, F. Acero, L. Baldini, J. Ballet, D. Bastieri, R. Bellazzini et al., *Incremental Fermi Large Area Telescope Fourth Source Catalog*, *ApJS* **260** (2022) 53.

[8] P. Goldoni, *Review of redshift values of bright AGNs with hard spectra in 4LAC catalog*, Apr., 2021. 10.5281/zenodo.5512660.

[9] C. Deil, J. Lefaucheur, R. Zanin, C. Boisson, B. Khelifi, R. Terrier et al., *Gammapy - A prototype for the CTA science tools*, *PoS* **ICRC2017** (2017) 766.

[10] L. Gréaux, J. Biteau and M. Nievas Rosillo, *A gamma-ray perspective on the cosmological optical controversy*, in *These proceedings*.

[11] O. Hervet, *The first skymap of the extragalactic background light from gamma-ray spectra, a new window on cosmological anisotropies*, in *These proceedings*.

[12] M. Nievas Rosillo, *A systematic study of the multi-wavelength properties of extreme gamma-ray blazars*, in *These proceedings*.

[13] G. Grolleron, J. Becerra Gonzalez, J. Biteau, M. Cerruti, R. Grau, L. Gréaux et al., *Variability studies of active galactic nuclei from the long-term monitoring program with the cherenkov telescope array*, in *These proceedings*.