Drifting instruments with Gammapy

The case of HAWC (and SWGO)

Gammapy User Call 27/07/2022

Laura Olivera-Nieto

A **Python** package for **gamma-ray** astronomy

HAWC
High Altitude Water Cherenkov
Gamma-Ray Observatory

Outline

1. Quick overview of HAWC data analysis

2. The GADF/VODF standard

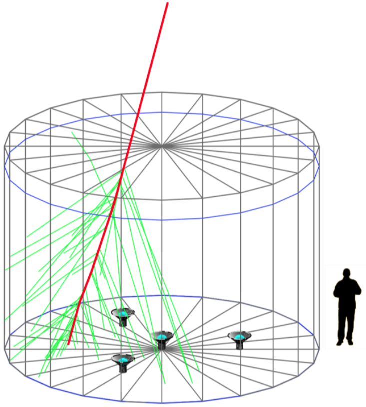
3. The Gammapy analysis workflow

4. What is different?

5. Validation of Gammapy for HAWC

6. What is missing?

The HAWC observatory



- An array of 300 tanks located in the flanks of the Sierra Negra in Mexico
- Very high duty cycle, with (almost) continuous observations
- Views 2/3 of the sky every day (including during the day!) → 24h long runs
- Designed for the range 100 GeV to 100 TeV
- Typical zenith range used < 60°

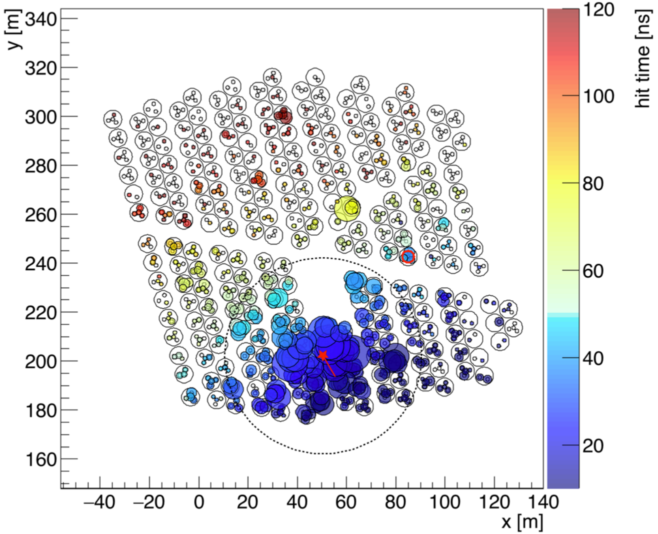
The fHit binning scheme

- Properties of reconstructed events in HAWC depend very strongly on the **number of tanks that were triggered** → **fHit** parameter

- Very different PSF and G/H separation threshold between bins

- Together with energy axis → 2D bins.

- different event classes

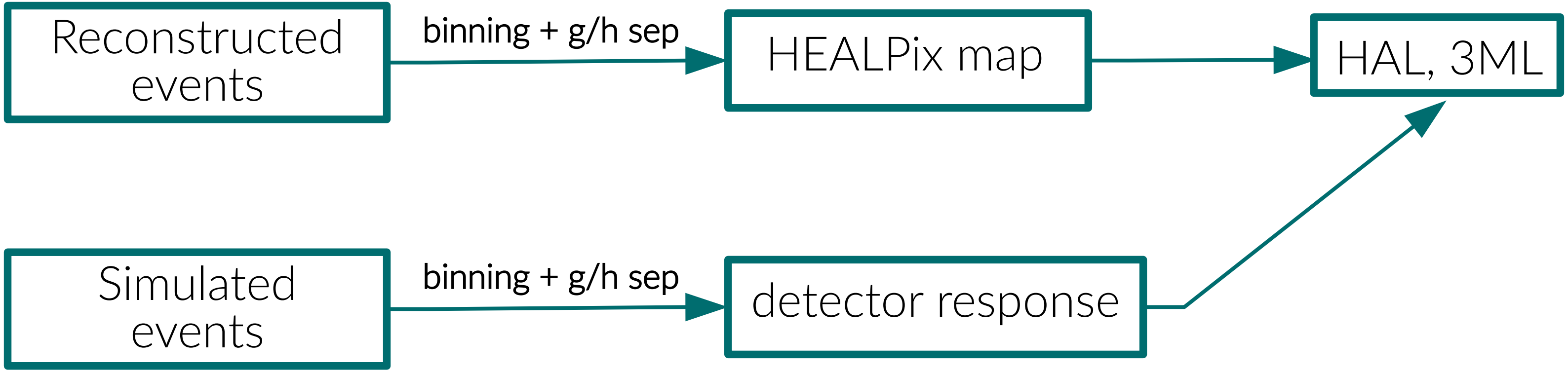


Bin number	Low fraction hit	High fraction hit
1	0.067	0.105
2	0.105	0.162
3	0.162	0.247
4	0.247	0.356
5	0.356	0.485
6	0.485	0.618
7	0.618	0.740
8	0.740	0.840
9	0.840	1.00

Bin	$\hat{E}$ energy range (TeV)
c	1-1.78
d	1.78-3.16
e	3.16-5.62
f	5.62-10.0
g	10.0-17.8
h	17.8-31.6
i	31.6-56.2
j	56.2-100
k	100-177
l	177-316

The standard HAWC analysis workflow

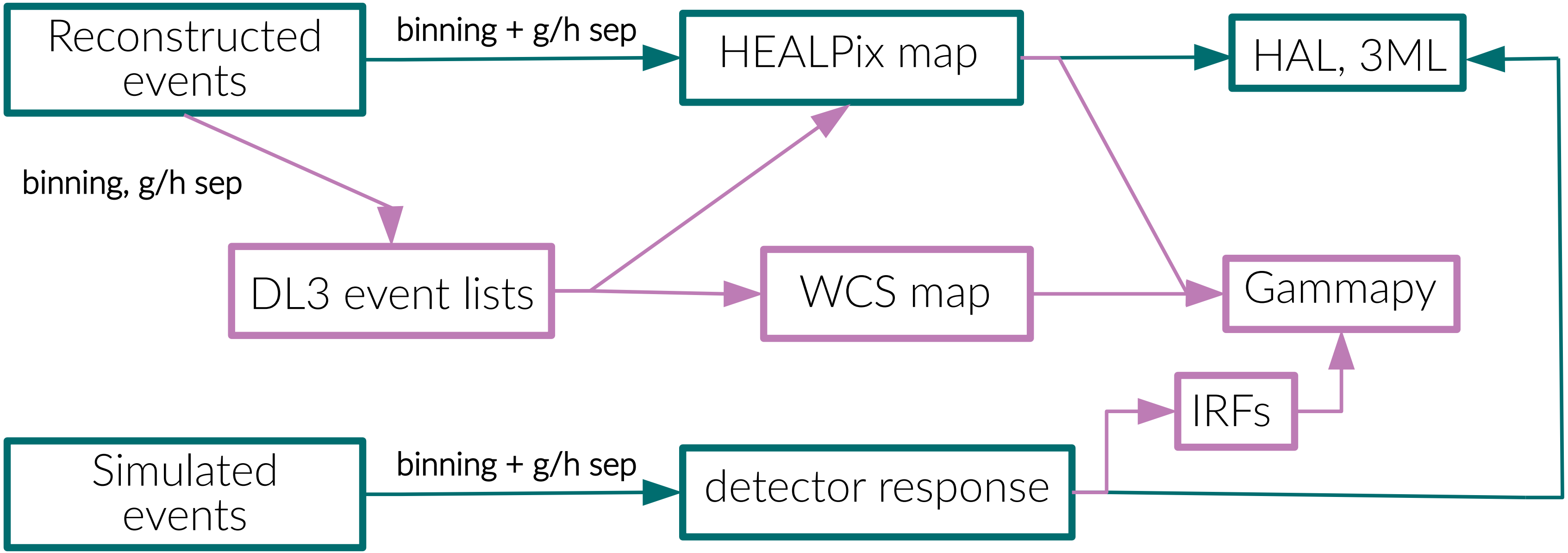
standard



Issues: time information lost, spatial binning fixed, map-making extremely computationally intensive process

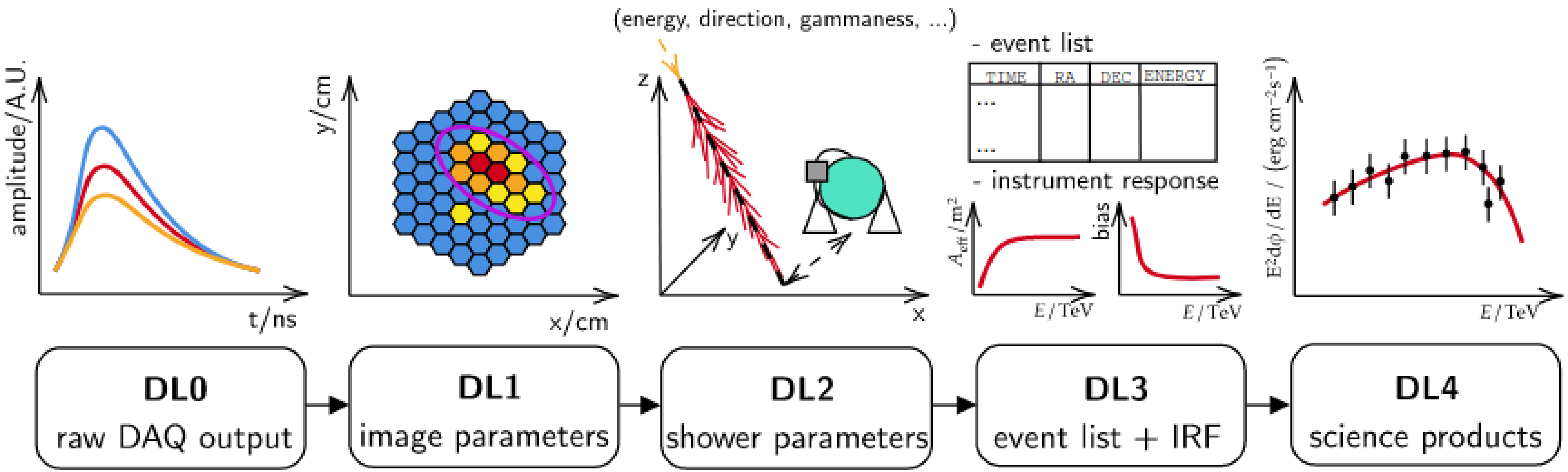
The expanded HAWC analysis workflow using Gammapy

expanded



DL3? The standard format for gamma-ray astronomy

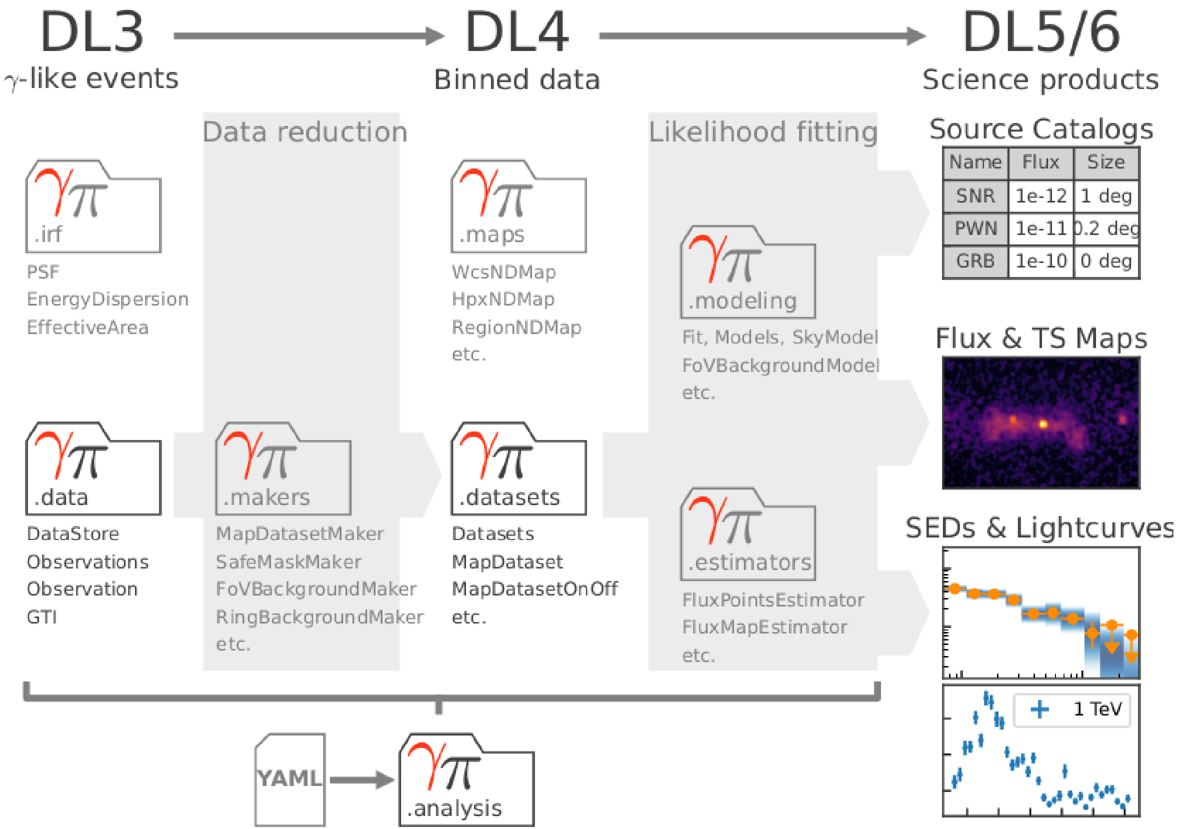
Nigro et al 2021



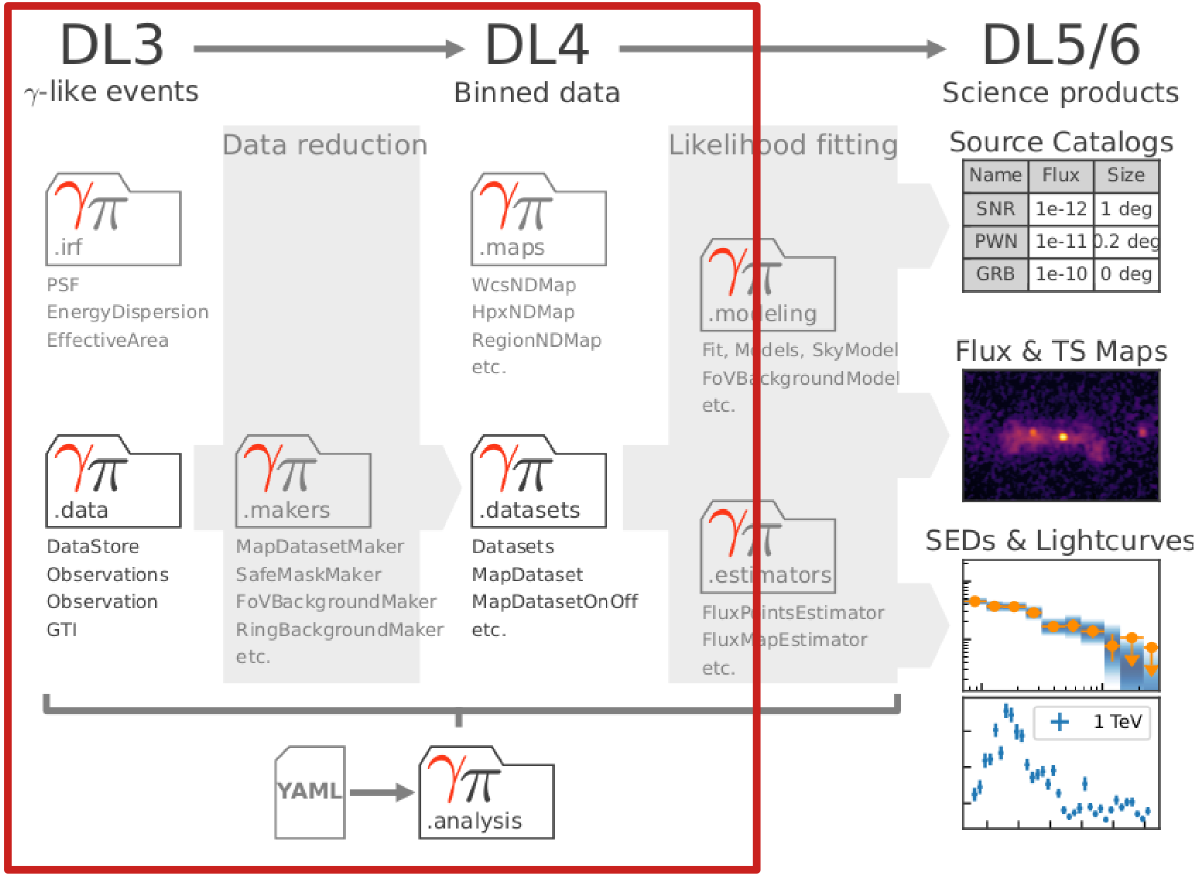
- DL3 is mostly based on FITS / Fermi-LAT data formats, which are established formats to deliver gamma-ray data to a larger user base
- Common format → common tool!

GADF/VODF… stay tuned!

Gammapy


DL3
γ-like events
DL4
Binned data
DL5/6
Science products
Data reduction
Likelihood fitting
Source Catalogs
Name | Flux | Size
SNR | 1e-12 | 1 deg
PWN | 1e-11 | 0.2 deg
GRB | 1e-10 | 0 deg
.irf
PSF
EnergyDispersion
EffectiveArea
.maps
WcsNDMap
HpxNDMap
RegionNDMap
etc.
.modeling
Fit, Models, SkyModel
FoVBackgroundModel
etc.
Flux & TS Maps
.data
DataStore
Observations
Observation
GTI
.makers
MapDatasetMaker
SafeMaskMaker
FoVBackgroundMaker
RingBackgroundMaker
etc.
.datasets
Datasets
MapDataset
MapDatasetOnOff
etc.
.estimators
FluxPointsEstimator
FluxMapEstimator
etc.
SEDs & Lightcurves
1 TeV
YAML
.analysis

Gammapy


DL3
γ-like events
DL4
Binned data
DL5/6
Science products
Data reduction
Likelihood fitting
.irf
PSF
EnergyDispersion
EffectiveArea
.maps
WcsNDMap
HpxNDMap
RegionNDMap
etc.
.modeling
Fit, Models, SkyModel
FoVBackgroundModel
etc.
.data
DataStore
Observations
Observation
GTI
.makers
MapDatasetMaker
SafeMaskMaker
FoVBackgroundMaker
RingBackgroundMaker
etc.
.datasets
Datasets
MapDataset
MapDatasetOnOff
etc.
.estimators
FluxPointsEstimator
FluxMapEstimator
etc.
YAML
.analysis
Source Catalogs
Name | Flux | Size
SNR | 1e-12 | 1 deg
PWN | 1e-11 | 0.2 deg
GRB | 1e-10 | 0 deg
Flux & TS Maps
SEDs & Lightcurves
1 TeV

HAWC analysis in Gammapy – event lists

- DL3: gamma-like events, with “science ready” parameters
- Mandatory columns: EVENT_ID, TIME, RA, DEC
- Can keep **any other field** that is relevant (e.g. fHit or core location)
- PRO: Time dimension isn’t lost!
- PRO: Map-making becomes much more flexible and faster
- CON: Depending on cut tightness, file sizes can be a problem → need to find a right balance for the smallest unit of time that defines an event list file

HAWC Collaboration: Validation of standardized data formats and tools for particle detector arrays.

Event ID	R.A (deg)	Dec. (deg)	Energy (TeV)	Time (s)	Core X (m)	Core Y (m)	Bin ID
1	296.401	18.649	6.698	1132183230.200404	50.4	212.8	7f
2	305.046	27.225	7.063	1132183236.213954	-30.7	214.9	7f
3	16.556	14.990	7.709	1132183250.7916136	-37.1	214.9	6f

Table 2: Simplified entries of an event list. Note that the real precision of the numbers has been reduced for formatting convenience.

HAWC analysis in Gammapy – event lists

- Existing standard was remarkably compatible with the needs of drifting instruments
- Only one change needed: make pointing information more flexible in file header
- Definition of observation mode "DRIFT"

In addition to the OGIP-defined values (`POINTING` , `RASTER` , `SLEW` and `SCAN`), we define the option `DRIFT` to accomodate ground-based wide-field instruments, in which local zenith/azimuth coordinates remain constant. In this case, the header keywords `RA_PNT` and `DEC_PNT` are no longer mandatory, and instead `ALT_PNT` and `AZ_PNT` are required.

- `OBS_MODE` **type: string**
  - Observation mode. See notes on OBS_MODE below.
- `RA_PNT` **type: float, unit: deg**
  - Pointing Right Ascension (see RA / DEC). Not mandatory if `OBS_MODE=DRIFT` , but average values could optionally be provided.
- `DEC_PNT` **type: float, unit: deg**
  - Pointing declination (see RA / DEC). Not mandatory if `OBS_MODE=DRIFT` , but average values could optionally be provided.
- `ALT_PNT` **type: float, unit: deg**
  - Pointing Altitude (see Alt / Az). Only mandatory if `OBS_MODE=DRIFT`
- `AZ_PNT` **type: float, unit: deg**
  - Pointing azimuth (see Alt / Az). Only mandatory if `OBS_MODE=DRIFT`

HAWC analysis in Gammapy – IRFs

- Instrument response functions describe the combined detection abilities and precision of an instrument data-taking and reconstruction procedure
- **Angular resolution (PSF)**: accuracy of reconstruction of the direction of the incident gamma-ray,
- **Energy resolution (EDISP)**: accuracy of reconstruction of the energy of the incident gamma-ray,
- **Effective area (AEFF)**: detection probability of the incident gamma-ray,
- **Residual hadronic background:** expected residual hadronic background by misclassified events
- All derived from simulations except sometimes background (when you have data!)

HAWC analysis in Gammapy – IRFs

- In IACTs, IRFs are typically given as a function of energy (reconstructed and/or true) and **offset from pointing position**
- The latter is obviously not very useful for an instrument like HAWC
- The HAWC response depends on energy and zenith
- Two options:

  1. **Zenith binned IRFs** $\rightarrow$ Need to be integrated with relevant source path before being used
  2. **Declination binned IRFs** $\rightarrow$ Case 1 but integrated for 1 source transit

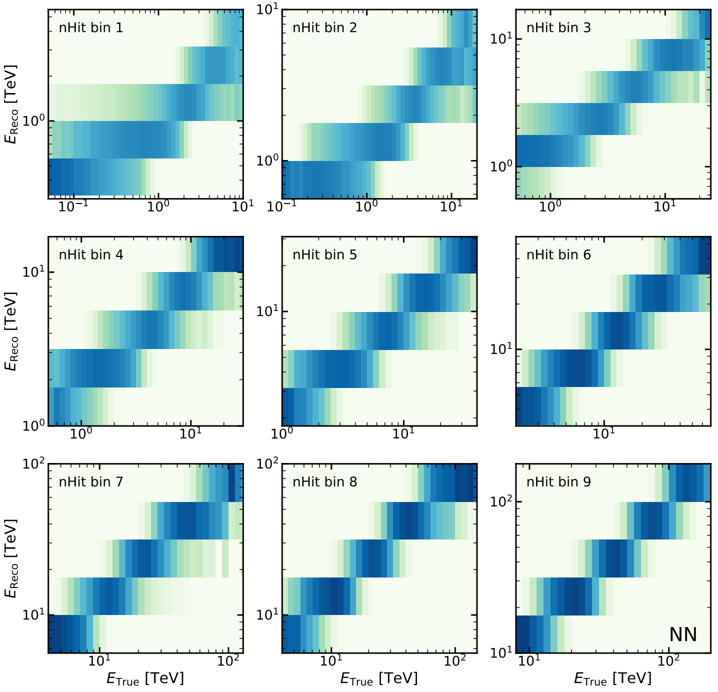
- Case 1 needed for transits, case 2 more practical for longer exposure analysis

HAWC analysis in Gammapy – declination binned-IRFs

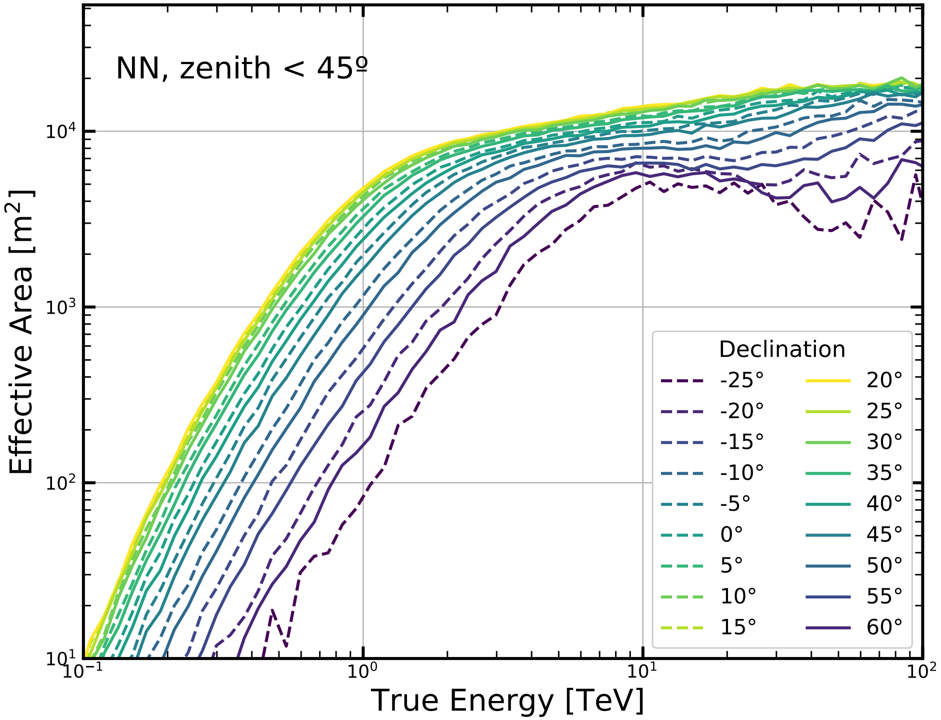
- For a given Earth location, the path that any source covers on the sky depends on its declination
- For example, at HAWC location, the Crab transits near zenith, but HAWC J1825-134 doesn't get very high
- The relevant quantity is the fraction of the time that a source spends in each zenith band → weights
- Use this weights to integrate the zenith-dependent response
- "Declination-binned" = full-sky map

```
class IRFMap:
    """IRF map base class for DL4 instrument response functions"""
```

HAWC analysis in Gammapy – declination binned-IRFs



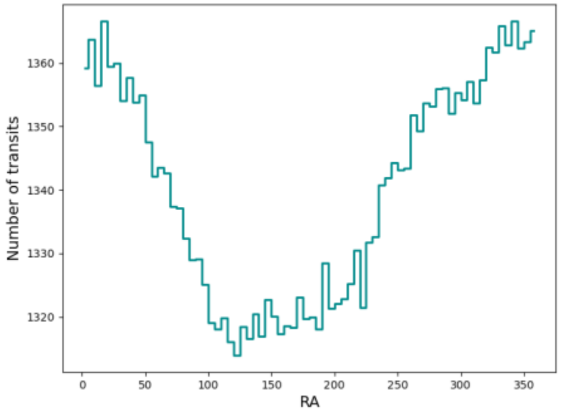
Crab location



HAWC analysis in Gammapy – exposure and GTI

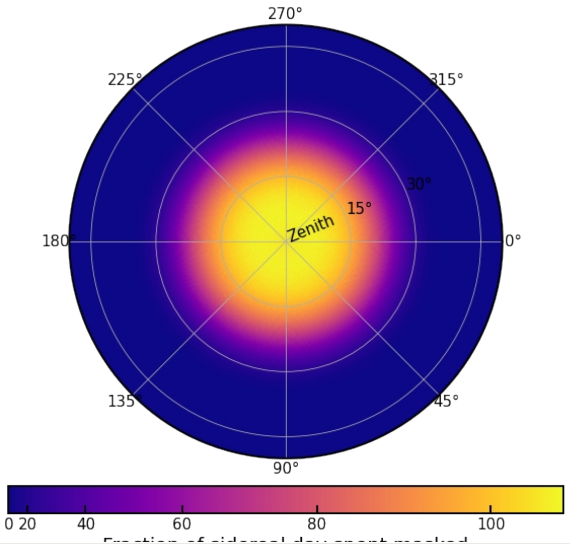
- Another useful quantity defined by the standard are the "Good-time intervals" or GTI
- Defined in the same way as in for the Fermi-LAT
- Time intervals during which the detector is on and taking data continuously
- In HAWC: we determine data quality at the "sub-run" level (125 seconds of data)
- Use GTIs to estimate the number of transits (depends on R.A.!) contained in an event list
- Combine with effective area for 1 transit

START	STOP
float64	float64
1168135330.0	1168135455.0
1168135455.0	1168135580.0



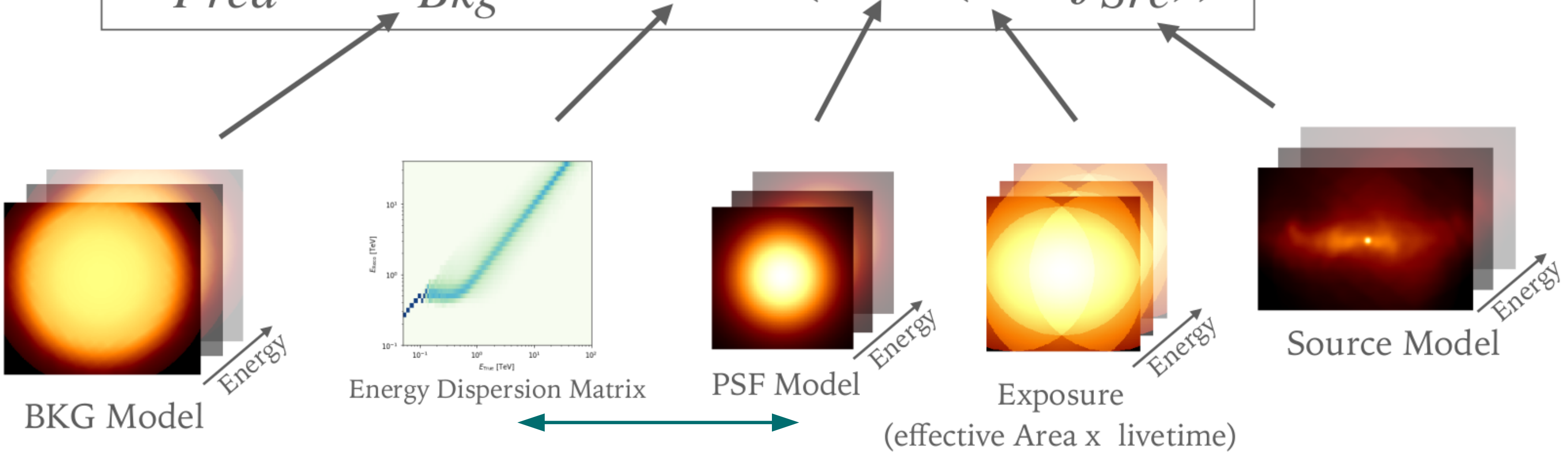
HAWC analysis in Gammapy – background

- Derived using data in regions where no gamma-ray sources are expected
- Again can be described as a rate as a function of zenith or as a full-sky map
- Which one is useful depends on the type of analysis



IACT "DL4" Data Model

$$N_{Pred} = N_{Bkg} + \mathrm{EDISP}(\mathrm{PSF}(\mathcal{E} \cdot f_{Src}))$$



HAWC PSF is given in reco energy

HAWC analysis in Gammapy – validation

Once all of the ingredients are ready, we set out to test the scheme in several steps

accepted to A&A!

1. Basic low level checks

2. Point source analysis

3. Extended source analysis

4. Time-domain analysis

5. Joint analysis with other gamma-ray instruments

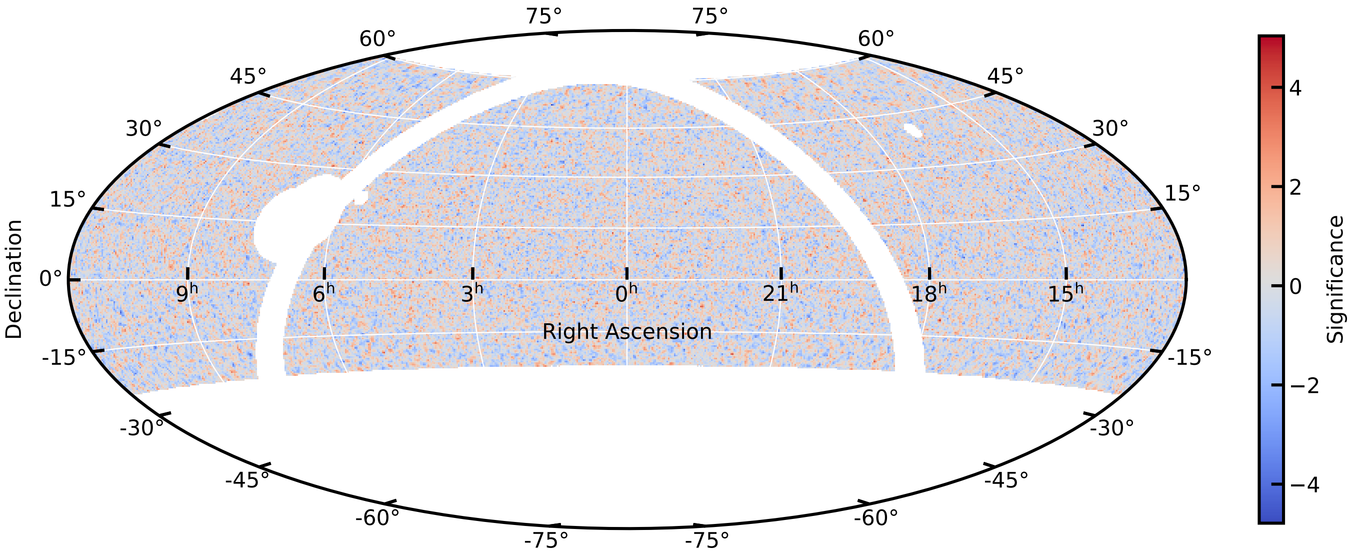
Validation of standardized data formats and tools for ground-level particle-based gamma-ray observatories.

A. Albert[11], R. Alfaro[13], J.C. Arteaga-Velázquez[14], H.A. Ayala Solares[6], R. Babu[2], E. Belmont-Moreno[13], C. Brisbois[4], K.S. Caballero-Mora[15], T. Capistrán[16], A. Carramiñana[17], S. Casanova[18], O. Chaparro-Amaro[19], U. Cotti[14], J. Cotzomi[20], S. Coutiño de León[5], E. De la Fuente[21], R. Diaz Hernandez[17], M.A. DuVernois[5], M. Durocher[11], C. Espinoza[13], K.L. Fan[4], M. Fernández Alonso[6], N. Fraija[16], J.A. García-González[22], H. Goksu[23], M.M. González[16], J.A. Goodman[4], J.P. Harding[11], J. Hinton[23], D. Huang[2], F. Hueyotl-Zahuantitla[15], P. Hüntemeyer[2], A. Jardin-Blicq[34,35,23], V. Joshi[24]*, J.T. Linnemann[1], A.L. Longinotti[16], G. Luis-Raya[25], K. Malone[12], V. Marandon[23], O. Martinez[20], J. Martínez-Castro[19], J.A. Matthews[26], P. Miranda-Romagnoli[27], J.A. Morales-Soto[14], E. Moreno[20], M. Mostafá[6], A. Nayerhoda[18], L. Nellen[28], M.U. Nisa[1], R. Noriega-Papaqui[27], L. Olivera-Nieto[23]*, E.G. Pérez-Pérez[25], C.D. Rho[29], D. Rosa-González[17], E. Ruiz-Velasco[23], D. Salazar-Gallegos[1], F. Salesa Greus[18,33], A. Sandoval[13], H. Schoorlemmer[23,31]*, J. Serna-Franco[13], A.J. Smith[4], Y. Son[29], R.W. Springer[3], K. Tollefson[1], I. Torres[17], R. Torres-Escobedo[32], R. Turner[2], F. Ureña-Mena[17], L. Villaseñor[20], X. Wang[2], I.J. Watson[29], E. Willox[4], H. Zhou[32], C. de León[14], A. Zepeda[30] (HAWC Collaboration), A. Donath[36]*, and S. Funk[24]

(Affiliations can be found after the references)

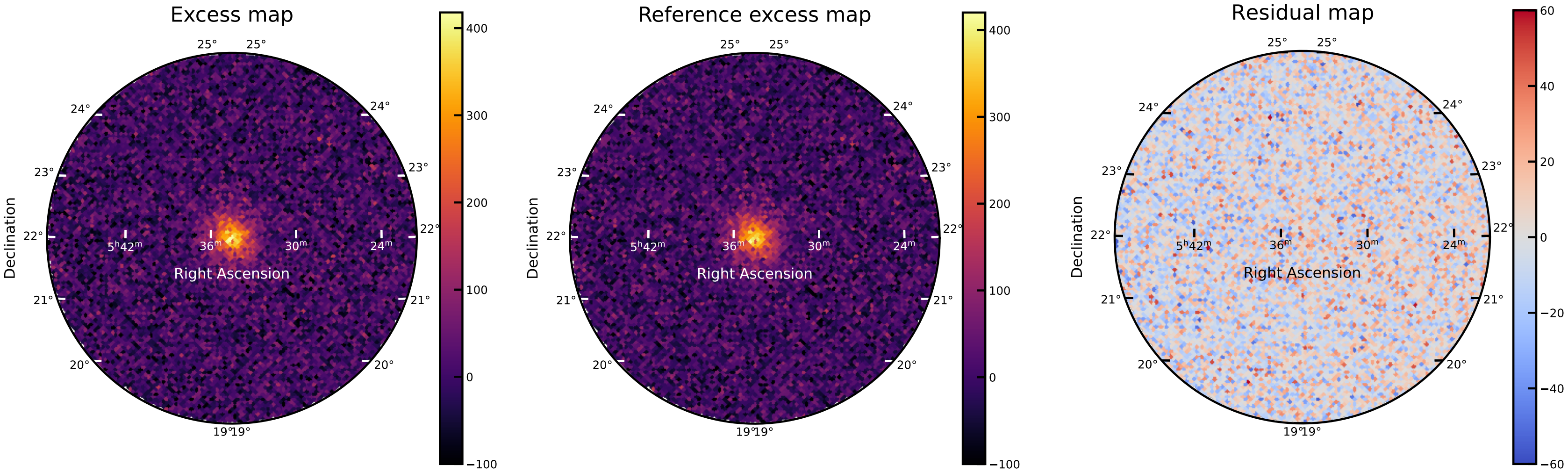
HAWC analysis in Gammapy – low level checks

- Checks of the background model: significance outside of exclusion mask



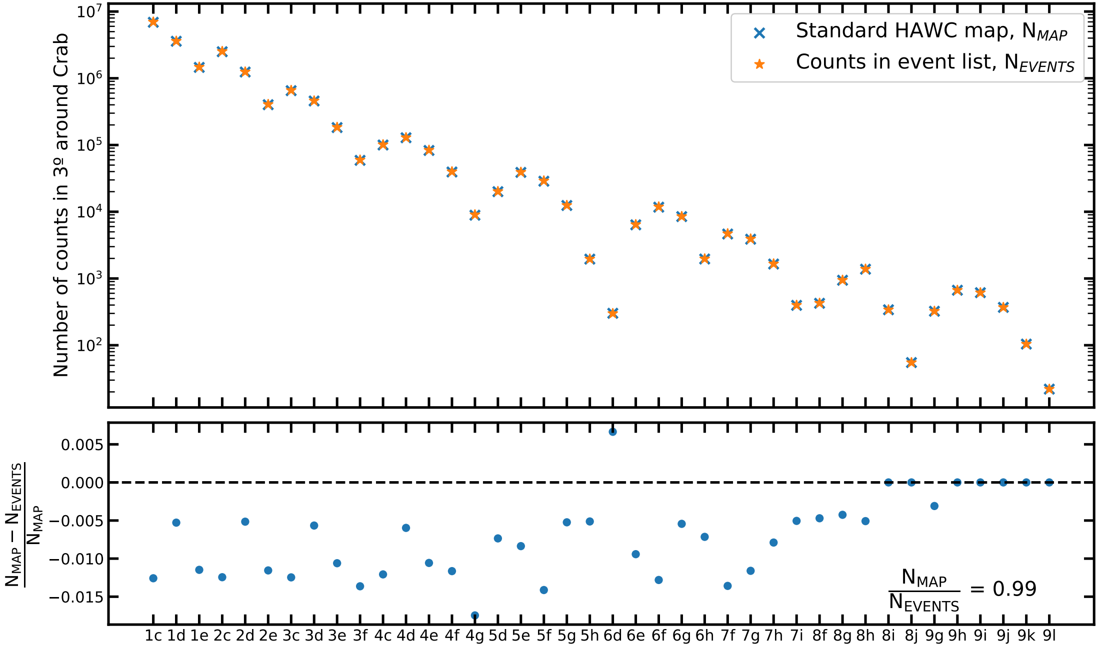
HAWC analysis in Gammapy – low level checks

- Checks of the event lists production: compare Crab counts with standard HAWC maps



HAWC analysis in Gammapy – low level checks

- Checks of the event lists production: compare Crab counts with standard HAWC maps

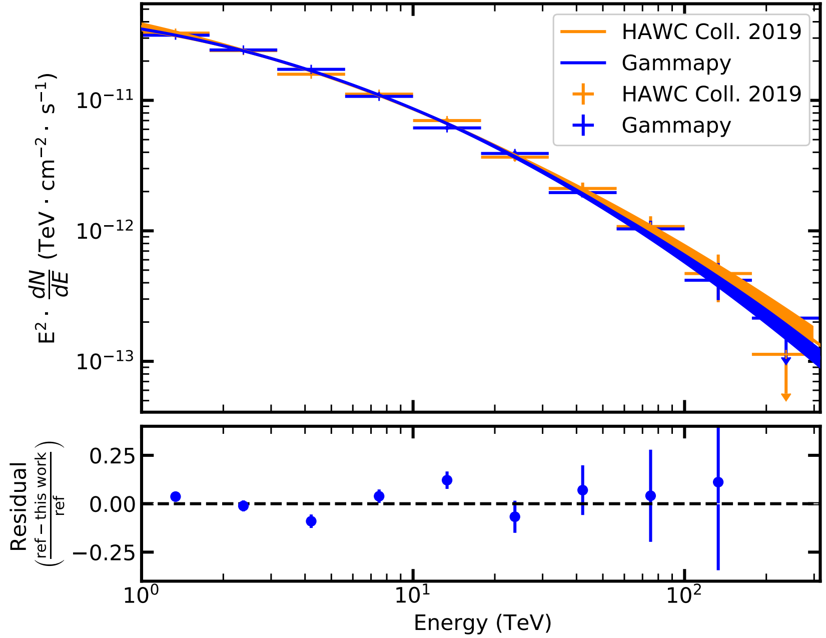


HAWC analysis in Gammapy – point source: the Crab

- Fit a combined spatial+spectral model jointly between all the relevant fHit bins
- Use both the dataset derived from event lists and the standard HAWC HEALPix map

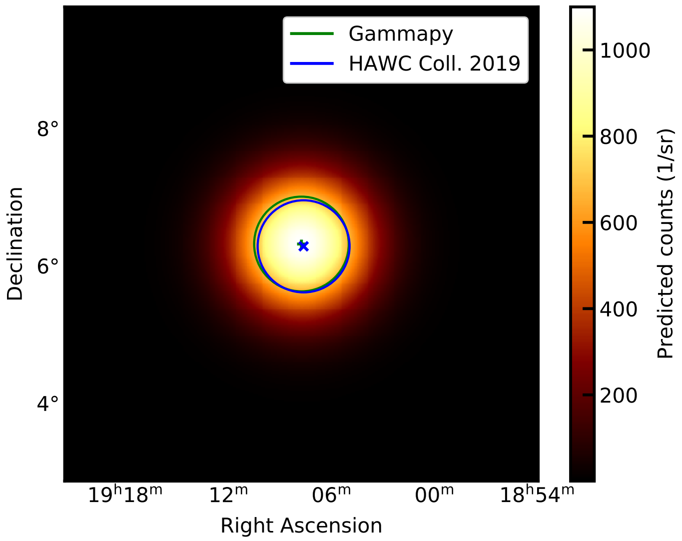
	ϕ_0 (10^{-13} TeV^{-1} cm^2 s^{-1})	α	β
From events	2.39±0.04	2.79±0.02	0.113±0.007
Reference	2.35±0.04	2.79±0.02	0.10±0.01
From map	2.35±0.05	2.79±0.02	0.12±0.01

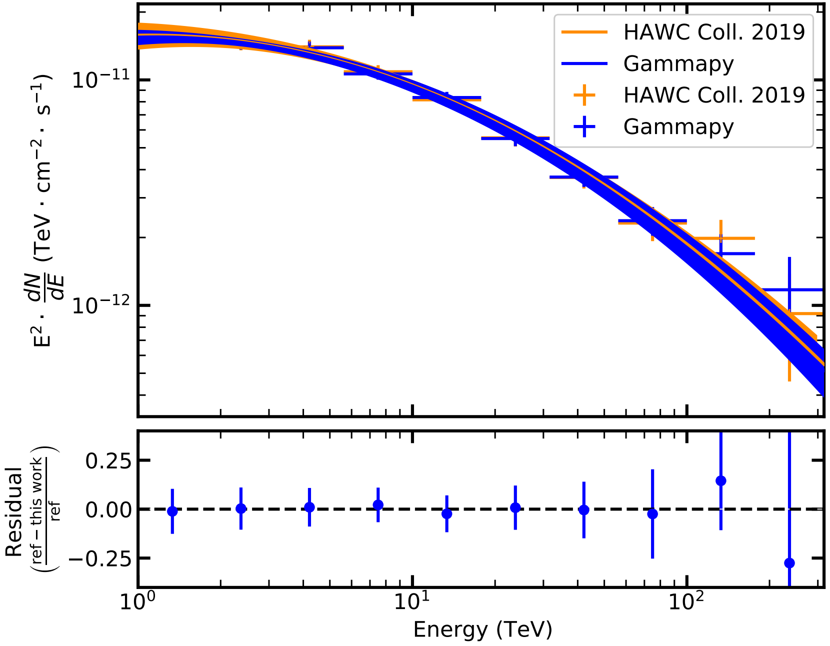
Table 4: Likelihood fit results for the Crab Nebula. The fit result obtained the DL3 products is given in the row labeled "From events". The fit result obtained using the standard HAWC map products is given in the "From map" row. The values in the "reference" column are taken from Abeysekara et al. (2019).



HAWC analysis in Gammapy – extended source: J1908+063

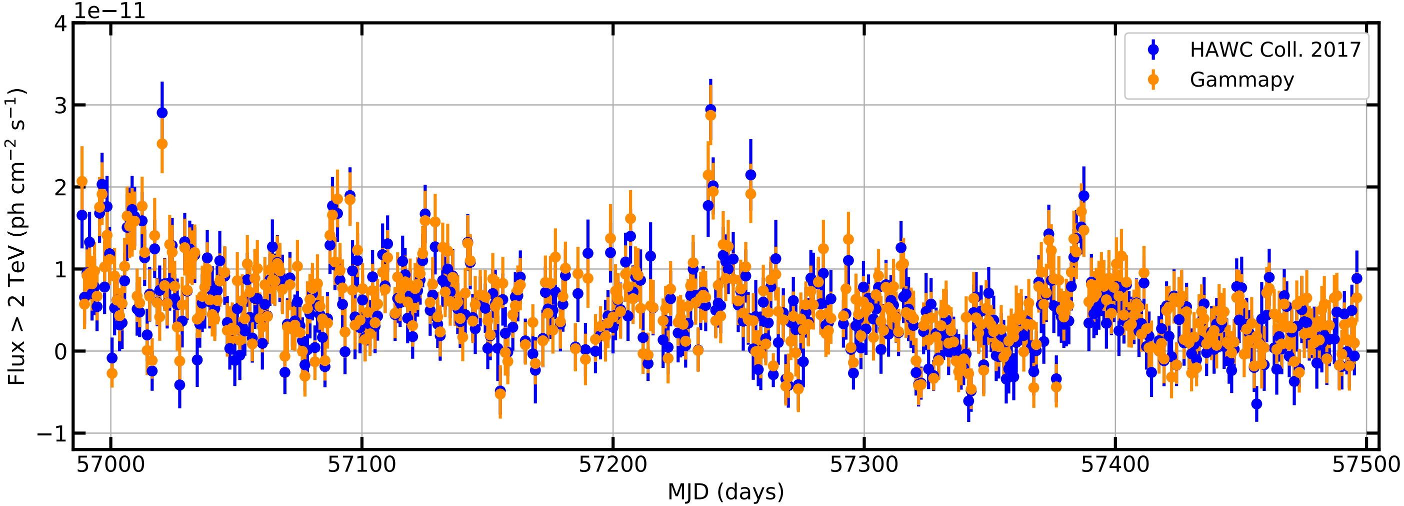
	R.A. (°)	Dec (°)	Extension (°)	ϕ_0 (10^{-13} TeV^{-1} cm^2 s^{-1})	α	β
From events	286.94±0.02	6.35±0.02	0.69±0.03	0.94±0.06	2.46±0.03	0.11±0.01
Reference	286.91±0.10	6.32±0.09	0.67±0.03	0.95±0.05	2.46±0.03	0.11±0.02
From map	286.96±0.03	6.36±0.03	0.68±0.03	0.94±0.06	2.45±0.04	0.12±0.02





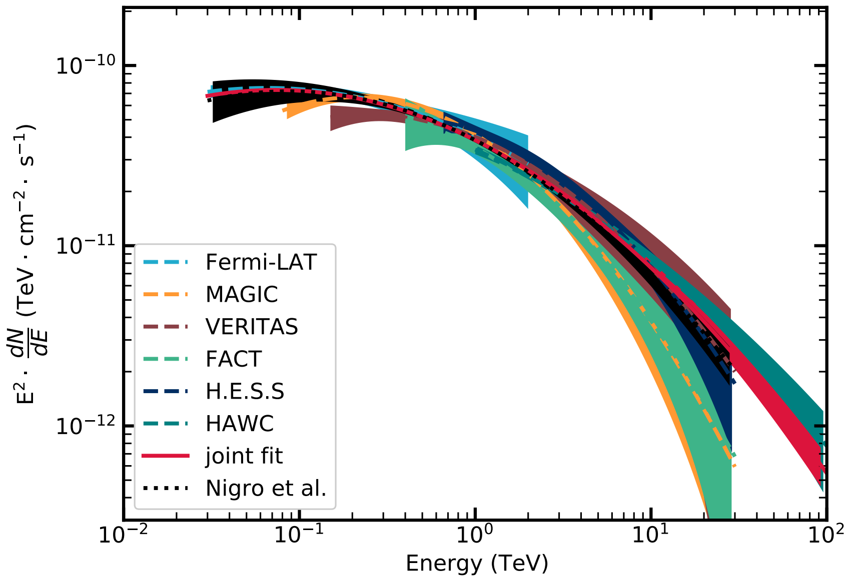
HAWC analysis in Gammapy – time domain: Mrk 421

- Light curve of a dataset **without energy estimators → also possible!**



HAWC analysis in Gammapy – joint fit

- Using public data from 5 other gamma-ray instruments
- Repeat the exercise in Nigro et al 2019 and perform a joint analysis of the Crab Nebula
- Most datasets very small, final result not the point → **proof of concept**
- The HAWC data and IRFs used for this plot **will be released publicly** when the paper is published (very very soon!)



HAWC analysis in Gammapy – what had to change?

- **PSF in reconstructed energy:** the HAWC PSF is computed as a function of reconstructed energy. This means it needs to be applied after the energy dispersion. This was implemented first as a flag on a model, and more recently, with the development of a new dedicated class

- **Pointing information not mandatory when reading (or using) events**

- **Support for IRFs given as a Map in the data reduction:** instead of projecting into the sky using a pointing position, map IRFs are interpolated into the relevant geometry

- **Increased support for HEALPix binning scheme:** allow the computation of predicted counts from a model in the HEALPix scheme

HAWC analysis in Gammapy – what is missing?

- **Full HEALPix support:** At the moment, the analysis is only fully supported for small (<10º) regions, and some of the methods don't support HEALPix geometries. This requires some maneuvering between WCS and HEALPix geometries, which can be tricky. Full-sky quantities are very important for drifting instruments, so this should improve.

- **Consistent event classes support:** the different bins are different "event classes" and should be treated separately. This works at the moment e.g. for fitting models but not, for example to compute a joint TS map

- Some of the functionality described here to integrate the zenith-binned IRF into declinations, compute exposure, etc are done using Gammapy but outside Gammapy. Given that these steps are necessary for a drifting instrument analysis, they could/should become a part of the data reduction process.

- (related to above) **Develop transit analysis support**

Outlook

- HAWC data analysis in Gammapy is possible and yields the same results as the original HAWC tools.

- Very few changes to the pipeline were needed for this, although some workarounds are still required

- None of the further required changes are a conceptual roadblock: they just need work

- The HAWC public data release will facilitate this work by providing a test dataset

- The Southern Wide-field Gamma-ray Observatory (SWGO) is planning to use Gammapy as its main tool, and already producing IRFs using Gammapy

- A shared tool has many advantages: large developer and user base, easy data sharing, easy joint analysis, no need to learn many different tools…

- The standard format is also further developing, now with dedicated input from drifting instruments