

1 Measuring High-Energy Spectra with HAWC

S. S. Marinelli^a and J. A. Goodman^{*b} for the HAWC Collaboration[†]

^aMichigan State University, United States

Email: marine20@msu.edu

^bUniversity of Maryland, United States

Email: goodman@umdgrb.umd.edu

The High-Altitude Water-Cherenkov (HAWC) experiment is a TeV γ -ray observatory located 4100 m above sea level on the Sierra Negra mountain in Puebla, Mexico. The detector consists of 300 water-filled tanks, each instrumented with 4 photomultiplier tubes that utilize the water-Cherenkov technique to detect atmospheric air showers produced by cosmic γ rays. Construction of HAWC was completed in March of 2015. The experiment's wide instantaneous field of view (2 sr) and high duty cycle (> 95%) make it a powerful survey instrument sensitive to pulsars, supernova remnants, and other γ -ray sources. The mechanisms of particle acceleration at these sources can be studied by analyzing their high-energy spectra. To this end, we have developed an event-by-event energy-reconstruction algorithm using an artificial neural network to estimate energies of primary γ rays at HAWC. We will present the details of this technique and its performance as well as the current progress toward using it to measure energy spectra of γ -ray sources.

*International Cosmic-Ray Conference
Bexco, South Korea
July 12–20, 2017*

*Speaker.

[†]A complete list of authors is available at <http://www.hawc-observatory.org/collaboration/icrc2017.php>.

1. Introduction

Among the fundamental questions in particle astrophysics are those of the sources and acceleration mechanisms of the high-energy particles that are observed to arrive at the Earth from space. Theories postulating different acceleration mechanisms have implications for the distributions of energies of particles emitted at these sources and therefore can be constrained by measurements of these sources' emitted energy spectra. Making such measurements requires the precise reconstruction of particle energies by observatories sensitive to these particles.

The High-Altitude Water-Cherenkov (HAWC) detector observes TeV γ rays from an altitude of 4100 m in the state of Puebla, Mexico. The experiment's 300 water-filled tanks are each instrumented with 4 photomultiplier tubes (PMTs) that detect the Cherenkov light produced by γ -ray-induced extensive air showers as they pass through the tanks. Particle arrival times and deposited charges are reconstructed from PMT data, allowing estimation of various shower parameters such as zenith and azimuth angles and primary energy. A detailed description of the detector's electronics and data-acquisition system can be found at [1]. By reconstructing primary energies with sufficient precision, HAWC can constrain the spectral energy distributions (SEDs) of γ -ray sources, thereby probing their mechanisms of particle acceleration. This paper introduces a new algorithm for reconstructing the energies of HAWC γ -ray events and establishes the performance of this technique via Monte Carlo (MC) simulations and on a calibration source, the Crab Nebula.

2. Energy-reconstruction technique

The algorithm described herein uses an artificial neural network (NN) to reconstruct energies of photons detected by HAWC. It was implemented using the Toolkit for Multivariate Analysis [2] NN implementation.

An NN is a complicated function mapping several quantities associated with an event (input variables) to some regression target or output variable, in this case $\log_{10} E$ where E is the primary energy of the shower. A detailed description of the functional form of the NN can be found in [2]. This function is characterized by many (479 in this implementation) free parameters called weights. The optimal values of the weights are determined in a process called training, in which the weights are optimized using an MC simulation of the detector. The training process is described in Section 2.2.

The energy algorithm consists of two NNs: one for low-multiplicity events and one for high-multiplicity events. This binning has been found to yield better performance than a single NN even though multiplicity variables are included among the inputs.

2.1 Input variables

Fifteen input variables are used to characterize the shower energy. These variables have been chosen in order to capture three qualities of the shower: the particle multiplicity in the detector, the fraction of the shower which landed outside of the detector, and the atmospheric attenuation of the shower.

The multiplicity in the detector is a crude measure of how much energy has been deposited in it and is quantified using the fraction of PMTs hit in the event, the fraction of tanks hit, and the

41 normalization factor (in logarithm) of the fit of the lateral distribution of the shower. This fit is
 42 described in more detail at [3].

43 The distance between the reconstructed shower-core location and the center of the HAWC
 44 array is included as an input variable in order to provide information about how much of the shower
 45 is not contained within the array.

46 Finally, the distance through which the shower has propagated in the atmosphere is quantified
 47 in two ways: using the reconstructed zenith angle and the lateral distribution of particles in the
 48 shower, which contains information about the shower’s age. The latter is provided to the NN
 49 in the form of ten input variables, the i th of which is the fraction of charge deposition in the
 50 event occurring in PMTs whose distances from the shower axis, measured in the shower plane, are
 51 between $(10\text{ m})i$ and $(10\text{ m})(i + 1)$ for i between 0 and 8 inclusively; the last variable is the fraction
 52 of charge deposition in PMTs more than 90 m from the axis.

53 2.2 Training

54 The goal of the training process is to choose the vector of NN weights \mathbf{w} that minimizes the
 55 error function, defined as

$$D(\mathbf{w}) \equiv \frac{1}{2} \sum_{i=1}^{N_{\text{events}}} u_i [\log_{10} \hat{E}(\mathbf{x}_i; \mathbf{w}) - \log_{10} E_i]^2 \quad (2.1)$$

56 where u_i is the weight of the i th MC event (the event’s relative importance in the training, unrelated
 57 to the NN parameters \mathbf{w} , which are unfortunately also called weights), E_i is the true energy of
 58 the i th event, \mathbf{x}_i is the i th event’s vector of input variables, and \hat{E} is the energy estimate for the
 59 event. This minimization is performed by repeatedly iterating over the sample of training events,
 60 with a small update to \mathbf{w} performed for each event. The details of the Broyden-Fletcher-Goldbarb-
 61 Shannon algorithm for updating the weights can be found in [2]. The MC events are weighted (the
 62 u_i are set) to resemble an E^{-2} power-law spectrum during the training. This weighting was chosen
 63 because it produces a relatively flat RMS error between 1 and 100 TeV (see Section 3.2).

64 3. Performance on Monte Carlo

65 Several other energy-reconstruction algorithms have been developed for the HAWC obser-
 66 vatory. These are described briefly in Section 3.1, and the performance of all energy variables
 67 including the NN is described in Section 3.2.

68 3.1 Other energy variables

69 The energy proxy that is presently used in the standard HAWC γ -ray analysis is the fraction
 70 of PMTs hit during the event, referred to as “fraction hit” or f_{hit} [3].

71 A likelihood-based algorithm for energy reconstruction has also been developed. This algo-
 72 rithm uses a likelihood function built using MC that takes into account the locations, arrival times,
 73 and deposited charges of the PMT hits that occur during the event. The hits are assumed to be
 74 statistically independent so that the likelihood is built by combining many single-hit likelihood
 75 functions. This algorithm has also been adapted for cosmic-ray energy reconstruction and is de-
 76 scribed in [4].

77 A third algorithm, called the “ground parameter” (GP), is an estimation of the photon primary
 78 energy as a function of reconstructed zenith angle and the value of the lateral distribution function
 79 at some optimal distance from the estimated shower axis. This is described in greater detail in [5].

80 3.2 Performance comparison

81 The performance of these energy-reconstruction techniques is evaluated using MC simula-
 82 tions of γ -ray showers interacting with the HAWC detector. These simulations are described in
 83 [3]. The MC events are weighted to simulate a point source transiting at 20° declination with an
 84 $E^{-2.63}$ power-law energy spectrum. Analysis cuts are applied, including an f_{hit} cut at 0.044 and a
 85 requirement that the reconstructed core be within the HAWC array. Background-rejection cuts are
 86 also applied in order to simulate a signal-event population resembling what would be used in an
 87 analysis of actual data.

88 Figure 1a shows the distribution of fraction hit as a function of true energy. Because higher-
 89 energy photons on average produce higher-multiplicity showers, fraction hit is positively correlated
 90 with primary energy. However, several factors prevent it from determining the energy precisely.
 91 Because showers are attenuated as they travel through the atmosphere, a shower coming from a
 92 higher angle from zenith or having its first interaction higher in the atmosphere will have a lower
 93 multiplicity than a shower of the same primary energy coming from a lower angle from zenith or
 94 interacting lower in the atmosphere. In addition, a shower whose core is not well centered on the
 95 detector will have a lower multiplicity in the data than one that is, again assuming equal primary
 96 energies. Furthermore, because fraction hit cannot exceed 1, at sufficiently high energies, the
 97 variable saturates and loses sensitivity to the energy.

98 Figure 1b shows the relationship between the NN variable and true energy. Because the NN
 99 includes input variables that contain information about the containment of the shower within the
 100 array and its atmospheric attenuation, it is able to achieve a better resolution than that of fraction
 101 hit. Since the NN takes into account information beyond just the particle multiplicity of the shower,
 102 it does not saturate within the range of energies thrown in the MC, which extends well beyond
 103 100 TeV.

104 Figure 2 shows the RMS error in log space of the three explicit energy estimates: the NN
 105 estimate, the likelihood estimate, and the GP estimate. The RMS error is defined as

$$\rho \equiv \sqrt{\langle (\log_{10} \hat{E} - \log_{10} E)^2 \rangle}. \quad (3.1)$$

106 Since fraction hit is not an explicit energy estimate, its \hat{E} is computed by choosing the energy that
 107 maximizes the likelihood of the observed f_{hit} for the event. The NN displays the best RMS error in
 108 all bins. The RMS error in the 46 to 68 TeV bin is around 0.12, indicating one tenth of one decade
 109 in energy, or about 32% in linear energy space. All of the reconstruction algorithms perform poorly
 110 below around 1 TeV. This is related to the f_{hit} cut: a substantial fraction of events below 1 TeV fail
 111 this cut, and those that pass it represent upwards multiplicity fluctuations. Thus their energies tend
 112 to be overestimated.

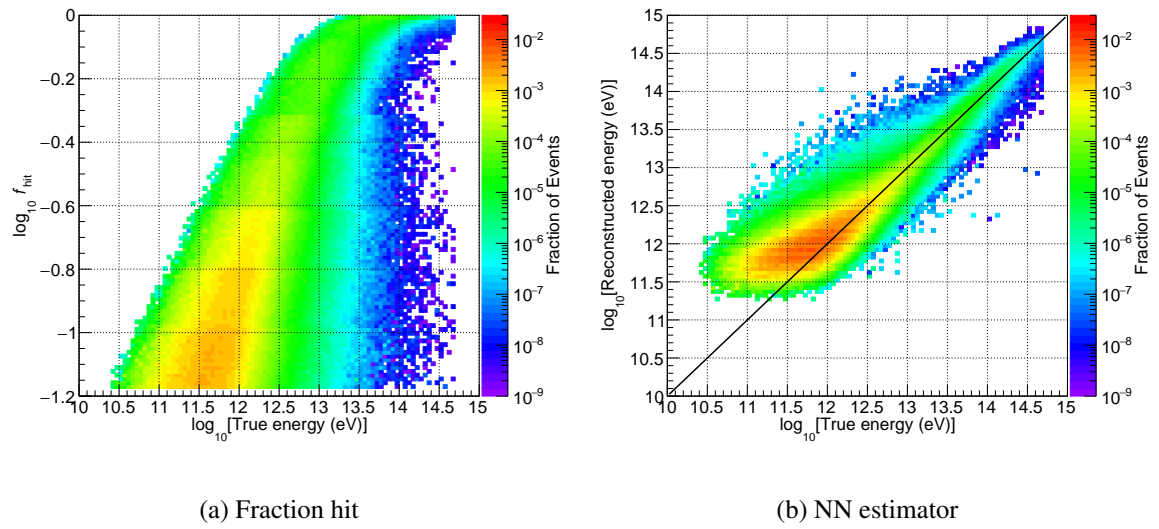


Figure 1: Joint distribution of each of two energy variables with true energy in MC.

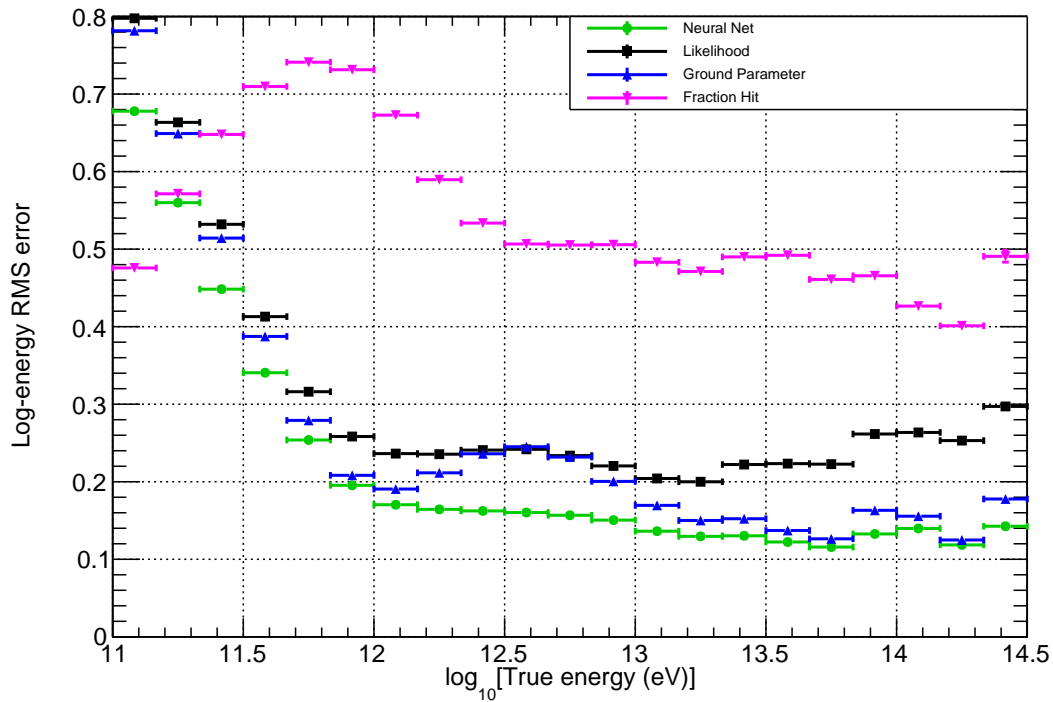


Figure 2: RMS error in log space of the NN, likelihood, and GP energy variables as a function of true energy in MC.

Parameter	Unit	Published HAWC		NN		GP	
		Value	Error	Value	Error	Value	Error
Φ_0	$10^{-13} (\text{cm}^2 \text{ s TeV})^{-1}$	2.51	$^{+0.12}_{-0.11}$	2.93	0.09	3.22	0.10
α		2.63	0.03	2.708	0.030	2.751	0.031
β		0.15	0.03	0.162	0.021	0.179	0.023

Table 1: Best-fit values and statistical uncertainties for log-parabola spectral parameters.

113 4. Calibration source: the Crab Nebula

114 In order to verify the ability of the new energy variables to accurately measure γ -ray energy
 115 spectra, a spectral fit of the Crab Nebula has been performed using the NN and GP variables. (The
 116 Likelihood energy variable is not currently used in γ -ray analyses.) This was carried out via the
 117 Multi-Mission Maximum Likelihood framework, described in [6], which interacts with HAWC’s
 118 Analysis and Event-Reconstruction Integrated Environment to perform a forward-folded fit of the
 119 true-energy distribution to the observed energy-estimate distribution, taking into account the joint
 120 distribution in Figure 1b. The Crab is modeled as a point source, with the expected distribution of
 121 reconstructed photon directions being the detector’s point-spread function centered on the source’s
 122 location. Its energy spectrum is assumed to be log-parabolic,

$$\frac{dN}{dE} = \Phi_0 \left(\frac{E}{E_0} \right)^{-\alpha - \beta \ln(E/E_0)}, \quad (4.1)$$

123 where dN/dE is the photon particle flux, Φ_0 , α , and β are free parameters, and the pivot en-
 124 ergy E_0 is chosen to be 7 TeV to decorrelate the estimates of Φ_0 and α . Events are sorted into
 125 two-dimensional bins of estimated energy and fraction hit in the analysis; fraction hit is included
 126 because it parameterizes the detector’s angular resolution better than any of the explicit energy
 127 estimates. A skymap is constructed for each bin, and a pixel-by-pixel fit is performed to find the
 128 log-parabola spectrum most compatible with the data. The result is plotted in Figure 3 along with
 129 H.E.S.S.’s reconstructed Crab spectrum from [7], and the fit spectral parameters’ values are given
 130 in Table 1 along with those from the published HAWC Crab spectrum in [3].

131 Table 1 shows that the NN spectral fit is better able to constrain the Φ_0 and β parameters than
 132 are the fit using only f_{hit} and the GP fit. The spectral index α is constrained comparably by all
 133 three analyses.

134 The SEDs computed using the two energy variables differ systematically from each other and
 135 from the H.E.S.S. result at some energies. An analysis of systematic errors has not yet been car-
 136 ried out and will be necessary in the future in order to understand this discrepancy. In particular
 137 the forward-folding technique is vulnerable to any differences in the distributions of the energy
 138 variables between data and MC, and such differences could result in the observed systematic dis-
 139 crepancy in the computed SEDs. The fractional systematic uncertainty on the published HAWC
 140 Crab spectral measurement is 50% [3], which, taken as a crude estimate of the systematic error on
 141 this analysis, is sufficient to account for the difference between the SEDs measured using the two
 142 new energy variables.

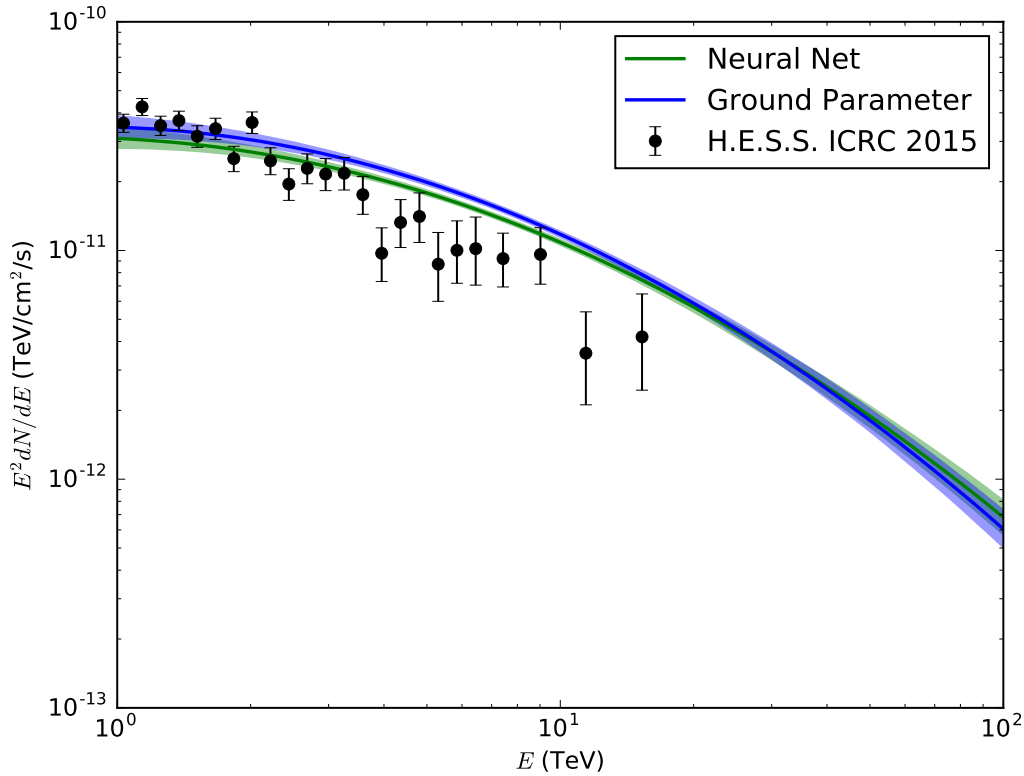


Figure 3: Preliminary fits of the Crab Nebula photon flux using the NN and GP energy variables. Shown errors on these fits are statistical only.

143 5. Conclusions

144 The development of the NN energy-reconstruction method allows HAWC to resolve energies
 145 up to 100 TeV. The technique has already been put into use in measuring energy spectra of TeV
 146 blazars; see [8]. Using the NN method, HAWC will be able to measure energy spectra of gamma-
 147 ray sources up to unprecedentedly high energies, providing a new window into the high-energy
 148 universe.

149 References

- 150 [1] A. U. Abeysekara et al. Search for TeV Gamma-Ray Emission from Point-like Sources in the Inner
 151 Galactic Plane with a Partial Configuration of the HAWC Observatory. *The Astrophysical Journal*,
 152 817(1):3, 2016.
- 153 [2] Andreas Hoecker et al. TMVA: Toolkit for Multivariate Data Analysis. *PoS, ACAT:040*, 2007.
- 154 [3] A. U. Abeysekara et al. Observation of the Crab Nebula with the HAWC Gamma-Ray Observatory.
 155 2017.

- 156 [4] J. C. Arteaga-Velázquez et al. Estimate of the energy spectrum of the light component of cosmic rays
157 in HAWC using the shower age and the fraction of hit PMT's. *Proceedings of the International*
158 *Cosmic-Ray Conference*, 2017.
- 159 [5] Kelly Malone. The gamma-ray sky above 50 TeV with the HAWC Observatory. April Meeting of the
160 American Physical Society, January 2017.
- 161 [6] G. Vianello et al. The Multi-Mission Maximum Likelihood framework (3ML). *Proceedings of the*
162 *International Cosmic-Ray Conference*, 2015.
- 163 [7] M. Holler et al. Observations of the Crab Nebula with H.E.S.S. Phase II. *Proceedings of the*
164 *International Cosmic-Ray Conference*, 2015.
- 165 [8] S. Coutiño de León et al. Spectral analysis of Markarian 421 and Markarian 501 with HAWC.
166 *Proceedings of the International Cosmic-Ray Conference*, 2017.