INVESTIGATING ICECUBE NEUTRINO ALERTS WITH THE HAWC γ-RAY OBSERVATORY

By

Ian Herzog

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics—Doctor of Philosophy

2026

ABSTRACT

Across the universe, highly energetic astrophysical sources release particles over timescales ranging from milliseconds to millennia. Gamma-ray and multi-messenger astronomy offer a powerful window into these phenomena, with the potential to shed light on some of the deepest unresolved questions in physics. The High Altitude Water Cherenkov (HAWC) gamma-ray observatory, an extensive air shower array with a large field of view and high duty cycle, is particularly well-positioned for both steady-state and transient searches. Two such searches are presented in this work: a complete analysis of the steady-state complex Cygnus Cocoon at TeV gamma ray energies to isolate a probable pulsar wind nebula called HAWC J2031+415 and a transient search of HAWC using public IceCube neutrino alerts. The former includes both a multi-source fit using 2400 days of HAWC data and then a multi-wavelength fit using radio and x-rays to conclude that HAWC J2031+415 is most certainly a pulsar wind nebula. The latter work combines HAWC and IceCube data to perform a multi-messenger search of HAWC data to search for potential coincidences. Approximately 25 are see, which corresponds to an expected detection rate of 5%. Both these projects contribute to understanding the origin of cosmic rays and what acceleration mechanisms sources use to produce the ultra high energies observed at Earth.

This thesis is dedicated to someone.

ACKNOWLEDGEMENTS

Your acknowledgements here.

funding: LANL and DOE, Kirsten HAWC collaboration members msu colleagues friends family

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC Akaike Information Criterion

AGN Active Galactic Nuclei

BBA Bayesian Block Algorithm

BIC Bayesian Information Criterion

HAWC High-Altitude Water Cherenkov Detector

DoG Difference of Gaussian

IC IceCube

DBE Diffuse Background Emission

LHAASO Large High Altitude Air Shower Observatory

Mrk Markarian

MTU Michigan Technological University

MSU Michigan State University

PWN Pulsar Wind Nebula

PSR Pulsar

VERITAS Very Energetic Radiation Imaging Telescope Array System

Fermi-LAT *Fermi-Large Area Telescope*

HEGRA High-Energy-Gamma-Ray Astronomy Observatory

IACT Imaging Atmospheric Cherenkov Telescopes

PS Point Source

EXT Extended Source

ROI Region of Interest

eV electron Volt

MeV Megaelectron Volt

GeV Gigaelectron Volt

TeV Teraelectron Volt

URM Unresolved Emission

ISM Interstellar Medium

TDCs Time to Digital Converters

FEB Front-End Board

DAQ Data Acquisition

HOMER HAWC Observatory Monitoring for Experiment and Reconstruction

SFCF Super Fast Core Fitter

NKG Nishimura-Kamata-Greisen

NN Neural Network

GP Ground Parameter

f_{hit} Fractional hit

DOM Digital Optical Modules

AMON Astrophysical Multimessenger Observatory Network

SWGO Southern Wide-field Gamma ray Observatory

HEP High Energy Physics

DOE Department of Energy

LANL Los Alamos National Lab

OBR Optical Backscatter Reflectometry

OFDR Optical Frequency Domain Reflectometry

ATHENA Advanced Tracking HAWC Experiment Notifications and Alerts

UMD University of Maryland

CHAPTER 1

AN INTRODUCTION TO HIGH-ENERGY ASTROPHYSICS

The study of particle astrophysics can help model the high-energy universe where the most extreme phenomena exist. There are many different ways that particles can be accelerated from the high to ultra-high-energy regimes, with many sources often combining mechanisms to create an energy spectrum spanning decades of energy. Sources can also exist across a wide range of time scales, from Gamma-Ray Bursts (GRBs) lasting seconds to SuperNova Remnants (SNRs) existing for tens of millions of years. Regardless, all sources at this energy range accelerate charged nuclei, primarily hydrogen ions, called cosmic rays. As will be discussed in Section 1.1.1, the origin of cosmic rays is quite challenging to ascertain, often requiring secondary interactions to be properly modeled. As such, even though the field of particle astrophysics has existed for over a 100 years, comparatively little is known about the origin of cosmic rays [29].

One known characteristic of cosmic-ray interactions is that one or several messenger particles can be created. Studies combining two or more of these messenger particles are called multi-messenger astronomy and will be discussed in Section 1.4. Combining observations from different detectors can aid in the search for the origin of cosmic rays and hint at their maximum potential energy.

In this work, gamma-ray astrophysics will be used to identify and analyze several different sources in the energy range of 300 GeV (1e9 electronvolts) to 300 TeV (1e12 electronvolts). First, a complex emission region known as the Cygnus Cocoon will be modeled (see Chapter 4). Next, gamma-ray and neutrino observations will be combined to perform a multi-messenger search across the sky to classify and constrain sources (see Chapter 5). Lastly, in Appendix B, a new search method will be introduced to construct a catalog of potential gamma-ray sources using image processing techniques.

1.1 Gamma-ray Astronomy

Before gamma-ray astronomy came the analysis of cosmic rays. Cosmic rays are high-energy charged particles that interact with Earth's atmosphere. They were first discovered in 1912 by

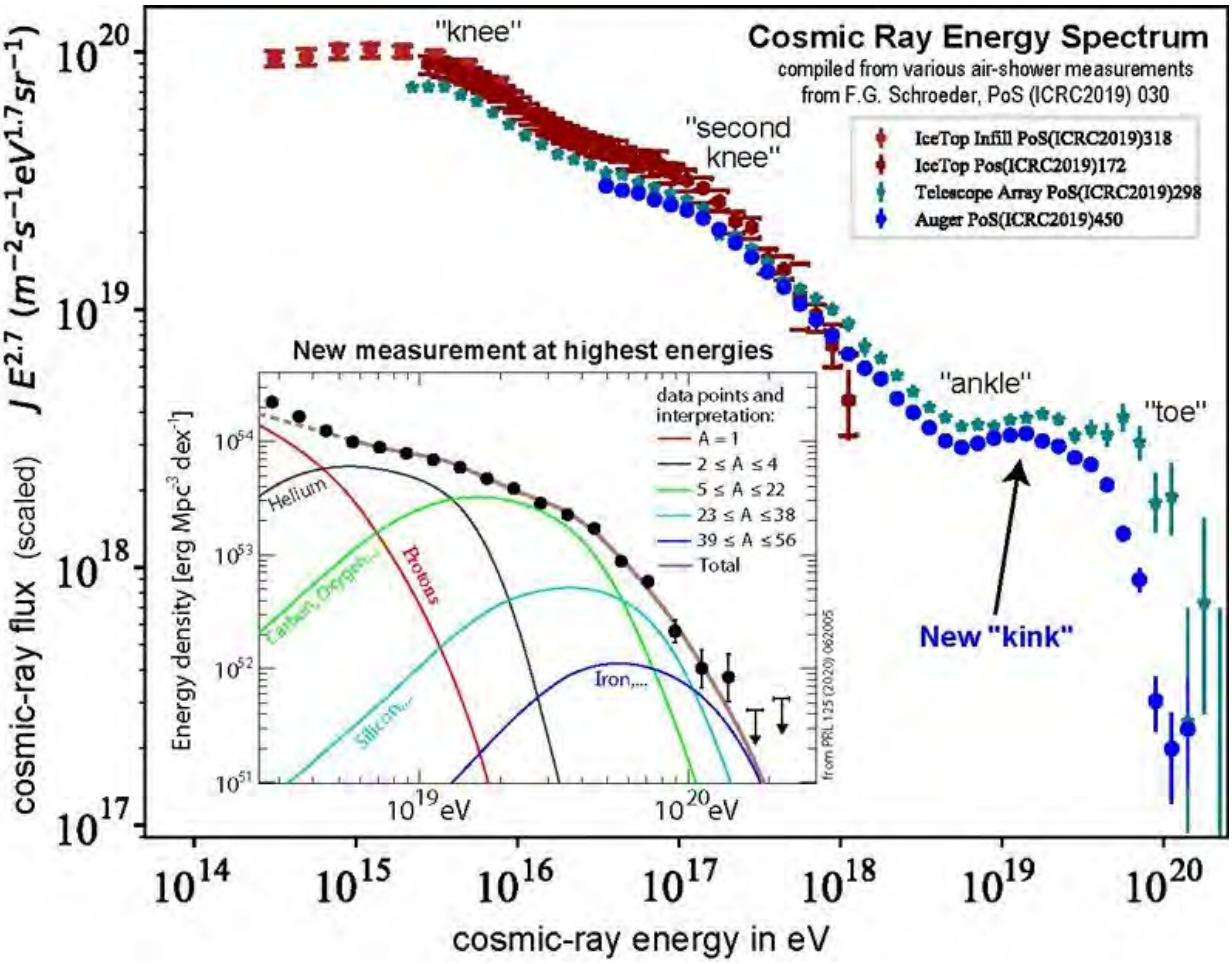

Figure 1.1 The high-energy component of the cosmic-ray spectrum. Note that the y-axis has energy scaled by $E^{2.7}$ (the expected spectrum) to emphasize subtle changes or structures that might be present. The primary particle for energies up to $\approx 10^{18}$ eV is the proton, with heavier nuclei contributing to the highest energies. Figure is from [4].

Victor Hess when he took a balloon high into the sky. He noticed that the radiation level observed increased with altitude, confirming that these particles were coming from outer space [30].

1.1.1 Cosmic Rays

Cosmic rays are primarily hydrogen ions (protons), but can have heavier nuclei like helium (beta particles) mixed in. Additionally, about 1% of cosmic rays are electrons. Of the remaining 99%, 87% are H^+, 12% are He^{2+}, and 1% are heavier nuclei [29]. In terms of energy, clear observations have been made from 1 GeV to 10 EeV (1e18 electronvolts). Broadly speaking, the spectrum follows a power law spectrum, shown below

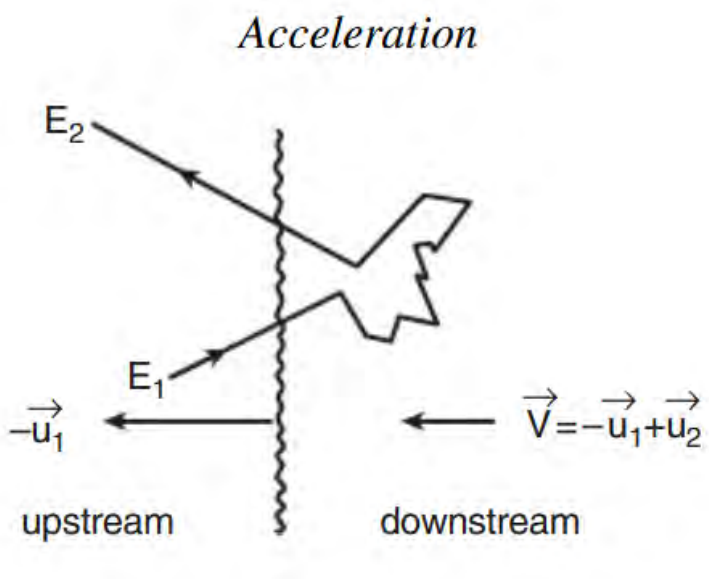

Figure 1.2 A diagram showing diffusive shock acceleration (DSA). The shock wave moves with speed $-u_1$, shocking all gas it encounters (the downstream region). As it encounters unshocked gas, it compresses and shocks it in the upstream region. The figure is from [5].

$$\frac{dN}{dE} = N_o \left(\frac{E}{E_p}\right)^{-\alpha} \tag{1.1}$$

where $\frac{dN}{dE}$ is the flux observed at Earth, N_o is the flux normalization (the flux at the source), E_p is the pivot energy, and α is the index of the spectrum. In log-log space, a power law spectrum appears as a straight line with slope equal to the index.

For the cosmic-ray spectrum, it roughly follows a power law with $\alpha = -2.7$ up to around 100 TeV [4]. From 100 TeV to 100 EeV, several structures appear in the spectrum, as seen in Figure 1.1. These structures correspond to index changes, occurring at 10^7 GeV (knee), 10^9 GeV (second knee), and at 10^{10} GeV (ankle). The spectrum rapidly decays around 10^{20} eV due to the Greisen-Zatsepin-Kuzmin (GZK) cutoff, where the cosmic rays begin interacting with the cosmic microwave background (CMB) radiation [31]. Tt is believed that galactic sources like the sun [32] produce rays up to the first knee while extragalactic sources produce most of the rays above the first knee [6].

1.1.1.1 Diffusive Shock Acceleration

To achieve these energies, cosmic rays are accelerated in extreme environments through a process called Diffusive Shock Acceleration (DSA). Derived fully in section 2.8 of [33] and section 12.1 in [5], DSA is the process of accelerating a given particle by having it pass through multiple shock fronts, gaining energy with each pass. A shock front is defined as an abrupt transition in temperature, pressure, or density of a given material's medium. A particle can become trapped in the strong magnetic fields present and be pulled repeatedly across different shock fronts before it eventually achieves enough energy to escape.

Consider a particle trapped inside a shock front like the scenario in Figure 1.2. Assuming an initial energy E_0 and fractional energy increase ζ with each pass, after n times, a particle's energy is defined as

$$E_n = E_0(1 + \zeta)^n. \tag{1.2}$$

If a particle has escape and containment probabilities P_{esc} and $1 - P_{\text{esc}}$, then after n encounters, the probability of a particle remaining inside the acceleration region is $(1 - P_{\text{esc}})^n$. To achieve a given energy E, a particle needs to achieve n encounters determined by

$$n = \ln\left(\frac{E}{E_0}\right)/\ln(1 + \zeta) \tag{1.3}$$

It follows that the proportion of particles with energies greater than E is given by finding all encounters with $m \geq n$:

$$N(\geq E) = \sum_{m=n}^{\infty}(1 - P_{\text{esc}})^m = \frac{(1 - P_{\text{esc}})^n}{P_{\text{esc}}} = \frac{1}{P_{\text{esc}}}\left(\frac{E}{E_0}\right)^{-\gamma} \tag{1.4}$$

with

$$\gamma = \ln\frac{1}{1 - P_{\text{esc}}}/\ln(1 + \zeta) \approx \frac{P_{\text{esc}}}{\zeta} = \frac{1}{\zeta} \times \frac{T_{\text{cycle}}}{T_{\text{esc}}} \tag{1.5}$$

This acceleration process gives the observed spectrum for cosmic rays. Equation 1.5 gives the characteristic acceleration time T_{cycle} for a given particle at energy E and the characteristic escape time T_{esc}. The ratio of these two gives the probability for a particle to escape per encounter. This implies two conclusions: higher energy particles take longer to accelerate (and to be observed) and that there is an energy cap for a given source. There are methods to increase a particle's potential energy (eg: see Chapter 12 in [5]), but those applications are beyond the scope of this work. Regardless, these two conclusions will be expanded further in Section 1.2 where gamma-ray sources are introduced.

1.1.2 Gamma Rays

While the cosmic-ray spectrum can be modeled, finding the origin of these particles is challenging. As they are charged, they interact with magnetic fields both to accelerate inside emission regions (see above) and with weaker background fields. These twist and bend particle paths, rendering it difficult or impossible to determine an origin. However, processes that accelerate cosmic rays also produce gamma rays and, as these are neutral photons, they are not curved by magnetic fields and can be traced back to their origin. These photons are produced from cosmic-ray particle interactions within a source. The production mechanisms can be broken into two categories: hadronic (proton) and leptonic (electron). There is one hadronic process: neutral pion decay, and there are 3 leptonic processes: synchrotron radiation, bremsstrahlung, and inverse Compton scattering. There is one additional method for leptons, called synchrotron self-Compton, that combines the lower energy synchrotron and higher energy inverse Compton processes. Often, a source can utilize several processes to produce a full Spectral Energy Distribution (SED), spanning from μeV to TeV. An example of such an SED is given in Figure 1.3, adapted from [6].

1.1.2.1 Pion Decay

When a proton collides with another nucleus or proton, they produce both neutral and charged pions, among other products. There are three pions that can be produced: $\pi^{+/-}, \pi^0$ [34], with each having unique decay products. These products are given by

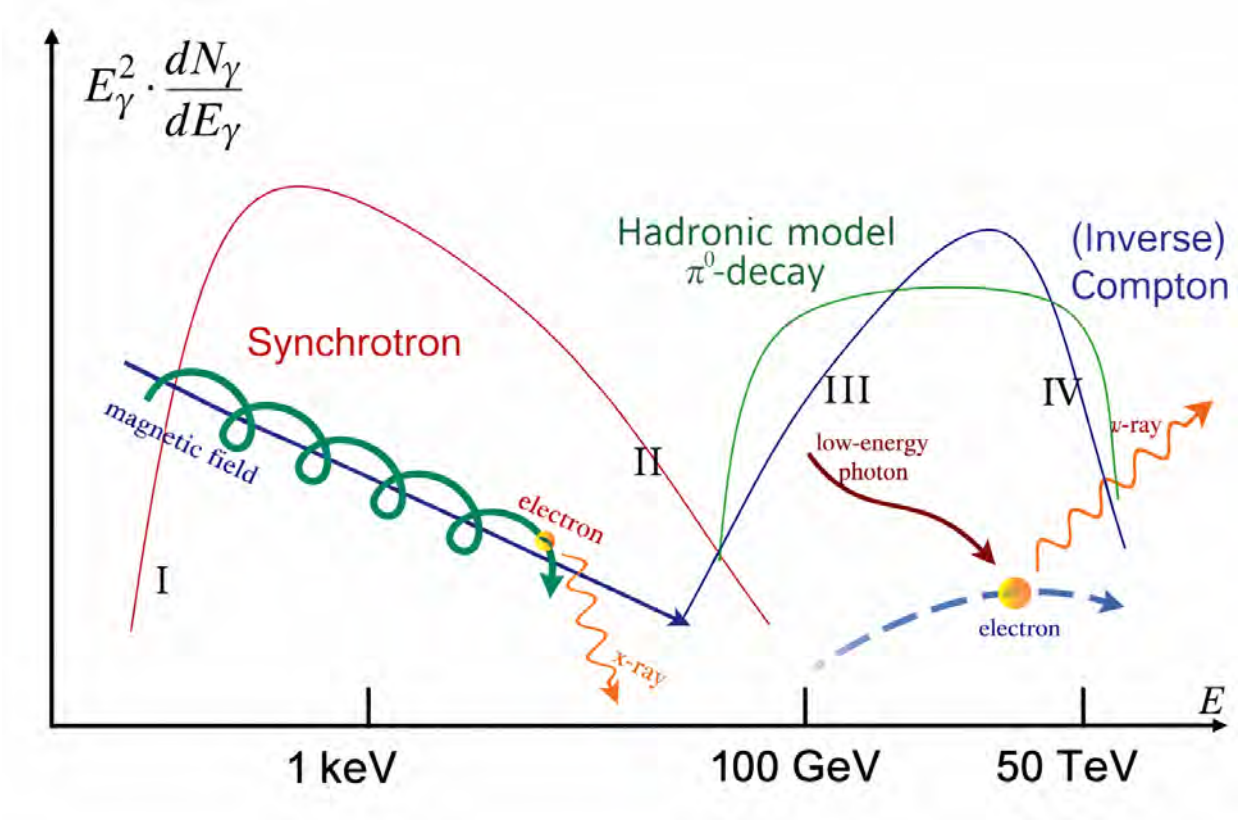

Figure 1.3 An example of a full energy spectrum for a potential source. At keV energies, the emission is dominated by the synchrotron process, with its lower end extending in μeV with radio sources. Once high MeV energies are reached, the hadronic pion decay (PD) steeply rises before plateauing while the leptonic inverse Compton (IC) has a more gradual increase before dropping off in the TeV region. Both IC and PD can extend into the 100's of TeV to, in extreme cases, PeV energies. The figure is from [6]

$$\begin{aligned} \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu \\ \pi^0 &\rightarrow \gamma + \gamma \end{aligned} \tag{1.6}$$

where the charged pions produce muons μ and neutrinos ν, while the neutral pions produce photons γ. In the rest frame of the pion, each photon has an energy of $E_\gamma = 67.5$ MeV but when boosted to the laboratory frame, this can be extended to TeV energies. This energy is given by

$$E_\gamma = \Gamma E_\gamma^* + \beta\Gamma p_\gamma^* \times \cos\theta^*, \tag{1.7}$$

where p_γ is the photon momentum, θ is the angle between the photons, $\Gamma = E_\pi/m_\pi$, $\beta = |p_\pi|/E_\pi$, and the $*$ indicates the values in the rest frame. Differentiating with respect to $\cos\theta^*$ and applying the chain rule gives the photon energy spectrum in the laboratory frame:

$$E_\gamma^2 \frac{dN}{dE} = \frac{1}{2\beta\Gamma p_\gamma^*}. \tag{1.8}$$

The E^2 comes from considering a source accelerating protons via DSA, which gives an index of 2. There are several characteristics of pion decay emission: it is only present in the 10's of MeV region due to the pion's rest mass, it rises very steeply in the low 100's of MeV before plateauing, and it traces the parent proton's distribution. These characteristics can be seen in Figure 1.3. This shape is often called a "pion bump" and is an immediate indication that the photons are primarily hadronic in nature. One note is that generally GeV-TeV data is needed to properly constrain the pion decay's index, as the pion bump can extend fully across both energies (further discussed in Chapter 4). Lastly, if a hadronic source extends to energies > 100 TeV, it is a potential PeVatron, or an accelerator capable of accelerating protons to over a PeV as there is a 1:9 ratio for photon:proton energies [35].

1.1.2.2 Synchrotron Radiation

Synchrotron emission occurs when an electron passes through a magnetic field. The electron will follow a helical path and radiate away its energy via photons, (see Figure 1.3). For a given energy E, an electron will emit a photon at a critical frequency v_c, given by

$$v_c = \left(\frac{E}{m_e c}\right)^2 \times \frac{eB}{2\pi m_e c} \tag{1.9}$$

where the magnetic field strength is given by B [6, 36]. As a charged particle, it will emit radiation with power P given by the Larmor formula

$$P = -\frac{dE}{dt} = \frac{2e^2}{3c^3}a^2 \tag{1.10}$$

with a being the particle's acceleration. Assuming circular motion, the relativistic acceleration of the particle can be substituted in, along with the Larmor radius, to give the energy loss due to synchrotron emission. This is given by

$$-\frac{dE}{dt} = \frac{(2EB)^2}{24\pi m_e^2 c}\sigma_T \tag{1.11}$$

with σ_T being the Thomson cross-section. If a power law distribution with normalization N_o and index α for the electron energies is considered,

$$\frac{dN}{dE}dE = N_o E^{-\alpha_e} dE, \tag{1.12}$$

then the flux density of the photons emitted from the electrons can be found. This is given by

$$E_\gamma^2 \frac{dN_\gamma}{dE_\gamma} = -\frac{dE}{dt}\frac{dN}{dE}dE = C \cdot E_\gamma^{-\alpha} B^{\alpha+1} \tag{1.13}$$

with $\alpha = (\alpha_e - 1)/2$ being the spectral index of a source. A given source's spectrum is also highly dependent on the magnetic field, with it often being fit with the source's spectrum to adequately constrain its emission [6, 37]. The synchrotron process can produce photons from μeV (radio) to keV (x-ray), with radio and x-ray data often being combined to constrain the magnetic field (see Chapter 4 for more details).

1.1.2.3 Bremsstrahlung

Bremsstrahlung, or braking radiation, occurs when a high-energy electron encounters a nucleus' electric field and slows down, emitting a photon in the process. This is shown in Feynman diagram form in Figure 1.4. The energy loss rate of the electrons is given by

$$-\frac{dE}{dt} = \left(\frac{cm_p n}{X_0}\right) E, \tag{1.14}$$

where m_p is the proton's mass, n is the number density of gas, X_0 is the radiation length (distance where the electron's energy falls to $1/e \approx 1/3$ its original value), and E is the electron's energy. The electron's lifetime can be calculated as

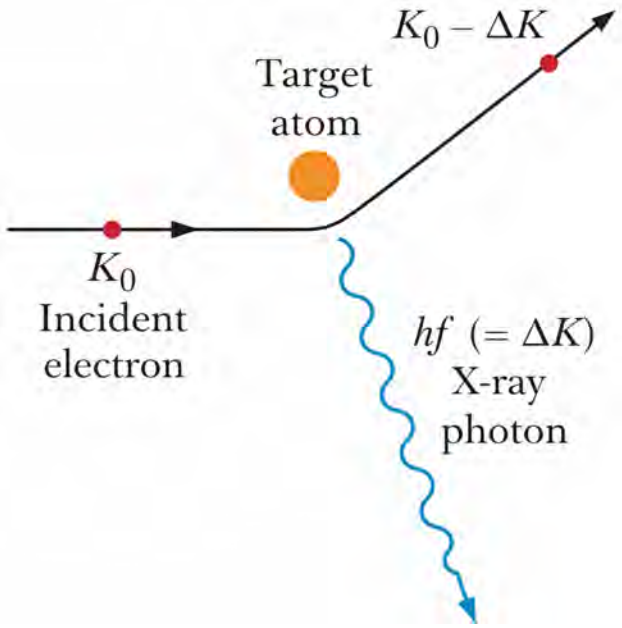

Figure 1.4 The bremsstrahlung process. An electron approaches an atom and has its path curved by the atom's electric field. This slows the electron down, emitting a photon in the process. Figure is from [7]

$$\tau = \frac{-E}{dE/dt} = 4 \times 10^7 (n/\mathrm{cm}^{-3})^{-1} \mathrm{yr}, \tag{1.15}$$

assuming the radiation length for hydrogen gas $X_0 = 60$ g/cm^2. It is clear that the energy loss due to bremsstrahlung is directly proportional its energy and will follow the same spectral shape. The result is a gamma-ray spectrum that will have the same index as the particle distribution [6]. While not shown in Figure 1.3, bremsstrahlung can be present from radio to GeV gamma rays, depending on the source.

1.1.2.4 Inverse Compton Scattering

If electrons are accelerated to high enough energies, they can scatter off background ambient photons from the CMB and boost the photons to GeV-TeV energies. This is known as inverse Compton (IC) scattering and is shown in Figure 1.5. Like with synchrotron emission, the IC photon index is given by $\alpha = (\alpha_e - 1)/2$. However, if the electrons become ultra-relativistic (like with IC), then they leave the Thomson regime and enter the Klein-Nishima (KN) regime. The KN cross-section is given by

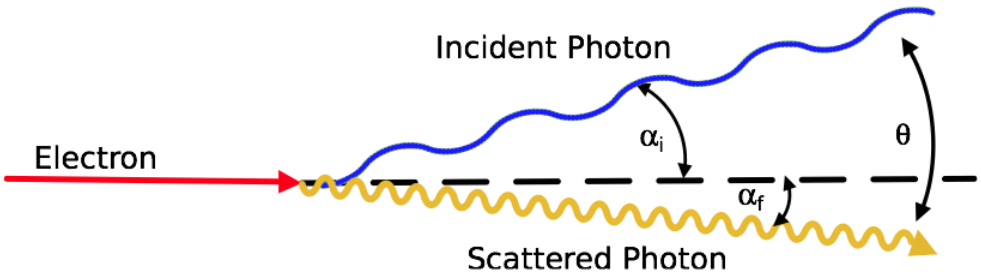

Figure 1.5 The inverse Compton scattering process. A high-energy electron collides with a low-energy photon. This collision boosts the photon's energy to the TeV range if the electron is at relativistic energies. The angles indicate the path that the photon travels before and after the collision. Figure is from [8].

$$\sigma_{\mathrm{KN}} = \frac{(\pi r_e)^2}{h\nu}\left(\ln(2h\nu) + \frac{1}{2}\right) \tag{1.16}$$

with h being the Planck constant, r_e is the electron radius, and ν is the photon frequency [6, 36]. In this regime, the SED becomes

$$E_\gamma^2 \frac{dN_\gamma}{dE_\gamma} \propto E_\gamma^{-\alpha_e} \ln E_\gamma \tag{1.17}$$

with the photon index being the same as the parent electrons. The typical spectral shape can be seen in the right of Figure 1.3.

1.1.2.5 Synchrotron Self-Compton

If there is a dense field of electrons interacting with a strong magnetic field, they will produce photons via synchrotron radiation. These photons then have the chance to interact again with high-energy electrons and then get up-scattered again with IC scattering. As the electrons are doing work twice, this form of emission is called Synchrotron Self-Compton (SSC) emission. [38]

It should be noted that this only becomes detectable if the electron and photon densities are sufficiently high, otherwise the SSC contribution to a source's SED will be sub-threshold [38, 36]. Broadly, this narrows the potential candidates to either very young or extremely energetic sources. An example of the former is the Crab Nebula, an incredibly bright and young pulsar wind nebula (PWN) that spans over 20 decades of energy, from radio to TeV gamma rays [1]. As for the latter,

some active galactic nuclei (AGN) emit via SSC [28]. Both source types will be discussed later in this chapter.

1.1.2.6 Extragalactic Background Light

As detectors have improved, we can peer further into the distant universe at increasingly higher energies. Near TeV energies, an interesting phenomenon was observed: something was suppressing TeV emission from distant extragalactic sources. These distances are measured in redshift, or how shifted the spectrum of a source becomes due to traveling through an expanding universe. This phenomenon is now known as Extragalactic Background Light (EBL) absorption, and it limits TeV gamma-ray experiments like the High-Altitude Water Cherenkov (HAWC) observatory to a range of $z = 0.3$, or around 4 billion light years from Earth [6, 28].

When a photon reaches a high enough energy E, it can interact with low energy (ϵ) background photons from the CMB to the near-ultraviolet (UV) wavelengths via pair production. This is shown by

$$\gamma_E + \gamma_\epsilon \rightarrow e^+ + e^- \rightarrow 2\gamma' \tag{1.18}$$

The two photons interact and produce an electron-positron pair, which then annihilate to produce two new photons at a much lower energy compared to the initial high-energy photons. The cross section for this process is a function of energy from the initial photons:

$$\sigma(E, \epsilon) \approx 1.25 \times 10^{-25} (\frac{(m_e c^2)^2}{E\epsilon}) \text{ cm}^2 \tag{1.19}$$

The cross-section is maximized when

$$\epsilon = \frac{2(m_e c^2)^2}{E} \approx \frac{520\text{GeV}}{E} \text{ eV} \tag{1.20}$$

with the threshold energy (where the pair production starts occurring) is a factor of two lower. As the energy of the principle photon increases, so does the cross-section for pair production, effectively suppressing ultra-high (> 100) TeV photons from high redshift sources [6]. This attenuates the source's spectrum according to

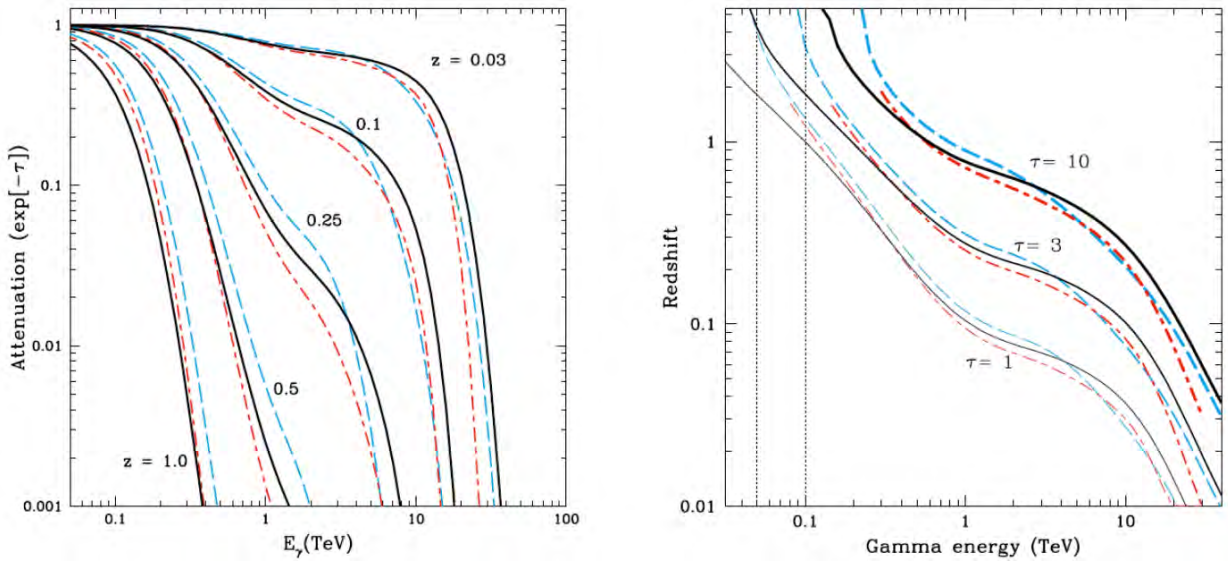

Figure 1.6 LEFT: This shows the attenuation of gamma rays at various distances with three different models. These models primarily consider the density of background photons that could absorb emission. RIGHT: This shows where the pair-production reaches certain optical depth thresholds for the same 3 models. Figures are from [9].

$$F(E,z) = F_0 e^{-\tau_{\gamma\gamma}(E,z)}, \tag{1.21}$$

where $F(E,z)$ and F_0 are the observed and emitted fluxes, and $\tau_{\gamma\gamma}$ is the optical depth function, dependent on the source's energy E and redshift z. Attenuation increases with distance, so the source becomes more suppressed the further away it is. This process becomes prominent at 1 TeV and makes it difficult or impossible to see sources in the 10's of TeV, depending on the distance. Figure 1.6 shows the attenuation of the gamma-ray energies as a function of distance along with the optical depth as a function of distance. It also shows how different EBL models effect the source's energy. Both figures are adapted from [9].

1.2 Gamma-ray Sources

There are an incredibly wide range of sources that emit in the TeV sky. The TeVCat is a catalog that collects all known TeV sources, with 318 known sources as of this writing [2]. These sources are primarily clustered along the galactic plane of the Milky Way, with a distribution of the sources shown in Figure 1.7. Two source classes emerge: galactic and extragalactic. The analyzes in

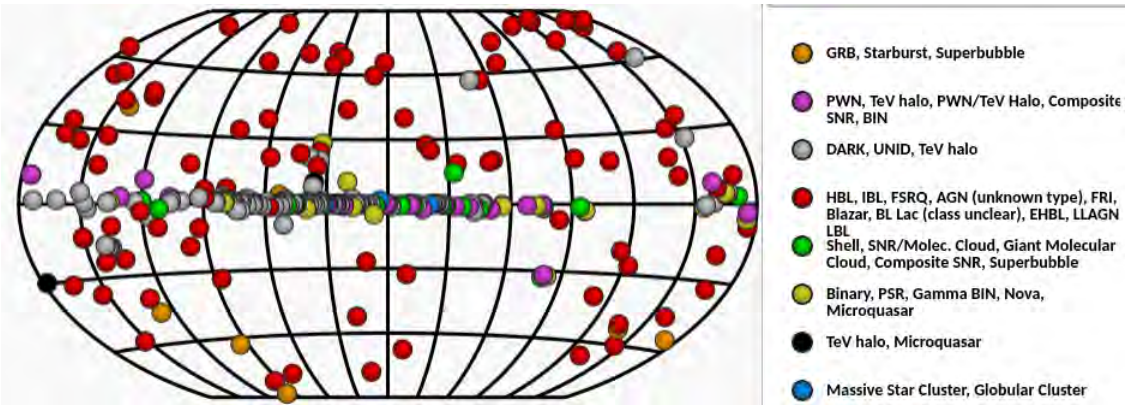

Figure 1.7 The current distribution of TeV gamma-ray sources. The dense band in the middle of the plot indicates the Milky Way's galactic plane, with the scattered sources off the plane primarily being extragalactic. Figure is from [2].

this work consider source classes from both, and will be introduced here. Additionally, further discussion on source types can be found in [39] and [36] for all the types discussed here.

1.2.1 Galactic Sources

These sources range from supernovae remnants (SNRs) to large multi-degree wide emission regions powered by star clusters. On the smaller scale, SNRs are caused when a massive star $M > M_{\text{sun}}$ exhausts its supply of fusionable material. When this occurs, a massive explosion obliterates the star and can outshine whole galaxies briefly. It releases approximately 10^{51} erg in kinetic energy which then forms shock fronts that can accelerate high-energy cosmic rays and, by extension, gamma rays. DSA occurs while the shell expands outward, typically lasting for a few thousand years. It is currently thought that most cosmic rays are accelerated by SNRs [6] and primarily emit hadronic gamma rays. Additionally, SNRs can produce neutron stars, pulsars, magnetars, and black holes, depending on the characteristics of the progenitor star.

Pulsar Wind Nebulae (PWN) are a subset of SNRs whose star turns into a pulsar when it dies. Pulsars are rapidly spinning neutron stars with incredibly strong magnetic fields that channel radiation into two jets at the north and south poles of the pulsar. Combined with the rapid rotation, they appear as lighthouses to Earth and are only visible when one of the jets is pointing to Earth. A pulsar has several metrics associated to it that determine if it can form a PWN: its age, energy budget (total energy emitted), and spin-down luminosity. As a pulsar rotates, it very slowly loses

energy via emitting from its poles and will eventually lose the ability to accelerate high-energy particles. This is primarily where the age of the pulsar becomes relevant; a Myr pulsar most likely cannot continue accelerating particles while an incredibly young pulsar (like the one powering the Crab nebula) is capable of accelerating particles to 100's of TeV [1]. Further discussion is in Chapter 4, where a PWN is completely modeled.

TeV halos are a subset of PWN whose central pulsar has either lost enough energy to no longer accelerate particles or has escaped the PWN itself. When a star goes supernova, the collapse and corresponding explosion is not perfectly symmetric due to variations in the star's density, and so the pulsar receives a kick velocity. This can often be 100's of km/s and will enable it to eventually reach and then pass the expanding shell. As was introduced with DSA, high-energy particles take much longer to accelerate compared to lower energies and will continue to be emitted long after x-ray and radio emission from a source disappears. This creates a source only visible in the TeV regime and whose class was first discovered by the HAWC observatory [40]. An example of a TeV halo would be the Geminga-Monogem system, which was the first TeV halo discovered [41].

Moving beyond SNRs, microquasars are a source class that relies on at least a binary system, with a compact object and at least one companion orbiting it. The compact object can be a neutron star, black hole, or white dwarf (product of a less massive ($M > 8M_{\text{sun}}$) stellar death) but, regardless, it slowly accretes matter from the companion star into a disk around its equator. As more matter is added to the accretion disk, friction begins to heat it, and it begins emitting photons across the electromagnetic spectrum. One other characteristic is that they can have variable emission, often tied to the orbital rate of the companion star. HAWC has observed several microquasars, with SS433 [42], V4641 [43], and LS5039 [44] being the most prominent.

The last galactic source class considered here are stellar clusters. These structures accelerate particles via the combined magnetic fields of dozens of stars within the cluster. Generally, the larger the number of massive stars, the higher energy the particles can achieve. Clusters like the OB2 in the Cygnus region have been proven to accelerate protons to PeV energies. This cluster, first observed in GeV gamma rays by the *Fermi-LAT* space telescope [45] and later in TeV-PeV by

HAWC [46] and LHAASO [47], is hundreds of light years across and contains dozens of extremely hot and massive stars. Additionally, the cluster contains at least 1 PWN that adds to its overall emission. This cluster will be discussed further in Chapter 4.

1.2.2 Extragalactic Sources

There are two primary extragalactic sources relevant to this work: Active Galactic Nuclei (AGN) and Gamma-Ray Bursts (GRBs). Broadly, AGN are galaxies with supermassive ($M > 10^6 M_{\text{sun}}$) black holes that accrete matter. The previously discussed microquasars have their name because they function as miniature versions of AGN, as the mechanisms are very similar, though at a much smaller scale. AGN are classified based off their orientation to Earth, with the various types being shown in Figure 1.8. Approximately 10% of AGN have jets and, if those jets are pointed towards Earth, the AGN is classified as a blazar. Unlike with microquasars, blazars are incredibly irregular, with flaring occurring whenever matter falls onto the accretion disk [6, 36].

The accretion disks for AGN are among the most efficient engines in the universe [48]. Depending on the disk, they can achieve a mass-to-energy conversion rate of $> 10\%$, far outpacing processes like fusion $\approx 0.5\%$. As mentioned above, AGN accrete matter onto their disk and compact it. As it compacts, it generates friction, heating up the matter and turning into a plasma. Additionally, the sections of the accretion disk close to the black hole's event horizon can move at near the speed of light. When large chunks of matter fall onto the disk, it becomes visible across the EM spectrum, even up to TeV energies if it is sufficiently close to Earth. HAWC sees two blazars, Markarians 421 and 501, both at redshifts of ≈ 0.03, and they can both flare to the brightest objects in the TeV sky [28, 3]. They will be discussed more in Chapter 5.

GRBs are extremely intense but short-lived phenomena, and are the most energetic events in the universe. They occur isotopically across the sky and are broadly extragalactic in nature. Depending on the event, they can range from keV to GeV, with extremely powerful GRBs initially theorized [49] and later confirmed [50] to emit in TeV. They are generally only visible in gamma rays for a few seconds, but have afterglows that can last days to months [6]. The origins of GRBs is still unknown, with events like supernovae, compact object mergers, AGN, and others all being

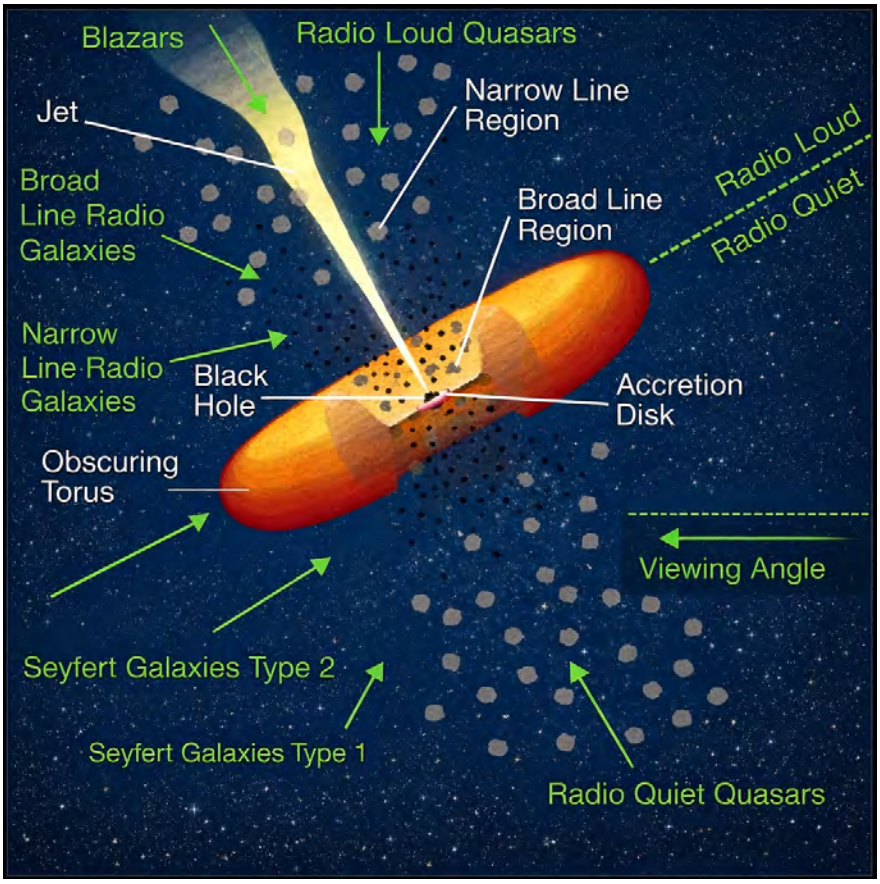

Figure 1.8 A cartoon that shows the various types of AGN with respect to Earth's view of them. The figure is from [10].

theorized to produce GRBs. As of writing, HAWC has not directly observed any GRBs.

1.3 Detection Techniques

To detect gamma rays, both direct and indirect methods are used. Direct measurements are taken when a gamma ray directly strikes a detector and must be located above the atmosphere while indirect methods study the secondary particles produced when gamma rays hit the Earth's atmosphere. Depending on the technique used, these detectors can be located between sea level and 5000 meters above sea level.

1.3.1 Direct Imaging

To directly detect gamma rays, satellites can have detectors mounted when they launch. These detectors are multi-layered to enable accurate directional and energy measurements of incident gamma rays. Shown in Figure 1.9, the incident gamma rays first pass through an anti-coincidence

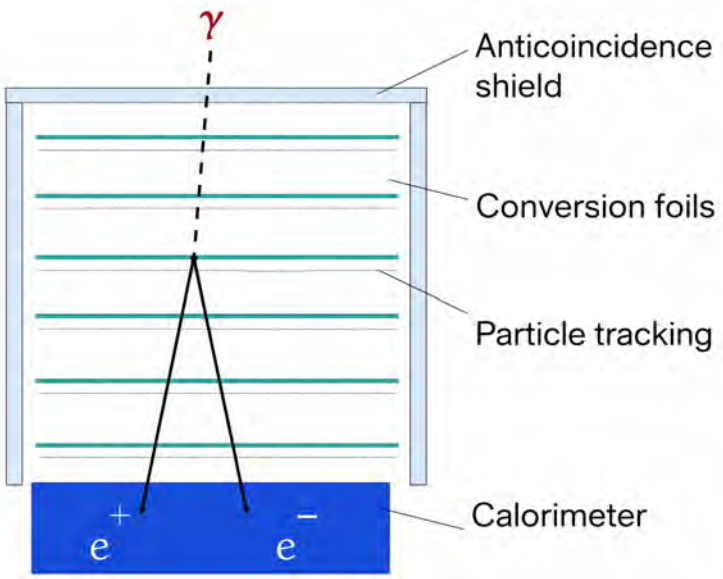

Figure 1.9 A diagram of the gamma ray detector mounted to the *Fermi-LAT* space telescope. Gamma rays are detected via pair-production inside the detector, with the direction and energy being recorded so an origin can be found. The figure is adapted from [11].

shield to reject cosmic rays before passing through a series of conversion layers. As the photons pass through, they interact with the conversion layers and pair-produce to produce electron-positron pairs. These pairs carry on through trackers built within the conversion layers to track timing for direction reconstruction before they finally strike a calorimeter, where their energies are recorded and the original gamma ray's energy can be reconstructed [6, 51].

The best current example of a space-based detector is the *Fermi-LAT* space satellite. Launched in 2008, *Fermi-LAT* is sensitive from 20 MeV to 300 GeV, with a dual-instrument suit to monitor both GRBs and collect long exposure data for source analysis. This has allowed *Fermi-LAT* to discover thousands of sources and GRBs [3], along with several unique structures like the Fermi bubbles, located above and below the galactic plane and may be related to the Milky Way's supermassive black hole [51].

Space-based telescopes have large fields of view and good angular resolutions. However, once gamma rays achieve energies over a few hundred GeV, the event rate becomes so small that it is highly improbable the gamma rays would strike the detector. If they do hit, they have such high energy that they can pass through the detector without interacting with it [51]. To see higher energy

gamma rays, land-based detectors like Imaging Air Cherenkov and extensive air shower arrays like Water Cherenkov Telescopes are needed.

1.3.2 Indirect Imaging

Indirect imaging refers to any process that does not directly capture a gamma ray. Instead, these techniques reconstruct secondary products generated from the gamma ray's interaction with the Earth's atmosphere.

1.3.2.1 Extensive Air Showers

When a gamma or cosmic ray hits Earth's atmosphere, it generates a cascade of particles that expands into a descending cone and can have footprints $> 100,000$ m^2. These are called Extensive Air Showers (EAS), and ground-based observatories use them to reconstruct the gamma or cosmic ray's initial energy and angle of incidence. An example of a cosmic-ray EAS is shown in Figure 1.10, with a proton being the incident particle. The primary product of a cosmic-ray shower are pions and other nuclear fragments, some of which then decay into lower energy gamma rays. By contrast, a gamma-ray EAS is primarily leptonic in nature, with electron-positron pair-production being the primary reaction. Regardless of the type, as an EAS progresses through the atmosphere, more interactions occur to fuel the cascade until the energy of the secondary particles drops below the ionization threshold, at which point the shower stops expanding [52]. Cosmic-ray EASs are the primary background for ground-base gamma-ray detectors, but they can be rejected by considering their unique shower shape compared to gamma-ray EASs. This will be further discussed in 2. Lastly, there are detected experiments like the Pierre Auger observatory in Argentina that can detect cosmic rays up to 100s of EeV to probe the absolute limits of particle acceleration [53].

1.3.2.2 Imaging Air Cherenkov Telescopes

Regardless of the origin, the products of an EAS can be studied by utilizing an Imaging Air Cherenkov Telescope (IACT). IACTs utilize mirrors that focus light to a central camera made of photon-detecting equipment like photomultiplier tubes (PMTs) and plastic scintillators. If the particles produced in an EAS are moving faster that the speed of light in a given medium (air in this case), they will emit blue light called Cherenkov light. Shown in Figure 1.11, if a particle's

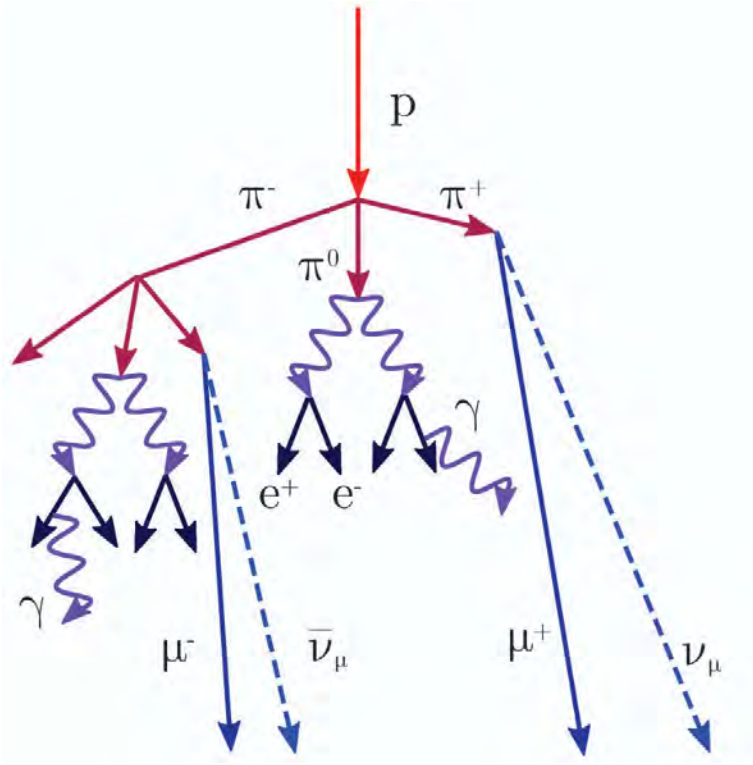

Figure 1.10 A example of a cosmic-ray EAS. In this instance, the secondary particles are hadronic and create a unique shower structure compared to gamma-ray showers. Figure is adapted from [12].

velocity v_p fulfills the condition $v_p > c/n$, with n being the refraction index of the material, it will emit a cone of radiation at an angle θ. Since the emitted waves travel at speed c/n, the total distance traveled for a given time t would be ct/n. Meanwhile, the particle travels a distance βct, $\beta = v_p/c$. Therefore, the emission angle is then

$$\cos\theta_C = \frac{1}{n\beta}. \tag{1.22}$$

For air, $\theta_C \approx 0.7°$, depending on the altitude as n changes as a function of altitude. Regardless, the light produced from a gamma-ray shower will form an ellipse-like region with a radius of around 120 meters, with the long axis pointing back towards the gamma ray's origin. The intensity of the Cherenkov radiation observed is roughly proportional to the initial photon's energy [6].

An IACT records these parameters and then constructs the EAS to determine the original gamma ray's energy and location. Current examples of IACTs are MAGIC [54], VERITAS [55],

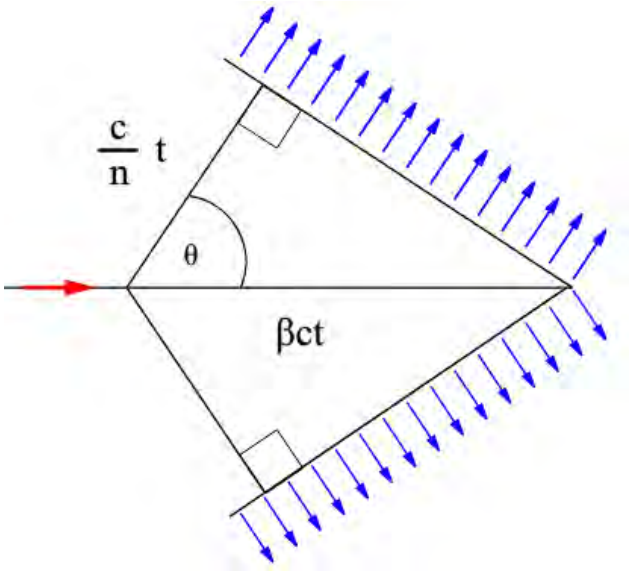

Figure 1.11 A diagram of how Cherenkov radiation (blue arrows) propagates with respect to the incident particle. Figure is adapted from [13].

and HESS [56]. All of these comprise several IACTs to both achieve superior background rejection and larger effective area, therefore allowing observations of higher energy (> 25 TeV) showers. These instruments also boast superior angular resolutions of $< 0.1°$, high sensitivities, and have about 15% energy resolution. However, they can only operate on clear, moonless nights due to data contamination caused by either the sun or moon light, and they have small instantaneous fields of view at $\approx 4°$. To improve the field of view and operation time, the much larger extensive air shower arrays are needed.

1.3.2.3 Extensive Air Shower (EAS) Arrays

EAS arrays are characterized by large numbers of detectors spread across a wide footprint and are generally located at higher altitudes to better capture an incoming EAS. As a result, gamma-ray EAS arrays can detect gamma rays from 100 GeV to > 1 PeV and cosmic-ray detectors like Pierre Auger can push into the EeV regime, depending on the configuration and size [53].

There are several different techniques that can detect air showers, with two of them being scintillation and water Cherenkov detectors. The former functions by having a scintillating material become energized by an EAS's products and emit a photon that is read by the detector while the latter works the same as IACTs except, instead of using air as the medium, they use water. An

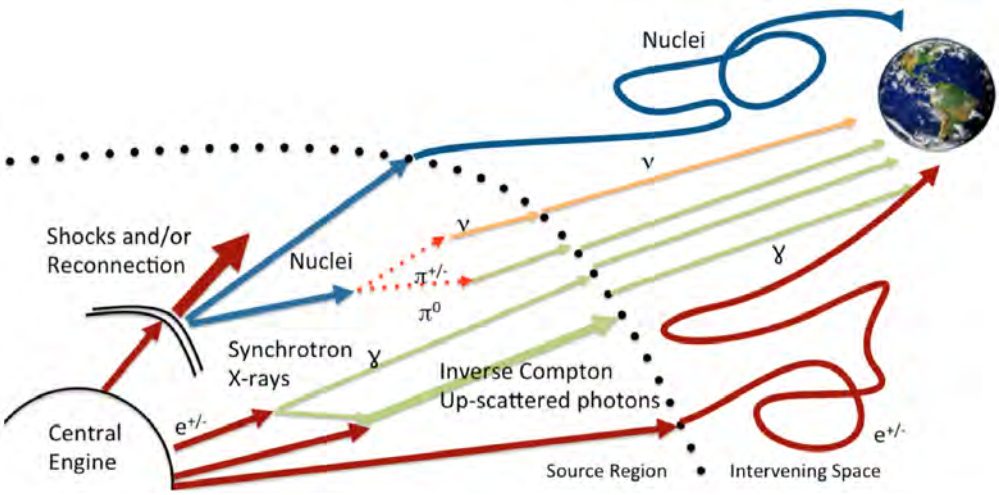

Figure 1.12 This shows the potential relation of cosmic rays, photons, and neutrinos. All three can be produced by a central engine and proper characterization of their emission can reveal what production mechanisms are present in a source. The figure is adapted from [14].

example of each would the Tibet-AS cosmic and gamma-ray observatory [57] for scintillation and HAWC [58] for water detectors, with HAWC being discussed further in Chapter 2. A newly built detector called LHAASO [35] combines these techniques and can achieve a much larger energy range compared to HAWC [35]. Broadly, while these instruments have comparatively poor angular resolution to IACTs, they have wide fields of view and almost continual operation, making them more suited for transient analyzes (see Chapter 5).

1.4 Multi-Messenger Astronomy

In the past few decades, many different observatories have been constructed to observe both the EM spectrum and other particles that astrophysical sources may emit. When combined, these "messengers" can completely characterize a source and determine the emission mechanisms present. There are pros and cons for each messenger, but their observation can be incredibly powerful when combined and can reveal mysteries of the universe.

1.4.1 Astrophysical Messengers

There are four types of messengers: cosmic rays, photons (radio to gamma rays), gravitational waves, and neutrinos. Three of these are shown in Figure 1.12 where cosmic rays, photons, and neutrinos are all emitted by a central engine, or source. The connection between cosmic rays

and photons has already been established above, but cosmic rays are charged particles that are accelerated in emission regions. They are bent by magnetic fields from both the source itself and ambient fields they encounter on their way to Earth, changing their direction and hiding the original source. Photons, specifically gamma rays, are also produced in these regions and are not affected by magnetic fields, allowing comparatively easy source identification. However, both redshift and EBL absorption at long distances make distant source analysis challenging.

Gravitational waves are produced by the acceleration of compact masses like neutron stars and black holes. When two objects orbit, they very slowly radiate energy in the form of gravitational waves. While all systems emit these waves, large masses are needed to produce detectable waves. They were first discovered in 2015 [59] when two stellar mass black holes ($M \approx 100 M_{\odot}$) collided. As they spiraled into each other, they emitted strong waves and, when they did collide, a strong GRB was observed a couple seconds later. This confirmed the origin of the waves and a potential explanation for GRBs. They are detected by kilometer-scale Michelson laser interferometers like LIGO [59].

Lastly, neutrinos are produced in the same processes as gamma and cosmic rays. From Equation 1.6, neutrinos are emitted as a product of pion decay; specifically, the charged pions decay into neutrinos and muons. Neutrinos interact only via the gravitational and weak forces and are incredibly light. Therefore, they can only be directly observed if they strike a nucleus. As such, neutrinos are not affected by the EBL or redshift and can travel the full length of the universe unless they directly hit another molecule. As such, massive detectors are needed to detect astrophysical neutrinos. One such detector, IceCube, will be briefly discussed in Chapter 2 [60].

1.4.2 Multi-Messenger Sources

As previously mentioned, neutrinos, cosmic rays and gamma rays can all be produced in the same source. Of the sources discussed previously, galactic candidates like PWN, SNRs, and microquasars are all predicted to produce neutrinos, but current detectors lack the sensitivity to observe significant neutrino emission. As of this writing, IceCube has just finished an upgrade and it is hoped to greatly improve sensitivity to galactic sources like the powerful PWN MRGO

1908+06 [61, 62].

Currently, to directly measure neutrinos, extreme extragalactic phenomena like blazars are required. Previously introduced, blazars are extremely bright and powerful active galactic nuclei with jets pointed towards Earth. One such blazar, TXS 0506+056 was detected by IceCube via a high-energy neutrino alert (discussed in Chapters 2 and 5) that was coincident with a gamma-ray flare visible at TeV energies [63]. To date, only a few sources have been identified from IceCube, but both its planned update to Gen2 [64] and other new observatories like P-ONE and KM3NET will discover more sources for follow-up [65].

CHAPTER 2

THE OBSERVATORIES

During my Ph.D. work, I have used three different observatories. My primary focus has been gamma ray analysis using the High-Altitude Water Cherenkov (HAWC) Observatory, with secondary applications with neutrino data from the IceCube Neutrino Observatory. Lastly, I performed weather data analysis in cooperation with the Southern Wide-field Gamma-ray Observatory (SWGO). I will discuss each observatory, with the primary focus of this chapter being on HAWC. Lastly, I am a member of both the HAWC and SWGO collaborations. These are large, multi-national groups, with HAWC primarily a USA-Mexico collaboration, and SWGO including many South American and European institutions as well.

2.1 The HAWC Gamma-Ray Observatory

The High-Altitude Water Cherenkov (HAWC) Gamma-ray Observatory is located in the state of Puebla, Mexico at an altitude of 4100 meters. First commissioned in March 2015, HAWC is sensitive to gamma rays between 300 GeV and approximately 300 TeV, though that can be higher, depending on the quality of the EAS HAWC detects. It has a wide Field of View (FOV) of 2 steradians and can see 2/3 of the sky each day. Located at a declination of 19°, HAWC can see any overhead event between the declinations of $-31°$ to 69°, or 50° off the 19° zenith declination. Additionally, HAWC has a high up-time of $> 95\%$. These characteristics make it ideally suited for both transient and steady-state (continuous emission) source searches like the ones described in this work [58]. A picture of HAWC is given in Figure 2.1. This figure shows both the main array and sub-array elements of the HAWC detector.

2.1.1 Water Cherenkov Detectors

As previously introduced, gamma or cosmic ray interactions with Earth's atmosphere produce extensive air showers (EAS) that cascade down to the surface (see Figure 1.10). At energies over 1 TeV, these showers can be ten's to hundred's of feet across and, at extremely high energies of EeV (e18 eV), their width can be tens of miles across [53]. Figure 2.2 shows two simulated EASs at 100 TeV, with the left being a gamma-ray shower and the right being a hadronic (cosmic-ray) shower.

Figure 2.1 The HAWC Observatory circa April 2024. The main array is a concentration of large tanks, with the outrigger array containing the much smaller tanks seen on the surrounding hillside. The mountain Pico de Orizaba is visible in the background. Photo taken by myself.

While the shower structure itself is different between the two, they both expand in size as they travel through the atmosphere before shrinking as they lose energy. Their footprints and products differ significantly and will be considered in Section 2.1.2.

When these shower products interact with a medium and are traveling faster than light in that medium, they emit Cherenkov radiation (see Chapter 1). For the HAWC observatory, the medium is water contained within 300 large metal tanks. Each tank is 7.3 meters wide and 5 meters tall. Inside each tank is a light-tight bladder that contains 200,000 liters of water with 4 photomultiplier tubes (PMTs) anchored to the bottom. There are three 8 inch PMTs surrounding a central 10 inch one, both made by Hamamatsu. An example of the tank as well as the main array layout is given in Figure 2.3. Lastly, a central counting house contains all data acquisition and processing computers for the observatory, indicated by the blank region near the center of Figure 2.3.

After commissioning, HAWC was upgraded with a sparse outer array of 345 small tanks. These tanks are 1.55 meters wide by 1.65 meters tall with a single eight inch PMT. These can be seen in Figure 2.1 with the small tanks on the hillside surrounding HAWC. Installed in 2018, they greatly expand HAWC's effective footprint by 4 times and are projected to significantly improve HAWC's

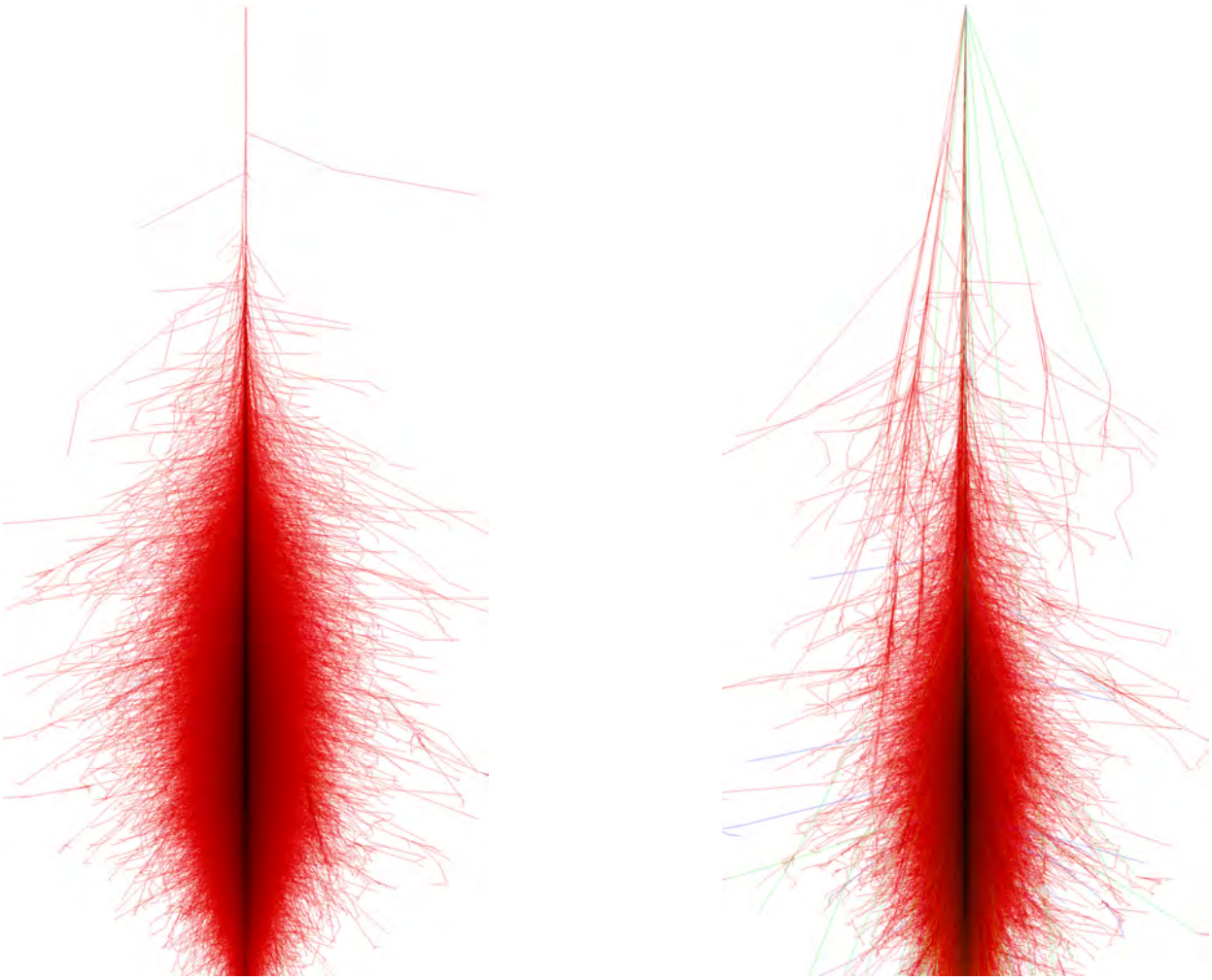

Figure 2.2 Examples of gamma (LEFT) vs hadron (RIGHT) showers at 100 TeV. From both images, the shower's width expands with decreasing altitude before attenuating in the atmosphere. For both plots, the red indicates electrons, blue are hadrons, and green are neutrons. The hadronic shower produces much heavier products, heavily affecting the shower's footprint. Both images were made using CORSIKA simulations [15].

sensitivity to > 10 TeV events [66]. While they have been collecting data since 2018, it has yet to be folded into the main data set and is not considered for this analysis. However, work is currently ongoing to include it with HAWC's next main data set.

2.1.2 Data Collection

Each PMT is connected to the counting house via a 610 ft long cables to maintain identical signal timing for all PMTs. These cables each connect to a front-end board (FEB) that processes the signal before running it through Time to Digital Converters (TDCs) to digitize the signals. The

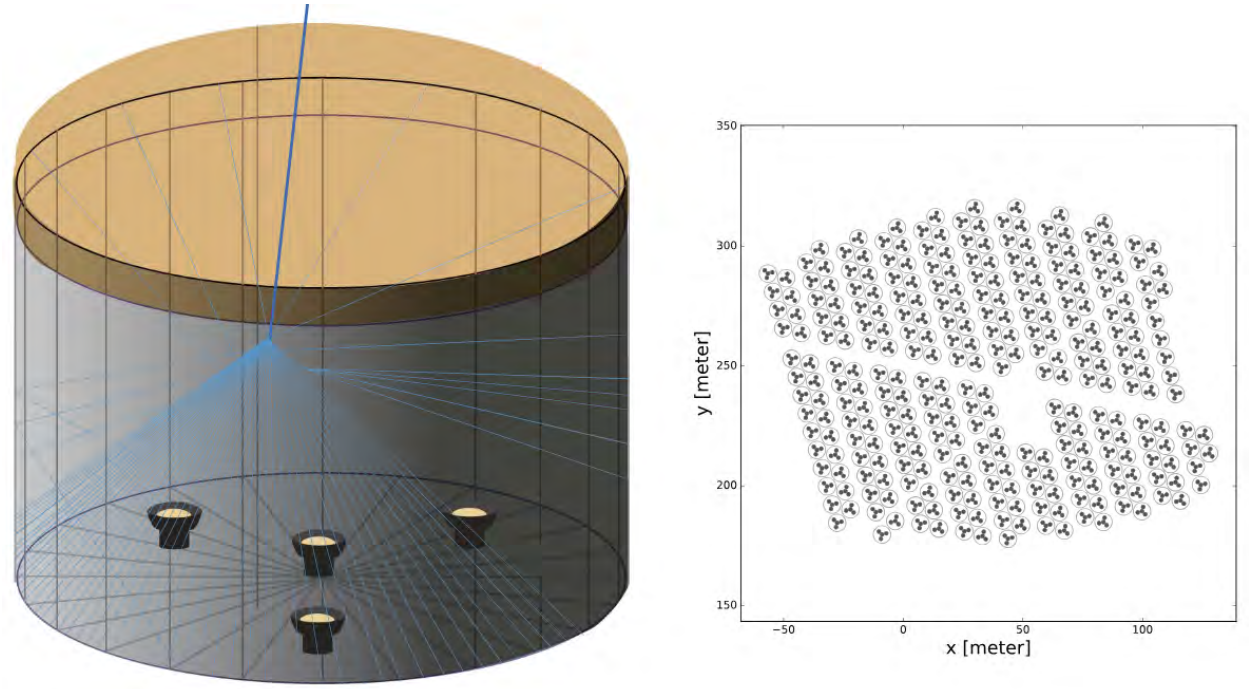

Figure 2.3 An example of one main array tank from HAWC. Each tank contains 4 PMTs that capture Cherenkov light emitted when a particle enters the tank. The right figure shows the layout of the main array. The figure is from [16].

FEBs also send high voltage to each PMT. There are 10 TDC modules, each with 128 channels. All TDC modules are synchronized by a GPS timing system that has a resolution of 98 picoseconds to ensure proper event correlations [58].

All data is then transferred to online systems that run the Reconstruction Client, which reconstructs the observed air shower from the raw PMT data. These allow for real-time detections and immediate follow-ups from other observatories. These include monitors for Gamma-Ray Bursts (GRBs), flare monitors for several Active Galactic Nuclei, gravitational waves flares, and high-energy photon alerts. On average, HAWC can follow-up within a minute of receiving an external alert like a GRB detection from *Fermi-LAT* or can release an alert if HAWC detects a flare. All other reconstruction is done off-site at the computing cluster at the University of Maryland or at the National Autonomous University of Mexico (UNAM) [17].

All processes included in the Data Acquisition system (DAQ) are controlled by a computer called Experiment Control. This system is responsible for starting and ending data runs, along

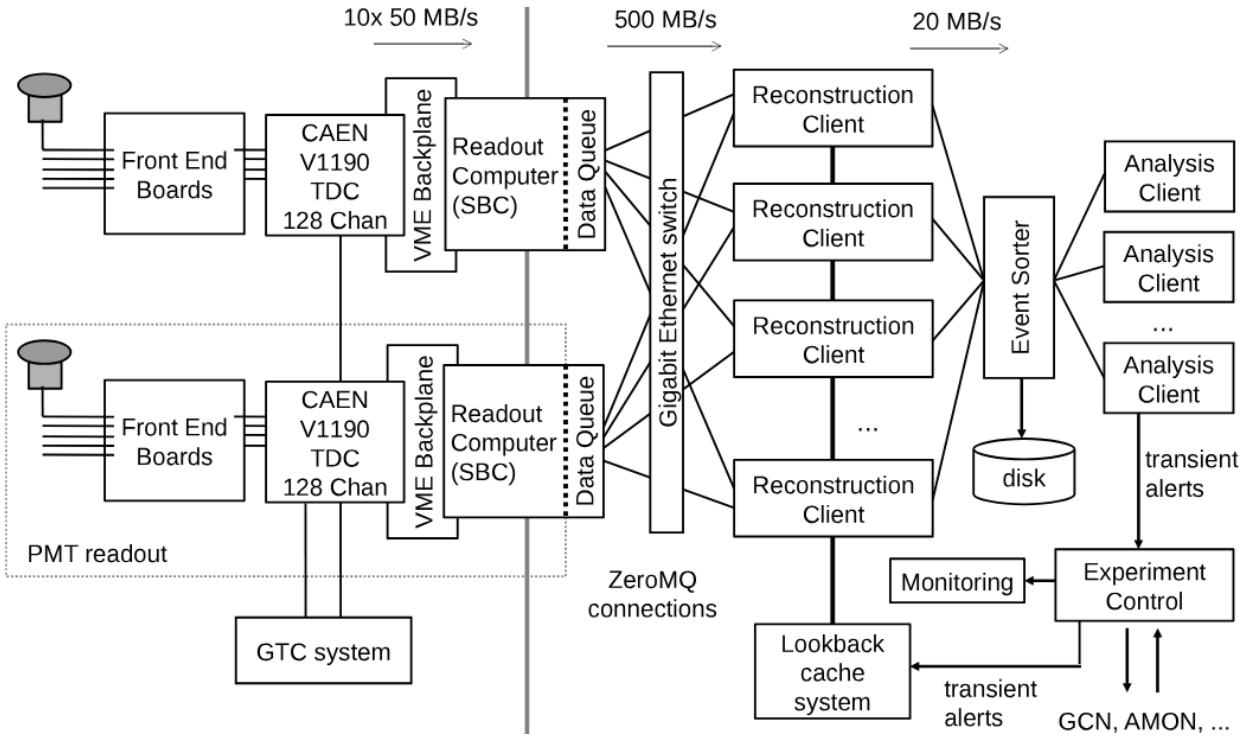

Figure 2.4 The pipeline for the DAQ system. On the left side of the gray vertical line, electronic components like the FEBs and TDCs collect data before passing it along to the right side of the diagram for data processing. Figure is from [17]

with collecting monitoring data for the experiment. If a run crashes or there is an issue at site (like power outages), it will attempt to automatically restart the run and otherwise alert the shifter of the issue. The main cause of run crashes generally are heavy storms or lightning activity, with recovery ranging from minutes to hours, depending on the severity [17]. The full DAQ pipeline is shown in Figure 2.4.

2.1.3 Data Reconstruction

In order to use the data that HAWC collects, the air shower needs to be reconstructed from the PMT measurements. HAWC uses the hit timings and charge deposited in the PMTs to reconstruct the direction and energy of the original particle. As previously mentioned, an EAS produced by a gamma or cosmic ray have very different structures, and so reconstructing HAWC's data allows for effective background rejection. Two example EAS reconstructions for gamma and hadronic showers is shown in Figure 2.5.

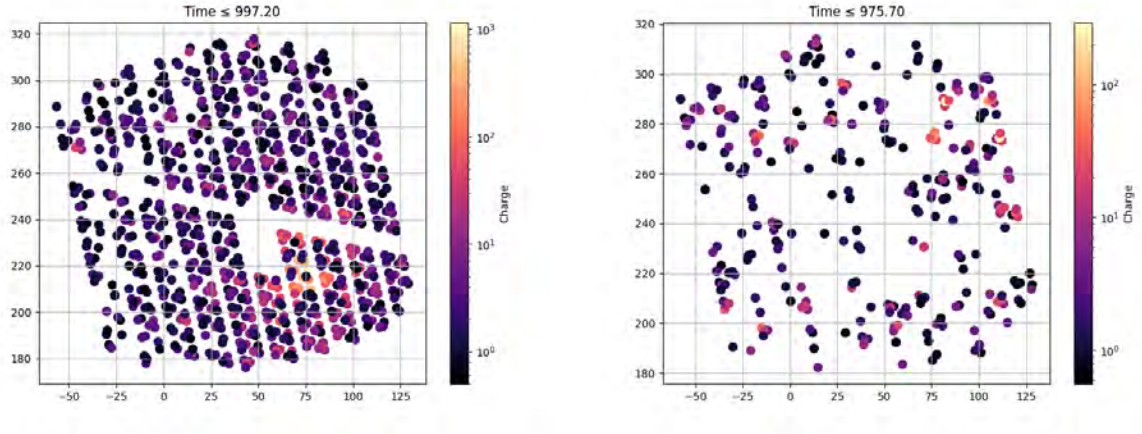

Figure 2.5 These two plots show the difference in footprint, timing, and charge deposition for gamma (LEFT) and hadronic (RIGHT) showers. Given the overall lighter shower components, gamma showers tend to be compact with a clear center while hadronic showers are more sparse with nodes of activity registered across the detector.

2.1.3.1 Angular Reconstruction

To reconstruct the original particle's origin, two steps must be taken: finding the location of the shower core and the angle the shower came from. The shower core's reconstruction will be discussed first.

The shower's core is defined as the location where the most charge is deposited in HAWC. The first core reconstruction step utilizes a simple center-of-mass calculation to find a rough estimate of the core's location. This takes the average PMT locations and weighting them by their recorded charges before passing the rough estimate along as the seed value for the next fitter: the Super Fast Core Fitter (SFCF).

The SFCF fits the lateral charge distributions shown in Figure 2.5 to a modified Nishimura-Kamata-Greisen (NKG) function:

$$S_i(A, r) = A\left(\frac{1}{2\pi\sigma^2}e^{-r_i^2/2\sigma^2} + \frac{N}{(0.5 + r_i/R_m)^3}\right) \tag{2.1}$$

with S_i is the signal from the ith PMT, $\sigma = 10$ m, $N = 5 \times 10^{-5}$, and $R_m = 120$ m is the Moliere radius of the atmosphere at HAWC's altitude [67]. This function has two components: a Gaussian

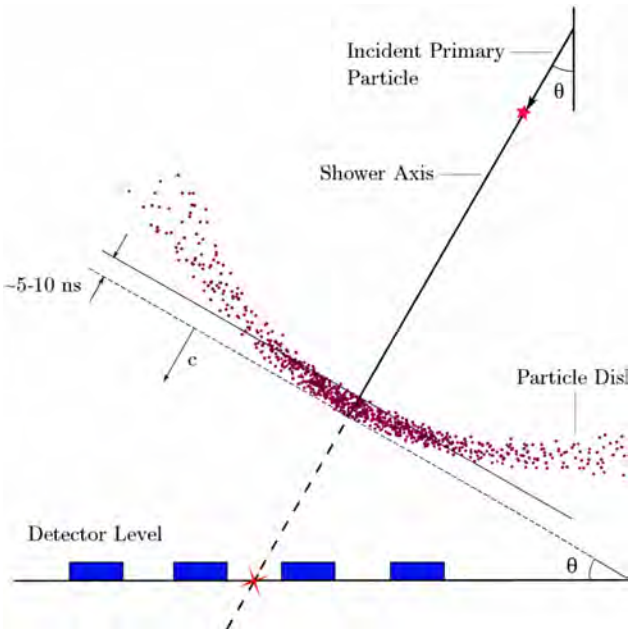

Figure 2.6 This shows the curved shower from an EAS. As secondary particles are produced, they are emitted at slight angles from the principle access, creating a curved shape. The curvature extent allows for angular reconstruction. Figure is from [12].

part to parameterize the charge distribution at the core and a tail with normalization N determined by HAWC simulations. The Moliere radius is defined as the radius of a given cylinder that contains 90% of the energy of the air shower. It is a function of the radiation length in air, the scattering energy, the critical energy, and the air density at HAWC. The fitted parameters are the core location r and amplitude A. The median error for the SFCF is between 2 and 4 meters for large and small events that land on HAWC, with significantly larger errors for showers whose cores are outside of the main array [67, 1].

Once the core's location is found, the direction can then be determined. This uses the relative timing information of all affected PMTs triggered from the EAS. Two factors need to be considered: the shower curvature and sampling of observed events. The former is shown in Figure 2.6 and is an effect caused by the secondary particles not all traveling directly along the shower axis. Particles emitted off-axis will spread the shower front and will arrive later than the principle component of the shower. The latter effect: sampling, is caused by the spread of arrival times of the secondary particles. There is about a 10 nanosecond spread for particles traveling along the shower axis, with

Table 2.1 Pass 5 f_{hit} binning scheme with median energies [1]

Pass 5	Hit PMT Fraction	Median Energy Crab On-array (TeV)	Median Energy Crab Off-array (TeV)
B0	2.7%–4.7%	0.28	0.57
B1	4.7%–6.8%	0.38	0.88
B2	6.8%–10.4%	0.53	1.29
B3	10.4%–16.1%	0.83	2.02
B4	16.1%–24.5%	1.37	3.66
B5	24.5%–35.1%	2.25	6.21
B6	35.1%–47.2%	3.68	10.27
B7	47.2%–59.9%	5.97	16.62
B8	59.9%–72.2%	9.54	25.78
B9	72.2%–82.2%	14.63	41.47
B10	82.2%–100.0%	30.46	73.91

this time interval increasing for off-axis particles. HAWC only tracks the arrival time of the first particle, so position-dependent timing corrections are applied to compensate. This correction is based off both simulations and data as simulations alone cannot accurately model the PMT's full response [67].

Both of these effects are used together to reconstruct the shower's incident angle. Once the necessary corrections are made, a χ^2 function is used to fit a plane to the recorded and corrected PMT hit times. This gives the angle with resolution between 0.1° and $> 1^\circ$, depending on the energy, incoming angle (also known as zenith angle), and how much of the shower front is on the main array [67].

HAWC uses three different methods to reconstruct its data. The first is called the Fractional Hit (f_{hit}) and purely relates the fraction of HAWC that was triggered by an EAS to energy. This breaks down for events greater than 30 TeV and so two energy estimators were created, both based off the f_{hit} estimator. f_{hit} will be discussed first.

2.1.3.2 Fractional Hit Estimator

The f_{hit} estimator takes a simple ratio of the PMTs included in the reconstruction compared to the total number of PMTs in the array. Typically, HAWC has around 1000 PMTs operational at any given time, allowing for a small number removed for maintenance at any given time. For increasing

Table 2.2 HAWC Energy Bin Definitions for GP and NN [1]. The first entry for each row is the log fraction of the energy while the parenthesis contains the TeV energy value.

Energy Bins	Min. $\log_{10}(E/\text{TeV})$	Max. $\log_{10}(E/\text{TeV})$
Ea	-0.50 (0.36)	-0.25 (0.56)
Eb	-0.25 (0.56)	0.00 (1.0)
Ec	0.00 (1.0)	0.25 (1.8)
Ed	0.25 (1.8)	0.50 (3.2)
Ee	0.50 (3.2)	0.75 (5.6)
Ef	0.75 (5.6)	1.00 (10)
Eg	1.00 (10)	1.25 (18)
Eh	1.25 (18)	1.50 (32)
Ei	1.50 (32)	1.75 (56)
Ej	1.75 (56)	2.00 (100)
Ek	2.00 (100)	2.25 (177)
El	2.25 (177)	2.50 (316)

energies, more of the main array is triggered, therefore creating a rough energy estimator. This method works both for on-array, where the shower core lands directly on the array, and off-array, where part of the shower lands on HAWC, but the core strikes nearby. The data is binned based off the fraction of the array triggered from a given shower, along with the estimated energy. These values are shown in Table 2.1 and are from [1].

This method has several limitations: energy range, spectral assumptions, and angular resolution. For energy range, once the main array is saturated, f_{hit} no longer can estimate the parent particle's energy. This occurs at 30 TeV for on-array and 75 TeV for off-array. Additionally, the fraction of the array triggered is dependent on the zenith angle of a given shower, with potentially low energy events triggering the whole array if they have a high zenith angle. Lastly, the angular resolution for low energy events can be quite poor, given the small fraction of the array triggered. For sub-TeV events, this can approach several degrees while 30 TeV events can have sub 0.2° angular resolutions [1]. This is the primary estimator used in Chapter 5 and Appendix B.

2.1.3.3 Energy Estimators

To address the limitations of the f_{hit}, two energy estimators were developed [68]. These are the Neural Network (NN) and Ground Parameter (GP) estimators. Discussed in both HAWC's 2019

[68] and 2024 [1] Crab papers, a brief summary of each is provided here.

The NN algorithm utilizes the Multivariate Analysis neural network implementation toolkit from [69]. Three general parameters are fed into the network: the energy amount saturating the array, the shower footprint in the array, and the attenuation experienced by the shower. These parameters are fed into a multilayer-perceptron network with two hidden layers and a logistic activation function. The two layers have 15 and 14 nodes, with 479 weights selected to minimize the internal error function.

The GP algorithm reconstructs the shower's energy by estimating the air shower's charge density at a fixed distance from the shower core. As discussed in section 3.1 in [68], simulations determined this distance to be 40 meters. To relate the charge density to the principle particle's energy, the equation

$$\log_{10} E = m(\theta) \log_{10} \rho_{40} + c(\theta), \tag{2.2}$$

is used, where $m(\theta)$ and $c(\theta)$ are linear and quadratic piecewise functions with respect to the zenith angle and ρ_{40} is the charge density at 40 meters. As with the NN estimator, the GP reconstructed data set greatly extends HAWC's energy range to over 100 TeV and can serve as a systematic cross-check for analyzes.

Both estimators greatly increase the sensitivity of the f_{hit} data set, allowing for 12 sub-bins to be created per each f_{hit} on-array bin. These bins are listed in Table 2.2 and range from 0.3 to 300 TeV. These bins are read off as "B0C0Ea", with B0 being the f_{hit} bin, C0 indicating it is on-array (C1 for off-array), and Ea being the energy estimator sub-bin. Each f_{hit} group of bins (B0, B1, etc) can be grouped by their energy bins, combined, and then output a single flux value. These are called flux points and are critical in both source identification and multi-messenger applications. All three of these estimators are used in Chapters 4 and 5.

2.1.4 Monitoring

HAWC utilizes an online monitoring website called HAWC Observatory Monitoring for Experiment and Reconstruction (HOMER) and is responsible for presenting all monitoring data collected

at site to off-site shifters. Michigan State University has a long history of maintaining this website, with my service work primarily related to working on HOMER. As such, this will be discussed further in the next chapter refsec:homer. Briefly, all data that was discussed above can be monitored using HOMER and its sub-pages, enabling effective remote monitoring of HAWC.

2.2 The IceCube Neutrino Observatory

The IceCube Neutrino Observatory is a cubic kilometer scale detector located deep within the glacier at the South Pole. It detects neutrinos from 10 GeV to PeV energies and it can view the entire sky. The Earth absorbs lower energy atmospheric neutrinos and so IceCube is more sensitive to the northern hemisphere. It is composed of 3 separate arrays: a main in-ice array of Digital Optical Modules (DOMs), a smaller dense cluster of DOMs inside the main array called DeepCore, and a cosmic ray observatory called IceTop. The whole observatory is shown in Figure 2.7. Unless otherwise noted, all information in this section is from the IceCube hardware paper [60].

The main array is composed of 5160 DOMs on 86 strings buried between 1450 and 2450 meters into the ice. Each string contains 60 DOMs, with the strings being arranged in a hexagonal pattern. Much like with HAWC, they detect neutrinos by Cherenkov radiation emitted when they interact with a nucleus in the ice via the charged-current weak interaction. This interaction produces a faster-than-light (in the medium) lepton, which then emits Cherenkov radiation. For 78 of the strings, the vertical spacing is 17 meters, with the remaining 8 having closer spacing of 7 meters. These 8 are the DeepCore array, and the tighter spacing allows increased sensitivity to lower energy events. The main array is sensitive to TeV-PeV events while the DeepCore array is primarily focused on GeV emission.

The IceTop array is located on the surface of the ice and functions as a cosmic ray detector. It comprises of 162 tanks in 81 stations that roughly mirror the string locations. Each tank contains two DOMs and is filled with 0.9 meters of ice. Like with the strings, there is a small, denser array above DeepCore. The main IceTop array is sensitive to cosmic rays from PeV to EeV energies, with the denser sub-component sensitive to 100 TeV rays. Its primary function is to study PeV energy gamma and cosmic rays, along with being a partial veto for downward heading neutrinos.

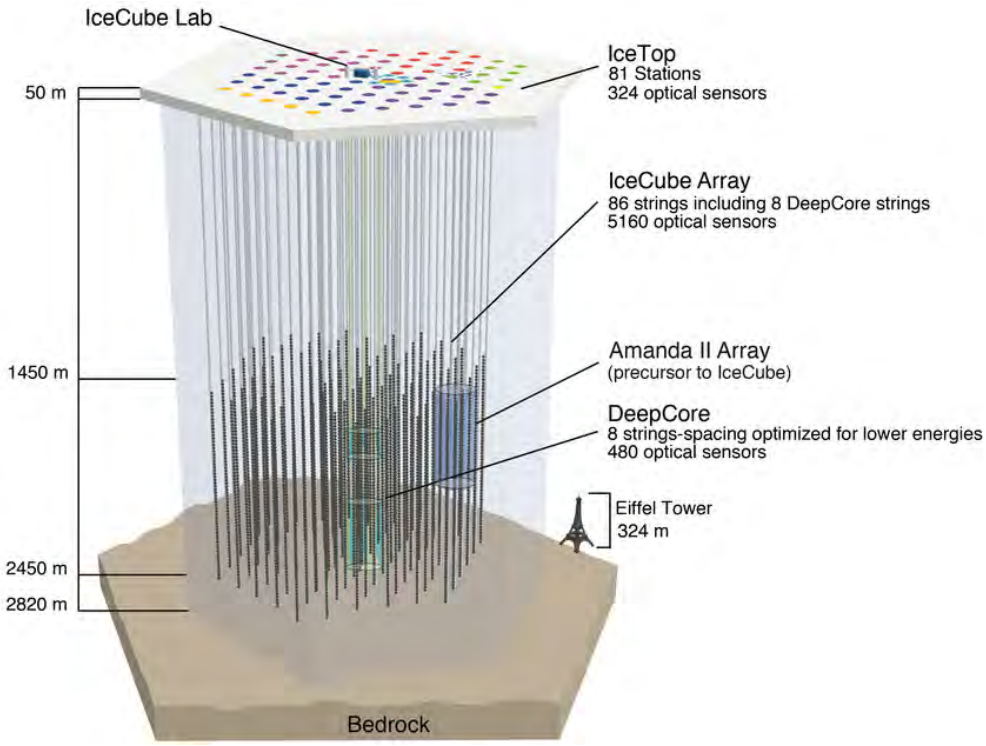

Figure 2.7 The IceCube Neutrino Observatory. The 3 main components are shown, with the Amanda II Array being the precursor to IceCube and merged into IceCube when construction finished. Figure is from [18].

Each DOM is a self-contained detector unit housed in a pressure-resistant glass sphere. Internally, it contains a 10-inch Hamamatsu (PMT), high-voltage circuitry, waveform digitization electronics, and a single-board data acquisition computer. The DOMs detect the Cherenkov radiation and digitize the resulting PMT waveforms, transmitting data to the surface counting house via the cable system. Due to a rigorous pre-deployment testing protocol, 98.4% of all deployed DOMs are operational and actively collecting data [60]. Unlike with HAWC, DOMs cannot undergo maintenance given their extreme location.

IceCube has broad applications to the field of neutrino astronomy. They have operated a real-time alert system from 2011 to present that reports probable high energy neutrinos via the Astrophysical Multimessenger Observatory Network (AMON) network [70]. Alerts from 2011 to 2023 were compiled, reconstructed, and then published in a catalog called IceCat 1 [71]. This alert system allows for follow-ups with HAWC, which is the focus of Chapter 5. Other scienific

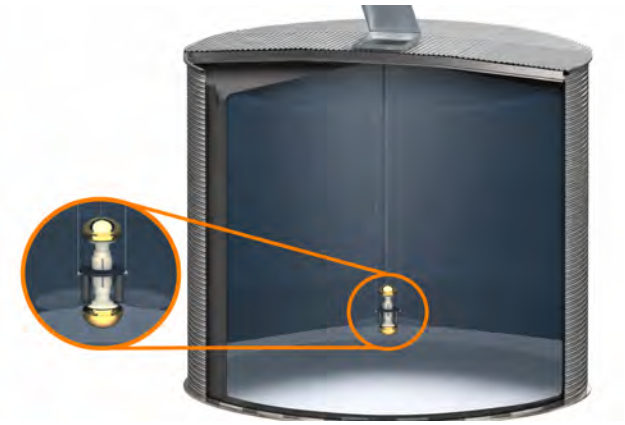

Figure 2.8 The design of SWGO's main array tanks. The internal volume is split into two spaces, with the upper chamber 4 meters tall and the lower being 0.8 meters tall. Each will have one 10 inch PMT, facing either up or downwards. Figure is from [19].

results of note are the detection of diffuse neutrino emission from galactic plane [72] and joint neutrino-gamma emission from a distance blazar, which was the first of its kind [63].

2.3 The Southern Wide-Field Gamma Ray Observatory

The Southern Wide-field Gamma ray Observatory (SWGO) is an upcoming observatory located in Pampa la Bola, Chile. While it is still in the planning stage, some prototype work has been completed and a tank design has been selected. As described in [19], SWGO will be built at an elevation of 4770 meters, significantly higher than HAWC, and will be approximately 1 square kilometer in area. The detector will comprise of three rings of tanks, with the first ring being the densest and HAWC-like, with the outer two rings being sparsely filled with smaller tanks. The fill factor, or percent of the area filled with tanks, will be approximately 70%, 4%, and 1.7% for the three layers.

As with HAWC, the main array will consist of many large tanks with, with potential measurements being a diameter of 5.2 meters and 4.1 meters tall [19]. Each tank will contain two stacked bladders and a double PMT setup. Inside each bladder is a 10 inch PMT, with one facing upwards and the other downwards. This design is shown in Figure 2.8, and is optimized for background rejection. The bottom chamber will reject cosmic rays via muon rejection [19].

The outer rings do not have final layout or design, but several prototypes are being tested.

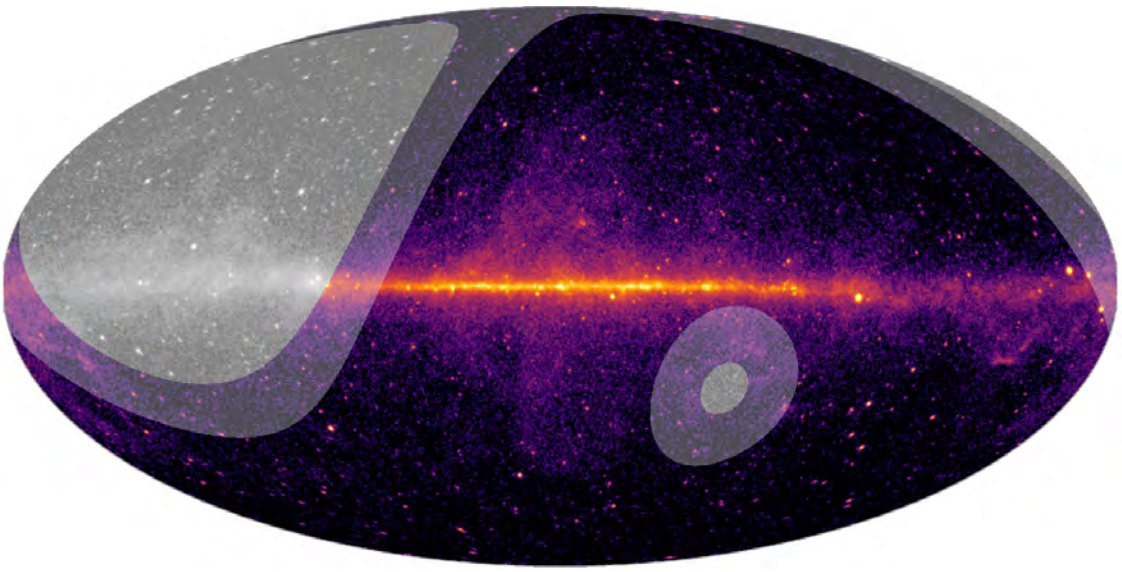

Figure 2.9 The projected field of view for SWGO. The bold grayed out regions are invisible and the lightly gray is high zenith for the detector. The figure is from [20].

These primarily focus on smaller, single or dual layer tanks with multi-PMT configurations. Three possibilities are currently considered for the sparse array configuration [20]:

1. an array of main-array tanks but with much further distances between the tanks
2. smaller tanks (approximately 1.7 meters tall) that have a single layer and a multi-PMT configuration
3. the same smaller tank but in the two-layer model configuration shown in Figure 2.8 of the main array tanks

The science goals of SWGO are many, with the key ones being the highest energy observatory in the Southern Hemisphere, an energy range from 100's GeV to > 1 PeV, and extending observations of both galactic and extragalactic sources. To date, there is only one other TeV observatory located in the Southern Hemisphere, the High Energy Stereoscopic System (HESS) Observatory [56]. SWGO's future field of view overlapped with *Fermi-Lat*'s full sky significance map is shown in Figure 2.9. Key regions are the inner galaxy, the galactic center, and many scattered blazars. This will allow extremely high energy observations with equivalent or superior angular resolution to other wide field of view instruments like HAWC and LHAASO.

CHAPTER 3

SERVICE WORK

This chapter highlights three projects during my Ph.D. that either were in service to a collaboration or to fulfill funding requirements. As part of HAWC, MSU currently maintains the online monitoring website for HAWC. When I joined the MSU HAWC group, I took over maintenance and feature development. For SWGO, I helped analyze weather data to assist in the site selection process. Finally, I received funding to travel to a national lab and perform an experiment related to high energy physics. In my case, I characterized fiber optic cables that were later used in measuring high energy particle interactions.

3.1 HAWC Monitoring

HAWC utilizes an online monitoring system to allow off-site shifters to monitor the experiment and report any issues to either expert shifters or the site crew as needed. This is done with two frameworks: one called the Advanced Tracking HAWC Experiment Notifications and Alerts (ATHENA) that collects and stores site data, and the HAWC Observatory Monitoring for Experiment and Reconstruction (HOMER) system to display it via a website.

3.1.1 ATHENA

ATHENA broadly handles three different functions: collecting on-site measurements, storing them in a database, and then pushing the data from the Mexican site to the computing cluster at the University of Maryland (UMD), where it can then be parsed and displayed by HOMER. The on-site measurements encompass most of the site systems, with examples including PMT rates, computer temperatures, and PMT high voltage. Since its inception, several significant additions like the outriggers have been added to the data collection suite [21]. The collection rate varies widely and is system dependent. The scaler rates are collected several times a minute while the water level in the tanks are polled every thirty minutes.

Once the data is collected, it is then pushed to the ATHENA server, a python-based library. This writes the data to a database using ZeroMQ, which each system receiving a table in the final MySQL database [21]. Approximately 8-12 gigabytes of data is written to this database every

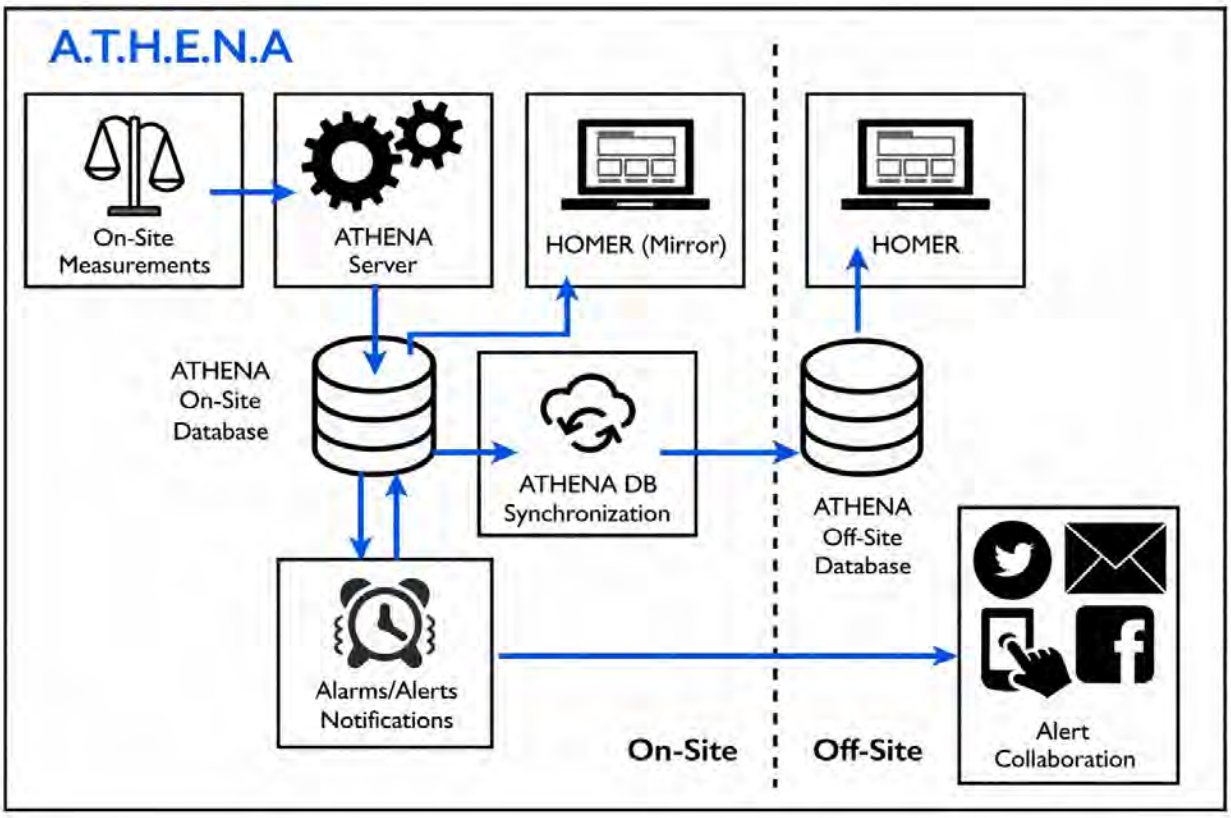

Figure 3.1 The layout for the ATHENA system. The on-site measurements are sent both to local and remote monitoring pages to allow the site crew to have real-time updates for any components that are being worked on. Figure is from [21].

year. There are two copies of the ATHENA database: one at site and one at UMD. One unforeseen issue that developed later was that MySQL scales poorly with large databases. This is discussed in Section 3.1.1.1.

The final step is to push the data from site to UMD. This process is split into three computational threads for different table groups: fast, slow, and computer. These labels correspond to how frequently data is collected from a given table (ex: TDC, online06, etc). Regardless of its label, each table is opened, the most recent 250 entries are selected, and then those entries are pushed to the remote database at UMD. In total, each table has approximately 20 kilobytes sent each time. The full flowchart of ATHENA is shown in Figure 3.1.

Whenever there is either a network or power interruption at site, the synchronization process gets

table_name	size_mb
measurements_TDCScalers_1	29770.52
measurements_TDCScalersMultiplicity_2	19443.22
measurements_scalers_21	15840.50
measurements_HVMon_1	3803.16
measurements_ORHighVoltage_1	2117.25
measurements_lv_1	1728.59
measurements_DataTransfer_4	1559.64
measurements_PDUs_1	1423.07
measurements_hv_1	1381.11
measurements_old_hv_1	1240.79
measurements_waterlevel_10	863.09
measurements_host_monitor_scaler_4	734.40
measurements_VMECrates_1	572.33
measurements_UPS_1	509.64

table_name	size_mb
measurements_TDCScalers_1	5464.63
measurements_TDCScalersMultiplicity_2	3675.38
measurements_ORHighVoltage_1	2743.04
measurements_DataTransfer_4	2155.43
measurements_scalers_21	1803.94
measurements_PDUs_1	948.82
measurements_VMECrates_1	698.32
measurements_HVMon_1	510.38
measurements_host_monitor_archive02_2	457.18
measurements_host_monitor_control_8	420.99
measurements_host_monitor_analysis01_3	405.26
measurements_UPS_1	368.06
measurements_host_monitor_scaler_4	364.97
measurements_host_monitor_service01_2	364.84

Figure 3.2 A comparison of the full (LEFT) vs culled (RIGHT) ATHENA databases. Significant space savings came from deleting old data from most tables in the ATHENA database. This greatly improved page loading times and stability.

delayed until normal operating conditions resume. The syncing scripts have automatic restarting functionality built-in, though manual intervention is occasionally needed. Generally, there are two points of failure: a syncing script fails to automatically restart, possibly due to too many failed attempts, or a network issue causes significant delays that the syncing script cannot overcome. When this happens, the relevant tables on the UMD database fall out of date and, by extension, the online monitoring website. Correcting this issue is MSU's responsibility and is discussed in Section 3.1.2.

3.1.1.1 Optimizing the ATHENA Database

One key project I worked on while maintaining the ATHENA database was optimizing it. This database is structured as an append only database; data is continuously added to the tables without ever removing it. While this is useful for long-duration tests for components like PMTs, this eventually caused significant issues syncing data from site to UMD, with pages on the online monitoring site frequently falling out of sync with the site. Additionally, page load times also started increasing with the database's size.

One hypothesis for these problems was the database's architecture. Initial testing found that MySQL has very poor performance when using large databases, so a plan was created to cull the database of old data. To this end, two steps were taken. First, a copy of the full ATHENA database was created that, if needed, could be reconstructed if a test culling failed. Second, a developer

space called HOMER-DEV was created to test for potential loading improvements. Additionally, this space could be used to test new HOMER features (see Section 3.1.2).

After extensive discussion within the collaboration, it was decided to keep a frozen copy of the full ATHENA database and delete any data from before April 20, 2023 from the live database. One issue that was encountered is that, while deleting data from a MySQL database is straightforward and can be run with

```
mysql -P $DB_port -h $DB_ip -u $DB_USER -p$DB_PASS $DB_NAME -e "DELETE FROM
$TABLE WHERE measurementTime < 168201960",
```

where the various $ parameters are specific to HAWC. It was found that executing this will not remove the physical size of the table. Testing revealed that, as the database is append only, it sequentially allocates space to store data. Crucially, even if data is deleted, the table "remembers" its previous size and keeps that space allocated in case it is needed again. To remove the excess space, each table then needed to be optimized with

```
mysql -P $DB_port -h $DB_ip -u $DB_USER -p$DB_PASS $DB_NAME -e "OPTIMIZE
TABLE $TABLE".
```

The results of this project are shown in Figure 3.2. Out of an initial size of 93 GB, the culled database cuts 57 GB out, with three years of data currently saved in the online database. There were significant stability and loading improvements with both the on and off site databases, with these being reflected in much better performance on the HOMER website.

3.1.2 HOMER

HOMER is the website that displays measurements from the site in near real-time. It utilizes the PHP language for pre-processing HTML [73] and uses SQL libraries to process the ATHENA database. The main page, shown in Figure 3.3, highlights the key information of the detector, such as the run duration, computer temperatures, high voltage status, and PMT rates. These are color-coded with green, yellow, and red for online shifters to quickly identify and report any issues.

Each of the main systems also have a dedicated page that contains more detailed information and plots of the data. An example of this is shown in Figure 3.4. For general array aspects like the

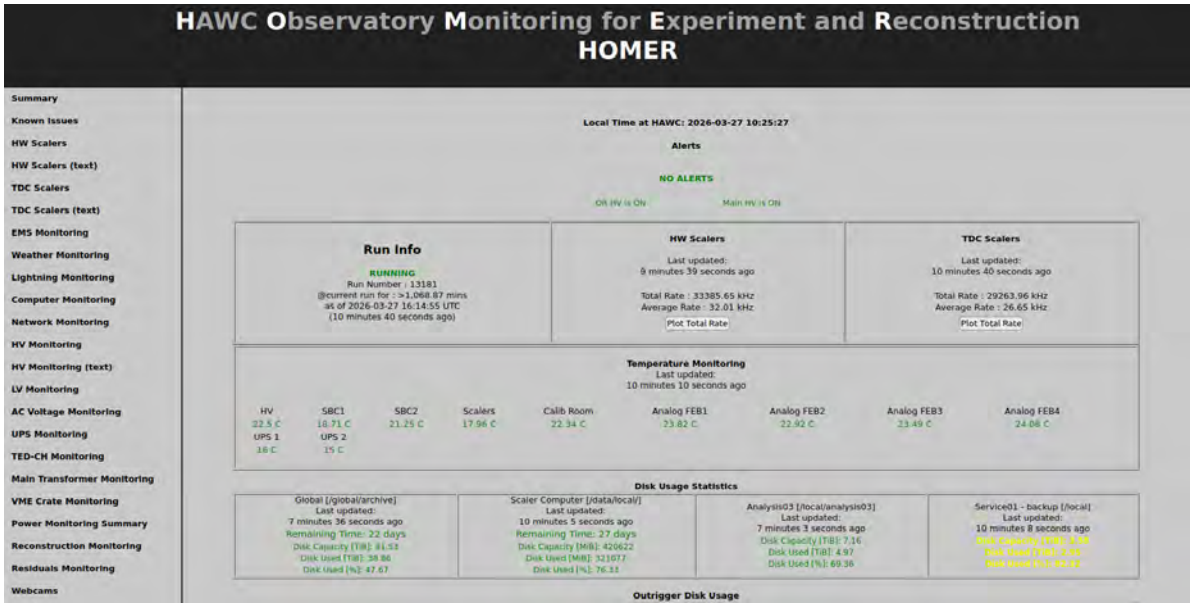

Figure 3.3 The main page of HOMER. This contains key information like the current data run status, the PMT response rates, temperature, and general computer information. The left column contains dedicated pages for various systems, each containing more detailed information or plots to indicate their behavior over the past two days.

TDC scalers, a top-down view of the main array is displayed with the status of each tank and PMT shown by a color scale. An example is shown in the left of Figure 3.4, with the operational PMTs shown in green. In this case, each PMT can be individually plotted by selecting it and entering in a desired date range. Other pages, like the VME crates, present tables of information for its current status.

The final feature of the monitoring system is the alert functionality. This polls various electronics and power inputs to see if they are within normal operating ranges. If they are not, a message is sent to HAWC's slack workspace indicating the problem. These messages range from temperature warnings to a full experiment shutdown if power is lost. HAWC has three backup battery sources that allow for the experiment to shutdown properly in the event of power failure. On start up, it is the role of expert shifters and the MSU HOMER team to ensure all monitoring and data acquisition scripts are functional.

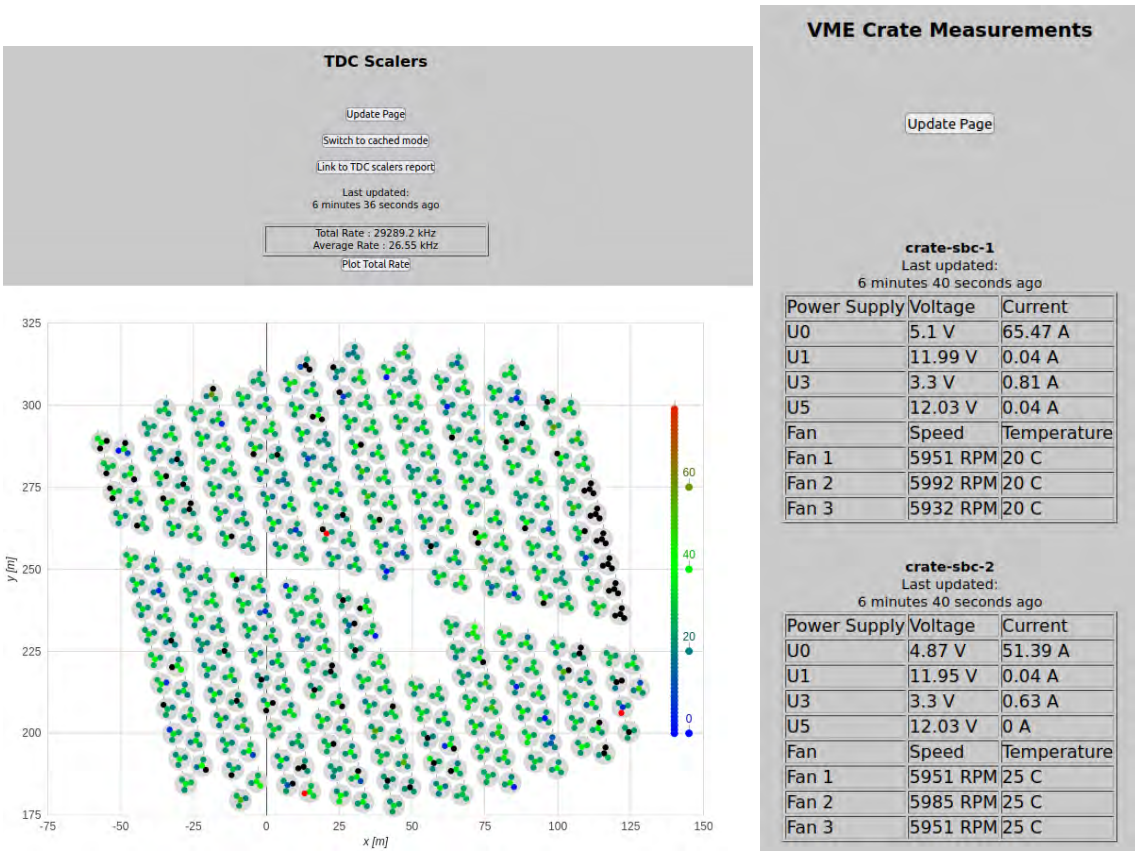

Figure 3.4 LEFT: the interactive TDC scaler plot for HAWC. Each PMT is shown in a color scale to indicate its status. PMTs not in the green are either down for maintenance or are slotted for future work. RIGHT: the status of the VME crates, a critical component of HAWC's operation.

3.2 SWGO Weather Analysis and Site Selection

This section focuses on the weather data analysis I performed for SWGO's site selection. There were three site options: Cerro Vecar in Argentina, Cerro Toco in Chile, and Imata in Peru. Each site had a weather monitoring station placed nearby, called AEROSITE, that captured temperature, humidity, atmospheric pressure, solar irradiation, wind speed and direction, electric field (efield) intensity and seismic activity. For Cerro Vecar and Imata, the AEROSITEs were placed at site, while Cerro Toco had its station placed a few miles away at Pamba la Bola. Of these, the temperature, wind speed, and efield are the most impactful to both construction and operation. Each of these sites were operational between one and two years, with data stored and accessed via MySQL databases (see above for a discussion on them). Additionally, 9 years of archival HAWC data was used for comparison, when necessary.

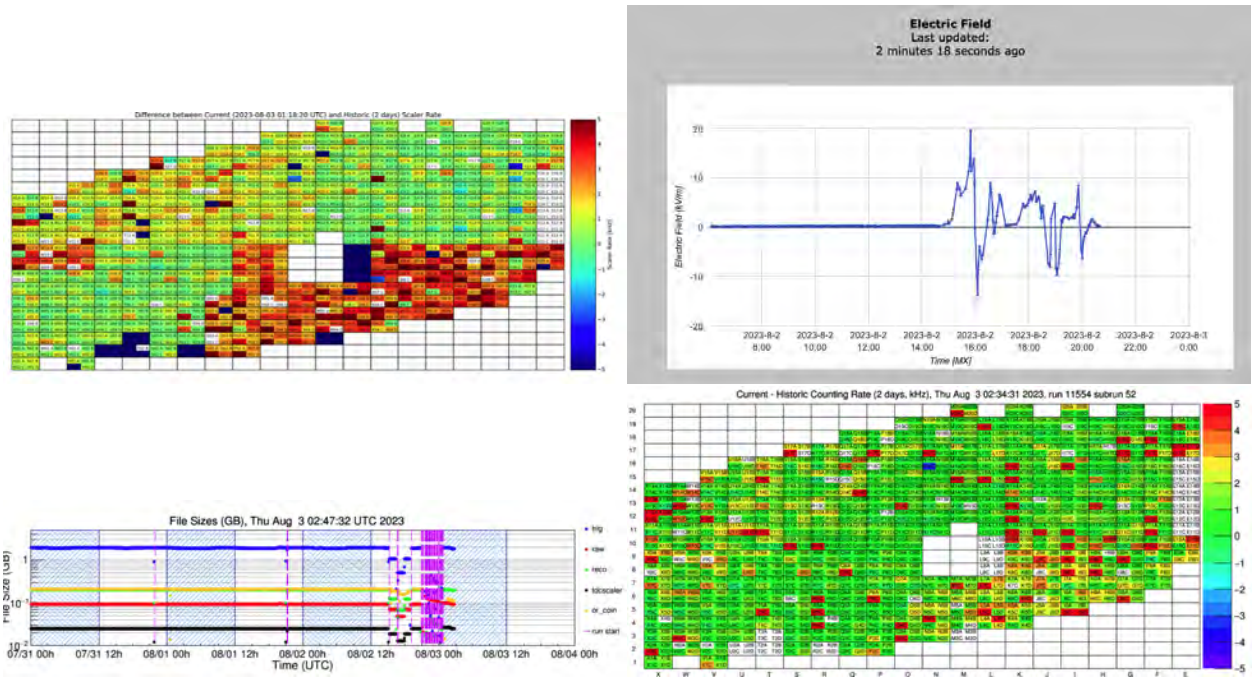

Figure 3.5 Different efield effects on HAWC. The upper left figure shows the scaler rate after a nearby lightning strike. Red and orange rates indicate the PMTs are fully saturated and are being overwhelmed by events. The upper right figure shows the direct efield measurement for this event. There is a large spike from the initial strike and then "echos" as the efield attempts to stabilize. The lower left figure shows the DAQ begin crashing from the over-saturation effects caused by the lightning strike. Once the efield normalizes, the experiment returns to normal operation. Lastly, the lower right shows what a high ambient efield looks like. While few PMTs are saturated, this still caused the DAQ to restart.

3.2.1 Electric Field

Arguably the most impactful weather parameter for operating closed tanks, electric field (henceforth "efield") fluctuations can and does (for HAWC) cause experiment restarts and loss of data. For HAWC and AEROSITEs, the efield sensor has a range of +/- 20 kV/m and, from HAWC's experience, experiment and data issues arise when the efield exceeds +/- 4 kV/m. While this normally happens when lightning strikes, it can also be caused by high ambient efields from clouds, nearby storms, or other atmospheric conditions. An example of what a lightning strike or general high ambient efield can do to HAWC data runs is shown in Fig. 3.5.

I looked at the following factors with these effects in mind: the number of events over the +/- 4 kV/m threshold, the separation between these two (are they more positive or negative), the percent

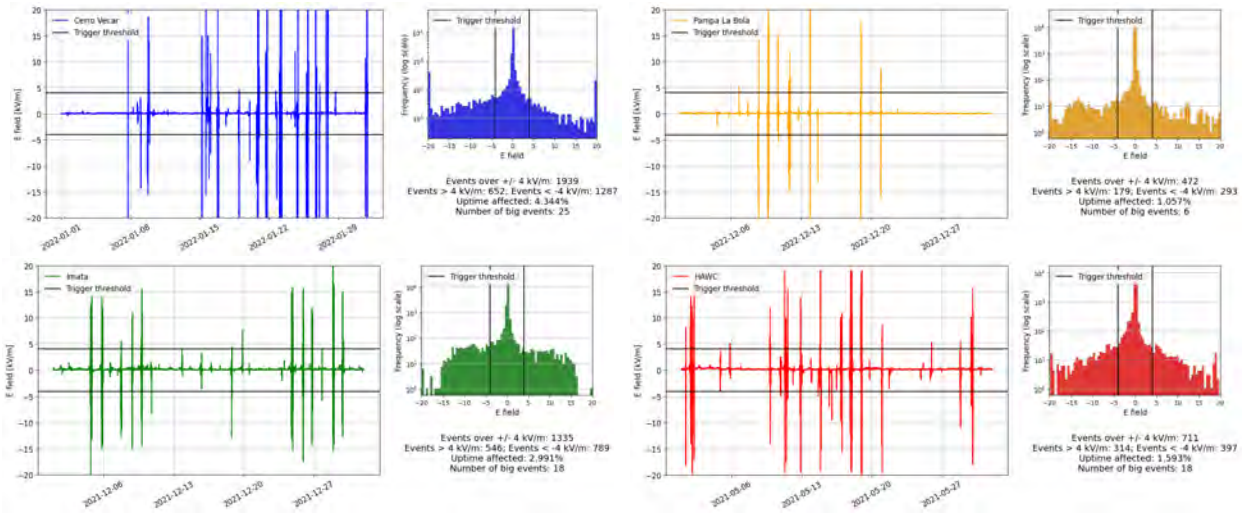

Figure 3.6 A typical active month for each site and HAWC. The left part of each plot shows the actual data collected for the month with the threshold boundaries labelled. The right parts are a log-histogram of that data to show the event rates of over-threshold events. As the sensors have a max range of 20 kV/m, there can be a larger excess there compared to other, nearby bins.

up-time affected in a given month, and the number of "big" events. The last parameter requires a bit more discussion.

A "big" event is anything that causes the variance of the data to go over a certain threshold value. This is to attempt to model lightning strikes in the data. A value of 60 was selected as it prevented small spikes from counting as a big event and seemed to catch most large events. It should be noted that this is an approximation and should only be used as such for these big events. An example for an "active" month for each primary site and HAWC is shown in Fig. 3.6.

The final combined estimated up time for each primary site is shown in Fig. 3.7. One note is that only Cerro Vecar has more that 2 years of efield data and, as such, none of the 3 sites have efield activity ranges like what HAWC does. All three sites + HAWC have notably more active summer seasons and, at least from the relatively short term data, Imata tends to be the most active.

3.2.2 Temperature

For temperature, the extra weather station data can be incorporated. To account for any systematic differences, they are plotted along side each other where the data is aggregated for every year a month has data. This can be seen in Fig. 3.8 where on the left is the hourly average and the

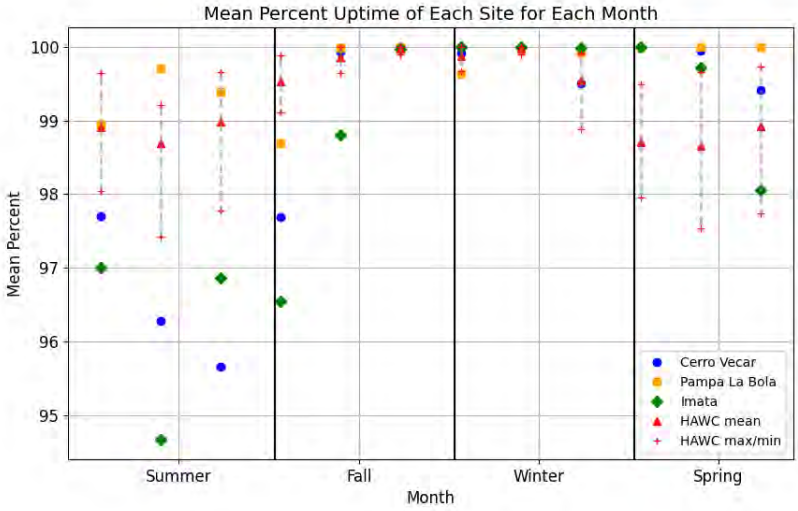

Figure 3.7 The estimated up-time for each site. It should be noted that a lack of data prevents the primary sites from having ranges for their up-times. The mean and extreme ranges of HAWC are included for context. The yearly average up times for each site are 98.8%, 99.7%, and 98.5% for Cerro Vecar, Pampa La Bola (Cerro Toco), and Imata respectively.

right is the minute resolution.

There are not many anomalies with the temperature range of each site, though Cerro Toco tends to be the coldest of the three (while also being at the highest altitude). As can be seen with the example month, the site consistently has lows below freezing the entire month of April. While this could be an issue of water potentially freezing inside the tanks, the massive amount of thermal energy the water holds should prevent any freezing from occurring.

3.2.3 Wind

Moving to wind speed, all sites have high wind speeds periodically. It should be noted that the extra weather station data may have sensor issues as there are many peaks of > 200 m/s so I applied a cut-off, where if the wind speed exceeded 30 m/s (67 miles per hour), it is excluded from the data set as construction and safe operation become impossible well before that. As these weather stations most likely have different sensors than the AEROSITE one, there is little hope of correctly determining what might cause this issue. As such, while extra data from nearby weather stations are shown, more weight should be given to the AEROSITE data as it was calibrated by the SWGO collaboration.

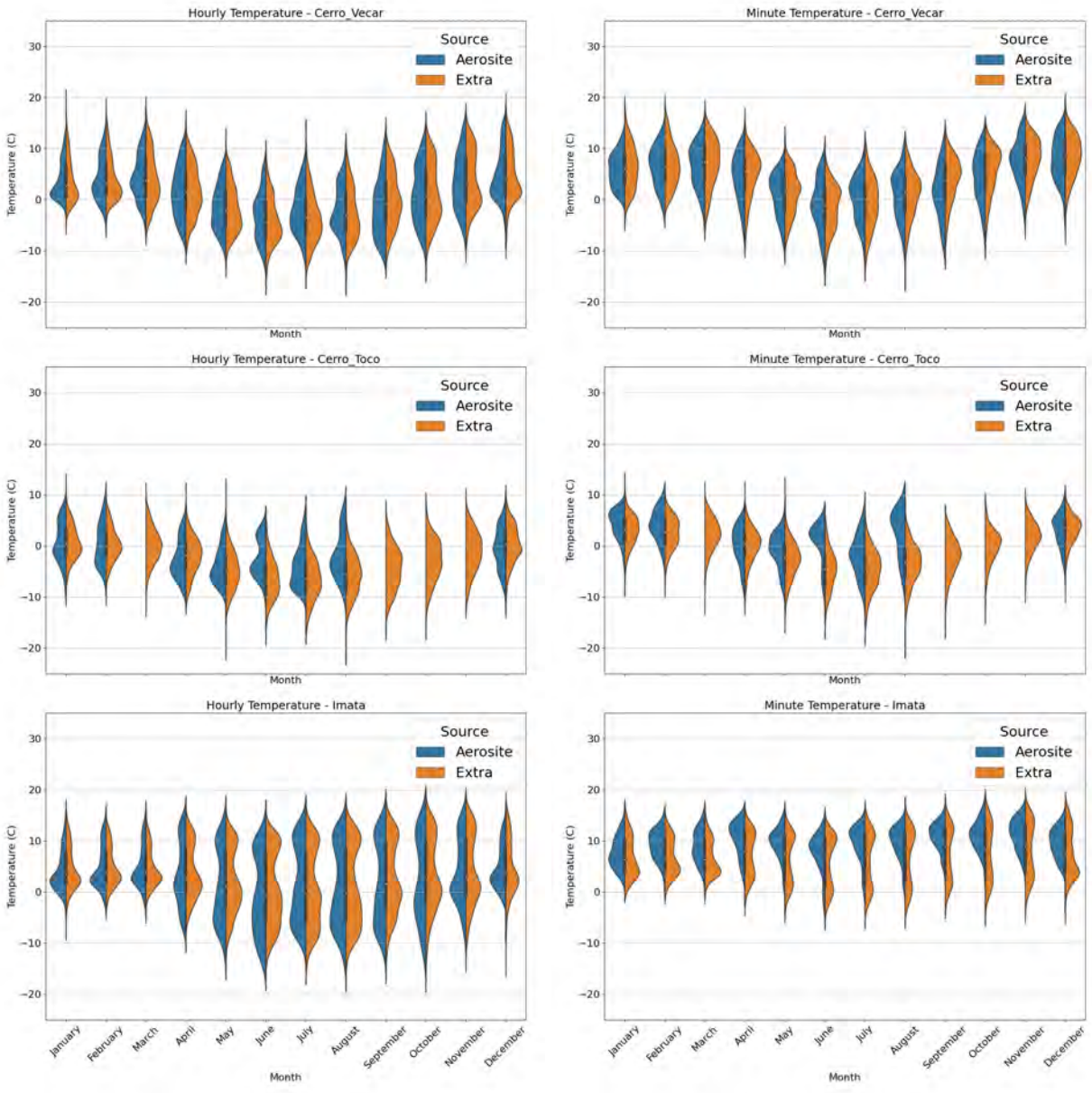

Figure 3.8 Temperature spread for all 3 sites with hour and minute resolutions. The left part of each violin is the AEROSITE data while the right has extra data from neighboring weather stations. Generally, both data sets observe comparable temperature distributions.

From Fig. 3.9, Cerro Vecar (see Table 3.1 as well) consistently has the highest wind speeds, routinely gusting over > 16 m/s, though the other two also have periods of high wind speeds as well. While this will most likely not have any significant impacts on experiment operation, it could cause issues with construction as high winds can blow over construction materials, especially if they are not anchored well. Per the Occupational Safety and Health Administration (OSHA) in the

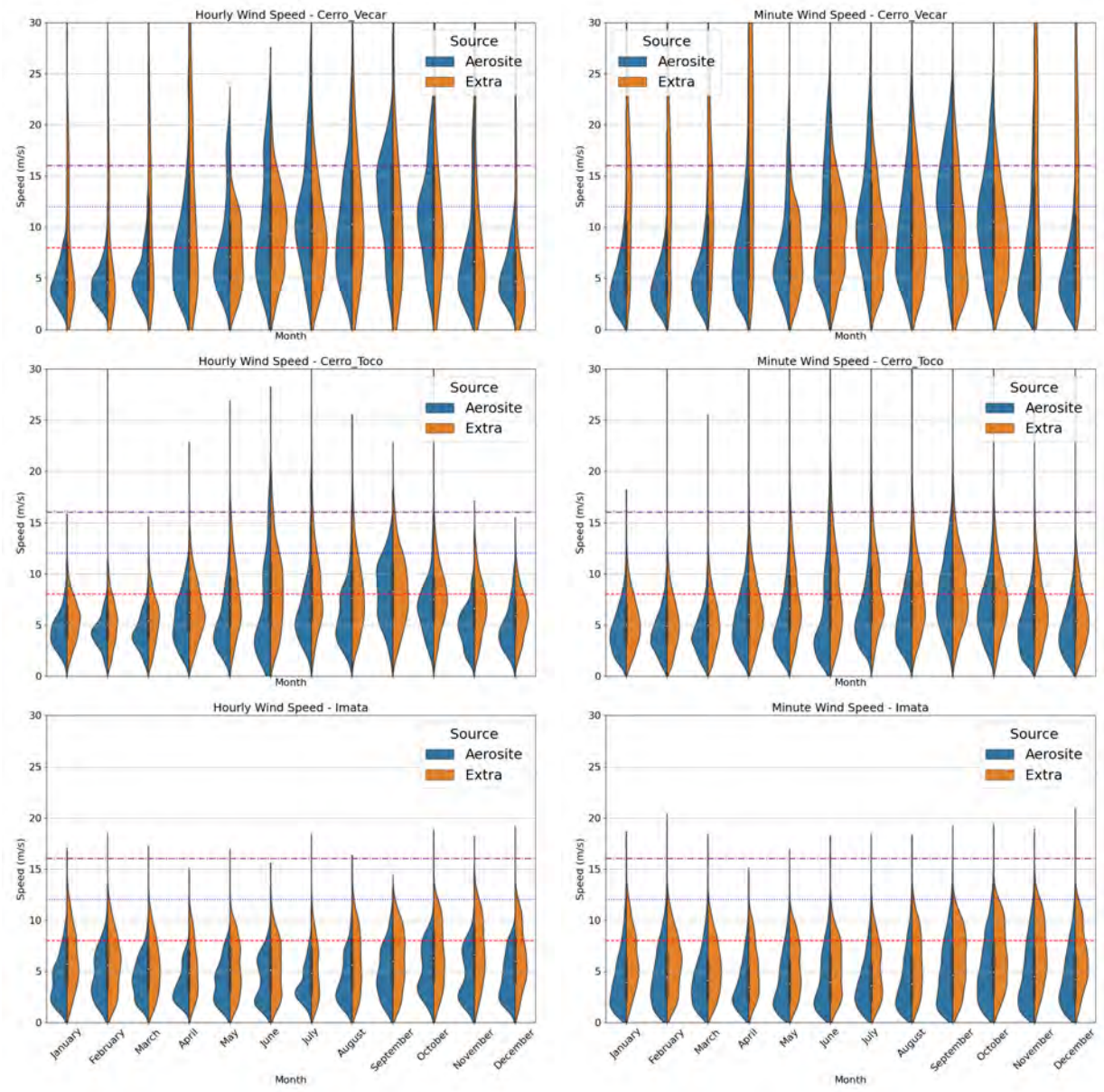

Figure 3.9 Wind spread for all 3 sites with hour and minute resolutions. The left part of each violin is the AEROSITE data while the right has the extra data from neighboring weather stations. Generally, both data sets observe comparable wind speed distributions, though the extra data seems to have odd spikes that may skew their data. The lines indicate 8, 12, and 16 m/s winds respectively and were used to generate the values in the tables.

United States [74], high winds for construction are considered to start around 20 mph or 8-9 m/s, so all sites will have challenges with building. At higher wind speeds, such as the gusts > 20 m/s (45 mph), it becomes dangerous or impossible to build from the force of the wind.

Tables 3.1, 3.2, and 3.3 give the respective percent of time spent above 8, 12, and 16 m/s for

Cerro Vecar, Cerro Toco, and Imata. This percent was found from the minute-based dataset and only considered AEROSITE data, given the inconsistencies encountered with the extra weather station data.

One last note is that while Cerro Toco was missing temperature data for 9 months in 2022, resulting in 4 missing months in the aggregation, the wind and efield sensors remained unaffected and continued collecting data.

3.2.4 Environment in the work hours

The times considered for normal work hours are 8 am (08:00) to 5 pm (17:00), Monday through Friday. Comparing to the full dataset, there is a clear trend towards warmer temperatures, as should be expected (the graph is in the appendix, Fig. A.1). Otherwise, it appears that, for all sites, there are still periods in the early morning when the temperatures are at or below 0 degrees Celsius so warmer clothing will be a requirement for all people working at the sites. Given the high elevation, this is not surprising.

As for the wind, broadly speaking, all sites see considerably less wind activity but Cerro Vecar still seems some high wind periods. The graphs are in Appendix A, Fig. A.2.

month	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Average
8 m/s	13.6	10.9	12.9	33.5	34.4	49.2	63.4	56.8	75.2	62.0	34.7	10.7	37.9
12 m/s	2.4	1.8	2.6	12.3	14.9	24.7	30.6	29.2	44.2	31.4	16.2	2.4	17.2
16 m/s	0.3	0.2	0.3	4.3	5.6	10.8	12.6	12.0	17.8	10.9	6.5	0.2	6.3

Table 3.1 Percent of time the wind speed went over each threshold for Cerro Vecar for each month

month	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Average
8 m/s	3.9	3.0	5.2	10.9	13.9	12.6	13.8	9.4	36.0	15.6	10.8	9.2	11.8
12 m/s	0.2	0.1	0.1	1.2	3.2	4.6	4.7	2.0	10.1	2.7	0.7	0.5	2.7
16 m/s	0.0	0.0	0.0	0.1	0.7	1.5	1.8	0.4	1.6	0.4	0.0	0.0	0.6

Table 3.2 Percent of time the wind speed went over each threshold for Cerro Toco for each month

month	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Average
8 m/s	4.3	5.9	4.5	2.2	4.4	5.0	3.2	5.6	8.8	3.3	7.3	5.4	4.9
12 m/s	0.3	0.4	0.3	0.1	0.1	0.3	0.2	0.4	0.9	0.3	0.6	0.3	0.3
16 m/s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 3.3 Percent of time the wind speed went over each threshold for Imata for each month

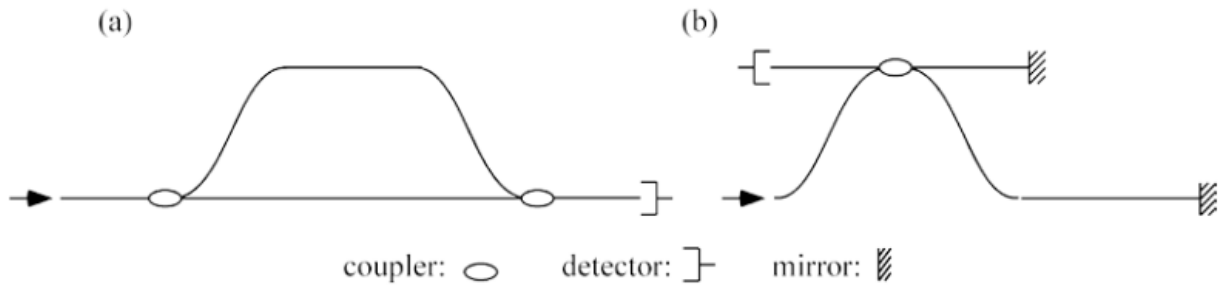

Figure 3.10 The layout for Mach-Zehnder (a) and Michelson (b) interferometers. Figure is from [22]

3.3 TRAIN-MI Hardware Project

Part of my Ph.D. was funded by the TRAIN-MI fellowship that was offered by the High Energy Physics (HEP) group at MSU and sponsored by the Department of Energy (DOE). The goal of this fellowship was to provide two years of funding and travel coverage to an authorized DOE lab to participate in a hardware project related to high energy instrumentation. My lab was the Los Alamos National Lab (LANL) in Los Alamos, New Mexico. I was under the guidance of Pat Harding and Kelly Malone. My project focused on Optical Backscatter Reflectometry (OBR) and its applications to cable characterization. I specifically used a Luna OBR 4600 to characterize fiber optic cables.

One additional note is that, due to the setup being located at LANL, I was not directly allowed to take any screenshots or photos of the setup. The photos I do have were taken by Pat Harding and were cleared for release by the lab. All other figures in this section are from reference materials and are cited accordingly.

3.3.1 Fiber Optic Interferometry

To characterize fiber optic cables, I used a Luna OBR 6400. This device uses interferometers to split beams down cables, with the reflected emission accurately mapping the entire cable. Two interferometers can be used for this: Mach-Zehnder and Michelson interferometers. Shown in Figure 3.10, the Mach-Zehnder design uses two couplers, one to split the laser beam to send the beams across two different paths and then one to recombine it before sending it to a detector. The

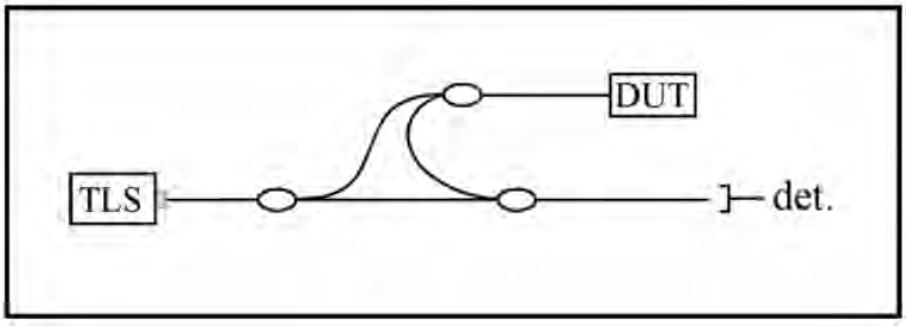

Figure 3.11 The layout for the Luna OBR 4600. The TLS is the tunable light source that sends light to both the detector and a device under test (DUT). The DUT is one one arm of a Mach-Zender interferometer. Figure is from [22].

Michelson design uses a single coupler to both split and recombine the beam after reflecting off mirrors.

Regardless, both designs can be treated with the same mathematics. Considering an input electric field as

$$E_{\text{in}} = E_0(t)E^{-i\omega(t)t}, \tag{3.1}$$

where $\omega(t)$ is the instantaneous frequency of the laser. Once the laser hits the decoupler, it is split into two fields that propagate down different length fibers before reflecting back and recombining. The resulting electric field at the detector is given by

$$E_{\text{out}} = E_0(t+\tau_1)e^{-i\omega(t+\tau_1)t} + E_0(t+\tau_2)e^{-i\omega(t+\tau_2)t} \tag{3.2}$$

where τ_1 and τ_2 are the time delays caused by the different paths in the interferometer. To find the electric output of the detector, the optical intensity is considered. This is proportional to the square of the electric field and, with $\tau = \tau_1 - \tau_2$ [22], is given by

$$I(\omega,t) = |E_0(t)|^2 + |E_0(t-\tau)|^2 + 2\,E_0(t)\,E_0(t-\tau)\cos\left[\omega(t)\tau\right]. \tag{3.3}$$

3.3.2 Optical Network for the Luna OBR 4600

The simplified layout of the Luna OBR 4600 is shown in Figure 3.11. A tunable laser sends light into decouplers that split the beam into a Mach-Zender interferometer. Part of the beam heads to a device under test (DUT) before heading to the detector, while the other fraction goes straight to the detector. The DUT is described by its frequency domain linear transfer function $H(\omega)$. This contains the amplitude $\rho(\omega)$ and phase $\phi(\omega)$ responses from the DUT, and is given by

$$H(\omega) = \rho(\omega)e^{i\phi(\omega)}. \tag{3.4}$$

Given the phase and amplitude responses, the detected power is once again proportional to the electric field intensity, and is given by [22]

$$I(\omega, t) = |E_0(t)|^2 + |E_0(t-\tau)|^2\rho(\omega)^2 + 2\rho(\omega)E_0(t)E_0(t-\tau)\cos[\omega(t)\tau - \phi(\omega)]. \tag{3.5}$$

This intensity is then put through a Fourier transformation to decouple the different response functions inside the intensity. The first two functions only contain low characteristic frequencies, therefore only being present around t=0. By contrast, the third term, called the interference term, oscillates in the time domain at the frequency $\omega(t)\tau$ and contains the DUT's information. Its time-domain location is determined by measuring the arm length difference τ. For a large enough τ, the device's response can be easily separated from the low frequency terms.

The Fourier decomposition for a given signal from a fiber optic cable is shown in Figure 3.12. As previously discussed, the DUT's response will oscillate in the time domain, so any peak can be analyzed to characterize the DUT. As the system gets more complex, more peaks will appear, separated by the fraction of the total distance they are from the laser.

Regardless of the system's complexity, an inverse Fourier transformation is then done to extract the desired time-domain data that contains either the entire or partial DUT response. This is done via Luna's software, and gives the losses from the test device as well as its group delay (length) [22]).

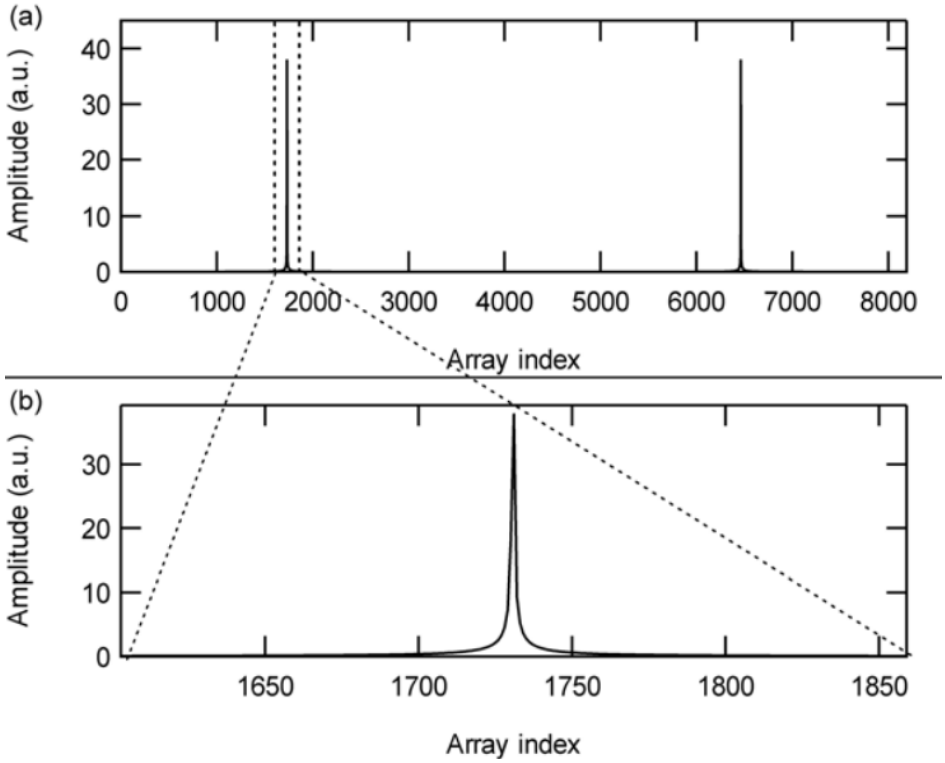

Figure 3.12 TOP: The Fourier transformation of the intensity function. The two peaks correspond to the oscillating response from the DUT. BOTTOM: The zoomed-in response from the device. Both plots are in the time domain (x-axis) with amplitudes of the response (y-axis). Figure is from [22].

3.3.3 Project Goal: Characterizing Cables

My project goal was to measure and characterize a batch of fiber optic cables that were needed for an experiment. My main objective was to measure 8 custom cables made internally at LANL and determine how close they were to one meter (the needed length). To do this, I utilized a large network switch cabinet (photo was not allowed to be taken) that contained each of the aforementioned cables and a reference cable known to be 1 meter long. I would then compare each cable to the reference, recording the loss and length for each.

The general process and user interface to the Luna are shown in Figure 3.13. For each cable, I would connect and select the "scan" option. This would start the measurement process and take approximately five seconds. Next, I located the start and end points of the cable in the time-domain space generated from the Fourier transformation described above. Shown in the bottom of Figure 3.13, the start is characterized by a large peak and the end having a large peak before dropping into

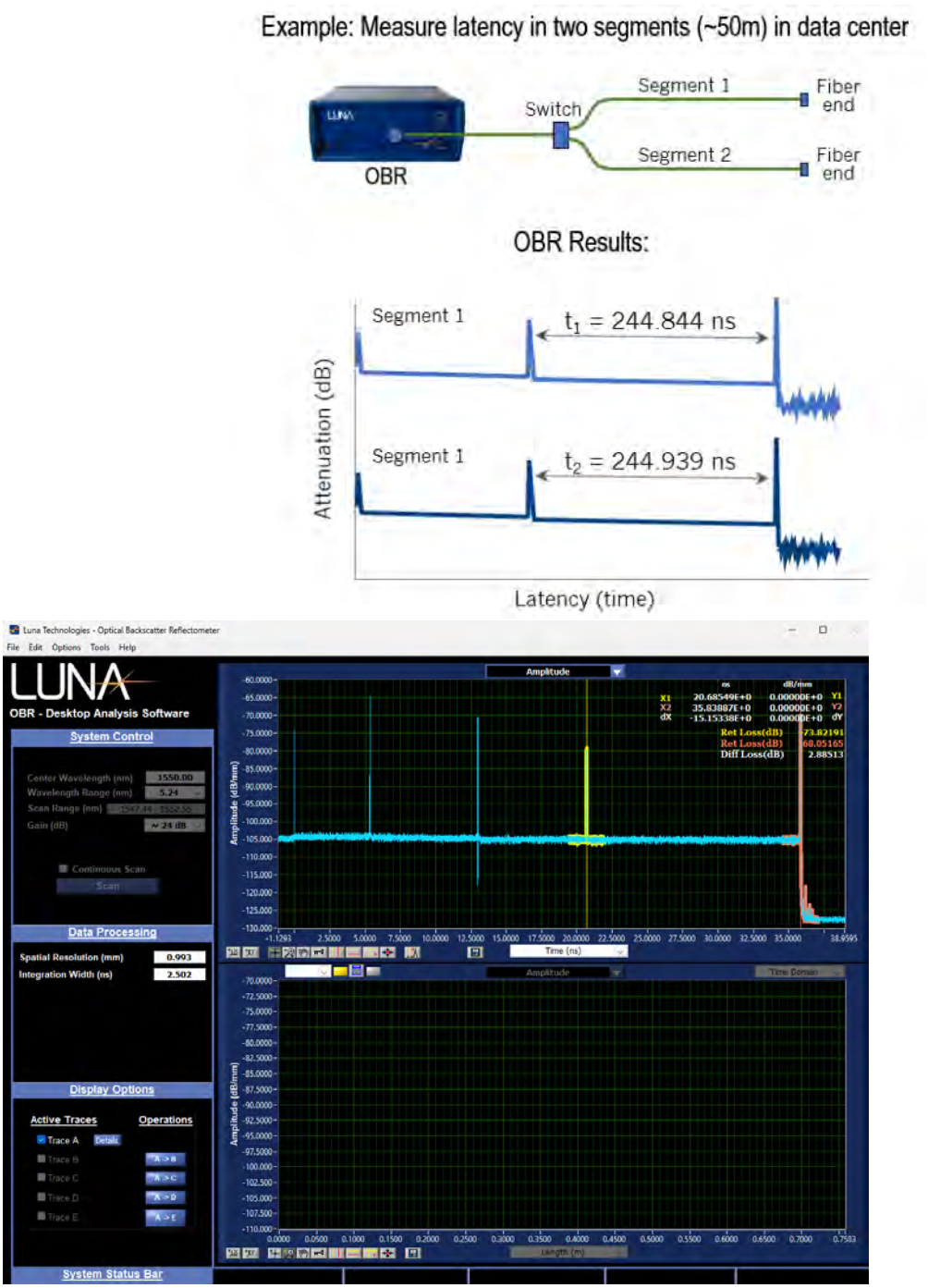

Figure 3.13 Top: An example of what two cable measurements look like with the OBR 4600. The length is determined by multiplying the time by the speed of light, with the difference in time (95 picoseconds) corresponding to a length difference of 19.3 millimeters. Figure is from [23]. BOTTOM: The actual user interface for the OBR. The yellow spike indicates the start of the cable with the orange region being the end. The software automatically finds the time-location of the highest amplitude and calculates the distance between them. Photo courtesy of Pat Harding at LANL.

noise. The other peaks before the cable start are internal reflections (first two terms in Equation 3.5) from the OBR and can easily be found by running a scan before connecting a cable to the Luna.

Once the reference cable was measured, I then switched to one of the unknown cables and ran the scan again. The Luna can store five cables simultaneously, with A defined as the reference and B-E the unknown cables. The upper figure in Figure 3.13 shows what the comparison between two cables looks like. They both have the same start point, but diverge with the end points based off their length difference. The length difference can then either be found by finding the individual lengths from the start and end peaks and taking the difference, or directly calculating it by selecting the desired display options in the bottom left corner of the bottom figure in Figure 3.13 and finding the end peaks there. Lastly, signal loss is defined as any loss in the amplitude of the light as it passes through the fiber.

The end result of this study found six of the cables to be within 0.005 meters of the reference, but cables 3 and 6 were significantly longer at 1.1 meters. No cables showed significant signal loss, with losses being measured at $< 1\%$. These two cables were later found to have come from a different batch, hence the significant deviation from the other six cables.

As previously stated, the goal of this project was to characterize fiber optic cables to determine both their length and any potential losses present that could affect data quality. Having precise length measurements is crucial to performing timing analysis for experimental setups. In this case, extremely accurate timing was required to reconstruct high energy particle interactions within a sealed chamber. Likewise, this process was done for the thousands of fiber optical and high voltage cables used when constructing HAWC. This project would not be possible without the facilities and expertise at LANL, along with the funding from the TRAIN-MI fellowship through MSU and the DOE.

CHAPTER 4

AN ANALYSIS OF PWN HAWC J2031+415

This work is a continuation of my Master's at Michigan Technological University (MTU). First, I will briefly introduce the history of HAWC J2031+415, then discuss what was done at MTU and the status of the project when I graduated, and then lastly highlight what I did at MSU. I have several publications on this source: my master's thesis [75], an Astrophysical Journal publication [37], along with two proceedings at the International Cosmic Ray Conferences in 2021 [76] and 2023 [77].

4.1 Background

First observed in 2005 by the High Energy Gamma Ray Astronomy (HEGRA) experiment, TeV J2032+4130 was the first very high energy (> 100 GeV) gamma-ray source in the TeV range with no lower energy counterpart [24]. TeV J2032+4130 is located in the *Fermi-LAT* Cygnus Cocoon region, a large extended source of GeV gamma-ray emission that contains the Cygnus OB-2 star cluster [45]. The extent of the original HEGRA detection was 0.11° and was comparatively dim to HEGRA's detection of the Crab above 1 TeV at 5% of the Crab's flux [78, 24].

Follow-up studies by the X-ray observatories Suzaku, Chandra, and XMM-Newton revealed significant diffuse non-thermal X-ray emission co-incident with TeV J2032+4130 [79, 80]. In [79], they revealed two sub-structures, one of which was co-incident with the pulsar PSR J2032+4127. In addition, they also observed a large diffuse excess measured across TeV J2032+4130's extent [79]. XMM-Newton's detection is roughly the same size as TeV J2032+4130, though no sub-structures were found [80]. Both measurements had fluxes significantly lower than that of the gamma-ray source. Two hypotheses were proposed for the emission: hadronic, driven by pion decay, and leptonic, produced via a combination of synchrotron and inverse Compton scattering. Both hypotheses are considered in this analysis and are discussed in Section 4.7.

Radio observations made using the Very Large Array (VLA) revealed a large number of radio sources in the direction of TeV J2032+4130's center of gravity (CoG). One of these sources was characterized by faint, non-thermal emission in roughly a half-circle around the CoG with a total

area of $\sim$ 27 arcmin^2 [81, 82]. The region had an estimated energy content of 6×10^{45} erg and seemed to indicate an efficient injector of nonthermal particles. Additionally, the semi-circular shape of the emission region seems to indicate an old supernova shell and may be the radio counterpart of TeV J2032+4130.

Later TeV observations by the Very Energetic Radiation Imaging Telescope Array System (VERITAS) in 2014 [83] and 2018 [27] found emission that corresponded to an asymmetric Gaussian within the energy range 0.5-50 TeV. In both [83] and [27], they hypothesize that the emission is from a pulsar wind nebula (PWN) and whose source is PSR J2032+4127. Furthermore, while they do not observe one, they predicted a cut-off near 10 TeV.

In the second [84] and third [85] catalogs published by the High-Altitude Water Cherenkov (HAWC) observatory, sources 2HWC and 3HWC J2031+415 were detected coincident with TeV J2032+4130. A follow-up dedicated analysis resolved two sources: HAWC J2031+415, which was associated with the probable PWN, and HAWC J2030+409, believed to be the TeV extension of the *Fermi-LAT* Cygnus Cocoon [46]. Though that analysis focused on the Cygnus Cocoon (henceforth referred to as the Cocoon), it was found that HAWC J2031+415 had an extension of 0.27° and a power-law with exponential cut-off spectral model with a cut-off on the order of $>$ 10 TeV and is consistent with VERITAS' observations [46, 83].

As asserted by VERITAS, PSR J2032+4127 is most likely the power source for the PWN. PSR J2032+4127 is an unique pulsar to power a PWN. First, it is old at an estimated characteristic age of $\sim$ 200 kyr and has an estimated spin-down luminosity $\dot{E} = 1.5 \times 10^{35}$ erg/s. Current estimates have moved it from 3.8 kpc in [79] to 1.33 ± 0.06 kpc in the most recent pulsar catalog published by the Australian Telescope National Facility (ATNF) [86]. This places it inside the Cocoon, which has a distance of $\sim$ 1.4 kpc [46, 45]. Additionally, it is a long period binary with the $\sim$ 15 $M_\odot$ star MT91 213 and has an orbital period of 50 years [87]. This makes the system unique, as TeV binary gamma-ray pulsar systems are rare. While not originally associated with observed X-ray emission [87], the pulsar is now believed to be responsible for it [83] and will be considered for the multi-wavelength analysis presented in Section 4.7. Additionally, in November 2017, it performed

its periastron with MT91 213 and flaring in both gamma and x-rays was detected [26]. This is investigated using HAWC data.

4.2 Previous Work on HAWC J2031+415

At MTU [75], I performed a source search of the Cygnus Cocoon region using 1434 days (pass 4) of HAWC data and resolved HAWC J2031+415 and J2030+409, and 3HWC J2020+403. I then found the best-fit spectra for each source: cut-off power law, power law, and power law, respectively. As my focus is the J2031, I then subtracted out J2030 and J2020, isolating the PWN.

With the PWN isolated, I performed an energy-dependent morphology study of HAWC J2031+415. Beside having a curved spectrum, another characteristic of PWN is having a strong energy-dependent morphology (changing shape with changing energy). For extremely bright sources like the Crab nebula, it is possible to take 1 energy estimator bin and perform a standard likelihood fit. That is not possible with HAWC J2031+415 as it is only a small fraction of the Crab's flux. Therefore, a different approach is needed; instead of directly fitting the data, I count the excess events.

This method, first used in analyzing HAWC J2019+368 [88], functions by first defining a rectangular ROI, defined as being 6° long, 1° wide, and rotated by 15° to align between HAWC J2031+415 and PSR J2032+4127. Next, the data set is divided into 4 bands, each with 3 energy estimator bins: abc, def, ghi, and jkl. These correspond to the energy ranges 0.316-1.77, 1.77-10, 10-56, and 56-316 TeV. Through careful calibration (see Section 5.3.2 in [75]) and bin selection (Section 5.3.3), the source size was found for bands 2, 3, and 4. No clear morphology was found as all 3 bands were within uncertainties.

4.3 HAWC J2031+415 and MSU

After I came to MSU, the HAWC collaboration released its fifth dataset (pass 5), that contained 2400 days of data. I performed some preliminary refits to the Cocoon region and found significantly different results, and so I redid the analysis during my first two years at MSU, although with several additions. This culminated in the work being published in 2024 [37]. The following sections are from the paper that I wrote (along with figures). Broadly, the spectral results, energy ranges, and

the energy-dependent morphology study all were greatly affected with the newer pass 5 data. For calibration details, please see Chapters 4 and 5 in my thesis [75] as they are the same as what was done in this chapter.

4.4 Source Search and Spectral Fitting

4.4.1 Region of Interest Considered

The data described above is subdivided into a Region of Interest (ROI). The ROI is a 6° circular region centered on ($l = 78.9°$, $b = 1.6°$) with a mask on 3HWC J2019+367. This is shown in Figure 4.1. As done in [46], the mask on 3HWC J2019+367 is used to prevent potential contamination caused by the brightest source in the Cygnus Cocoon region.

4.4.2 Method

To fit the gamma-ray data, I used both the Multi-Mission Maximum Likelihood framework [89, 3ML][1] and the HAWC Accelerated Likelihood (HAL)[2] [90] plugin. This implementation allows extensive multi-source fitting for complex regions. The framework considers a test statistic (TS) that evaluates the statistical significance of a given model with a given number of free parameters. The TS is used to compare an alternative hypothesis with a null hypothesis. It is defined as

$$\mathrm{TS} = 2\ln\left(\frac{L_{\mathrm{alt}}}{L_{\mathrm{null}}}\right) \tag{4.1}$$

If two alternate nested hypotheses are compared, $\Delta\mathrm{TS} = \mathrm{TS}_2 - \mathrm{TS}_1 = 2\ln(L_2/L_1)$ can be used to determine which model is preferred [67]. If the difference in free parameters between the models is 1, then Wilks' theorem can be used to give a pre-trial significance that follows $\sigma = \sqrt{\mathrm{TS}}$ [91].

To fully model the emission in the ROI, I performed a source search method similar to that of the *Fermi*-LAT extended source catalog [92]. All models considered are from the *Astromodels*[3] Python package. This search method can be broken into 3 broad sections: a Point Source (PS) search, an Extension (EXT) text, and then a spectral test. The process is defined by the following steps:

[1] `https://github.com/threeML/threeML`
[2] `https://github.com/threeML/hawc_hal`
[3] `https://github.com/threeML/astromodels`

1. First, contributions from both unresolved source emission and the diffuse background emission is modelled using the Unresolved Radiation Model (URM). This model uses a 2D Gaussian template centered at 0° along the galactic plane. The spectral model is defined as a power law (PL) defined by

$$\frac{dN}{dE} = N_o \left(\frac{E}{E_p} \right)^{-\alpha}, \tag{4.2}$$

where N_o and α are the flux normalization and index for the source. E_p is the fixed pivot energy for the source and is selected to minimize the correlation between the flux normalization and index. As done in [93], a spectral index of 2.7 and a pivot of 7 TeV is assumed while N_o is free to float with the fit. After the fit, the TS is calcuated and the ROI is searched for any remaining positive excess.

A floating PS is added to the URM model (1 PS + URM) at the pixel of highest remaining emission. The ROI is then refitted with this new model (PS + URM), the TS is calculated, and then compared to the URM only model. If the ΔTS > 25, then the 1 PS + URM is the preferred model and is kept as the new preferred model.

2. Another floating PS is then added to the model at the next pixel of highest significance and the process is repeated. This continues until adding a PS does not give a ΔTS > 25, in which case the PS search is completed. The final PS is excluded from this last model as it failed to cross the necessary TS threshold.

3. All sources must initially be detected as point sources, regardless if they are extended or not. An extended source (EXT) is 2D symmetric Gaussian model defined by

$$\text{EXT} = \left(\frac{180}{\pi} \right)^2 \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\vec{\theta}^2}{2\sigma^2} \right) \tag{4.3}$$

where the $\left(\frac{180}{\pi}\right)^2$ factor converts degrees to radians and the $\vec{\theta}$ and σ parameters are the location and radius of the Gaussian, respectively. The brightest PS is converted to an EXT

and the region is refit with the location of all sources fixed but the index and flux normalization free. There are two possible outcomes:

a) If ΔTS > 25, then the extension is preferred and kept. After the fit, all sources that have a TS < 25 are dropped from the model.

b) If ΔTS < 25, the extension is rejected and the model reverts to the pre-extension test.

4. Regardless of the previous EXT test, the next brightest PS is then tested. This process repeats until all sources have been tested. Once the extension test is completed, all the remaining sources in the ROI are refit with free locations.

5. In a similar fashion, the spectrum of each source is now tested. There are 3 models considered: the PL model in Equation 4.2, a power law with an exponential cutoff (PLC), and a log-parabola (LP) model. The latter two are given by Equations 4.4 and 4.5 respectively.

$$\frac{dN}{dE} = N_o \left(\frac{E}{E_p}\right)^{-\alpha} \exp\left(-\frac{E}{E_c}\right) \tag{4.4}$$

$$\frac{dN}{dE} = N_o \left(\frac{E}{E_p}\right)^{-\alpha-\beta\ln(E/E_p)} \tag{4.5}$$

The extra terms E_c and β are the cut-off energy and curvature of the spectrum respectively.

6. To compare the spectral models, the Akaike Information Criterion (AIC) is used and is given by [94]

$$\text{AIC} = 2k - 2\ln(\hat{L}) \tag{4.6}$$

The AIC adds a term that penalizes more complex models given by the k number of parameters for a given maximized likelihood $\hat{L}$. As outlined in [95], the lower the AIC, the more preferred a model is. A ΔAIC between 0 and 2 indicates both models fit well and the simpler should be considered and, for ΔAIC > 3, the higher AIC model becomes decreasingly likely.

4.4.3 Results

Once the source search process has been completed, the final source model contains the following: URM and three extended sources: HAWC J2031+415, HAWC J2030+409, and 3HWC J2020+403 and is in agreement with [46]. HAWC J2030+409 is associated with the *Fermi*-LAT Cygnus Cocoon [92]. HAWC J2031+415 is near co-incident with TeV J2032+4130 and is assumed to be its higher energy counterpart. 3HWC J2020+403 is coincident with the supernova Gamma Cygni and has been also observed by *Fermi-LAT* [3].

The results from the fitting process are given in Table 4.4.3 and the preferred spectral models for the sources being as follows: HAWC J2031+415 is a PLC, HAWC J2030+409 is a LP, and 3HWC J2020+403 is a PL. A brief comparison to the previously published work is discussed in Section 4.4.4.

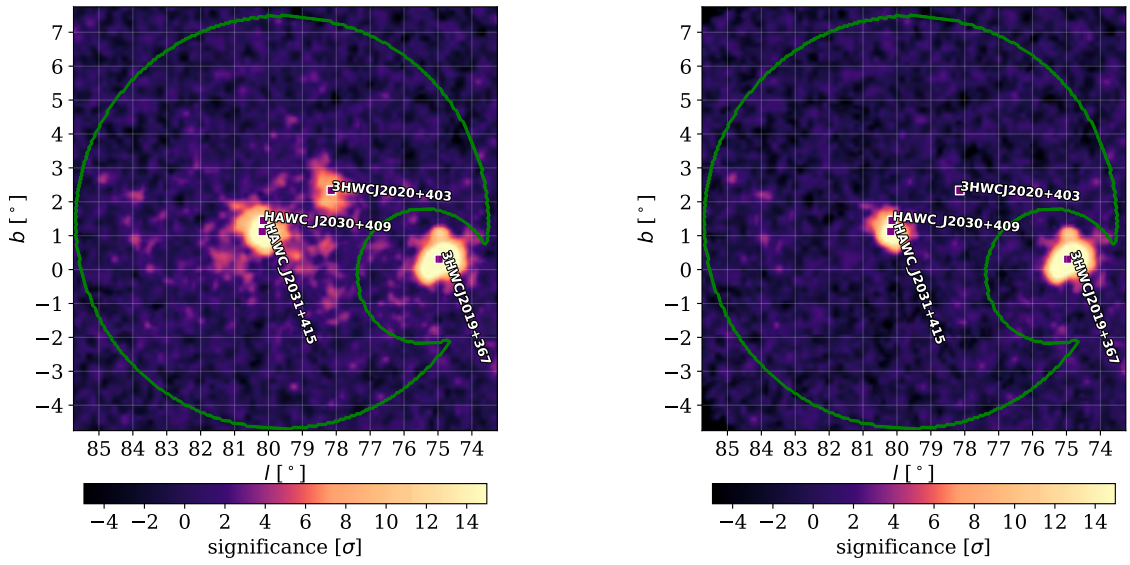

Figure 4.1 LEFT: the significance map of the ROI (green contour) with the source associations found in Section 4.4. A mask is placed on 3HWC J2019+367 to avoid contamination from its emission for this analysis. RIGHT: HAWC J2031+415's emission is shown after contributions from HAWC J2030+409 and 3HWC J2020+403 were subtracted from the data map.

The systematic uncertainties given in Table 4.4.3 are found by performing a series of fits with

Source Name	Spectral Parameters	Morphology
HAWC J2031+415	$\phi_{4.9\,\mathrm{TeV}} = 1.29^{+0.14+0.15}_{-0.12-0.25} \times 10^{-13}$ $\alpha = 1.94^{+0.10+0.10}_{-0.10-0.19}$ $E_c = 32^{+7+5}_{-7-4}$	$\sigma = 0.255^{+0.016+0.015}_{-0.016-0.019}$ RA = $307.92^{+0.02+0.01}_{-0.02-0.01}$ Dec = $41.48^{+0.02+0.01}_{-0.02-0.01}$
HAWC J2030+409	$\phi_{4.2\,\mathrm{TeV}} = 1.1^{+0.12+0.20}_{-0.11-0.09} \times 10^{-12}$ $\alpha = 2.59^{+0.07+0.08}_{-0.07-0.18}$ $\beta = 0.11^{+0.04+0.02}_{-0.04-0.05}$	$\sigma = 2.50^{+0.26+0.25}_{-0.26-0.48}$ RA = $307.54^{+0.22+0.10}_{-0.22-0.24}$ Dec = $41.64^{0.24+0.10}_{-0.24-0.31}$
3HWC J2020+403	$\phi_{1.1\,\mathrm{TeV}} = 4.3^{+0.8+0.6}_{-0.7-0.3} \times 10^{-12}$ $\alpha = 2.91^{+0.07+0.04}_{-0.07-0.09}$	$\sigma = 0.36^{+0.05+0.03}_{-0.05-0.02}$ RA = $305.05^{0.07+0.05}_{-0.07-0.03}$ Dec = $40.52^{+0.05+0.03}_{-0.05-0.03}$

Table 4.1 Fit results from the systematic source search. The first uncertainty listed is statistical and the second is systematic. The units are as follows: ϕ_{E_p} is the flux normalisation with units 1/(TeV cm^2 s); RA, Dec, and σ are given in degrees; and E_c has units of TeV. E_p was found for each source independently.

detector response files that describe different detector configurations. Further details are given in [68] and [96] and are briefly summarized here. These response files are generated assuming different PMT response to showers (efficiency over time, response, etc) and are then compared to HAWC's standard response file. The fitting process is then repeated with these new response files, the difference between the new fit values and those in Table 4.4.3 are found, and then are added in quadrature to produce the total systematic uncertainties.

I determine the energy range of each source using the following procedure as in [97]. Each spectral model is independently multiplied by a step function that models an abrupt cut-off in the spectrum. The only free parameters are the cut-off values; all other parameters are fixed at their best fit values. These values float until the TS of significance of the sources drops by 1σ. This gives the 1σ energy limits for the sources as follows in units of TeV: 0.4-151 for HAWC J2031+415, 0.5-250 for HAWC J2030+409, and 0.26-100 for 3HWC J2020+403.

With the multi-source fitting process complete, I then isolate the emission of HAWC J2031+415 by subtracting out the modelled emission of the URM, HAWC J2030+409, and 3HWC J2020+403 from Figure 4.1 LEFT. The result is shown in Figure 4.1 RIGHT.

Figure 4.2 shows the SED of HAWC J2031+415 compared to selected other observations. These observations represent the most current detections of TeV J2032+4130 from their respective

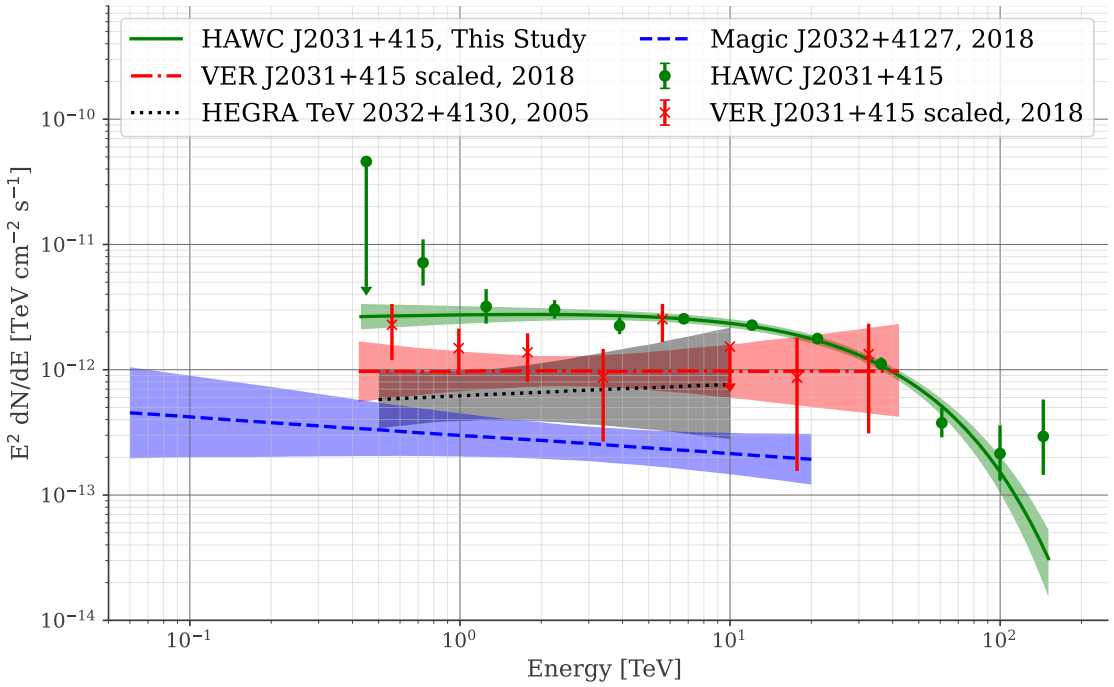

Figure 4.2 SED of HAWC J2031+415. The other observations are from [24, 25, 26, 27] respectively and were selected as the most current independent observations available. Additionally, all uncertainties are statistical only.

observatories. It can be seen that HAWC's observation is in conflict with all other observations. This can be explained by HAWC J2031+415's much larger extension compared to previous studies. For example, in HEGRA's initial discovery, TeV J2032+4130 had an extent of 0.11° compared to this work's 0.26°, thus leading to a much larger flux. Of special note is the scaling done to VERITAS's measurement. I follow the procedure outlined in [88] and is briefly described below.

The spectrum reported by VERITAS was found from a smaller region than the full observed region presented in [27]. VER J2031+415's morphology is described as an asymmetric Gaussian with extents $0.15 \pm .03°$ and $0.07 \pm 0.01°$ for the semi-major and semi-minor axes with a 63° rotation to the northwest. By contrast, the flux calculation uses a circular region with radius 0.23° centered on VER J2031+415. This method is different to what is used in this analysis, where the flux calculation is computed concurrently with the morphological fit. As such, the flux measurements of VERITAS and HAWC may be systematically offset for larger extended sources like HAWC J2031+415. To account for this offset, the flux reported by VERITAS is scaled by assuming a larger

integration region. This gives a scaling factor of 1.49 and is used in Figure 4.2.

4.4.4 Comparison to Previous Work

This work found identical (within uncertainties) morphological and spectral models for HAWC J2031+415 and spectral model for 3HWC J2020+403 as in [46] but there is tension with HAWC J2030+409's spectral model. In [46], the preferred model was a power law with $\phi_{4.2\,\mathrm{TeV}} = 9.3^{0.9+0.93}_{-0.8-1.23} \times 10^{-13}$ TeV/(cm^{2}s) and $\gamma = 2.64^{+0.05+0.09}_{-0.05-0.03}$ while this work found a log parabola model to be strongly preferred with ΔAIC = 90. This is explained by the superior background rejection utilized with the newer data set that reveals a curvature at the highest energies. The spectral and morphological fits for HAWC J2031+415 and 3HWC J2020+403 are comparable to [46]. One additional note is the recently published LHAASO result of the Cygnus region [98] where they find a PL spectral model for HAWC J2030+409. An in-depth comparison is beyond the scope of this work, but the two models are compatible within systematic uncertainties.

For 3HWC J2020+403, a minor difference in morphology compared to [46, 99] was found. In [99], its model was found to be a disk rather than a symmetric 2D Gaussian. For a disk model, the emitted flux is held constant over a fixed radius rather than decreasing radially with a 2D Gaussian. I tested this by creating a disk model with a fixed radius of 0.63° from [99] and the whole model was refitted. The result was a negligible difference in TS (ΔTS < 1) and, while the Gaussian model is used for this analysis, a dedicated work on 3HWC J2020+403 is need to determine its true morphology.

4.5 Energy-Dependent Morphology Study

4.5.1 Methodology

In order to study any possible energy-dependent morphology of HAWC J2031+415, I utilized the method described in [88, 100]. This method uses the longitudinal profiles of discrete energy bands over the source to count the number of excess events observed. Six energy bands are selected in TeV units: 0.3-1.0, 1.0-3.2, 3.2-10, 10-32, 32-100, and 100-316. Each energy band consists of NN 2 energy bins, similar to [100] and whose specific energy definitions are defined in [96].

The longitudinal profile region is defined as a rectangle of dimensions 6° long by 1° centered

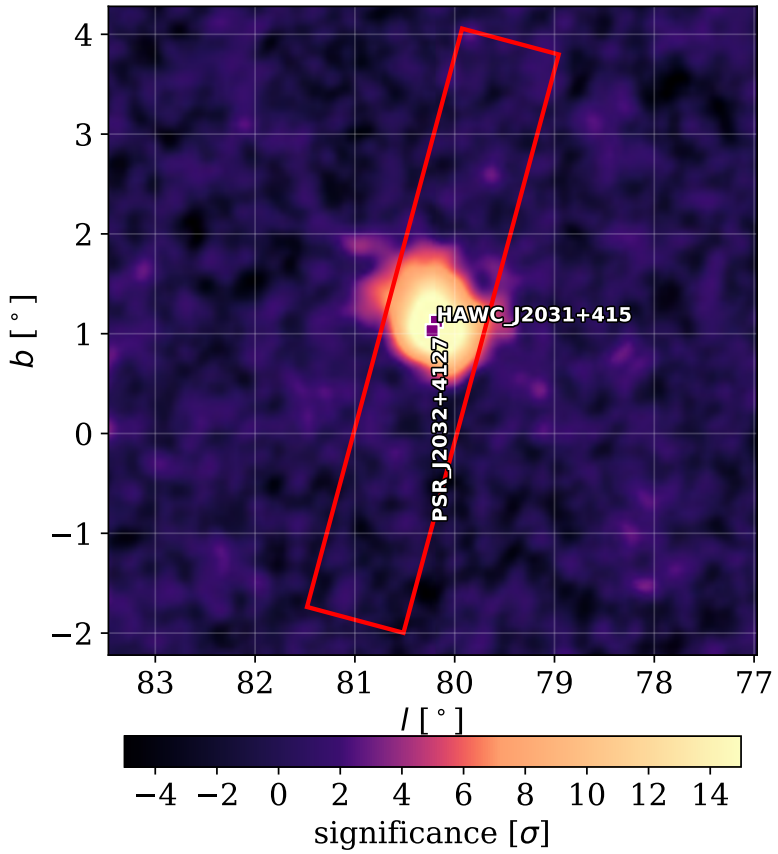

Figure 4.3 The significance map used for the energy morphology study. The rectangle highlights the longitudinal profile region used for the energy-dependent morphology study.

at the pulsar's location with HAWC J2030+409, 3HWC J2020+403, and the URM subtracted out. Furthermore, the rectangle is rotated by 15° to lie on the line connecting HAWC J2031+415's centroid and PSR J2032+4127 in Galactic coordinates. This is to determine whether the observed emission trends toward the pulsar's location with changing energy. This can be seen in Figure 4.3.

To determine the true size of the extended emission in each energy band, the following procedure is used. First, the rectangular regions is divided into 50 bins, each with a width of 0.12°. Then the excess counts of each bin are summed and plotted. This is shown by the data points in Figure 4.4. To measure the intrinsic width of the extension, a 1D Gaussian is fit to each band, as indicated

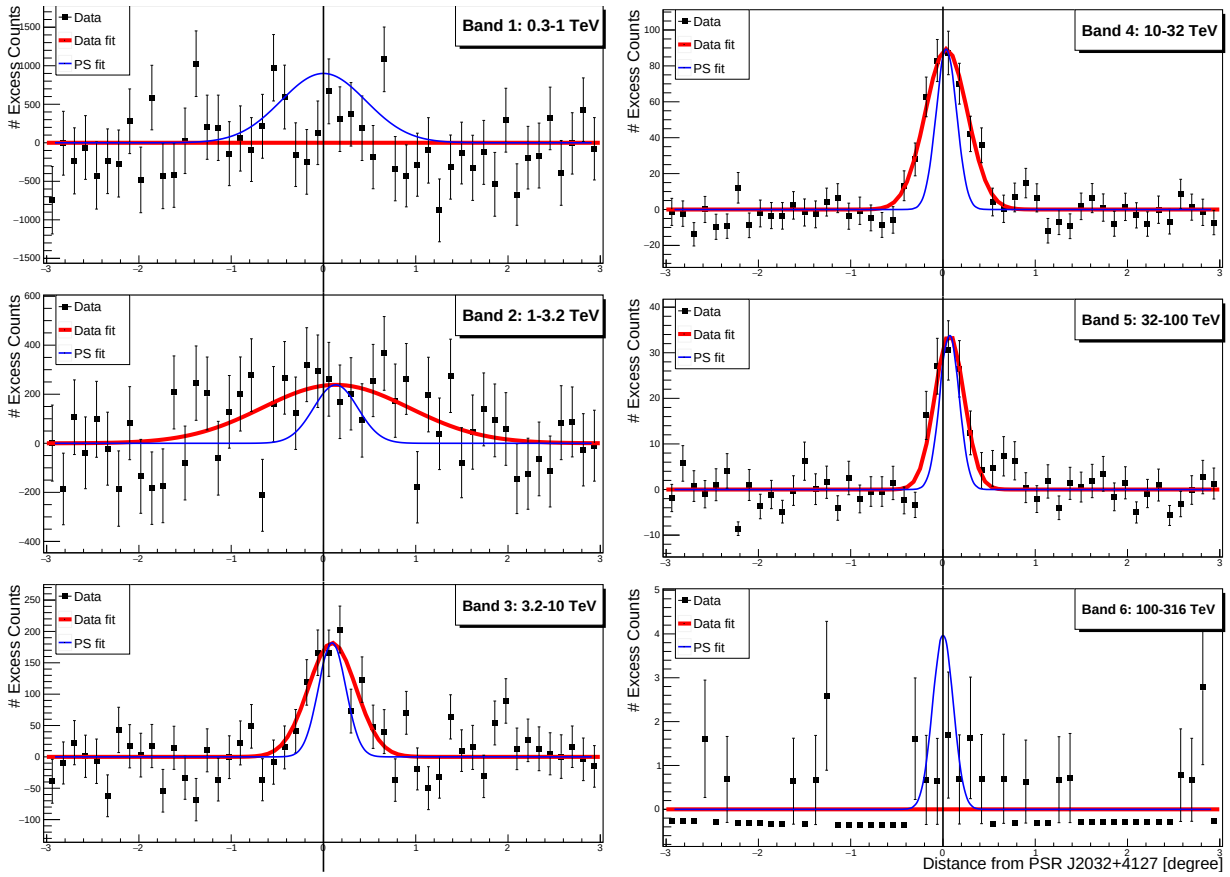

Figure 4.4 The longitudinal profiles for the excess count maps for HAWC J2031+415. The red fitted lines correspond to the fitted Gaussians of each band while the blue dashed lines are the simulated point source Gaussians discussed in Section 4.5. The location of PSR J2032+4127 is indicated by the vertical line. The distance between HAWC J2031+415's best fit centroid location and the pulsar's location is 0.13° or about 3 pc.

by the red lines in Figure 4.4. However, as discussed in [76] and [100], there is a smearing effect caused by point sources not appearing point-like with this method.

To rectify this, a point source with HAWC J2031+415's index of 1.94 is simulated at PSR J2032+4127's location. This simulated source is then handled in the same method described above. The 1D Gaussians found with the simulated source are the "smearing" effect and are shown by the dashed blue line in Figure 4.4. This effect can now be subtracted out in quadrature with the observed data fits.

If a band is fit and a source cannot be identified, then it is considered a null result and a flat line is plotted. This indicates that either I cannot resolve the source from the background or the source

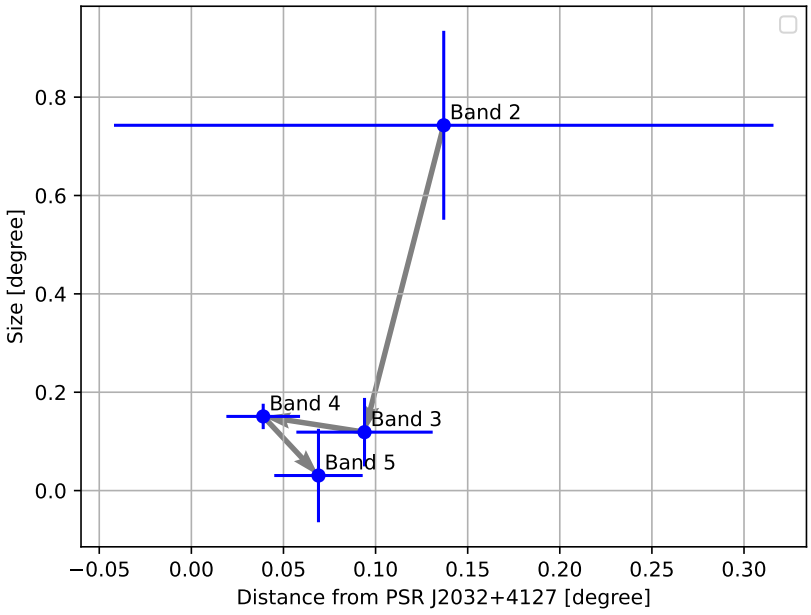

Figure 4.5 The results from the energy morphology study as described in Section 4.5. HAWC J2031+415's true size is presented on the y-axis and the distance to the pulsar is on the x-axis.

is not visible to HAWC in that energy regime. Additionally, if the excess counts are negative, this means that, for the selected bin, there are more background events compared to data events.

4.5.2 Results

From Figure 4.4, there are no significant detections in bands 1 and 6; this is most likely caused by the spectrum of HAWC J2031+415. This source is not significantly detected in GeV or high TeV energies, which is what these two bands primarily comprise of. While there are fits for all other bands, band 2 requires more investigation. Its fit is diffuse and was checked against the diffuse background emission model to ensure the observed emission was from HAWC J2031+415 and not a large background fluctuation. The emission is observed at a 5σ level and is confirmed as a positive detection of HAWC J2031+415. Bands 3, 4, and 5 all show significant detections.

Figure 4.5 shows the size of TeV emission with increasing energy and its location with respect to PSR J2032+4127. The size is the true width of emission after subtracting both the point source smearing and any systematic offsets between data and simulation. This true width is given by

$$\sigma_{\text{true}} = \sqrt{\sigma_{\text{fit}}^2 - (\sigma_{\text{sim}} + \sigma_{\text{offset}})^2} \tag{4.7}$$

While some faint energy-dependent morphology is present, particularly in the morphology shift from bands 2 to 3, there is no discernible trend at higher energies. Likewise, while there is a faint trend towards the pulsar's location at lower energies, there is nothing conclusive at higher energies. Additionally, this study was done on the best-fit location of HAWC J2031+415 to observe any potential shifting. The best fit locations for each band trend towards the location in Table 4.4.3 and a slight though inconclusive change in morphology was detected.

4.6 Periastron of PSR J2032+4127

On November 13, 2017, PSR J2032+4127 completed its periastron of MT91 213. For approximately 170 days before periastron, gamma-ray flaring was observed by both VERITAS and MAGIC [26] from the binary system. HAWC has been proved to observe flares on a much shorter timescale [28] and so I investigate this flaring. To do this, a 172-day section of HAWC data from May 26 to November 14 in 2017 along with a smaller ROI of 3 degrees were selected. The significance map is shown in Figure 4.6.

1 hot spot appears in the data with a significance of 6.10. This is compared to a separate map made before any flaring was detected. This map has 170 days and runs from September 30, 2016 to March 19, 2017. A similar hot spot of 6.37σ is found. Next, the source search method introduced in Section 4.4 is used on both these maps to determine both the number of sources and their fluxes to see if any flaring can be resolved. The results for both maps produce a single point source located near HAWC J2031+415. While the extension test for both sources was rejected with a $\Delta\text{TS} \approx 15$, the extension for both sources was 0.35°, significantly larger than HAWC J2031+415's extension. This may be due to contamination from HAWC J2030+409.

It is clear that, while HAWC does resolve a source in the 170-day maps, I lack sufficient data to properly resolve any flaring from the periastron and instead only detect diffuse emission from the PWN. Potentially, future background rejection algorithms may yield more concrete results but currently HAWC does not detect the periastron of PRS J2032+4127 and MT91 213.

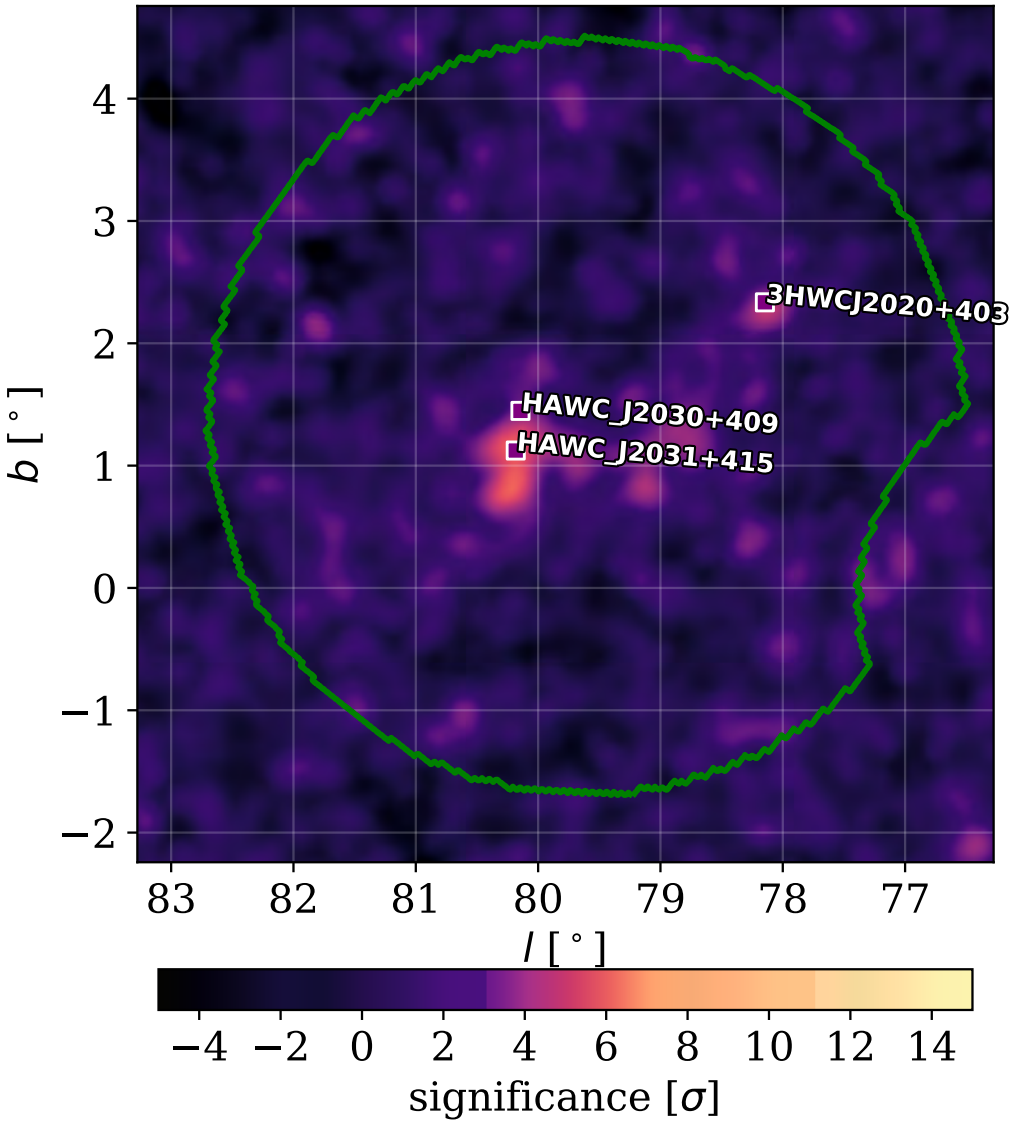

Figure 4.6 172 days of HAWC data from May 26 to November 14 in 2017. The maximum significance is located near HAWC J2031+415 and is 6.10 for this interval.

4.7 Multi-wavelength Fitting

4.7.1 NAIMA Framework

The Naima software is a non-thermal modeling framework that utilizes Markov chain Monte Carlo calculations [101]. It has both leptonic and hadronic models that take flux points like the ones shown in Figure 4.2 as inputs and fits different emission models to them. At the gamma-ray regime, the leptonic model considers inverse Compton (IC) scattering of relativistic electrons off low energy photons from the cosmic microwave background (CMB) radiation and the far and near infrared (FIR, NIR), while the X-ray regime considers synchrotron (SYN) emission released by high energy electrons moving through magnetic fields. The hadronic model considers π^0 decay (PD) from proton-proton collisions.

4.7.2 Methodology

I break the multi-wavelength data introduced in Section 4.1 into two groups: TeV data fitted to the IC and PD models, and the lower energy range data that is fit to the synchrotron model. The former is the HAWC flux points presented in Figure 4.2 and serve as the high energy constraint for the models while the latter requires a more in-depth discussion.

The two x-ray observations from Suzaku [79] and XMM-Newton [80] both detected diffuse x-ray emission inside the CoG of HEGRA's and, by extension, HAWC's source extent. As discussed in [79], they were able to resolve 2 small ($\simeq$ 0.01 deg) distinct sub-substructures, one of which they associate as the x-ray PWN produced by PSR J2032+4127 and produced a spectrum for it. In contrast, [80] only detects diffuse emission and not the small structures found by Suzaku. Currently, it is not clear which data corresponds to the x-ray PWN but, for this analysis, I consider the Suzaku structure 1 detection to be the probably x-ray pwn and the XMM-Newton diffuse detection to be the upper limit for the x-ray emission.

Similar to the XMM-Newton detection, the radio observation from the VLA [81] found faint diffuse radio emission in HEGRA's CoG but did not resolve any sub-structures associated with the PWN. While some small emission co-incident with Suzaku's structure 1 was observed in Figure 1 of [81], no detailed analysis was presented on that emission. Therefore, as with XMM-Newton, the

diffuse emission found by the VLA is considered as the upper limit of the radio emission produced by the PWN.

To model this data, 4 separate scenarios are considered: a leptonic model with IC handling the HAWC data and SYN modeling the lower energy data, a TeV only fit using IC and assuming different magnetic fields for the SYN model, a hadronic model only using PD to fit the HAWC data, and then another PD model with a fixed index at 2. This is to test whether there is enough TeV data to adequately constrain the hadronic index to the expected value of 2, assuming the primary acceleration mechanism comes from diffusive shock acceleration. The seed photon fields considered for the leptonic model are the CMB and values for the NIR and FIR approximated from [102]. The column density to find the proton density is $7.7 \times 10^{21} \mathrm{cm}^{-2}$ from [103] which, considering HAWC J2031+415's extent of ~ 5.9 parsecs, gives a density of 417 protons/cm^3.

For both models, cut-off power-law spectra are assumed for proton and electron populations. This comes from observed flux points for HAWC J2031+415 and its best-fit spectral model. The parameters are the same as in Equation 4.4 with E_p being set to 20 TeV for both models. Additionally, the energies required for the two models are given as W_e for leptonic and both W_e and W_p for lepto-hadronic. The results from these models are shown in Figure 4.7 and Table 4.7.2.

Model	log(ϕ)(1/TeV)	E_p (TeV)	α	E_c(TeV)	B(μG)	W$_{e,p}$ > 1TeV (erg)	AIC
IC + SYN	42.8 ± 0.1	20	2.1 ± 0.2	59 +13 -22	1.48 ± 0.24	$(1.8 \pm 0.5) \times 10^{46}$	19.5
IC only	42.7 ± 0.1	20	2.1 ± 0.3	53 +12 -19	—	$(1.5 \pm 0.6) \times 10^{46}$	17.5
PD free index	43.1 ± 0.06	20	1.6 ± 0.3	139 +36 -52	—	$(3.1 \pm 0.5) \times 10^{46}$	17.9
PD fixed index	43.2 ± 0.03	20	2 (fixed)	265 +28 -40	—	$(4.6 \pm 0.3) \times 10^{46}$	19.2

Table 4.2 The fit values that correspond to the fitting process discussed in Section 4.7. As in Section 4.4, the parameters are the flux normalization, pivot energy, index, and the cut-off. The magnetic field for the synchrotron is then presented along with the separate electron and proton population energies. Finally, the AIC for all fits is presented. All uncertainties presented are statistical only.

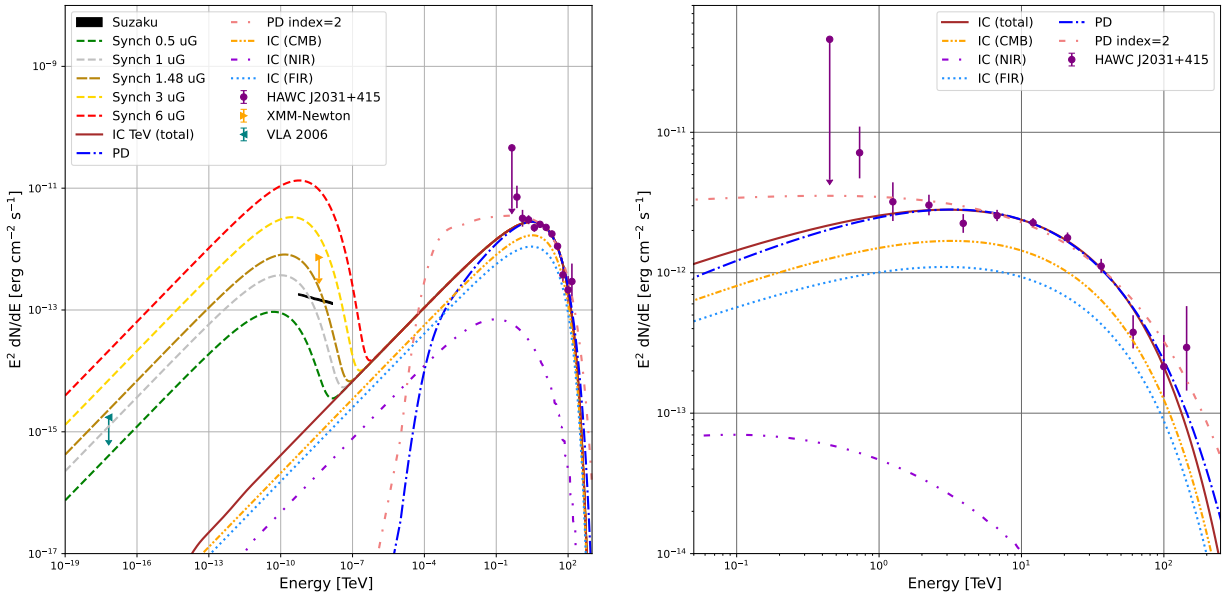

Figure 4.7 LEFT: the full multi-wavelength fit results. The different SYN fit bands indicate the expected x-ray and radio fluxes for different magnetic fields from the IC fit using HAWC data. The 1.48μG fit indicated by the thicker dashed line is the best fit magnetic field using the Suzaku spectrum. Additionally, the IC HAWC data fit is broken into the CMB, FIR, and NIR components. RIGHT: the zoomed in TeV fit results. All 3 TeV models fit the data well, though it is clear that lower energy data is needed to properly constrain the PD model index.

4.7.3 Leptonic Results

From Figure 4.7, it is immediately clear that, while the TeV fits for the IC and the two PD models fit the HAWC data well, the magnetic field resulting from the combined IC and SYN field is quite low at 1.48μG. Furthermore, applying different magnetic field values to the TeV only IC fit (the 0.5, 1, 3, and 6μG lines) corroborates this low magnetic field. This leads to several different possible conclusions that are discussed below.

The first potential conclusion is that the x-ray source observed by Suzaku is the x-ray counterpart to HAWC J2031+415. This would make the x-ray PWN approximately 13 times smaller than the TeV PWN. While significantly smaller than the TeV region, this may be expected for older systems. As discussed in [104] and [105], x-ray bright PWN require powerful young pulsars and, as this is an old system, it could explain the low energy and small size observed in x-rays. Such a low magnetic field may indicate that the synchrotron nebula is nearing the end of its life. In this case, more detailed diffuse radio measurements are needed to properly constrain the magnetic field. A

more detailed discussion is presented in Section 5 of [105].

An alternative possibility is that the x-ray PWN is much more diffuse than Suzaku's approximately 0.02° diffuse measurement. If I consider XMM-Newton's larger source and flux, then the fitted magnetic field increases to $\simeq 3\mu$G which, while still low, does approximately equal the galactic average. While more diffuse x-ray emission observations, it could indicate that the x-ray PWN is significantly larger than Suzaku's detection and, by extension, give a larger average magnetic field.

One final consideration is the binary nature of this system. As discussed in [105], stellar winds from the companion star collide with the pulsar wind and creates highly variable emission. It has been shown [106] that x-ray emission may be produced by IC off these dense photon fields and may significantly contribute to the observed x-ray emission. A detailed x-ray analysis would be needed to determine if this is occurring in the PSR J2032+4127/MT91 213 system but it could explain the small emission region and lower x-ray flux.

Another consideration is the energy budget E_T of PSR J2032+4127 and the relationship it has with the observed electron population. An approximation of E_T can be found by multiplying $\dot{E} = 1.5 \times 10^{35}$ erg/s by the pulsar's characteristic age of $\sim$ 200kyr to give an energy budget of $\sim 9 \times 10^{47}$ erg (see. [107]). While the E_T found is most probably lower than the actual energy budget, a comparison can still be drawn between W_e and the observed E_T and reveals that W_e is $\sim 2\%$ that of the energy budget for PSR J2032+4127. This value is reasonable [41] and indicates that PSR J2032+4127 is capable of producing the observed electron population.

4.7.4 Hadronic Scenario

On the hadronic models, both fit the HAWC data well visually and statistically. There are two key factors to note: the value of the free index fit compared to the fixed index fit and the found E_c for both models. Both are discussed below.

The rather low value of the fitted index at 1.6 ± 0.3 compared to the expected value of 2 indicates that HAWC data alone may not be sufficient to properly constrain the hadronic index. It is clear from the fixed index fit that additional data, particularly in the GeV range, would greatly aid in

constraining the index. There are GeV flux points from *Fermi-LAT* presented in [27] that are matched to a 0.14° extended source but they were obtained using a simplified model for the Cygnus Cocoon and may contain contamination from the larger source. As such, they are excluded from this analysis.

Considering the E_c for both fits along with the observed maximum energy of 151 TeV, this would make HAWC J2031+415 a potential PeVatron with both models having $E_c > 100$ TeV. Indeed, recent observations from the LHAASO observatory have detected > 1 PeV events from this region [35]. However, it is extremely doubtful that the PWN is capable of producing such high energy events. A follow-up paper analyzing the PeV LHAASO sources [108] found that PSR J2032+4127's low spin-down luminosity is incapable of producing such high energy events. The binary nature of the PWN may help contribution to the hadronic model, but that is beyond the scope of this paper. It is clear that, though the HAWC data can be fitted well with a simple hadronic model, it is impossible that such a model is sufficient to power such high energy events.

4.7.5 Scenario Comparison

While both the leptonic and hadronic models fit HAWC's data well, there are concerns with both models. For the leptonic models, while the TeV data is fit well with the IC and IC+SYN models, the lack of significant x-ray and radio data makes it difficult to properly constrain the ambient magnetic field. The fitted value of 1.48μG is significantly lower than the lower end of expected magnetic fields for PWN [104] but this also could be a symptom of the x-ray PWN fading away [105]. More radio data would help constrain the magnetic field of the PWN.

For the hadronic model, while it fits the HAWC data well, it is also physically impossible for the PWN to produce PeV protons given its low spin-down luminosity [108]. Therefore, while it is included in this analysis for completeness, it is clearly not the preferred model for HAWC J2031+415's TeV emission.

4.8 Conclusions

The morphological studies I presented here reveal HAWC J2031+415 to be an extended emission region modelled as a symmetric Gaussian. As predicted by [83] and [27], it has a spectral shape of

a power law with exponential cut-off energy E_c = 32 TeV and is seen to 151 TeV. Given its close proximity to TeV J2032+4130, HAWC J2031+415 is most probably the high energy extension of this unidentified source. An energy-dependent morphology study found that, while there might be a downward trend in size, no significant conclusive was found. Additionally, a periastron study was done to observe whether HAWC saw any flaring, but none was detected.

I then performed a multi-wavelength analysis considering radio and x-ray data using the NAIMA framework. I considered 4 models, two leptonic (one with lower energy data, one with only TeV data) and 2 hadronic (one with a fixed index and one with a free index). All 4 models fit the available data well, but the hadronic models are physically incompatible with the PWN scenario. The leptonic models found a very low magnetic field of 1.48μG. This could indicate that the x-ray counterpart of the PWN may be fading, but more x-ray and radio data is needed to confirm this hypothesis. In conclusion, I present the highest energy observations of the PWN candidate HAWC J2031+415 and, by extension, TeV J2032+4130.

CHAPTER 5

TRANSIENT γ-RAY SEARCHES WITH ICECUBE NEUTRINO ALERTS

This chapter details my main PhD project: a joint HAWC-IceCube search for coincident γ-ray and neutrino emission. Section 5.1 introduces the data sets that I consider from both observatories, along with selection criteria that I used to modify the data. Section 5.2 introduces the fitting algorithm I apply to the HAWC data to search for transient events; this being the Bayesian Block Algorithm (BBA). Sections 5.3 and 5.4 discuss the results of the BBA on HAWC data and how sensitive HAWC might be to transient events. Section 5.5 discusses the potential multi-messenger applications for this work by considering two Active Galactic Nuclei (AGN) that HAWC sees. Lastly, Section 5.6 considers systematic uncertainties from the BBA and discusses potential improvements to future studies.

5.1 Data Considered

The IceCube neutrino data considered are astrophysical neutrino alerts from [71] from 2011 to 2022 and from [70] onwards. This analysis considers alerts up to July 8, 2025, with 368 alerts in HAWC's FOV. The alerts contain the alert name, time, location (including 95% containment), energy, and probability of being astrophysical. They are split into two categories: bronze if the probability is $\geq$ 30% and gold if the probability is $\geq$ 50%. These are indicated by either B or G, respectively.

For HAWC data, two data sets are considered: the full time-integrated Pass 5 data set containing 2565 full days of reconstructed data, discussed in [1], and individual daily maps. These daily maps utilize the same reconstruction method as the Pass 5 data set and run from March 20, 2015, to January 15, 2024. Accounting for detector downtime (complete power down), there are 3175 days of potential data, though in practice that number is lower depending on what IceCube alert is considered. This is due to small-scale maintenance that may affect the detector for part of the day. It should be noted that the Pass 5 data set takes this into account and excludes those days, while the daily maps do not. This explains the difference in days in the respective data sets. The implications of this are discussed in Section 5.6. The full data set is used to search for potential neutrino

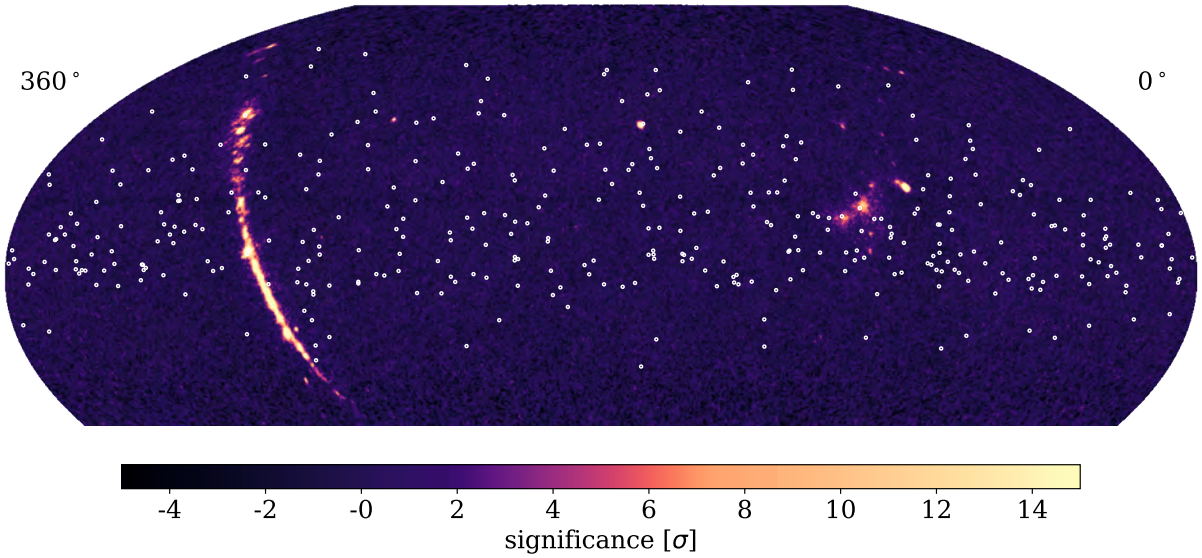

Figure 5.1 HAWC's full sky data map with the 368 visible IceCube alerts. The white dots show the alerts that lie within the declination interval −20° to 60°. The plot is drawn in celestial coordinates with an index of 2.7 and a pivot energy (E_p) of 7 TeV.

coincidences with HAWC steady-state sources while the daily maps can search for transient events like potential flares. All the IceCube alerts visible to HAWC are shown in Figure 5.1.

5.2 Methodology

5.2.1 Data Collection

Searching the HAWC data using the IceCube alerts is done as follows: first, the location and uncertainty region from each alert is extracted and used to construct a Region of Interest (ROI). Then, the time integrated HAWC data set is scanned to find the brightest significance spot inside the ROI to determine if a steady γ-ray source is present. These results are used to draw potential correlations between the alerts and known HAWC sources, and are discussed further in Section 5.5.

Next, both the 2.0 and 3.0 daily map sets are considered. For each day, the the pixels in the data maps within the ROI are scanned to search for the pixel with the maximum flux value, corresponding uncertainty and its location (RA, DEC). If a flare (event above background) is present and visible to

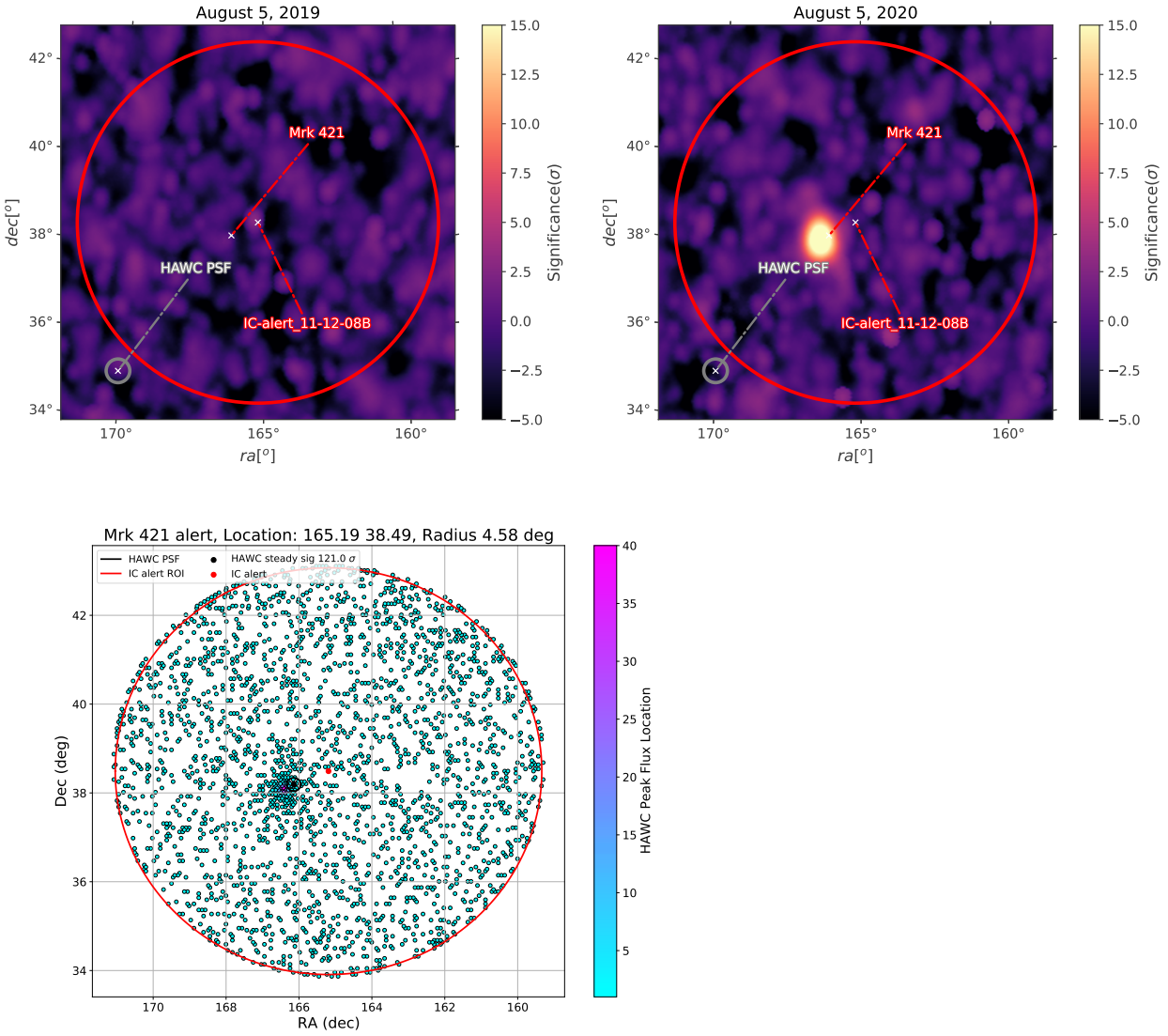

Figure 5.2 The top two plots indicate two daily maps of Mrk 421, one where it is quiet and when it is flaring. The red circle indicates the IceCube alert's ROI. HAWC's point spread function (PSF) is indicated in black. The bottom plot shows the frequency that a given RA, Dec has the max flux in the daily map. For a recurrent flaring source like Mrk 421. HAWC sees Mrk 421 to 121σ with Pass 5 data.

HAWC, its flux would appear above the background for the given day. The fluxes and uncertainties are extracted, and then a light curve is made for the ROI. The TeV-bright AGN Mrk 421 is located 0.8 degrees from the neutrino alert IC111208B and used as an example for this procedure. Example significance maps and resulting maximum flux distribution are shown in Figure 5.2.

5.2.2 The Bayesian Block Algorithm

Once the light curves have been made, they can then be fit with a Bayesian Block Algorithm (BBA). While described thoroughly in [109], a summary is given here. The goal of the algorithm is to determine statically significance signal from background/steady-state emission. It accomplishes this by fitting the data with a series of piece-wise functions (other models are available, if needed), where the discontinuities occur at "change points". At minimum, any fit will have two changes points, the first and last points of the data set, and at maximum, include every point, depending on the sensitivity (see Section 5.2.2.3 for more details).

Once fit, the quality of the fit needs to be judged. To do this, both a global and local "fitness function" are defined. These are given by

$$F[P(T)] = \sum_{k=1}^{N_{blocks}} f(B_k) \tag{5.1}$$

where $F[P(T)]$ is the total fitness function of partition P with interval T (ie, a configuration of 6 blocks across the whole data set) and $f(B_k)$ is the fitness function for each block. The individual fitness functions are model independent and can be selected based off the data. Each block is defined by two change points across the time interval. These points are where the data undergoes an abrupt, statistically significant, change. The statistical significance of this change is dependent on how sensitive the calibration of the algorithm is (see Section 5.2.2.3). These points are defined at the start of the change and the point prior to the next change. If there are no changes detected, then the start and end points are the beginning and end of the data set.

5.2.2.1 The Algorithm

Traditionally, testing every combination of blocks in a data set sized N is computationally expensive and time-consuming as there are 2^N possible configurations. However, the BBA is exceptionally efficient and can arrive at a solution in a time of order N^2. This is because it follows a similar principle to mathematical induction, and is outlined below.

Consider $P^{opt}(R)$ be the optimal partition for some R data points. For R=1, there is only 1 option, for R=2, there are 2 options, and so on. Each optimal partition is stored, along with the

location of the last change point. Assume that $P^{opt}(R)$ has been found for some R. The next step is to find $P^{opt}(R+1)$. For some r, consider set of all possible partitions whose last block starts at point r and, by definition, ends at R+1. This block's fitness is defined by $F(r)$, and the optimal partition of the rest of the configuration is defined as $P^{opt}(r-1)$.

Because the fitness functions are additive, the final fitness function of this new configuration is given by

$$A(r) = F(r) + \begin{cases} 0 & r = 1 \\ P^{opt}(r-1) & r = 2, 3, \ldots, R+1 \end{cases} \tag{5.2}$$

where A(1) is the special case where only 1 block is made. As the optimum configurations for all the previous steps have been found already, the new best-fit fitness function can be found by simply maximizing F(r), before this to is saved in $P^{opt}(r)$. This repeats for R=N, and then the change points are extracted to indicate where the blocks are formed.

5.2.2.2 The Fitness Function

The only thing left before running the BBA is defining what the fitness function is. For this analysis, I use the "Point Measurements" fitting type [109]. This means the data considered has a time stamp, a value, an uncertainty, and that there are no duplicate time entries. Additionally, the data points and uncertainties are assumed to be independent (ie, no correlation) and that the observation error (how far the data points are from a model) are statistically independent. This allows the observed data to be written as

$$x_n = x(t_n) = s(t_n) + z_n,\ n = 1, 2, ..., N \tag{5.3}$$

where the observed data is a function of a signal s and some observational error z. The error z is generally only known through its statistical distribution. If the case where z follows a normal distribution with mean 0 and variance σ is considered, then the probability distribution follows

$$P(z_n)dz_n = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(\frac{-z_n^2}{2\sigma_n^2}\right) dz_n \tag{5.4}$$

If a steady-state model is assumed so $s = \lambda$, the likelihood of a given measurement becomes

$$L_n = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(\frac{-(x_n - \lambda)^2}{\sigma_n^2}\right) \tag{5.5}$$

If the block k is considered, then the likelihood of k is the product of the likelihoods of the data inside. To determine the maximum of the likelihood, the parameter function

$$\frac{-1}{2} \sum_n \frac{(x_n - \lambda)^2}{\sigma_n^2} \tag{5.6}$$

needs to be maximized. Taking the derivative with respect to λ and solving, the maximum is at

$$\lambda_{\max} = \sum_n \frac{x_n}{\sigma_n^2} / \sum_{n'} \frac{1}{\sigma_{n'}} \tag{5.7}$$

If the following quantities are defined:

$$a_k = \sum_n \frac{1}{2\sigma_n^2}, \; b_k = -\sum_n \frac{x_n}{\sigma_n^2} \tag{5.8}$$

and they are plugged back into the log likelihood function, the fitness function becomes

$$\log L_{\max}^k = b_k^2 / 4a_k \tag{5.9}$$

The final fitness function for a given configuration only depends on the data and its uncertainty. The only role the time component plays is to determine which block the data gets placed in.

5.2.2.3 Calibrating the BBA

To prevent overfitting, where each block is fit to each data point, a geometric prior is introduced that assigns smaller probabilities to models containing large number of blocks. This is shown with

$$P(N_{\text{blocks}}) = \frac{1 - \gamma}{1 - \gamma^{N+1}} \gamma^{N_{\text{blocks}}} \tag{5.10}$$

where γ is a tunable parameter. The log of this quantity, $-\log(\gamma)$, is called the "ncp_prior" and determines the sensitivity of the algorithm. The smaller the value, the more sensitive the algorithm becomes, and will allow or prefer more blocks to be fit. The opposite is true for large ncp_prior

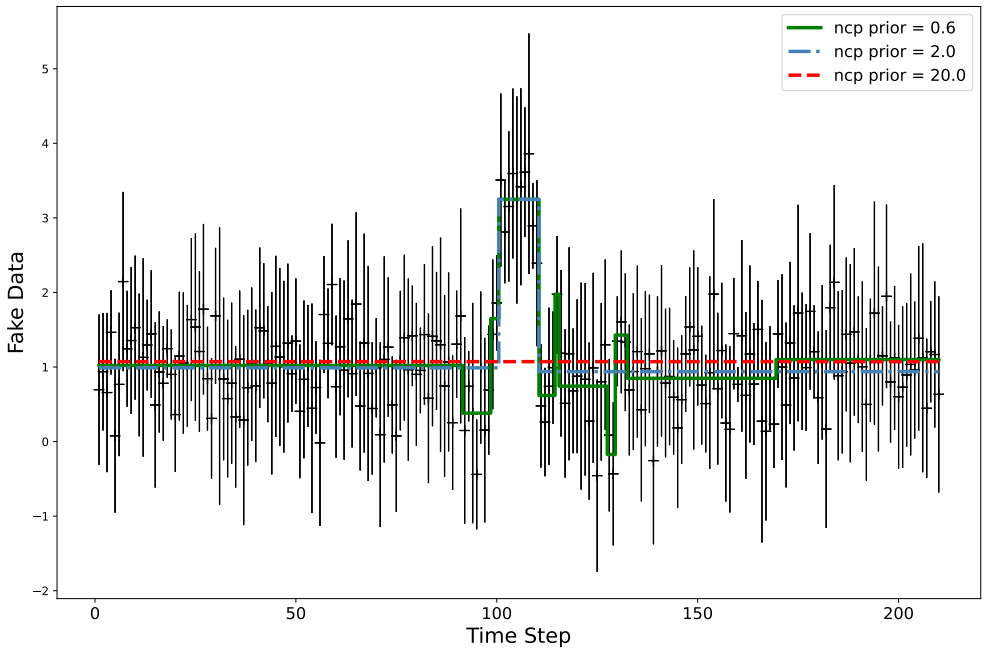

Figure 5.3 An example of 3 fits with the BBA using different ncp_prior values. As ncp_prior increases, the BBA fits fewer fluctuations before missing the simulated flare.

values. An example of this is shown in Figure 5.3. As clearly shown, selecting the wrong value for the ncp_prior can result in missed detections or large number of false positives. This penalty factor is added to the fitness function of each block (shown below) and requires calibration.

$$\log L^k_{\max} = b_k^2/4a_k - \log \gamma \tag{5.11}$$

To calibrate the BBA, I looked at both declination and index dependence. The former requires the definition of declination bands and the latter requires repeating the former's calibration. I selected 10 degree wide bands ranging from -20 to 60 degree declination, with each band centered on the halfway point. So, I have 8 bands centered on [-15, -5, 5, 15, 25, 35, 45, 55]. While attempts to calibrate each band's ncp_prior were successful, implementations of these declination-specific values resulted in an unacceptably high FPR of $> 40\%$. Therefore, I decided to focus on only the core declination bands where HAWC is the most sensitive and calibrate off of those. These are [5, 15, 25, 35]. I then find 1 value for each index assumption that fulfills the 5% false positive rate.

In order to calibrate the algorithm and find my desired ncp_prior values, I need to first generate

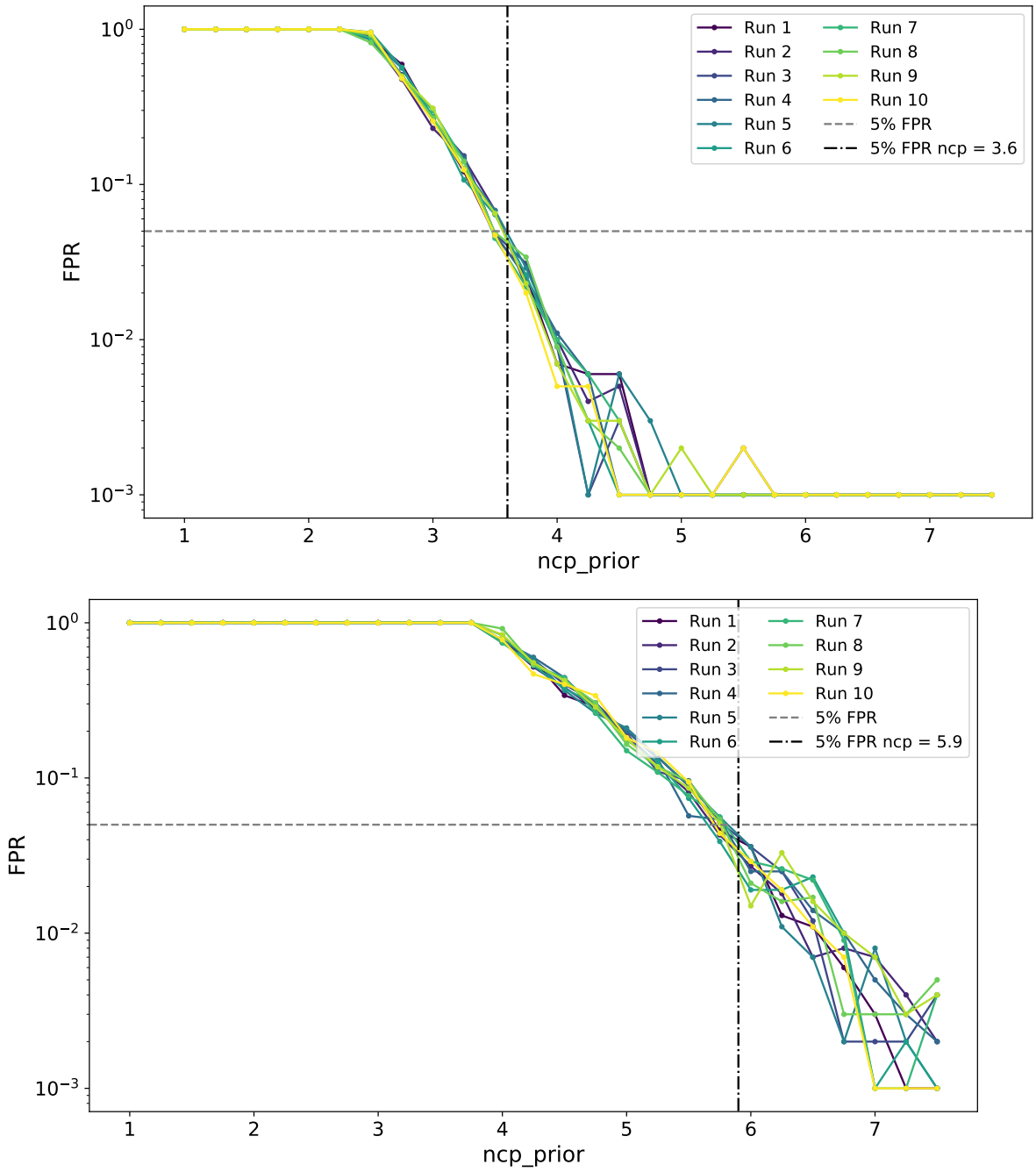

Figure 5.4 Comparing the two index ncp_prior values for 10 runs. As above in the above figure, the dashed line indicates the 5% FPR. The dashdot line indicates what ncp_prior achieves that.

fake light curves. I sampled real data from the respective declination bands outlined above. I first defined 36 10 degree wide circular ROIs for each declination band, centered on RA = [5, 15, ..., 355]. Then I performed the max flux search for each ROI, yielding 36x3000 csv files containing

the relevant max flux values and errors. For declination bands 25 and 35, I removed the csv files including the Crab and Mrk 421 to prevent known source contamination. Then, I concatenated the remaining files to create massive array containing the max flux, error, and locations.

With the data set created, I could now randomly generate virtually infinite numbers of data-like light curves. While there is some concern of including potential flares, The BBA in this analysis generally requires multiple days of flare-like behavior (see Section 5.4), and the probability of that occurring in randomly sampled data is virtually 0. This data can then be fit with the BBA with a ncp_prior value (say 2.0) before being regenerated and fit again.

The range for ncp_prior values tested ranged from 1.0 to 7.5. Each ncp_prior was tested against 1000 simulated light curves. The FPR for each run is the total number of change points (boundary of blocks) minus 2 for the given 1 block fit (null result). So a fit with 3 blocks would have 4 - 2 change points, or a 200% FPR. The lowest FPR a run can have while detecting a random fluctuation is 100%. All 1000 of the runs' FPR is then summed and averaged to find the average FPR. The goal is to find the ncp_prior that achieves a 5% FPR. If the total FPR is over 1 (100%), then it gets set to 1. The results from the 10 runs for both maps are given in Figure 5.4. The final ncp_prior values are 3.6 and 5.9 for the 2.0 and 3.0 maps.

5.3 Results

The results for the BBA fits to the HAWC data are presented in Tables 5.1 and 5.2 for the 2.0 and 3.0 maps, respectively. The alerts listed in the two tables correspond to "coincident detections", where the BBA finds a non-null ($n_{cp} > 0$) result for a given alert's light curve. There are 16 (4.3%) and 17 (4.6%) coincident detections for the two maps. This is in-line expected background detection rate of 5%. Of the detected alerts, 3 contained HAWC steady-state $\sigma > 5$ hotspots in the full 2565-day map. This is in contrast to the results found in [71], where 8 3HWC sources are found. This is because they scanned all IceCube alerts for known sources, while here I only consider the alerts that triggered the BBA. 2 of the 3 detections correspond to the AGNs Mrk 421 and 501 (11-12-08B and 24-03-07G) and are discussed further in Section 5.5. The third alert, 16-03-07B captures part of the Geminga TeV Halo [85] at the edge of its containment radius. All

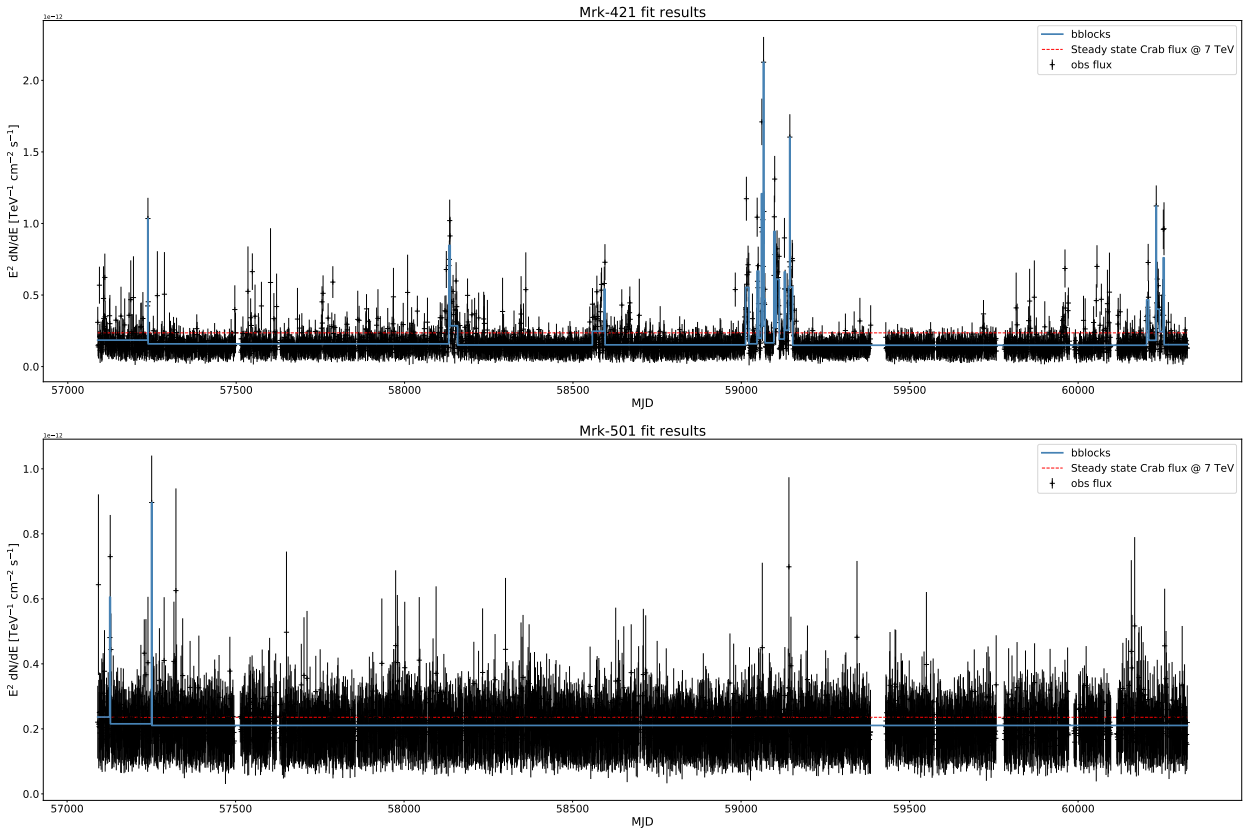

Figure 5.5 The light curves and BBA fits to the HAWC daily light curves for the neutrino alerts containing Mrks 421 (upper plot) and 501 (lower plot). The alerts that captured the two Markarians are 11-12-08B and 24-03-07G, respectively. Mrk 421's alert occurred before HAWC started operations, and Mrk 501's alert occurred just after the cut-off for the daily maps.

other detections are possible transient events and are worth investigating.

I will consider the 2.0 results first. 14 light curves have only one change point in the BBA analysis, while 2 have greater than 1. These single change point alerts are all located within the first year of HAWC's operation and show a transition from higher to lower flux. The higher average flux duration lasts from 100 to 300 days, depending on the alert. Given the common date range, the higher flux is most likely due to detector maintenance shortly after operations started as construction finished shortly after HAWC came online in March 2015. An example is shown in the upper figure in Figure 5.6.

As for the 3.0 results, they broadly do not suffer from the same systematic issue the 2.0 results. Aside from two, all light curves are fit by at least 2 change points (3 blocks) to the data. These

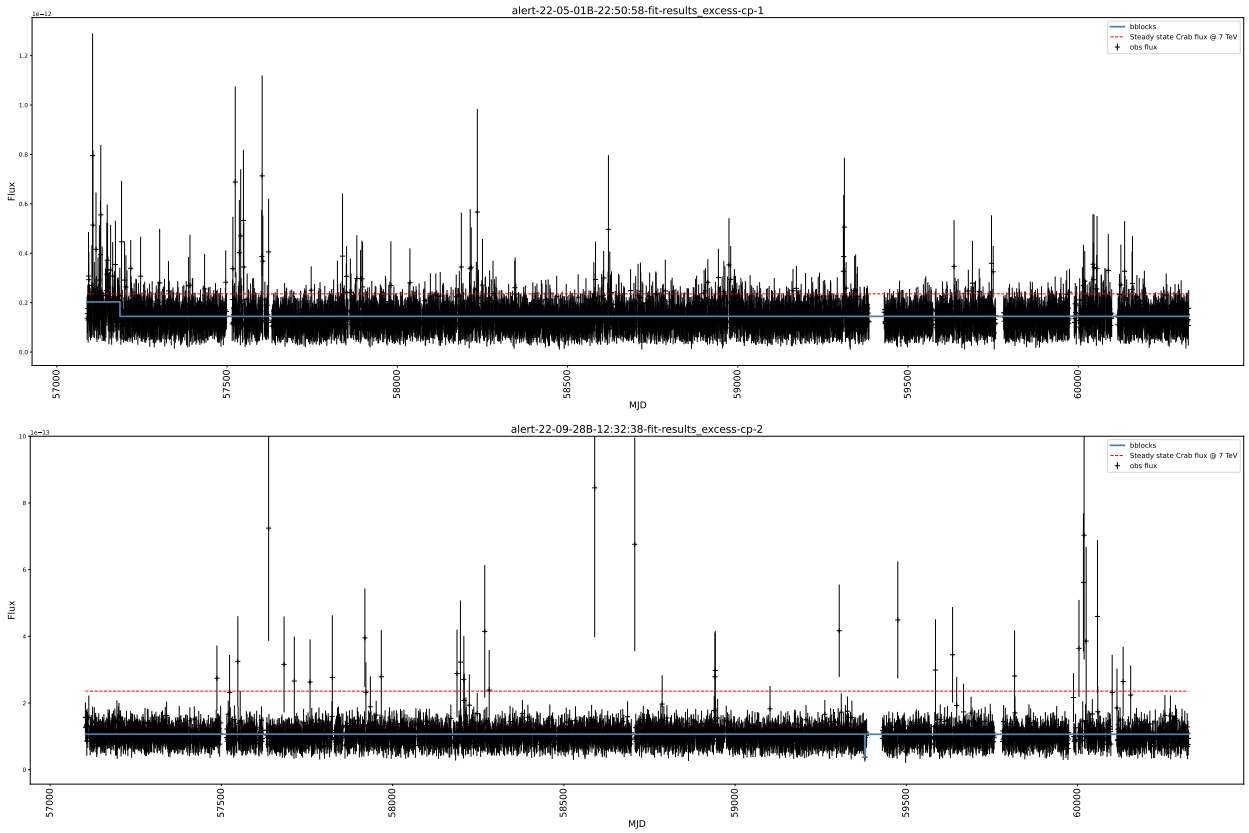

Figure 5.6 Examples of false positives for the 2.0 (upper plot) and 3.0 (lower plot) maps. Aside from the two Markarians, the 2.0 maps all share this slight excess in the first year of data. For the 3.0 maps, there's generally a low flux value around MJD 59400, though it moves slightly, depending on the map.

points are scattered across the whole date range (2015-2024), though they are being fit to under rather than the over fluctuations indicative of flares. These points are slightly below the mean flux distribution and are well constrained. An example is shown in the lower figure in Figure 5.6.

IceCube Alert Parameters					HAWC Results				Nearest γ-ray Source
IceCube alert Name	RA (deg)	DEC deg	Radius (deg)	s	n_{cp}	σ_{SS}	RA (deg)	Dec (deg)	Name (Distance) (deg)
11-12-08B-17:15:15	165.19	38.49	4.58	0.49	61	121	166.16	38.21	Mrk 421 (0.8)
14-05-03B-22:58:06	162.3	46.57	7.17	0.40	1	4.5	166.01	47.46	1ES 1028+511 (5.2)
15-05-26B-00:24:27	139.79	-1.49	1.84	0.28	1	3.2	139.61	-0.82	4FGL J0914.1-0202 (1.4)
15-09-18B-13:05:49	49.83	-2.95	2.23	0.28	1	3.4	49.48	-3.99	–
15-09-23G-00:38:34	103.23	3.96	0.80	0.33	1	2.8	103.8	4.22	–
17-08-19G-06:37:09	26.98	18.88	1.77	0.51	1	3.1	27.42	19.51	–
18-04-10G-18:38:35	218.5	0.56	0.88	0.60	1	2.5	219.42	0.45	–
18-11-21G-13:55:09	132.19	32.93	5.52	0.65	1	4.7	134.08	29.14	RX J0850.5+3455 (2.0)
19-04-13B-15:57:01	245.57	21.98	1.34	0.38	1	3.4	245.65	21.66	–
19-07-30G-20:50:41	226.14	10.77	1.36	0.67	1	3.4	227.29	10.77	PKS 1502+106 (0.3)
21-02-10G-11:53:55	206.06	4.78	0.85	0.65	1	2.1	205.49	5.27	4C +05.57 (0.5)
21-12-16B-07:07:38	316.05	15.79	1.83	0.34	1	2.7	317.64	16.61	4FGL J2100.0+1445 (1.4)
22-05-01B-22:50:58	311.57	18.68	0.93	0.39	1	3.6	311.84	19.12	–
24-02-29B-15:49:06	72.25	15.79	1.14	0.48	1	2.7	72.42	15.68	4FGL J0445.7+1535 (0.8)
24-03-07G-07:45:46	239.63	39.94	15.4	0.61	4	26.1	253.54	39.75	Mrk 501 (13.9)
25-02-07B-02:07:55	132.93	20.66	1.66	0.46	1	3.6	134.08	19.63	–

Table 5.1 The IceCube alerts and location parameters that were detected with this transient search on the index 2.0 maps are given in the first 5 columns. The s value is the astrophysical probability for that alert. The next 4 columns give the change point number found by the BBA, the HAWC max steady-state significance σ_{SS}, and its location. The last column indicates whether a GeV or TeV γ-ray source exists within the alert ROI. These sources are from either from TeVCat [2] or 4FGL [3]. These detections correspond to a 4.3% coincident detection rate

IceCube Alert Parameters					HAWC Results				Nearest γ-ray Source
IceCube alert Name	RA (deg)	DEC deg	Radius (deg)	s	n_{cp}	σ_{SS}	RA (deg)	Dec (deg)	Name (Distance) (deg)
11-12-08B-17:15:15	165.19	38.49	4.58	0.49	44	121	166.16	38.21	Mrk 421 (0.8)
12-08-07G-04:58:17	330.07	1.42	0.68	0.74	2	2.1	330.6	1.79	–
14-01-14B-21:04:09	337.59	0.71	0.83	0.34	1	2.1	337.5	-0.04	PWN J2227+0037 (0.6)
15-01-20B-23:37:54	95.89	14.13	0.89	0.34	2	4.3	139.92	40.57	–
16-01-04G-10:39:08	79.41	5.0	0.85	0.57	2	2.5	79.63	4.89	TXS 0513+054 (0.8)
16-03-07B-16:43:11	91.32	10.47	5.73	0.28	1	5.8	97.16	10.73	NVS J0660015+124344 (2.1)
16-06-14B-12:37:43	214.76	40.82	3.65	0.41	2	3.8	210.8	42.9	H1426+428 (2.5)
16-10-01G-10:32:23	192.57	37.12	2.14	0.64	2	3.1	190.77	37.36	Ton 116 (1.6)
18-04-10G-18:38:35	218.5	0.56	0.88	0.60	2	2.5	219.42	0.45	–
19-10-01G-20:09:18	313.99	12.79	3.17	0.59	2	2.7	315.83	10.43	PKS 2047+098 (3.1)
20-01-09G-23:41:39	165.45	11.8	2.62	0.77	2	3.4	167.17	11.15	TXS 1100+122 (0.4)
20-11-15G-02:07:26	195.12	1.38	1.26	0.46	2	3.4	194.94	2.05	–
22-09-28B-12:32:38	207.42	10.43	1.42	0.38	2	2.6	206.98	11.53	RX J1351.3+1115 (0.9)
22-12-24B-00:55:09	335.74	1.42	0.58	0.28	2	2.4	335.35	1.27	4FGL J2223.3+0122 (0.4)
22-12-29G-07:25:27	31.9	4.18	1.33	0.54	2	2.7	32.39	4.33	–
23-04-01B-16:14:18	8.17	1.94	3.06	0.27	2	3.3	5.84	2.84	NVS J003007-000008 (2.0)
24-02-29B-15:49:06	72.25	15.79	1.14	0.48	2	2.7	72.42	15.68	4FGL J0445.7+1535 (0.8)

Table 5.2 The IceCube alerts and location parameters that were detected with this transient search on the index 2.0 maps are given in the first 5 columns. The s value is the astrophysical probability for that alert. The next 4 columns give the change point number found by the BBA, the HAWC max steady-state significance σ_{SS}, and its location. The last column indicates whether a GeV or TeV γ-ray source exists within the alert ROI. These sources are from either from TeVCat [2] or 4FGL [3]. These detections correspond to a 4.6% coincident detection rate.

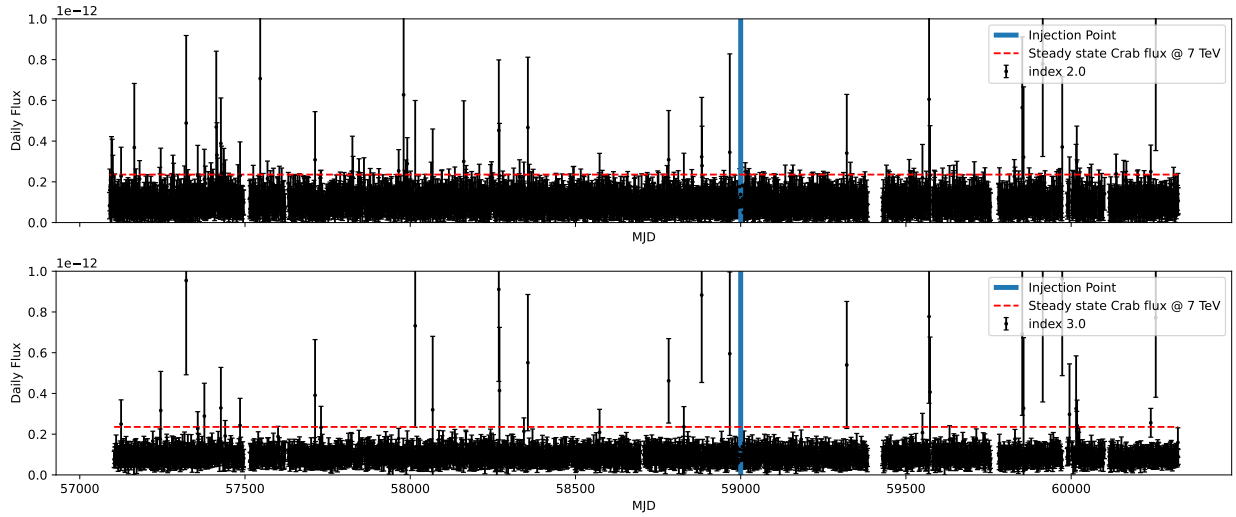

Figure 5.7 Comparing the 2.0 (upper plot) and 3.0 (lower plot) data sets for the IceCube alert 24-04-24G used in the injection studies. The simulated Crab events are injected at MJD 59000, indicated by the vertical blue line.

5.4 Sensitivity Studies

To test the sensitivity of the BBA, I perform a series of injection tests using one IceCube alert for both index maps. The selected alert is 24-04-24G, located at 327.08, 3.06, with an alert radius of 1.61 degrees. This alert is selected as it had a null detection for both maps. The light curves for the 2.0 and 3.0 maps are shown in Figure 5.7.

After selecting the light curve, I injected fake flares of varying fluxes (in Crab units) with durations of 1, 3, 5, and 10 days. The source is injected at date MJD 59000 (May 31, 2020). The flux scaling factor ranges over 0.5 to 2.0, and is multiplied by the daily Crab flux for the given day. To find what the actual flux value for the Crab would be, I performed 10 Crab fits on the relevant daily maps. These fits assumed a simple power law to match our maps' assumption along with a fixed index (either 2 or 3). The flux normalization and location (ra, dec) were free for the fit. The days fit were MJDs 59000 to 59009. Once these fluxes are found, they are multiplied by the desired scaling value to get the final scaled flux.

To simulate the fake flares, the Analysis and Event Reconstruction Integrated Environment (AERIE) framework is used. Described in [110], this is a modular framework that HAWC uses

for reconstruction and data manipulation. I used AERIE to construct a series of simulated sources inside the IceCube alert's ROI. To simulate the sources, AERIE used the best fit flux value from the Crab fits and then scaled it by the desired fraction of Crab flux. Lastly, the max flux search described in Section 5.2.1 is run on the simulated sources to construct the new, simulated light curve.

With the light curve made, the BBA is then fit to determine how sensitive HAWC is to transient events. The results of this are presented in Table 5.3. While the two maps differ slightly, in general, HAWC would need to observe multi-day events for fluxes less than the Crab. If the event is higher energy, then Crab-like events only need 1 day to be detected in the 2.0 maps while lower energy events require 1.25 x Crab flux. The key conclusion is that HAWC requires the event to at least be 0.75 x Crab to be detected in either map. Anything dimmer and HAWC cannot detect it.

Flare Duration	2.0 Crab Flux Fraction							3.0 Crab Flux Fraction						
–	0.5	0.75	1.0	1.25	1.5	1.75	2.0	0.5	0.75	1.0	1.25	1.5	1.75	2.0
1 day	x	x	✓	✓	✓	✓	✓	x	x	x	✓	✓	✓	✓
3 days	x	✓	✓	✓	✓	✓	✓	x	x	x	✓	✓	✓	✓
5 days	x	✓	✓	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓
10 days	x	✓	✓	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓

Table 5.3 Detection rate of the BBA with the injected maps. For the 2.0 maps, the lower ncp_prior allows for faster detection with lower relative flux compared to the 3.0 maps.

	Mrk 421	Mrk 501
$\sqrt{TS}$	121	26
RA (deg)	166.14 +/- 0.03	253.54 +/- 0.01
Dec (deg)	38.19 +/- 0.02	39.76 +/- 0.01
$N_0 * 10^{-12}$ [TeV^{-1} cm^{-1} s^{-1}]	6.85 + /−0.23	1.07 + /−0.05
α	−2.437 + /−0.027	−2.57 + /−0.05
E_c[TeV]	9.9 + /−1.2	–
E_{max}	16.3	15.0

Table 5.4 Fit parameters for the two Markarians. The fits assumed a pivot of 2 TeV. The max energy is presented at the 2σ threshold. Additionally, the uncertainties are statistical only.

5.5 Multi-Messenger Analysis with IceCube Alerts

As AGNs have been proven to emit neutrinos [63] and two of the neutrino alerts include Mrks 421 (11-12-08B) and 501 ((24-03-07G)) inside their 90% containment radii, I perform an analysis to determine if the Markarians could be the source of the neutrino alerts. Both Markarians were quiescent during these events, so a steady-state emission model is assumed.

While HAWC does have a long-term analysis of the two Markarians [28], that analysis only used approximately 1000 days of data. For this analysis, I use the 2565-day Pass 5 data set. Additionally, I used the neural network energy estimator introduced in Section 2.1.3.3. This data uses a superior background rejection algorithms that are discussed in [111].

To fit the two Markarians, the spectral models from [28] are used. These are a cut-off power law and a power law for Mrks 421 and 501 respectively. Additionally, the Extragalactic Background Light (EBL) must be considered. The EBL absorbs high energy γ-rays and re-emits them at lower energies. The spectral models fitted for the Markarians are

$$\frac{dE}{dN} = N_o(\frac{E}{E_p})^{-\alpha} * \exp(-\tau(E,z)) \tag{5.12}$$

$$\frac{dE}{dN} = N_o(\frac{E}{E_p})^{-\alpha} * \exp(\frac{-E}{E_c}) * \exp(-\tau(E,z)). \tag{5.13}$$

The pivot energy is fixed to 2 TeV and the EBL model is given by $-\tau$. It is a function of the energy and redshift. The model considered is [112] with redshifts of 0.03 and 0.031 for Mrks 421 and 501, respectively [28]. I fit the models using the Multi-Mission Maximum Likelihood (3ML)

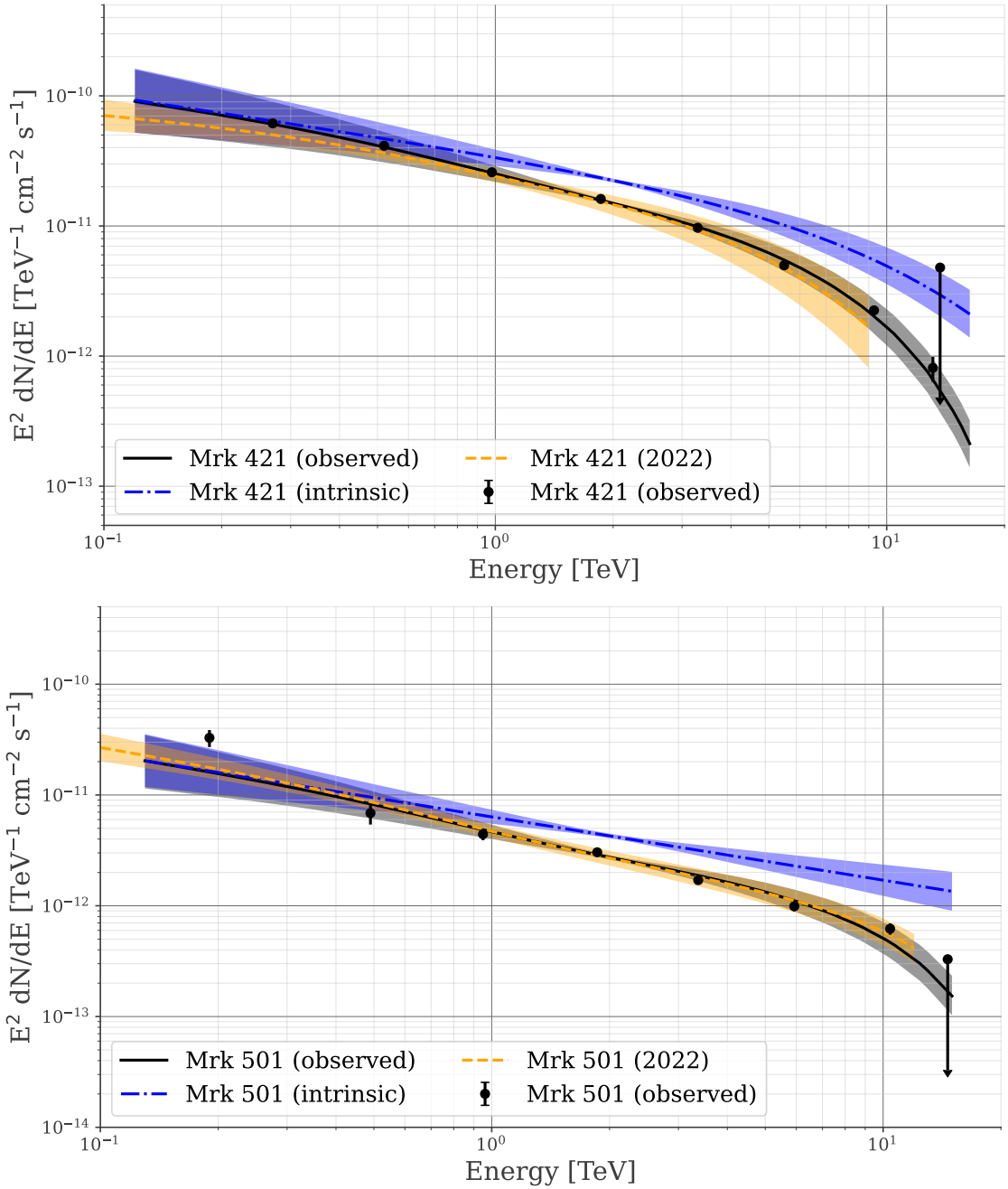

Figure 5.8 The spectral energy distributions for Mrk 421 (right) and Mrk 501 (left). The fits for this work are solid black with their intrinsic spectra indicated in dash dotted blue. Lastly, the orange dashed lines are the previous results from [28].

framework [89][1] with the HAWC Accelerated Likelihood (HAL) plugin [90][2]. The results of these fits are given in Table 5.4 and their spectra are shown in Figure 5.8.

[1] `https://github.com/threeML/threeML`
[2] `https://github.com/threeML/hawc_hal`

I then model their emission. As discussed in Section 1, neutrinos are produced via hadronic mechanisms which, at TeV energies, is primarily driven by pion decay. Any γ-rays produced via pion decay will have approximately a 2:1 energy relation with neutrinos [34]. The reported energies of the two neutrinos are 123 and 191 TeV for Mrks 421 and 501, respectively. Therefore, I would expect to see 250 to 400 TeV photons (before EBL attenuation) from the Markarians, making them Pevatrons. It should be noted that multi-wavelength fits for both Mrks 421 and 501 have been described in [28] with a self-scattering Compton (SSC) leptonic (electron-based) model. This will be discussed later this section

To test this, the Naima [101] and gammapy [113] python packages are used. Naima is non-thermal modelling framework while gammapy is a flexible fitting package for γ-ray data. I use gammapy to fit a Naima model of pion-producing protons to HAWC data, assuming a cut-off power law spectrum, and consider the EBL attenuation model of γ-rays with the model by [112] EBL model.

Fitting this model to the HAWC data provides inconclusive results. The indices and cut-offs are -1.23 +/- 0.02, 250 +/- 5 TeV for Mrk 421 and -0.257 +/- 0.04, 4600 +/- 910 TeV for Mrk 501. These cut-off energies indicate that the Markarians may be capable of producing the IceCube neutrinos, and the indices are non-physical. Additionally, the fit quality is poor, with reduced χ^2 values of 0.6 and 0.2 for the two Mrks 421 and 501. Two potential explanations for the poor quality of the fit results are given below.

The first and more important component is that the preferred particle fitting for Mrks 421 and 501 is the self-scattering Compton model. That is the model used in HAWC's previous results [28], and accurately describes all observed data from radio to TeV γ-rays. However, that does not rule out pion decay being present. Certain models, such as a proton-synchrotron model discussed in chapter 10 of [36], allow for hadronic emission at TeV energies. An alternate option is that the γ-rays produced from pion decay get absorbed by the dense environment in the accretion disk while the neutrinos can easily escape. In this case, the available γ-ray could not be used to model the hadronic process. Testing these scenarios requires further multi-wavelength investigations with

more data, as is beyond the scope of this work.

Last of note are the fit results for the pion decay parameters. While the cut-off energies have some promise for producing γ-rays > 200 TeV, the fit values for the indices are not compatible with physical models. Considering the proton-synchrotron model mentioned above, that predicts an index of 2 [36], as does shock acceleration. These poor fit results hint that HAWC's data disfavors a simple pion decay model, though more complex models and external data from lower energies need to be considered to conclusively rule this out. Lastly, the probabilities for these alerts being astrophysical are 49% for Mrk 421 and 61% for Mrk 501, along with a 90% containment radius of 4.6 and 15.4 degrees respectively. The aforementioned factors further increase the probability of false association.

5.6 Systematic Uncertainties

Considering the HAWC-IceCube coincident search first, there are three factors that affect the results from the BBA: the calibration of the BBA and potential declination effects, and a potentially low signal/noise ratio. As the calibration and declination dependence are correlated, they will be considered first. The method used for calibrating the BBA only considered four of the possible 8 declinations bands available. These declination bands, centered on [5, 15, 25, 35] degrees, represent the region nearest to HAWC's zenith of 19 degrees. As shown in Figure 3 in [1], HAWC's PSF worsens the further off zenith the events are. This effect is reflected in the average higher flux and larger uncertainty values that can be seen in Mrk 501's light curve in Figure 5.5. Broadly speaking, alerts closer to zenith have a more constrained uncertainty distribution.

Several methods were tested to address this issue. First, the overall declination range was set to [-20, 60] degrees to help reduce this systematic. Then, several configurations for declination bands were tested. These ranged from 8 bands, each 10 degrees wide, to the 3 bands defined in [1]: [0, 26], [26, 37], [37, 46] degrees off zenith. While calibrating these different configurations would succeed, applying them to real alerts usually resulted in unreasonably high detection rates of > 10%. A compromise was made to focus on where HAWC is most sensitive while still being able to see a high zenith flare, if it was sufficiently strong. This is the primary reason why most of

the detections presented in Tables 5.1 and 5.2 are relatively close to HAWC's zenith.

It is also worth noting is that, while the BBA is noted to be relatively agnostic to ncp_prior values for moderate to high signal-to-noise ratios, it can have issues in low signal-to-noise environments [109]. This partially motivated the decision to restrict calibration to HAWC's core declination bands. This decision maximizes the possibility of finding flares within 20 degrees of HAWC's zenith, but penalizes the probability of finding high declination events. Declination-dependent calibration was attempted, but was too inconsistent to be applied across the relevant declination band.

CHAPTER 6

CONCLUSIONS

In this work, two analyses were performed, one using just HAWC data while the other was a joint HAWC-IceCube analysis. These explored some of the fundamental questions in particle astrophysics, like the origin of cosmic rays and neutrinos. The key goal was to determine the potential production methods these sources use to emit their respective messenger.

In Chapter 4, a detected analysis of the probable PWN HAWC J2031+415 was performed using 2400 days of HAWC data. This entailed first performing a multi-source fit of the Cygnus Cocoon to disentangle HAWC J2031+415 from the larger Cocoon and the SNR Gamma Cygni. This revealed a slightly extended source with a cut-off power law spectral model and an energy range of 0.4 to 151 TeV. An energy-dependent morphology study was then performed to see if HAWC J2031+415's location would shift towards the pulsar theorized to be powering it. While no clear locational shift was observed, a faint decreasing in size with increasing energies was observed. Next, the periastron of the binary pulsar system was done, with no clear flaring observed in HAWC data. Lastly, a multi-wavelength fit using radio, x-ray, and gamma ray data was considered. This found slight evidence for the observed HAWC data to be produced leptonicly with inverse Compton scattering producing TeV photons from the PWN.

Chapter 5 describes a transient search of HAWC daily maps using public IceCube neutrino alerts. Daily maps from March 20, 2015 to January 15, 2024, totaling approximately 3000 days of HAWC data when accounting for detector downtime, and 365 neutrino alerts were used. A Bayesian Block algorithm was utilized to search for potential gamma ray flares that might occur within a given IceCube neutrino alert. Finding such emission would imply that the hadronic acceleration mechanism pion decay would be responsible for producing the TeV gamma rays and neutrinos. Of the 365 neutrino alerts, HAWC detected gamma ray flaring in approximately 25, depending on the spectral assumptions. This in line with a false positive rate of 5%, but two of the alerts contained the AGN Markarians 421 and 501. Full source fits were performed to test for hadronic emission, with the results heavily disfavoring this conclusion. The upcoming second IceCube alert

catalog promises superior reconstruction algorithms and the just completed IceCube Gen 2 upgrade will greatly improved its angular resolution. Such enhancements, combined with the successful implementation of HAWC's outrigger array, would greatly improve this analysis in the future.

Beyond HAWC and IceCube, there are several other observatories, both completed and planned, that will greatly expand the field of multi-messenger astronomy. Gamma ray observatories like LHAASO and the planned CTA and SWGO will both deliver superior angular resolution and extended energies into the PeV regime. Likewise, future neutrino experiments like KM3Net and P-ONE will significantly improve on IceCube's already impressive sensitivity and energy range. Improved communication among the various instruments will allow for near-instantaneous follow-ups of detections, therefore revealing source acceleration mechanisms and, potentially, new physics. This field is still in its infancy and has a bright future.

BIBLIOGRAPHY

[1] A Albert, R Alfaro, C Alvarez, A Andrés, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, R Babu, E Belmont-Moreno, A Bernal, et al. Performance of the hawc observatory and tev gamma-ray measurements of the crab nebula with improved extensive air shower reconstruction algorithms. *The Astrophysical Journal*, 972(2):144, 2024.

[2] S. P. Wakely and D. Horan. TeVCat: An online catalog for Very High Energy Gamma-Ray Astronomy. *International Cosmic Ray Conference*, 3:1341–1344, 2008.

[3] Soheila Abdollahi et al. Fermi large area telescope fourth source catalog. *The Astrophysical Journal Supplement Series*, 247(1):33, 2020.

[4] Christian Spiering. Status and perspectives of astroparticle physics in europe. *arXiv preprint arXiv:0804.1500*, 2008.

[5] Thomas K Gaisser, Ralph Engel, and Elisa Resconi. *Cosmic rays and particle physics*. Cambridge University Press, 2016.

[6] Maurizio Spurio et al. *Particles and astrophysics*, volume 1799. Springer, 2014.

[7] Jearl Walker, David Halliday, and Robert Resnick. Halliday & resnick fundamentals of physics. *(No Title)*, 2014.

[8] Ilya Obodovskiy. *Radiation*. Elsevier, 2019.

[9] Joel R Primack, Alberto Dominguez, Rudy C Gilmore, and Rachel S Somerville. Extragalactic background light and gamma-ray attenuation. In *AIP Conference Proceedings*, volume 1381, pages 72–83. American Institute of Physics, 2011.

[10] D Raban. Infrared interferometric obser ation of dust in the nuclei of acti e galaxies. *Leiden University Scholarly Publications*, 2009.

[11] David J Thompson. Space detectors for gamma rays (100 mev–100 gev): From egret to fermi lat. *Comptes Rendus Physique*, 16(6-7):600–609, 2015.

[12] Zigfried Hampel-Arias. *Cosmic ray observations at the TeV scale with the HAWC observatory*. The University of Wisconsin-Madison, 2017.

[13] Horvath, Arpad. Own work (no machine-readable source provided). `https://commons.wikimedia.org/w/index.php?curid=637092`, n.d. Wikimedia Commons, CC BY-SA 2.5.

[14] John Pretz. Highlights from the high altitude water cherenkov observatory. *arXiv preprint arXiv:1509.07851*, 2015.

[15] Dieter Heck, Johannes Knapp, JN Capdevielle, G Schatz, T Thouw, et al. Corsika: A monte carlo code to simulate extensive air showers. *Report fzka*, 6019(11), 1998.

[16] AU Abeysekara, A Albert, R Alfaro, C Alvarez, JD Álvarez, R Arceo, JC Arteaga-Velázquez, HA Ayala Solares, AS Barber, N Bautista-Elivar, et al. Observation of the crab nebula with the hawc gamma-ray observatory. *The Astrophysical Journal*, 843(1):39, 2017.

[17] AU Abeysekara, R Alfaro, C Alvarez, JD Álvarez, R Arceo, JC Arteaga-Velázquez, HA Ayala Solares, AS Barber, BM Baughman, N Bautista-Elivar, et al. Data acquisition architecture and online processing system for the hawc gamma-ray observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 888:138–146, 2018.

[18] Karen Andeen and Matthias Plum. Latest cosmic ray results from icetop and icecube. In *EPJ Web of Conferences*, volume 210, page 03005. EDP Sciences, 2019.

[19] Hazal Göksu. Status of the southern wide-field gamma-ray observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, page 171304, 2026.

[20] P Abreu, R Alfaro, A Alfonso, M Andrade, EO Angüner, EA Anita-Rangel, O Aquines-Gutiérrez, C Arcaro, R Arceo, JC Arteaga-Velázquez, et al. Science prospects for the southern wide-field gamma-ray observatory: Swgo. *arXiv preprint arXiv:2506.01786*, 2025.

[21] J. Pretz. A.T.H.E.N.A. advanced tracking of HAWC experiment notifications and alerts., 2015.

[22] Luna. Luna Optical Backscatter Reflectometer 4600 User Guide, 03 2026.

[23] Luna. OPTICAL BACKSCATTER REFLECTOMETRY (OBR) - OVERVIEW AND APPLICATIONS, 03 2026.

[24] Felix Aharonian et al. The unidentified tev source (tev j2032+ 4130) and surrounding field: Final hegra iact-system results. *Astronomy & Astrophysics*, 431(1):197–202, 2005.

[25] J Albert et al. Magic observations of the unidentified γ-ray source tev j2032+ 4130. *The Astrophysical Journal*, 675(1):L25, 2008.

[26] AU Abeysekara et al. Periastron observations of tev gamma-ray emission from a binary system with a 50-year period. *The Astrophysical Journal Letters*, 867(1):L19, 2018.

[27] AU Abeysekara, A Archer, Taylor Aune, Wystan Benbow, Ralph Bird, Robert Brose, M Buchovecky, V Bugaev, Wei Cui, MK Daniel, et al. A very high energy γ-ray survey toward the cygnus region of the galaxy. *The Astrophysical Journal*, 861(2):134, 2018.

[28] A Albert, R Alfaro, C Alvarez, JR Angeles Camacho, JC Arteaga-Velázquez, KP Arunbabu, D Avila Rojas, HA Ayala Solares, V Baghmanyan, E Belmont-Moreno, et al. Long-term spectra of the blazars mrk 421 and mrk 501 at tev energies seen by hawc. *The Astrophysical Journal*, 929(2):125, 2022.

[29] P A Zyla et al. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8):083C01, 08 2020.

[30] Victor Hess. On the observations of the penetrating radiation during seven balloon flights. *arXiv preprint arXiv:1808.02927*, 2018.

[31] Georgi T Zatsepin and Vadem A Kuz'min. Upper limit of the spectrum of cosmic rays. *Soviet Journal of Experimental and Theoretical Physics Letters*, 4:78, 1966.

[32] R Alfaro, C Alvarez, JC Arteaga-Velázquez, KP Arunbabu, D Avila Rojas, R Babu, E Belmont-Moreno, KS Caballero-Mora, T Capistrán, A Carramiñana, et al. Exploring the coronal magnetic field with galactic cosmic rays: the sun shadow observed by hawc. *The Astrophysical Journal*, 966(1):67, 2024.

[33] Stephan Rosswog and Marcus Brüggen. Introduction to high-energy astrophysics. *American history*, 1861(1900), 2007.

[34] David Griffiths. *Introduction to elementary particles*. John Wiley & Sons, 2020.

[35] Zhen Cao et al. Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray galactic sources. *Nature*, 594(7861):33–36, 2021.

[36] Felix A Aharonian. *Very high energy cosmic gamma radiation: a crucial window on the extreme Universe*. World Scientific, 2004.

[37] R Alfaro, C Alvarez, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, R Babu, E Belmont-Moreno, KS Caballero-Mora, T Capistrán, A Carramiñana, et al. Analysis of the emission and morphology of the pulsar wind nebula candidate hawc j2031+ 415. *The Astrophysical Journal*, 975(2):198, 2024.

[38] Gabriele Ghisellini. *Radiative processes in high energy astrophysics*. Springer, 2013.

[39] Anna Lia Longinotti, HAWC Collaboration, et al. The first 10years of the hawc gamma-ray observatory: science results. *arXiv preprint arXiv:2512.03938*, 2025.

[40] AU Abeysekara, A Albert, R Alfaro, C Alvarez, JD Álvarez, R Arceo, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, AS Barber, et al. Extended gamma-ray sources around pulsars constrain the origin of the positron flux at earth. *Science*, 358(6365):911–914, 2017.

[41] Mattia Di Mauro, Silvia Manconi, and Fiorenza Donato. Detection of a γ-ray halo around

geminga with the fermi-lat data and implications for the positron flux. *Physical Review D*, 100(12):123015, 2019.

[42] AU Abeysekara, A Albert, R Alfaro, C Alvarez, JD Álvarez, R Arceo, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, E Belmont-Moreno, et al. Very-high-energy particle acceleration powered by the jets of the microquasar ss 433. *Nature*, 562(7725):82–85, 2018.

[43] R Alfaro, C Alvarez, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, R Babu, E Belmont-Moreno, KS Caballero-Mora, T Capistrán, A Carramiñana, et al. Ultra-high-energy gamma-ray bubble around microquasar v4641 sgr. *Nature*, 634(8034):557–560, 2024.

[44] R Alfaro, M Araya, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, R Babu, P Bangale, E Belmont-Moreno, A Bernal, KS Caballero-Mora, et al. Orbital modulation of gamma rays up to 100 tev from ls 5039. *The Astrophysical Journal Letters*, 987(2):L42, 2025.

[45] Markus Ackermann et al. A cocoon of freshly accelerated cosmic rays detected by fermi in the cygnus superbubble. *science*, 334(6059):1103–1107, 2011.

[46] AU Abeysekara et al. Hawc observations of the acceleration of very-high-energy cosmic rays in the cygnus cocoon. *Nature astronomy*, 5(5):465–471, 2021.

[47] Cong Li. Detection of emission from cygnus cocoon above 100tev with lhaaso. In *37th International Cosmic Ray Conference. 12-23 July 2021. Berlin*, page 843, 2022.

[48] Yergali Kurmanov, Kuantay Boshkayev, Talgar Konysbayev, Orlando Luongo, Nazym Saiyp, Ainur Urazalina, Gulfeiruz Ikhsan, and Gulnara Suliyeva. Accretion disks properties around regular black hole solutions obtained from non-linear electrodynamics. *Physics of the Dark Universe*, 46:101566, 2024.

[49] Razmik Mirzoyan. Major change in understanding of grbs at tev. *arXiv preprint arXiv:2005.03641*, 2020.

[50] Yong Huang, Shicong Hu, Songzhan Chen, Min Zha, Cheng Liu, Zhiguo Yao, Zhen Cao, et al. Lhaaso observed grb 221009a with more than 5000 vhe photons up to around 18 tev. *GRB Coordinates Network*, 32677:1, 2022.

[51] William B Atwood, Aous A Abdo, Markus Ackermann, W t Althouse, B Anderson, M Axelsson, Luca Baldini, J Ballet, DL Band, Guido Barbiellini, et al. The large area telescope on the fermi gamma-ray space telescope mission. *The Astrophysical Journal*, 697(2):1071–1102, 2009.

[52] F Aharonian, J Buckley, T Kifune, and G Sinnis. High energy astrophysics with ground-based gamma ray detectors. *Reports on Progress in Physics*, 71(9):096901, 2008.

[53] Ingo Allekotte, AF Barbosa, P Bauleo, C Bonifazi, B Civit, CO Escobar, B García, G Guedes, M Gómez Berisso, JL Harton, et al. The surface detector system of the pierre auger observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 586(3):409–420, 2008.

[54] Daniel Ferenc, MAGIC collaboration, et al. The magic gamma-ray observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 553(1-2):274–281, 2005.

[55] Jamie Holder, RW Atkins, HM Badran, G Blaylock, SM Bradbury, JH Buckley, KL Byrum, DA Carter-Lewis, O Celik, YCK Chow, et al. The first veritas telescope. *Astroparticle Physics*, 25(6):391–401, 2006.

[56] James Anthony Hinton, Hess Collaboration, et al. The status of the hess project. *New Astronomy Reviews*, 48(5-6):331–337, 2004.

[57] M Amenomori, Z Cao, LK Ding, ZY Feng, K Hibino, N Hotta, Q Huang, AX Huo, HY Jia, GZ Jiang, et al. Search for steady emission of 10-tev gamma rays from the crab nebula, cygnus x-3, and hercules x-1 using the tibet air shower array. *Physical review letters*, 69(17):2468, 1992.

[58] A.U. Abeysekara et al. The high-altitude water cherenkov (hawc) observatory in méxico: The primary detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1052:168253, 2023.

[59] Benjamin P Abbott, Richard Abbott, TD Abbott, MR Abernathy, Fausto Acernese, K Ackley, C Adams, T Adams, Paolo Addesso, RX Adhikari, et al. Properties of the binary black hole merger gw150914. *Physical review letters*, 116(24):241102, 2016.

[60] Mark G Aartsen, M Ackermann, J Adams, JA Aguilar, M Ahlers, M Ahrens, D Altmann, K Andeen, T Anderson, I Ansseau, et al. The icecube neutrino observatory: instrumentation and online systems. *Journal of Instrumentation*, 12(03):P03012, 2017.

[61] Viviana Niro. Detecting the brightest hawc sources with icecube in the upcoming years. *Physical Review D*, 103(10):103020, 2021.

[62] Andrea Albert, R Alfaro, C Alvarez, JD Álvarez, JR Angeles Camacho, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, R Babu, E Belmont-Moreno, et al. Hawc study of the ultra-high-energy spectrum of mgro j1908+ 06. *The Astrophysical Journal*, 928(2):116, 2022.

[63] Rasha Abbasi, Markus Ackermann, Jenni Adams, Sanjib Kumar Agarwalla, Juanan Aguilar, Markus Ahlers, Jean-Marco Alameddine, Najia_Moureen Binte Amin, Karen Andeen, Gisela Anton, et al. Txs 0506+ 056 with updated icecube data. *Proceedings of Science*, 2023.

[64] Mark G Aartsen, R Abbasi, M Ackermann, J Adams, JA Aguilar, M Ahlers, M Ahrens, C Alispach, P Allison, NM Amin, et al. Icecube-gen2: the window to the extreme universe. *Journal of Physics G: Nuclear and Particle Physics*, 48(6):060501, 2021.

[65] Paweł Malecki. Pacific ocean neutrino experiment. *Universe*, 10(2):53, 2024.

[66] Vincent Marandon, Armelle Jardin-Blicq, and Harm Schoorlemmer. Latest news from the hawc outrigger array. *arXiv preprint arXiv:1908.07634*, 2019.

[67] Abeysekara et al. triggered. *The Astrophysical Journal*, 843(1):39, 2017.

[68] AU Abeysekara et al. Measurement of the crab nebula spectrum past 100 tev with hawc. *The Astrophysical Journal*, 881(2):134, 2019.

[69] P Speckmayer, A Höcker, J Stelzer, and H Voss. The toolkit for multivariate data analysis, TMVA 4. *Journal of Physics: Conference Series*, 219(3):032057, apr 2010.

[70] Hugo Alberto Ayala Solares. Amon multimessenger alerts: Past and future. *Galaxies*, 7(1):19, 2019.

[71] R Abbasi, M Ackermann, J Adams, S_K Agarwalla, J_A Aguilar, M Ahlers, J_M Alameddine, N_M Amin, K Andeen, G Anton, et al. Icecat-1: the icecube event catalog of alert tracks. *The Astrophysical Journal Supplement Series*, 269(1):25, 2023.

[72] IceCube Collaboration*†, R Abbasi, M Ackermann, J Adams, JA Aguilar, M Ahlers, M Ahrens, JM Alameddine, AA Alves Jr, NM Amin, et al. Observation of high-energy neutrinos from the galactic plane. *Science*, 380(6652):1338–1343, 2023.

[73] PHP homepage.

[74] Occupational Safety and Health Administration in the United States. Safety and health regulations for construction, 07 2024.

[75] Ian Herzog. A combined spectral and energy morphology analysis of gamma ray source hawc j2031+ 415 in the cygnus constellation. Master's thesis, Michigan Technological University, 2022.

[76] Anushka Udara Abeysekara et al. Spectral and Energy Morphology Analysis Study of HAWC J2031+415. In *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*, volume 395, page 836, 2021.

[77] Ian Herzog, A Albert, R Alfaro, C Alvarez, A Andrés, JC Arteaga-Velázquez, D Avila Rojas, HA Ayala Solares, R Babu, E Belmont-Moreno, et al. A spectral, morphological, and emission analysis of gamma ray source hawc j2031+ 415. *Proceedings of Science*, 444:757, 2024.

[78] GP Rowell et al. The new unidentified tev source in cygnus (tev j2032+ 4130): Hegra iact-system results. In *International Cosmic Ray Conference*, volume 4, page 2345, 2003.

[79] Hiroshi Murakami, Shunji Kitamoto, Akiko Kawachi, and Takeshi Nakamori. Detection of x-ray emission from the unidentified tev gamma-ray source tev j2032+ 4130. *Publications of the Astronomical Society of Japan*, 63(sp3):S873–S878, 2011.

[80] D Horns, AID Hoffmann, A Santangelo, FA Aharonian, and GP Rowell. Xmm-newton observations of the first unidentified tev gamma-ray source tev j2032+ 4130. *Astronomy & Astrophysics*, 469(1):L17–L21, 2007.

[81] Josep M Paredes, Josep Martí, CH Ishwara Chandra, and Valentí Bosch-Ramon. The population of radio sources in the field of the unidentified gamma-ray source tev j2032+ 4130. *The Astrophysical Journal*, 654(2):L135, 2006.

[82] Josep Marti, Josep M Paredes, CH Ishwara Chandra, and Valentı Bosch-Ramon. Deep radio images of the hegra and whipple tev sources in the cygnus ob2 region. *Astronomy & Astrophysics*, 472(2):557–564, 2007.

[83] E Aliu et al. Observations of the unidentified gamma-ray source tev j2032+ 4130 by veritas. *The Astrophysical Journal*, 783(1):16, 2014.

[84] AU Abeysekara et al. The 2hwc hawc observatory gamma-ray catalog. *The Astrophysical Journal*, 843(1):40, 2017.

[85] A Albert et al. 3hwc: The third hawc catalog of very-high-energy gamma-ray sources. *The Astrophysical Journal*, 905(1):76, 2020.

[86] R N Manchester, G B Hobbs, A Teoh, and M Hobbs. The australia telescope national facility pulsar catalogue. *The Astronomical Journal*, 129(4):1993, 2005.

[87] AG Lyne et al. The binary nature of psr j2032+ 4127. *Monthly Notices of the Royal Astronomical Society*, 451(1):581–587, 2015.

[88] A Albert et al. Spectrum and morphology of the very-high-energy source hawc j2019+ 368. *The Astrophysical Journal*, 911(2):143, 2021.

[89] Giacomo Vianello et al. The multi-mission maximum likelihood framework (3ml). *arXiv preprint arXiv:1507.08343*, 2015.

[90] Abeysekara et al. Characterizing gamma-ray sources with hal (hawc accelerated likelihood) and 3ml. In *37th International Cosmic Ray Conference.*, 2022.

[91] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60 – 62, 1938.

[92] Ackermann et al. Search for extended sources in the galactic plane using six years of fermi-large area telescope pass 8 data above 10 gev. *The Astrophysical Journal*, 843(2):139, 2017.

[93] A Albert, R Alfaro, C Alvarez, JC Arteaga-Velázquez, D Avila Rojas, R Babu, E Belmont-Moreno, A Bernal, M Breuhaus, KS Caballero-Mora, et al. Tev analysis of a source rich region with hawc observatory: Is hess j1809-193 a potential hadronic pevatron? *arXiv preprint arXiv:2407.08849*, 2024.

[94] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[95] D Anderson and K Burnham. Model selection and multi-model inference. *Second. NY: Springer-Verlag*, 63(2020):10, 2004.

[96] A Albert, R Alfaro, C Alvarez, A Andrés, JC Arteaga-Velázquez, D Avila Rojas, HA Solares, R Babu, E Belmont-Moreno, KS Caballero-Mora, et al. Performance of the hawc observatory and tev gamma-ray measurements of the crab nebula with improved extensive air shower reconstruction algorithms. *arXiv preprint arXiv:2405.06050*, 2024.

[97] A. U. Abeysekara et al. Extended gamma-ray sources around pulsars constrain the origin of the positron flux at Earth. *Science*, 358(6365):911–914, November 2017.

[98] LHAASO Collaboration. An ultrahigh-energy γ-ray bubble powered by a super pevatron. *Science Bulletin*, 69(4):449–457, 2024.

[99] Henrike Fleischhack. Modeling the non-thermal emission of the gamma cygni supernova remnant up to the highest energies. In *36th International Cosmic Ray Conference*, 2019.

[100] Vikas Joshi. *Reconstruction and analysis of highest energy γ-rays and its application to pulsar wind nebulae*. PhD thesis, Heidelberg University, Germany, January 2019.

[101] Víctor Zabalza. naima: a python package for inference of relativistic particle energy distributions from observed nonthermal spectra. *arXiv preprint arXiv:1509.03319*, 2015.

[102] CC Popescu, R Yang, RJ Tuffs, Giovanni Natale, M Rushton, and F Aharonian. A radiation transfer model for the milky way: I. radiation fields and application to high-energy astrophysics. *Monthly Notices of the Royal Astronomical Society*, 470(3):2539–2558, 2017.

[103] Fernando Camilo, Paul S Ray, Scott M Ransom, Marta Burgay, TJ Johnson, M Kerr, EV Gotthelf, JP Halpern, J Reynolds, RW Romani, et al. Radio detection of lat psrs j1741–2054 and j2032+ 4127: No longer just gamma-ray pulsars. *The Astrophysical Journal*, 705(1):1, 2009.

[104] Stephen P Reynolds, BM Gaensler, and Fabrizio Bocchino. Magnetic fields in supernova

remnants and pulsar-wind nebulae. *Space science reviews*, 166:231–261, 2012.

[105] Barbara Olmi and Niccolò Bucciantini. The dawes review 11: from young to old: the evolutionary path of pulsar wind nebulae. *Publications of the Astronomical Society of Australia*, 40:e007, 2023.

[106] Andrzej A Zdziarski, Andrii Neronov, and Maria Chernyakova. A compact pulsar wind nebula model of the γ-ray-loud binary ls i+ 61 303. *Monthly Notices of the Royal Astronomical Society*, 403(4):1873–1886, 2010.

[107] H. Abdalla et al. The population of tev pulsar wind nebulae in the h.e.s.s. galactic plane survey. *A&A*, 612:A2, 2018.

[108] Emma de Oña Wilhelmi, Rubén López-Coto, Elena Amato, and Felix Aharonian. On the potential of bright, young pulsars to power ultrahigh gamma-ray sources. *The Astrophysical Journal Letters*, 930(1):L2, 2022.

[109] Jeffrey D Scargle, Jay P Norris, Brad Jackson, and James Chiang. Studies in astronomical time series analysis. vi. bayesian block representations. *The Astrophysical Journal*, 764(2):167, 2013.

[110] AU Abeysekara, R Alfaro, C Alvarez, JD Álvarez, R Arceo, JC Arteaga-Velázquez, HA Ayala Solares, AS Barber, BM Baughman, N Bautista-Elivar, et al. Data acquisition architecture and online processing system for the hawc gamma-ray observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 888:138–146, 2018.

[111] R Alfaro, C Alvarez, A Andrés, E Anita-Rangel, M Araya, JC Arteaga-Velázquez, D Avila Rojas, HA Solares, R Babu, P Bangale, et al. Hawc performance enhanced by machine learning in gamma-hadron separation. *arXiv preprint arXiv:2506.18277*, 2025.

[112] Alberto Franceschini, Giulia Rodighiero, and Mattia Vaccari. Extragalactic optical-infrared background radiation, its time evolution and the cosmic photon-photon opacity. *Astronomy & Astrophysics*, 487(3):837–852, 2008.

[113] Axel Donath, Régis Terrier, Quentin Remy, Atreyee Sinha, Cosimo Nigro, Fabio Pintore, Bruno Khélifi, Laura Olivera-Nieto, Jose Enrique Ruiz, Kai Brügge, et al. Gammapy: A python package for gamma-ray astronomy. *Astronomy & Astrophysics*, 678:A157, 2023.

[114] Elena Amato. The theory of pulsar wind nebulae: recent progress. *arXiv preprint arXiv:2001.04442*, 2020.

[115] S Groetsch. The fourth hawc catalog of very-high-energy gamma-ray sources. In *39th International Cosmic Ray Conference*, page 665, 2026.

[116] Andrea Zonca, Leo Singer, Daniel Lenz, Martin Reinecke, Cyrille Rosset, Eric Hivon, and Krzysztof Gorski. healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python. *Journal of Open Source Software*, 4(35):1298, March 2019.

[117] Krzysztof M Gorski, Eric Hivon, Anthony J Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759, 2005.

[118] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[119] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, 6 2014.

[120] Alison Peisker. *Transient Searches with the HAWC Gamma-Ray Observatory*. Michigan State University, 2021.

[121] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[122] Q Remy, YA Gallant, and Matthieu Renaud. Prospects on high-energy source searches based on pattern recognition object detection in the hess galactic plane survey and catalogue cross-matches. *Astroparticle Physics*, 122:102462, 2020.

[123] Rishi Babu, Palmer Wentworth, Ian Herzog, Dan Salazar, and Mehr Un Nisa. Improving gamma-ray source search with image processing. *arXiv preprint arXiv:2507.10307*, 2025.

[124] OpenAI. Chatgpt, 2025. Version: [Feb 18 version], [Large language model].

[125] Jason P Moore. Basic theory and operating principles of optical frequency domain reflectometry measurement systems as applied to fiber bragg grating sensors. Technical report, NASA, 2023.

[126] Sarah M Wagner, Paul R Burd, Daniela Dorner, Karl Mannheim, Sara Buson, Andrea Gokus, Greg Madejski, Jeffrey D Scargle, Axel Arbet-Engels, Dominik Baack, et al. Statistical properties of flux variations in blazar light curves at gev and tev energies. *arXiv preprint arXiv:2110.14797*, 2021.

APPENDIX A

SWGO EXTRA WEATHER ANALYSIS

These are the extra plots I made for the analysis that highlight the working hour conditions for the various sites.

For temperature during the working hours, the times considered for normal work hours are 8 am (08:00) to 5 pm (17:00), Monday through Friday and are seen in Figure A.1.

For the wind during working hours, the times considered for normal work hours are 8 am (08:00) to 5 pm (17:00), Monday through Friday and are seen in Figure A.2.

While the least important of the 4 parameters discussed, humidity data can still be used to infer the potential amount of rain each site might get. If the humidity is at or near 100 percent, it most likely is or just finished raining and, given the arid climates for all the sites, the average humidity should be low.

In general, there does not appear to be any significant issues with humidity and both construction and operation. While some correlation can be drawn between high humidity and precipitation, that is dubious at best and shouldn't be heavily weighted. One interesting point is that there may be high build-ups of static electricity at all sites (including HAWC) from the dry environment.

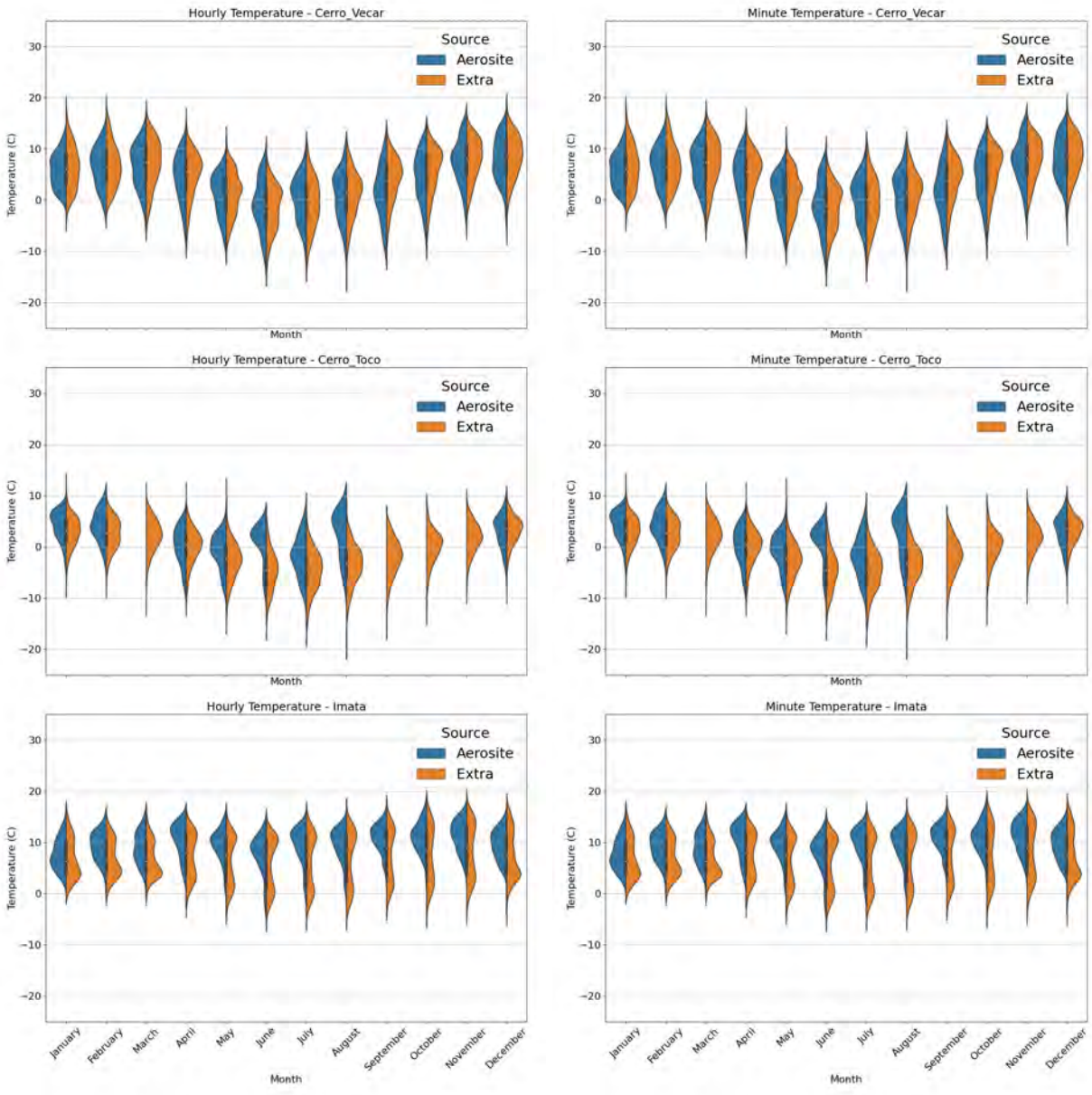

Figure A.1 Temperature spread for all 3 sites with hour and minute resolutions during the work hours. The left part of each violin is the AEROSITE data while the right has the extra data. Generally, both data sets observe comparable temperature distributions.

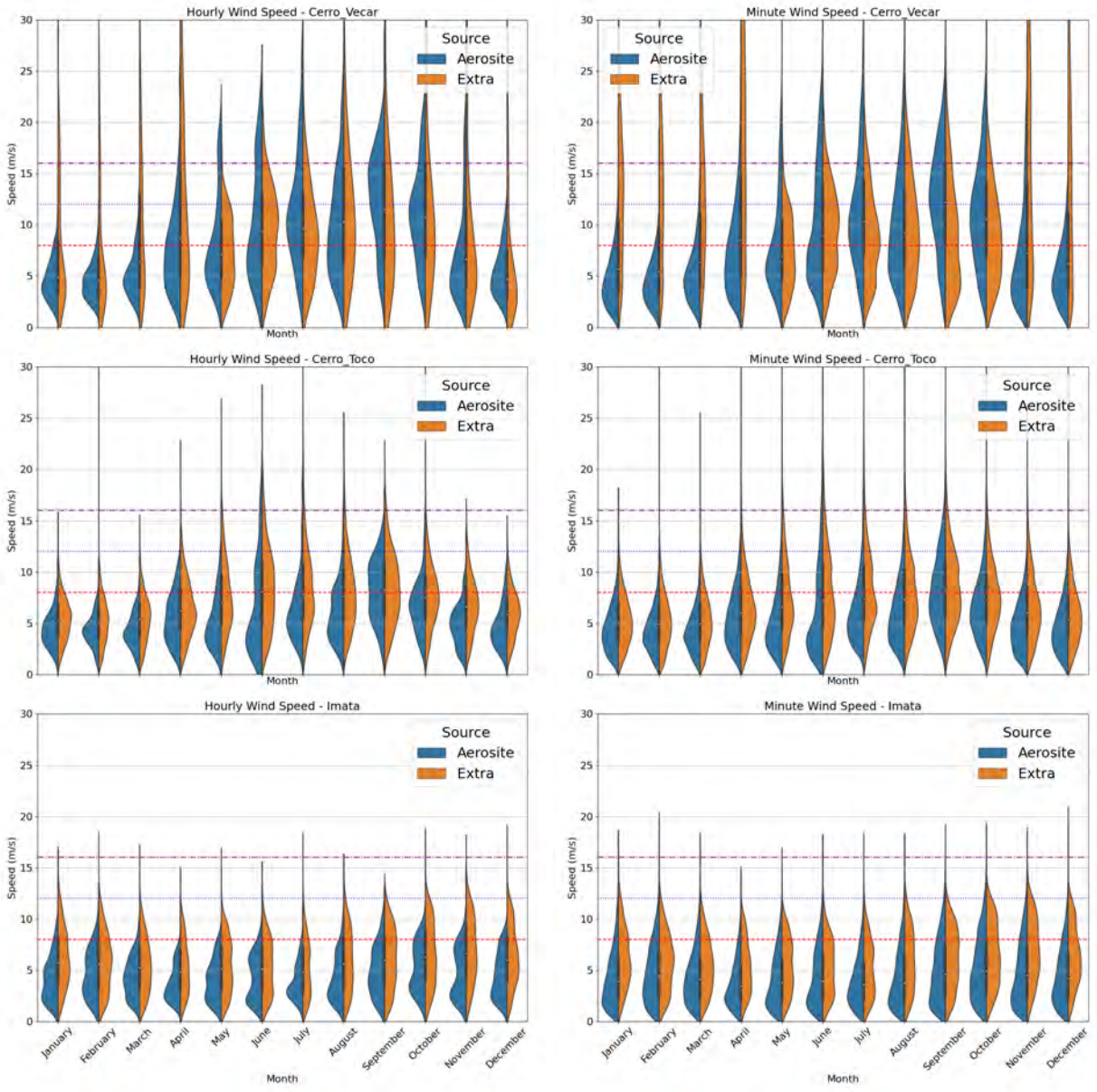

Figure A.2 Wind spread for all 3 sites with hour and minute resolutions during the work hours. The left part of each violin is the AEROSITE data while the right has the extra data. Generally, both data sets observe comparable wind speed distributions, though the extra data seems to have odd spikes that may skew their data. The lines indicate 8, 12, and 16 m/s winds respectively and were used to generate the values in the three site wind speed tables.

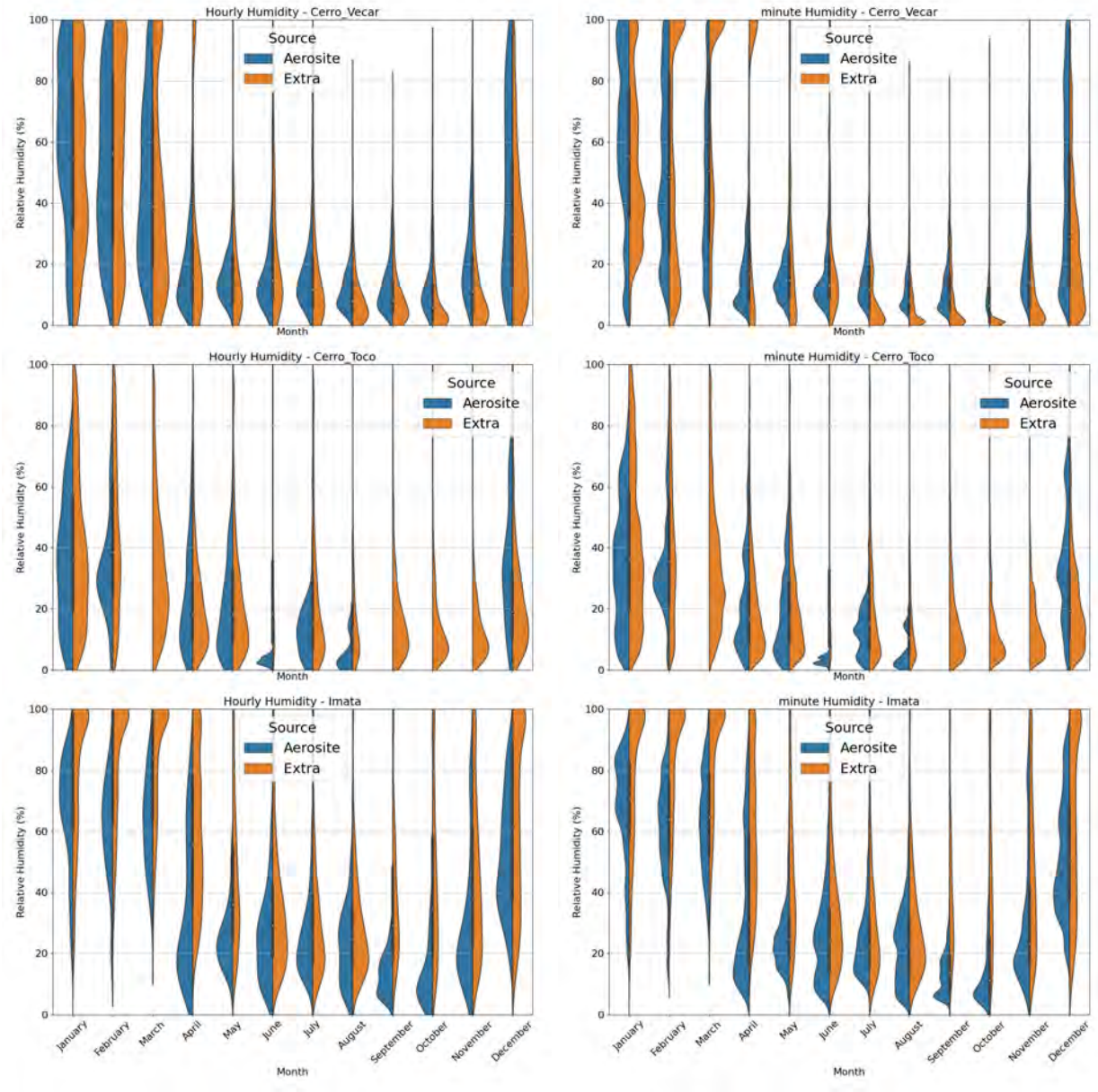

Figure A.3 Humidity spread for all 3 sites with hour and minute resolutions. The left part of each violin is the AEROSITE data while the right has the extra data. Generally, both data sets observe comparable humidities and all 3 sites follow a similar cycle, most likely due to the monsoon season.

APPENDIX B

THE FIFTH HAWC CATALOG: IMAGE ANALYSIS

During my analysis of HAWC J2031+415 (see Chapter 4, I helped developed the source search method outlined in Section 4.4. This method iteratively adds point sources to a map before testing each source's extension and spectral curvature. Once complete, it gives an accurate breakdown of the sources in a region, though at the cost of extremely long computation time in complex regions, sometimes taking over a month to finish adding point sources. Additionally, it has stability issues if there are sources at the edge of the ROI, where emission potentially bleeds outside the ROI. Regardless, this algorithm has been expanded to perform a full sky survey and is being used to create the fourth HAWC catalog (4HWC) [115].

While working on my joint HAWC-IceCube transient analysis (see Chapter 5), I worked with several people at Michigan State to help develop a new algorithm that improves the 4HWC with regards to both accuracy and computation time. The group members are the postdocs Dan Salazar and Rishi Babu, myself, and Palmer Wentworth, an undergraduate student. The project goal is to generate a seed model that an end-user can do perform follow-up studies with in a fraction of the time of the 4HWC method. Additionally, I presented this work at the International Cosmic Ray Conference (ICRC) 2025 in Geneva, Switzerland [123]. This chapter discusses the current status of the project, along with preliminary results. Lastly, this work based off the work from [122].

B.1 Introduction

As discussed in the HAWC J2031+415 analysis, a normal HAWC fit utilizes both the 3ML and HAL plugins to perform a joint-likelihood fit by fitting a given model (source location, morphology, spectral parameters, etc) to a specified dataset, and return a $-\log(\text{likelihood})$ value. A best fit model is determined when a global minimum value is found in the likelihood curve. This curve is constructed sequentially via a stochastic gradient descent to prevent the fitting process getting stuck in a local minimum. The computation time required scales heavily with the number of sources and free parameters, with a single point source with 4 free parameters (ra, dec, index, flux) taking around 10 minutes, but a larger model like the one in the HAWC J2031+415 with 3 extended sources

taking 3-6 hours, depending on the parameters being tested. The 4HWC method can introduce dozens of point sources before moving to the extension test, where most of these point sources will be dropped. This adds massive overhead to the process and can result in stability issues, with the fits frequently crashing and having to restart due to its complexity.

To this end, the method introduced here aims to significantly speed up the source search algorithm by utilizing an image process pipeline to parse through the data. It finds significance peaks and places seed sources for later testing. As there is no fitting directly included in the image processing, it can parse through data incredibly quickly compared the 4HWC method, taking minutes instead of days to recover similar models. Once a seed model is generated, the end user can do manual testing to either verify or edit the number and type of sources.

B.2 Detection pipeline

The first step of this pipeline is to generate a significance map using HAWC data. This is done using the Analysis and Event Reconstruction Integrated Environment (AERIE) framework [110] and can take in different source assumptions to test different hypothesizes. Generally, the default assumption is a point source PSF of 0.3 degrees and a 2.7 spectral index to match the expected diffuse background. Regardless, once the map is created, a large of ROI of order $5^\circ \times 5^\circ$ or larger is defined.

B.2.1 Image Pre-processing

Before the source search algorithm is run on the image, the significance map is pre-processed to remove large under-fluctuations. These regions occur in pixels where there is not enough background statistics or at the extreme edges of HAWC's field of view, and appear as large negative significance excesses. These artifacts can cause issues with the blob finding and source association steps (described below) and appear as fake sources. Therefore, a floor search is performed on the data and all significance $< -5\sigma$ is set to -5σ. An example of a floored significance map is given in the upper left figure in Figure B.1. This allows for negative significances to be present in the map without impeding the source searching algorithm.

Once the significance map has been floored, a search is done for any pixels with $\sigma > 5\sigma$. This

is HAWC's standard detection threshold and corresponds to a 99.98% chance of a source detection. If there are peaks found, the pipeline moves to the next step.

B.2.2 Image normalization

While the significance map has been floored, the ROI may still contain incredibly bright sources that may suppress fainter or more diffuse sources. This biases the algorithm to find only the brightest sources in the ROI. To this end, the significance inside the ROI is then normalized to range from [0,1], thus bringing out fainter sources and improving contrast. Additionally, this allows feature enhancement via a difference of Gaussians method (see next section). This is shown in the upper right figure in Figure B.1.

B.2.3 Feature Enhancement

The normalized map by itself is not sufficient to bring out underlying structures in a given region, nor does it properly account for HAWC's PSF. It should be noted that, while HAWC's PSF is incorporated when making the map with AERIE or fitting with 3ML, the significance map itself is made of pixels with width $\approx 0.05^\circ$ and neighboring pixels can have significantly different values. To this end, a Gaussian kernel with a smearing radius equal to HAWC's average PSF of 0.3° is applied to the map. This takes a given pixel and convolves it with neighboring pixels to smooth out sharp gradients in significance and enhance dimmer regions more. The smoothed map is shown in the lower left figure in Figure B.1.

To search for sources, a Difference of Gaussian (DoG) method is applied [121]. The normalized map has the smeared map subtracted from it (see lower right figure in Figure B.1). This removes large scale structures like the Diffuse Background Emission (DBE) and extremely large extended sources while preserving point and smaller extended sources. The remaining sources appear as sharp peaks in the pixel significance distribution. These peaks can be visualized in both the 1 and 3 dimensional plots shown in the left and right figures in Figure B.2, respectively. The 1D intensity histogram (green in LEFT Figure B.2) resembles a 0 mean Gaussian with significant positive and negative tails; both are due to the extremely bright emission present. Therefore, the distribution is fit with a Gaussian function (indicated by red) and the 3σ deviation is calculated. This will be used

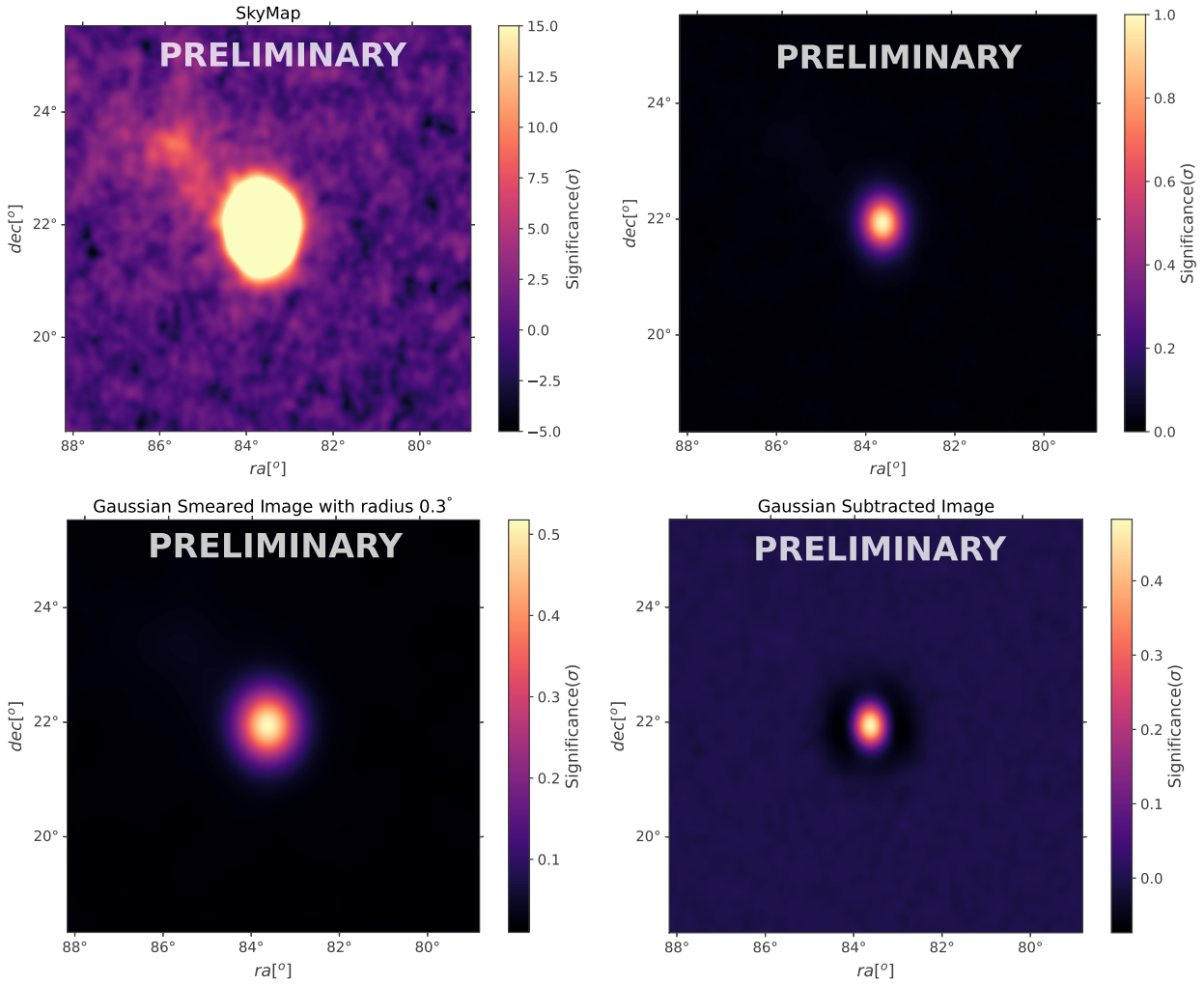

Figure B.1 An example of the DRIPS algorithm in action. First, the initial significance map is shown (upper left). Then, the map is normalized for processing (upper right). This normalized map is then smeared with HAWC's PSF (lower left) before subtracting from the initial normalized map (lower right).

in the intensity threshold search for point sources. A 3D version of the peak intensity distribution is shown on the right in RIGHT Figure B.2.

B.2.4 Peak Detection and Intensity Threshold

This analysis utilizes a blob detection algorithm that is based off the DoG method for blob searching in images (see [121] for more details) and scikit-image for the image processing [119]. The algorithm functions by first building an array of images of the DoG image with increasingly large smearing radii and taking the difference between consecutive images. This is done twice, first

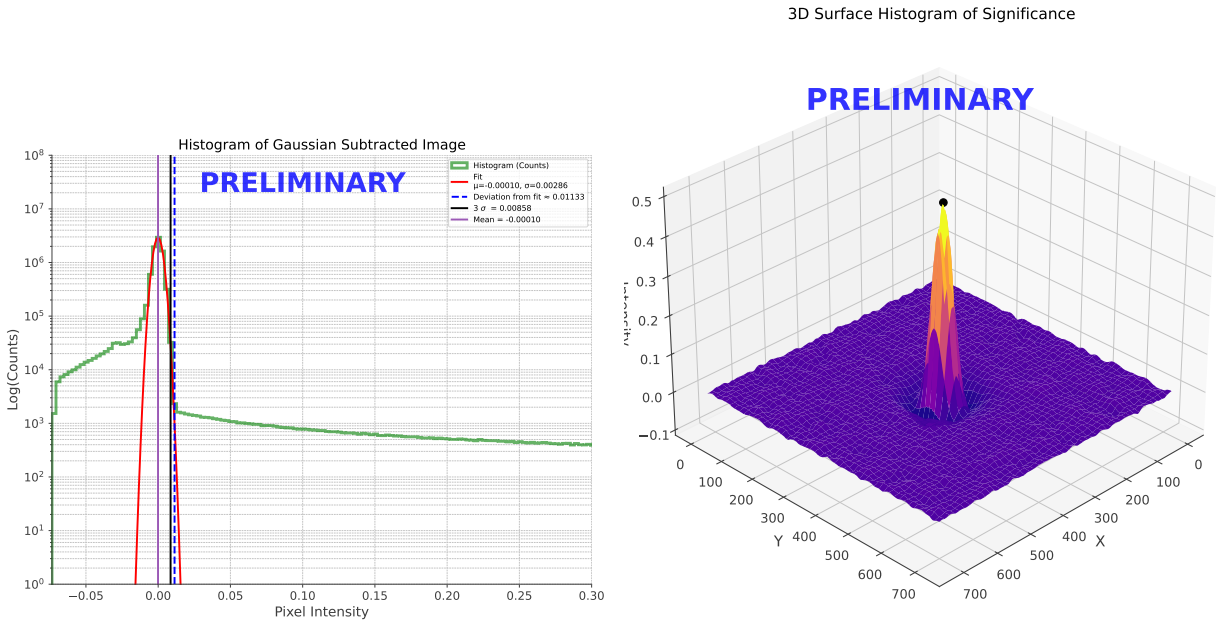

Figure B.2 LEFT: The histogram distribution of the pixel intensities from the subtracted emission region (lower right in Figure B.1 are shown in green. The red curve indicates the best fit result for a Gaussian model, with the solid purple line indicating the mean of the fit. If there is no signal present, the binned data should follow a Gaussian with a mean of 0 and a width of . The blue dashed line indicates where the fit begins deviating from the data, and the black line is the 3σ deviation of the Gaussian fit. RIGHT: This is the 3D pixel intensity distribution of the subtracted emission region. There is clearly a single peak present.

for point and small extended sources, and then for larger extended sources. For the small source search, two radius ranges are defined, one from [0.1, 0.3] for point sources and [0.3, 0.5] for small extended sources.

The difference comparison for all images. are then stored in a separate 3D array. Next, the pixels of each image in the 3D array is scanned to search for local extrema (maxima/minima). For each pixel scanned, a secondary scan is run on the neighboring 8 pixels to determine if it is a local extrema. If it is, that pixel is then compared to its corresponding pixel in all other images in the 3D array. If it is an extrema across all images, it is tagged as a blob and the x,y coordinates are saved, along with how wide the peak's distribution is.

The DoG blob search algorithm returns many potential sources, both real and fake, that must be categorized and, if necessary, rejected. To this end, several filtering steps are taken. First, any blobs within 1° of the ROI are removed as there is risk of potential emission bleeding either in or out of the

ROI from neighboring sources. This will also help reduce duplicate sources from appearing during a catalog search. Next, an internal limit on blobs that do not reach the 3σ threshold calculated in LEFT Figure B.2. Then the remaining blobs then have their pixels' corresponding peak significance in the initial significance map checked. If the pixel's σ is less than 5, the blob is dropped as it falls below HAWC's threshold for new sources [37]. Lastly, an overlap test is conducted between the point and small extended sources. If a source is found in both maps, their respective extensions are checked. If the extension is equal to 0.3°, then the extended source is dropped.

The process is broadly the same with the large extended source search, except for the interval ranges checked and the seed map used. The interval range considered is between [0.4,0.8], and then [0.8, 1.5] to differentiate between medium and large extended sources. Like with the point and small extended source check, if there a source is found in both maps and has an extension equal to the limit (0.8°), the larger source is dropped. As for the map, a new normalized map is created. This emphasizes

With all seed sources collected, the last step is to remove any overlapping sources. Considering the extended sources first; if there are two extended sources (one from the point and larger extended source searches) co-located, the larger one is dropped. In a similar manner, if a point and extended source are within 0.5° of each other, the point source is kept while the extended source is dropped. The point source is then assigned a new parameter P_{ext}, which is the probability that point source is extended. It is weighted linearly, with $P_{ext} = 1$ if the point and extended source locations were coincident and $P_{ext} \rightarrow 0$ for distances near 0.5°.

B.3 Source Fitting and Association

Once the image analysis pipeline is complete, a seed model is created. This takes in the sources identified in the pipeline and creates corresponding point and extend source models for them (see Chapter 4 for model definitions) with simple power laws. All parameters are free (location, flux normalization, index) and the final result of that fit constitutes the toy model for the given ROI. In the example above, the ROI considered is the Crab Nebula and the image analysis successfully identifies both the Crab and Crab on Fire, an extended source that has recently become visible to

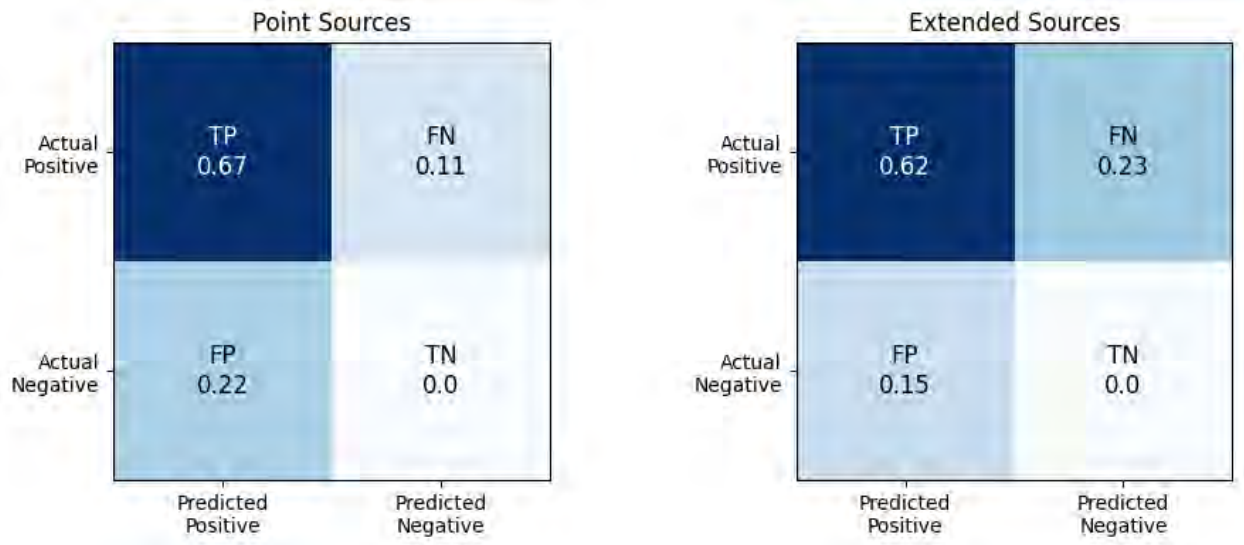

Figure B.3 The confusion matrices for point and extended sources from the simulation studies. DRIPS achieves 67% and 62% detection rates for point and extended sources.

HAWC. The final fit result from this algorithm can then be taken and new source hypotheses like new spectral models can be tested. For example, the Crab is well-known to have a log-parabola spectrum instead of a power law [96] and an end user can easily test this.

B.3.1 Sensitivity Studies

To test the reliability of the algorithm, it is calibrated against 250 randomly injected models that span across the sky and have randomly designed source parameters. Each model contains at least 1 source that can be either point or extended, have an index between -1.5 and -3.5, and a flux between 0.1 to 100 percent of the Crab flux. This is the same test suite that was used in testing the 4HWC [115] and can serve as a useful metric for success.

Each simulated region is run through the algorithm, with the resulting found sources being compared to the initial injected model file. Point vs extended sources are compared separately. For a point source, a true positive detection is if an injected source is within half a degree of a found source. For an extended source, it is a true positive if a found source lies within the extension of an injected source. Each found source can only be matched with one injected source, they are both then removed from consideration for follow-up source comparisons. False positives and negatives

are defined the same for both source classes: false negative is where the DRIPS algorithm missed an injected source and false positives are where DRIPS found a source that was not injected. True negatives are set to zero by convention to match the 4HWC method [115].

The results are shown in Figure B.3. Currently, DRIPS achieves a 67% detection rate of point sources and 62% for extended sources. For point sources, there is a significant rate of false positives at 22%. This is primarily correlated with the high false negative rate of 23% for extended sources as dimmer extended sources can be mistaken for one or several point sources, depending on the injected source's size and flux. Work is currently ongoing to tweak the sensitivity of the algorithm to handle large dim sources, so this is expected to improve with coming revisions.

One additional note is the inconsistency of the injected models. Due to their randomly generated nature, there are several cases of injected models not achieving 5σ in HAWC's significance maps. These then are ignored by the algorithm as one step is dropping any found sources with $\sigma < 5\sigma$. This causes a large false negative rate, especially with extended sources, and is the primary reason why they are often min-categorized as point sources. Small background fluctuations can add to the injected source to make small regions of 5σ, thus keeping them in the final result.

I am currently creating a new batch of simulations with more physical sources to better translate the simulation studies to what HAWC sees. The algorithm is proven to detect bright ($> 10\sigma$) sources, so this will focus on generating sources between 5σ and 15σ. The final maps will have two to four sources, with a combination of point and extended sources. The index, flux, and extension ranges are all still in discussion, but preliminary ranges are as follows:

- point source:
  - index: 1.8 - 3.1
  - flux: 0.01 - 0.1 Crab flux
  - small extended source
    - index: 1.8 - 3.1
    - flux: 0.03 - 0.1 Crab flux

        * extension: 0.1 - 0.5

    - large extended source

        * index: 1.8 - 2.5

        * flux: 0.08 - 0.12 Crab flux

        * extension: 0.5 - 1.0

These represent a more accurate spread of current HAWC sources and should reduce the number of low significance sources that normal HAWC fitting would reject.

B.4 Conclusions

Going forward, the main goal of this algorithm is to construct a catalog search method to produce the next HAWC catalog. It presents a rapid and decently accurate method compared to the long 4HWC process, and recovers a sizable portion of known HAWC sources. Furthermore, it is planned to be publicly available with a HAWC data set. More refinement is needed, but this is an extremely impactful project and will be a valued asset to the collaboration when complete.