# Joint Likelihood Fits for the Study of Galactic Objects with HAWC

**Henrike Fleischhack**[*]

*Michigan Technological University*

*E-mail:* hfleisch@mtu.edu

**Petra Huentemeyer**

*Michigan Technological University*

**for the HAWC collaboration** [†]

Studying gamma-ray emission by Galactic objects is key to understanding the origins and acceleration mechanisms of Galactic cosmic ray electrons and hadrons. The HAWC observatory provides an unprecedented view of the gamma-ray sky at TeV energies and is particularly suited for the study of Galactic objects. However, the interpretation of the measured data poses several challenges. The high density of sources and source candidates can cause source confusion and make it harder to disentangle the origin of the emission. The relatively low angular resolution of HAWC, compared to instruments in optical or radio wavelengths, can further cause the emission of neighboring sources to bleed into each other or even make them look like one extended source. On the other hand, with its wide field of view, HAWC is uniquely suited for the study of extended sources. However, this requires the simultaneous modeling of both their morphology and emission spectrum. Joint likelihood fits to data taken over a larger range of energies can help overcome these challenges and achieve the full potential of the HAWC detector. In this presentation, we will discuss how systematic uncertainties related to joint likelihood fits can affect the measurements.

---

[*]Speaker.

[†]See http://www.hawc-observatory.org/collaboration/icrc2017.php for full author list.

## 1. Introduction

The HAWC detector has been observing the very-high energy (VHE, $E > 100\,\mathrm{GeV}$) sky for more than two years. The first catalog, comprising 39 sources, was released in early 2017 [1]. With a large dataset collected, HAWC is starting to become systematics-limited for some studies. In this paper, we will investigate two sources of systematic uncertainty: the correlation between assumed source location and flux measurements as well as the uncertainty in flux measurements due to incomplete modeling of the point spread function (PSF).

## 2. The HAWC Detector

The High Altitude Water Cherenkov Observatory (HAWC) [1] is a wide-field-of-view gamma-ray detector located at an altitude of 4.1 km in the volcán Sierra Negra, Mexico. It is sensitive to gamma-ray showers above 300 GeV and is able to observe $\gamma$-ray sources at declinations between $-20°$ and $+60°$, with an instaneous field-of-view of nearly 2 sr.

HAWC consists of 300 water cherenkov detectors (WCDs), each instrumented with four photomultiplier tubes (PMTs). Charged particles in extensive air showers emit Cherenkov light in the WCDs, which is detected by the PMTs. The readout system uses discriminators with two different thresholds, with the lower (higher) threshold corresponding to a pulse height of about 0.25 (4.0) photo-electrons. When a readout is triggered, usually requiring 28 PMTs to be hit within a time window of 150 ns, the timestamps of the threshold crossings for each PMT are written to disk.

With the full detector online and current trigger settings, the array trigger rate is about 20 kHz. The event rate is dominated by cosmic-ray showers by several orders of magnitude, even for strong gamma-ray sources. Most of this background can be removed later during data analysis using cuts on the shape of the lateral distribution function of the shower.

## 3. Data Analysis

### 3.1 Event Reconstruction

The event reconstruction proceeds in several steps. First, the deposited charge and arrival time in each PMT are calculated from the threshold crossing times.

The shower core position and the shower direction are calculated in an iterative procedure. The shower core is found by fitting the measured charge distribution with a simplified NKG function. The first iteration of the fit uses a simple estimate of the center-of-mass as a starting value for the core position.

The arrival direction is found by fitting a plane to the shower arrival times in each PMT. The arrivai times are corrected for the shower curvature, which causes particles far from the shower core to be delayed due to a longer traveling distance, and for the fact taht PMTs that see a high rate tend to measure an earlier arrival time than PMTs that see a low rate of particles (sampling bias).

2

The core and direction fits are repeated, using the results from the previous fit as starting values and restricting the PMTs used to hits recorded within 50 ns of the previously fitted shower plane.

HAWC currently uses two photon/hadron separation parameters, which are both related to the shape of the lateral distribution function.

The HAWC analysis currently uses the fraction $f_{hit}$ of hit PMTs relative to the number of available, i.e. installed and working PMTs as an energy proxy. As the detector response (e.g., point spread function and detector efficiency) depends on $f_{hit}$, events are binned according to $f_{hit}$ for the following analysis steps. Photon/hadron cuts are optimized separately for each bin to yield the maximum significance on the Crab Nebula. More details on the event reconstruction, $f_{hit}$ bins, and photon/hadron cuts can be found in [2].

For the results presented here, we used 706 days of HAWC data. The recorded events were binned according to $f_{hit}$ and filled into maps using the HealPix pixelation scheme with $N_{side} = 1024$.

## 3.2 Likelihood Analysis

The 3ML framework[1][3] was used for data analysis of the HAWC data. It relies on the HAWC likelihood analysis framework *LiFF* [4]. For the likelihood analysis, the predicted counts from a source plus background model in each bin, accounting for the detector response, are compared to the observed events. The parameters of the source model are adjusted to maximize the likelihood function.

In order to probe the statistical significance of a source candidate, the likelihood of the best-fit source plus background model, $\mathcal{L}^{max}$, is compared to the background-only hypothesis, $\mathcal{L}^{BG}$. We use a test statistic defined as

$$TS = 2 \cdot \log\left(\frac{\mathcal{L}^{max}}{\mathcal{L}^{BG}}\right).$$

In case of nested models with one free parameter the square root of the test statistics corresponds to a gaussian significance. This is the case for the skymaps shown here, where for each pixel one pointsource was fit on top of the background with the frewe normalization free, but fixed spectral shape.

The source model generally consists of a spatial model and a spectral model. The models are convolved with the appropriate detector response functions and the exposure time to calculate the predicted number of signal events per bin. The number of background counts in each bin is estimated using the *direct integration* method [5].

Skymaps (flux maps or significance maps) are created via a putative source search: For each pixel in the map, a single point-source fit is performed with the location of the point source centered on the pixel in question. The spectrum is fixed to a power law with a pivot energy of 7 TeV and an index of $-2.63$ (Crab-like). The only free parameter is the flux normalization. The flux map is produced from the best-fit flux normalization values.

---

[1]`https://github.com/giacomov/3ML`

The significance value for each pixel is calculated from the test statistics comparing the best-fit model to the background-only hypothesis.

### 3.3 Detector Resolution

The likelihood analysis described above requires a model of the detector response, in order to calculate the number of expected counts in each spatial and $f_{hit}$ bin. In particular, it is necessary to know HAWC's point spread function (PSF) to predict how far from the source a gamma-ray event may be reconstructed. HAWC's PSF depends on the fraction of PMTs hit as well as on the source declination. It is determined from simulations of gamma-ray showers, which are reconstructed in the same way as data, and can be described by a double Gaussian function.

HAWC's angular resolution improves dramatically with the fraction of PMTs hit. For the Crab Nebula and other sources at $22°$ declination, the angular resolution ($68\%$ containment radius of the PSF) is about $1°$ in the first $f_{hit}$ bin (between $6.7\%$ to $10.5\%$ of PMTs hit) and drops to $0.17°$ in the highest $f_{hit}$ bin (more than $84\%$) of PMTs hit) [2]. The predicted angular resolution from simulations was compared to the width of the gamma-ray emission from the Crab Nebula, whose true extent is assumed to be much smaller than HAWC's PSF. The two agree within the statistical uncertainties on the measured PSF, which correspond to $5\%$ to $10\%$ [2].

The absolute pointint of the HAWC detector was calibrated with the known position of the Crab Nebula, and cross-checked using the known positions of two other bright point sources, the blazars Mrk 421 and Mrk 501. The residual uncertainty in the absolute pointint was estimated to be about $0.1°$ [2].

The standard HAWC analysis currently uses the fraction of PMTs hit as an energy proxy. The resolution of this energy proxy is very poor and there is large overlap in the energy distributions of the different $f_{hit}$ bins. Still, there is sufficient correlation between $f_{hit}$ and energy that HAWC is sensitive to the spectral shape via the $f_{hit}$ dependent likelihood analysis. Energy estimators with improved resolution have been developed and are currently being evaluated [6].

### 3.4 Source Injection Studies

In order to study the effect of an inadequately modeled PSF on the analysis results, the width of the HAWC PSF model was scaled up and down within the uncertainties mentioned above. Nine 'scaled' PSF models were produced, with scaling factors between 0.8 and 1.2. The nominal HAWC PSF is modeled as the sum of two Gaussians for each $f_{hit}$ bin. For the scaled models, both Gaussian width parameters were multiplied by the same scaling factor.

For each scaled PSF model, "fake" maps were produced by injecting point sources on top of the background (from direct integration) and applying Poissonian fluctuations. The spectrum of the injected source was fixed to the measured Crab spectrum from [2], with a log-parabola shape:

$$\frac{\mathrm{d}N}{\mathrm{d}E} = N_0 \cdot \left( \frac{E}{E_0} \right)^{-\alpha + \beta \cdot \log(E/E_0)},$$

with pivot energy $E_0 = 7\,\mathrm{TeV}$, flux normalization $N_0 = 2.512 \cdot 10^{-13}\,\mathrm{TeV^{-1}\,s^{-1}\,cm^{-2}}$, spectral index $\alpha = 2.63$ and curvature parameter $\beta = -0.15$.

Different offsets from the nominal position of the Crab Nebula were tried. For each value of the PSF scaling factor and each value of the source offset, twenty maps were produced with different random seeds for the fluctuations. A Crab-like source injected with the nominal PSF is detected with $\sqrt{TS} > 130$ compared to the background-only model.

## 4. Correlation Between Source Offset, PSF, and Measured Spectrum

The injected source maps produced according to the prescription above were fit with a point source model at the nominal position of the Crab Nebula and a log-parabola spectrum. The spectral curvature parameter $\beta$ did not show any significant dependence on either the source offset or the injected witdh. The resulting distributions of the best-fit flux normalization and spectral index can be seen in Fig. 1. Within the statistical uncertainties, the spectral fit does not change significantly if the source is offset by less than $0.1°$ from the nominal position. However, there is a dependence on the width of the injected source: For an injected PSF that is $10\,\%$ narrower than the PSF used for the analysis, the reconstructed spectrum is (on average) softer by $0.03$ and has a $6\,\%$ higher flux normalization. If the injected PSF is $10\,\%$ wider than nominal, the flux is underestimated and the index becomes harder. These effects are already accounted for in the systematic uncertainties. HAWC assesses a $20\,\%$ uncertainty on the measured flux normalization and an uncertainty of $0.1$ on the spectral index due to incomplete modeling of the point spread function.

## 5. Spatial Residuals

For each injected source map, a point source model was fit to the injected source and the predicted counts from the best-fit source model were subtracted. Significance maps for these residuals were produced according to the procedure outlined in Section 3.2. Some example skymaps can be seen in Fig. 2.

As expected, if the PSF used for injecting the source is narrower than the one used for the analysis, a ring-like deficit appears around the source position (c.f. Fig. 2a). In contrast, if the injected PSF was wider than the one used for the analysis, a ring-like excess structure appears around the source (c.f. Fig. 2b). Small shifts in the source offset can cause a significant excess, which appears like a second point source (c.f. Fig. 2c). The flux from this 'extra' source corresponds to $2\,\%$ of the injected flux (see Fig. 2d).

A combination of a small source offset and a mis-modeled PSF can cause an even more significant excess to appear (c.f. Fig. 2e).
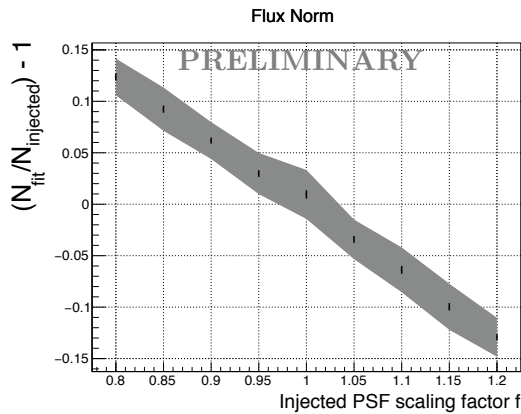
We also tested effects of a possibly asymmetric PSF. For this test, the PSF was reduced by $0.1°$ in each $f_{hit}$ bin and a 2-dimensional Gaussian source with width $0.1°$ and length
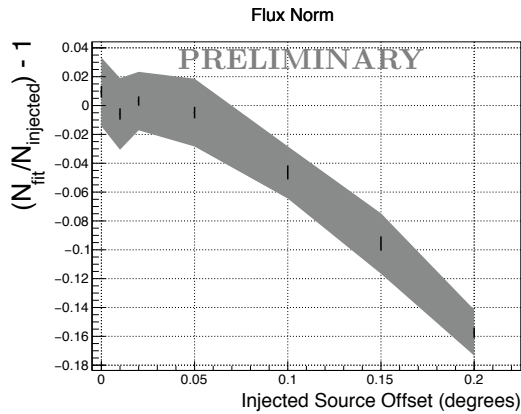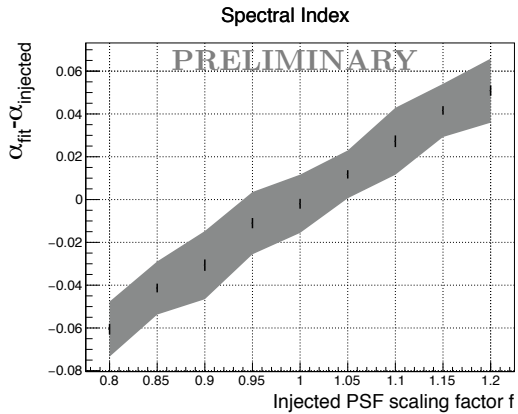
(a) Dependence of the flux normalization on the offset of the injected source.
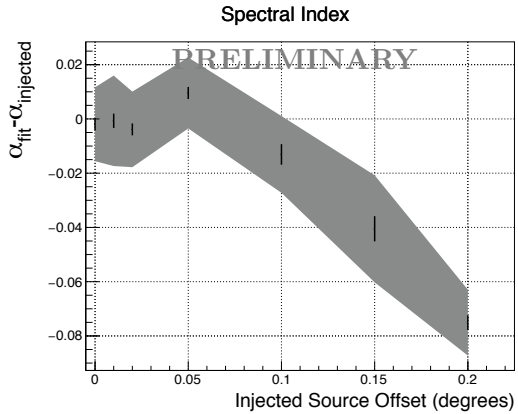


(b) Dependence of the spectral index on the offset of the injected source



(c) Dependence of the flux normalization on the width of the injected source.



(d) Dependence of the spectral index on the width of the injected source

**Figure 1:** Spectral parameters depending on the scaling factor applied to the PSF and the offset of the injected source from the nominal Crab position. Black bars: Statistical uncertainty on the mean, grey band: width of the distribution.

$0.1°$ and $0.2°$ respectively was injected. One or two significant excesses may appear (c.f. Fig. 2f).

In many example skymaps, the excesses cross the $5\sigma$ significance criterium usually used to claim discovery of a new source. The flux normalizations associated with these excesses correspond to a small percentage (up to $5\%$) of the injected source flux.

## 6. Conclusions

With the HAWC instrument providing unprecented coverage of the TeV gamma-ray sky, source modeling and source subtraction are important tools for the analysis, for example to disentangle complex regions such as the Cygnus region. Measured point source fluxes in the 2HWC catalog may have contaminations of up to $30\%$ from Galactic diffuse emission [1], and the study of diffuse emission with HAWC is currently under way.
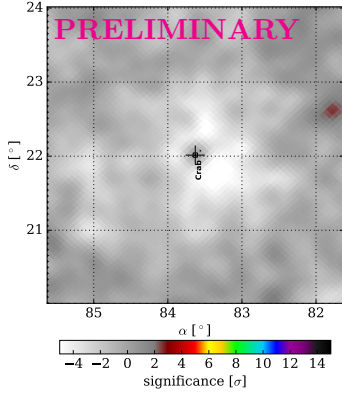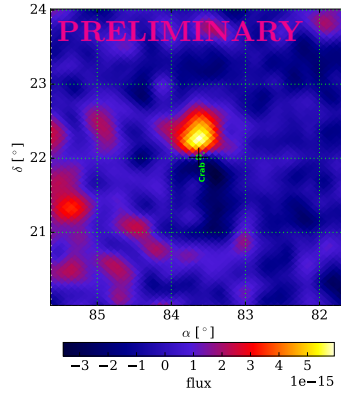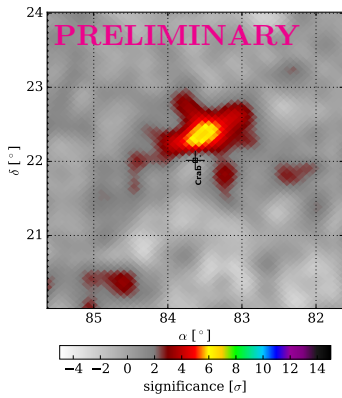
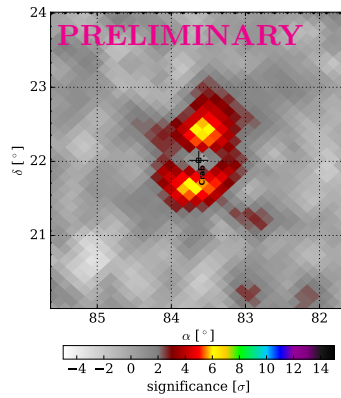(a) PSF scaled by 0.85.
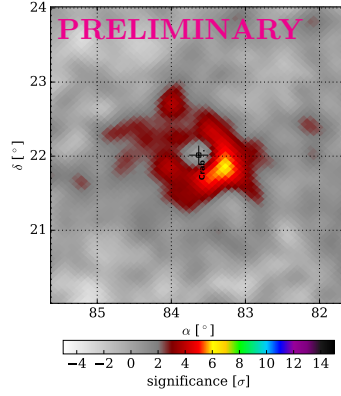
(b) PSF scaled by 1.15.
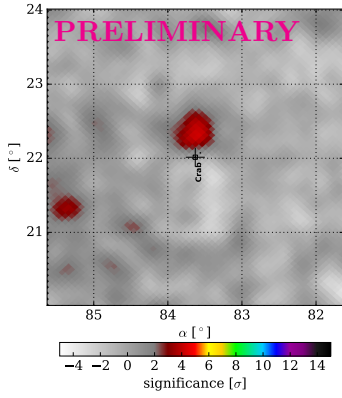
(c) Source offset by 0.1°.

(d) Source offset by 0.1°.

(e) Source offset by 0.1° and PSF scaled by 1.1.

(f) Source elongated by 0.1°.

**Figure 2:** Significance and flux maps of the residuals after subtracting the best-fit point source model from the injected source. The flux values correspond to the differential flux at 7 TeV, for a power-law spectum with index $-2.63$.

All these studies rely on modeling the response of the HAWC detector to gamma-ray emission. In particular, knowledge of the point spread function is crucial. We have shown here that a systematic error of 10 % in the angular resolution has a measurable effect on the fitted spectral parameter for strong sources like the Crab Nebula, but these effects are accounted for in the systematic uncertainties.

Additionally, inadequate modeling of the PSF can cause excesses or deficits to appear in the excess maps after subtracting the best-fit source model predictions. For the current unceratinties in the angular resolution and pointing offset, these excess correspond to up to 5 % of the subtracted source. For strong sources like the Crab Nebula, event this relatively small fraction can lead to significant ($> 5\sigma$) excesses.

Thus, all studies that rely on source subtraction or fitting multiple sources close together should take these effects into account. Efforts to improve the angular resolution and the modeling of the PSF are currently under way, which will improve the performance of the HAWC observatory, especially regarding studies of diffuse emission or of regions with overlapping sources.

## Acknowledgments

## References

[1] A. U. Abeysekara et al., *The 2HWC HAWC Observatory Gamma-Ray Catalog*, *The Astrophysical Journal* **843** (2017) 40.

[2] A. U. Abeysekara et al., *Observation of the Crab Nebula with the HAWC Gamma-Ray Observatory*, *The Astrophysical Journal* **843** (2017) 39.

[3] G. Vianello, R. J. Lauer, P. Younk, L. Tibaldo, J. M. Burgess, H. Ayala et al., *The Multi-Mission Maximum Likelihood framework (3ML)*, *PoS* **ICRC2015** (2016) 1042.

[4] P. W. Younk, R. J. Lauer, G. Vianello, J. P. Harding, H. A. Ayala Solares and H. Zhou, *A high-level analysis framework for HAWC*, *PoS* **ICRC2015** (2016) 948.

[5] R. Atkins et al., *Observation of TeV Gamma Rays from the Crab Nebula with Milagro Using a New Background Rejection Technique*, *Astrophys. J.* **595** (2003) 803.

[6] S. Marinelli et al., *Measuring High-Energy Spectra with HAWC*, *these proceedings* (2017) .