

New gamma/hadron separation parameters for a neural network for HAWC

E. Bourbeau^a, T. Capistrán^{*b}, I. Torres^b, E. Moreno^c for the HAWC Collaboration^d

^aMcGill University, 45 Sherbrooke Street West, Montreal, Quebec, Canada

^bInstituto Nacional de Astrofísica, Óptica y Electrónica, Luis Enrique Erro 1, Tonantzintla, Puebla 72840, México

^cFacultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Ciudad Universitaria, Colonia San Manuel Puebla, México

^dFor a complete author list, see <http://www.hawc-observatory.org/collaboration/icrc2017.php>.

Email: etienne.bourbeau@gmail.com, tcapistran@inaoep.mx, ibrahim@inaoep.mx, emoreno@fcfm.buap.mx

The High-Altitude Water Cherenkov experiment (HAWC) observatory is located 4100 meters above sea level. HAWC is able to detect secondary particles from extensive air showers (EAS) initiated in the interaction of a primary particle (either a gamma or a charged cosmic ray) with the upper atmosphere. Because an overwhelming majority of EAS events are triggered by cosmic rays, background noise suppression plays an important role in the data analysis process of the HAWC observatory. Currently, HAWC uses cuts on two parameters (whose values depend on the spatial distribution and luminosity of an event) to separate gamma-ray events from background hadronic showers. In this work, a search for additional gamma-hadron separation parameters was conducted to improve the efficiency of the HAWC background suppression technique. The best-performing parameters were integrated to a feed-forward Multilayer Perceptron Neural Network (MLP-NN), along with the traditional parameters. Various iterations of MLP-NN's were trained on Monte Carlo data, and tested on Crab data. Preliminary results show that the addition of new parameters can improve the significance of the point source at high-energies (\sim TeV), at the expense of slightly worse performance in conventional low-energy bins (\sim GeV). Further work is underway to improve the efficiency of the neural network at low energies.

*35th International Cosmic Ray Conference — ICRC2017
10–20 July, 2017
Bexco, Busan, Korea*

*Speaker.

1. Introduction

The High-Altitude Water Cherenkov, HAWC, is a gamma-ray observatory that is able to detect primary particles in the energy range from 500 GeV to 100 TeV. HAWC is composed of 300 water Cherenkov detector (WCDs), each instrumented with 4 photomultiplier tubes (PMTs) that detect the Cherenkov light from charged particles in extensive air showers. Given that 99 % of the events detected by HAWC come from cosmic-ray air showers, the study of gamma-ray sources requires the development of efficient techniques to achieve gamma-hadron separation [1, 2, 3].

Currently, HAWC uses two parameters (compactness “ \mathcal{C} ” and PINCness “ \mathcal{P} ”) that help distinguish between gamma events and hadron events, this way of identifying particles is called standard cut (for more details see the section 2.6 of [3]). Both parameters depend upon the spatial distribution of the charge deposited by the secondary particles of the shower inside of the array. All events are separated into bins (\mathcal{B}), that depend on the fraction of available PMTs, f_{hit} , that record light during the event. The events satisfying certain optimized cut values on \mathcal{P} and \mathcal{C} are then considered photon candidates (Table 1 shows the optimal cut of each bin). In this work, we used neural networks to improve the gamma-hadron separation, in addition, we investigated the performance of the network when \mathcal{C} , \mathcal{P} are used and when including new parameters in the inputs that could help distinguish between the two types of showers.

\mathcal{B}	f_{hit}	\mathcal{P} maximum	\mathcal{C} minimum
0	0.044 - 0.067	<1.4	>16
1	0.067 - 0.102	2.2	7.0
2	0.105 - 0.162	3.0	9.0
3	0.162 - 0.247	2.3	11.0
4	0.247 - 0.356	1.9	15.0
5	0.356 - 0.485	1.9	18.0
6	0.485 - 0.618	1.7	17.0
7	0.618 - 0.740	1.8	15.0
8	0.740 - 0.840	1.8	15.0
9	0.840 - 1.000	1.6	3.0

Table 1: The definition of the size bin \mathcal{B} is given by the fraction of available PMTs, (f_{hit}), that record light during the event. Each bin has its own optimal cut on Compactnes (\mathcal{C}) and PINCness (\mathcal{P}). Events with a \mathcal{P} less than indicated and a \mathcal{C} greater than indicated are considered photon candidates [3].

2. Training data set

The MC events were used to train the networks. Two types of particles were used: gamma and proton events, and they were simulated with a HAWC-250 configuration, that is, only 83 % of the array was activated. We used protons to simulate cosmic rays because they constitute nearly 99 % of the cosmic ray intensity. The conditions for selecting the events were:

- The event is well reconstructed, that is, the core and plane fit have been correctly computed.
- The event falls within HAWC.

In addition, a uniform distribution of nHits¹ (an energy proxy) was chosen to better sample the high-energy range. Because if the natural distribution ($E^{-2.7}$) is used [5], the network will have more experience to recognize events with low energy and have a poor performance at high energy.

2.1 Finding new parameters

The performance of neural networks is heavily dependent on the input parameters used [6]. Therefore, several parameters were analyzed. For the first attempt, the parameters that are in the HAWC data stream, were investigated. For the next attempt, new parameters were computed based on the timing information of the events. The following parameters are some of the ones that were evaluated for their G/H separation potential:

- Parameters that belong to the HAWC data stream
 - `planefitChi2`: The χ^2 of a Gaussian plane fit to the shower front.
 - **SFCFChi2**: The χ^2 of a faster EAS core fit.
 - **logmaxPE**: The PMT that has the maximum Photo-Electron (PE) charge during the event.
 - `logNPE`: The logarithm of the total PE charge calculated in an event.
- New time-based parameters
 - `deltaTime`: It is the time difference between the trigger time of the first PMT and the last PMT that are triggered during the event
 - `TimeVariance`: It is a simple measure of the scatter in PMT trigger time for a given event.

The parameters with bold font (SFCFChi2 and logmaxPE) are used in this analysis because of their different profiles for gamma and protons, for example, the histogram of SFCFChi2 (Figure 1). The gamma histogram is skewed left, whereas the hadron histogram has a more symmetric shape. Therefore they can be recognised with a cut. We don't use the other parameters due to their histograms having the similar pattern cannot be separated with a cut like, such as `deltaTime` parameter (Figure 2).

3. Training stage

In the training stage, HAWC simulations were used (see section 2) for training the Neural Networks (NNs) to identify between primary particles. The NN used is a Multilayer Perceptron [7]. We defined a NN output value of 1 for gamma-like events, and 0 for hadron events: that is, the network must produce a value of 1 if the event is a gamma and 0 if it is a hadron (Figure 3b shows

¹nHit is the number of PMTs activated during the event.

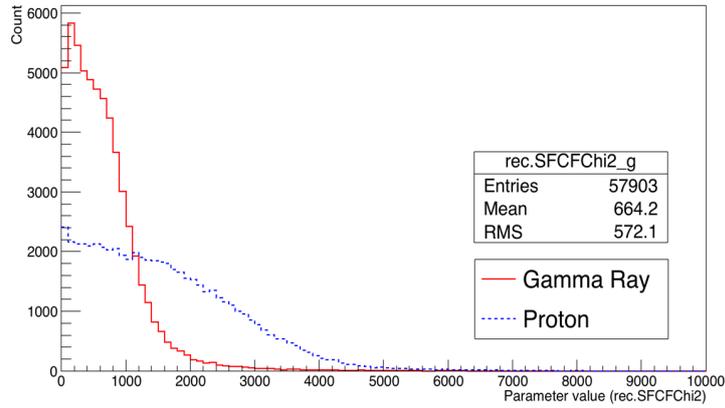


Figure 1: The histograms of the χ^2 of the core fit (SFCFChi2) using MC data. It is a potential parameter to use as input of the network for distinguishing between gammas and hadrons events since they have different shape and the histogram of gamma events is skewed right.

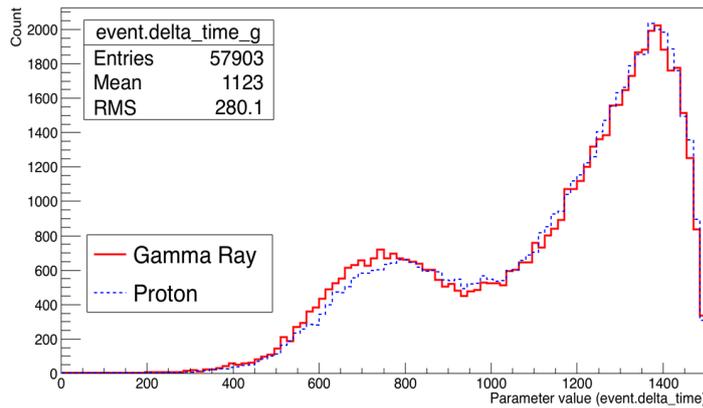


Figure 2: The histograms of the deltatime for gamma and proton induced EAS using MC data, that is the time between first particle and last particle detected in the event. As both histograms have the same pattern (skewed left), this parameter is not useful as an input to the network.

an example of the output neural network). In order to identify the type of the events, a cut is used, in other words, if the output of the network is equal or greater than the cut value the event is considered like gamma event, in another case, it is a hadron event. The optimal cut is when it has an excellent hadron rejection but an acceptable gamma efficiency, it is when the Q factor has the maximum value, where is defined as TPR / \sqrt{FPR} , where TPR is the fraction of correctly classified gamma events, also called gamma efficiency, and FPR is the fraction of badly classified hadron events (hadron rejection).

Before adding the new parameters described in section 2.1 to the network, one hundred training iterations of a simple 2-parameter network (Compactness and PINCness) were made, to see if the resulting optimal cut value for distinguishing between particles was sensitive to the random seed

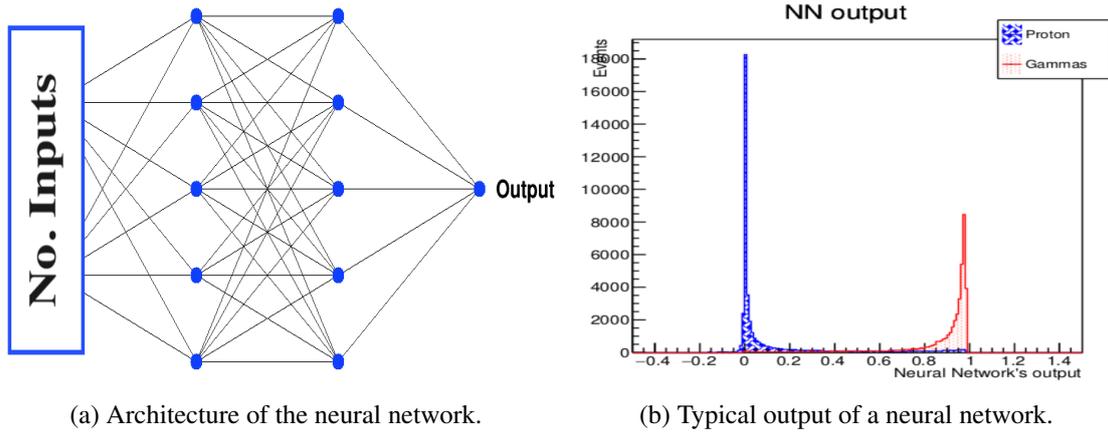


Figure 3: In (a) shows the architecture of the network. The number of neurons in the first layer depends on the number of inputs that the network uses. In (b) is shown the histogram of the output of the NN2 after it was trained.

values given to the NN weights. The histogram on figure 4a shows the optimal cut values obtained from these trainings. The above figure shows that the training process appears to be giving stable cut values, which do not fluctuate too much. The slight variation of the maximum Q-factor (see figure 4b) can be attributed to the random nature of the initial NN weight seeding in the training process.

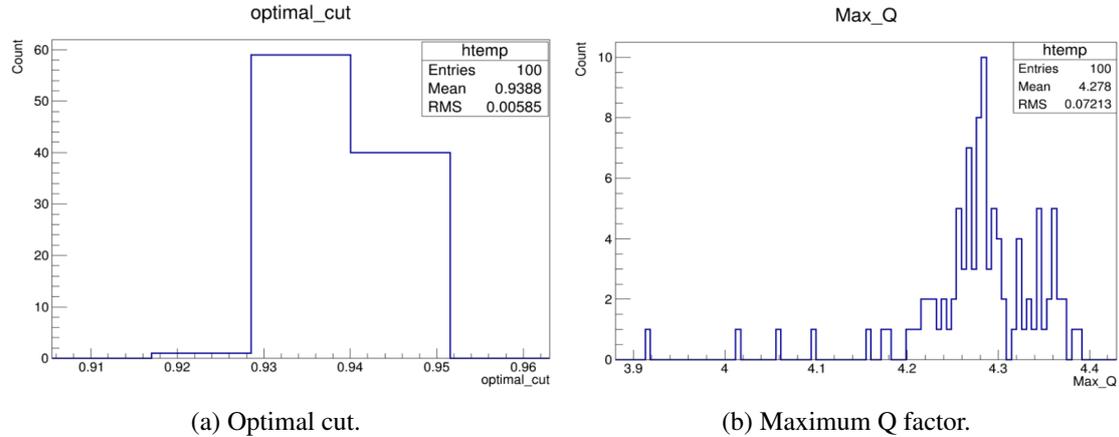


Figure 4: One hundred NNs were trained in order to investigate the fluctuation on the optimal cut and the Q factor in the training stage.

For exploring the increase on G/H separation efficiency, three NNs were trained with similar architecture but different inputs (NoInputs:5:5:1, see Figure 3a). The output layer defines the probability of being a gamma event or a hadron event. The networks that are used in this work are shown below:

- NN1: It uses three input parameters: \mathcal{C} , \mathcal{P} and SFCFChi2 .

- NN2: It uses three input parameters: \mathcal{C} , \mathcal{P} and logmaxPE.
- NN3: It uses four input parameters: \mathcal{C} , \mathcal{P} , SFCFChi2 and logmaxPE.

4. Testing stage

The best way to compare the different performance between Standard Cut and the NNs is using HAWC data. We chose a set of well-reconstructed events² within $\pm 6^\circ$ of the Crab Nebula in the period from November 2014 to December 2015. One significance map per energy bin was made with a Simple Analysis Tools that is implemented in the official HAWC-software. From each map, the maximum significance was extracted, and given in the Table 2. The results in the bin 0 (low-energy) show that the NN2 has better performances than the standard cut, suggesting that the logmaxPE parameter is useful for distinguishing events with low energy. Meanwhile, SFCFChi2 appears to improve the high-energy performance of the G/H separation process (bin 8 and 9), with NN1 having the best results in this energy range. The combination of both parameters in the neural network, on the other hand, worsens the separation performances in almost every energy bin.

\mathcal{B}	Standard cut	NN1	NN2	NN3
0	4.49	2.76	5.14	3.15
1	9.60	3.92	8.24	2.67
2	15.39	6.89	14.28	5.15
3	22.95	15.48	21.74	12.29
4	30.13	29.75	30.98	25.44
5	32.25	34.71	33.46	32.53
6	32.75	34.77	31.88	31.74
7	29.37	29.82	27.02	27.82
8	28.73	31.25	28.04	28.89
9	28.21	31.64	26.28	27.38

Table 2: The maximum significance in the Crab map obtained from each neural network trained in this work and from the standard cut.

5. Conclusion

In this work, we trained 3 neural networks with \mathcal{C} , \mathcal{P} and additional parameters, that had a potential for gamma-hadron separation in HAWC. In each network, a cut value on the NN's normalized output was selected by finding the maximal quality factor in the training phase of the analysis. Following the training on simulated shower events, the neural nets were tested using sample data from the Crab, and evaluating the maximum significance obtained on the source for all three types of NNs. The SFCFChi2 and logmaxPE parameters, which were selected for their higher gamma-hadron separation potential, have allowed improve the high-energy and low-energy ranges of the

²The core and plane are correctly computed.

data bins, respectively. The combination of the two parameters inside the same neural network has, on the other hand, worse performances than the standard cuts, a behaviour which requires further investigations.

Acknowledgments

We acknowledge the support from: the US National Science Foundation (NSF); the US Department of Energy Office of High-Energy Physics; the Laboratory Directed Research and Development (LDRD) program of Los Alamos National Laboratory; Consejo Nacional de Ciencia y Tecnología (CONACyT), México (grants 271051, 232656, 260378, 179588, 239762, 254964, 271737, 258865, 243290, 132197), Laboratorio Nacional HAWC de rayos gamma; L'OREAL Fellowship for Women in Science 2014; Red HAWC, México; DGAPA-UNAM (grants IG100317, IN111315, IN111716-3, IA102715, 109916, IA102917); VIEP-BUAP; PIFI 2012, 2013, PROFOCIE 2014, 2015; the University of Wisconsin Alumni Research Foundation; the Institute of Geophysics, Planetary Physics, and Signatures at Los Alamos National Laboratory; Polish Science Centre grant DEC-2014/13/B/ST9/945; Coordinación de la Investigación Científica de la Universidad Michoacana. Thanks to Luciano Díaz and Eduardo Murrieta for technical support.

References

- [1] A. U. Abeysekara et al., *The HAWC real-time flare monitor for rapid detection of transient events*, *ArXiv e-prints* (Apr., 2017) [[arXiv:1704.0741](https://arxiv.org/abs/1704.0741)].
- [2] A. U. Abeysekara et al., *Daily Monitoring of TeV Gamma-Ray Emission from Mrk 421, Mrk 501, and the Crab Nebula with HAWC*, *The Astrophysical Journal* **841** (June, 2017) 100, [[arXiv:1703.0696](https://arxiv.org/abs/1703.0696)].
- [3] A. U. Abeysekara et al., *Observation of the Crab Nebula with the HAWC Gamma-Ray Observatory*, *ArXiv e-prints* (Jan., 2017) [[arXiv:1701.0177](https://arxiv.org/abs/1701.0177)].
- [4] J. G. J. G. Wilson, *Cosmic rays*. London ; Wykeham Publications ; Andover : [Distributed by] Chapman and Hall, 1976. Distributed in the Western Hemisphere by Springer Verlag, New York.
- [5] P. Grieder, *Extensive Air Showers: High Energy Phenomena and Astrophysical Aspects - A Tutorial, Reference Manual and Data Book*. Astrophysics and space science library. Springer Berlin Heidelberg, 2010.
- [6] Geoffrey Hinton, “Neural Networks for Machine Learning | Coursera.” <https://www.coursera.org/learn/neural-networks>. Accessed January, 2016.
- [7] Christophe Delaere, “class TMultiLayerPerceptron.” <http://root.cern.ch/root/html/TMultiLayerPerceptron.html>. Accessed January, 2014.