Gamma-Hadron Separation Update

Aous Abdo
Jim Linnemann

August 29, 2005

Abstract

In this memo we study in more details the variables introduced by Abdo in [1]. As mentioned in [1] , with these new variables Q-Factors greater than 2.5 are obtainable. The distributions of these variables for gamma MC, proton MC, and data are compared. Different cuts on these variables were developed and tested on 416 days of Crab data for which a statistical significance of up to 5.05 compared to 3.64 from the compactness cut were obtained.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Variables and Event Sample

As mentioned in [1], in the search for new variables many different 1D and 2D distributions of the following parameters, nOut, nTop, nBot, nFit, nb2, mxPE, and cxPE were studied, where:

nOut	:	number of outriggers hit.
nTop	:	number of PMTs hit in the top layer(air shower layer).
nBot	:	number of PMTs hit in the bottom layer(muon layer).
nFit	:	number of PMTs entered in the fit.
nb2	:	number of muon layer tubes with at least 2PEs.
mxPE	:	number of PEs in the muon layer tube with the highest number of PEs.
cxPE	:	number of PEs in the muon layer tube with the highest number of PEs where a region of 10m around the core is excluded from consideration[2].

These distributions were generated for gamma MonteCarlo, proton MonteCarlo, and data. The MC events were thrown out to 1 km from the detector with a -2.4 and -2.7 energy spectrum indices for gammas and protons, respectively. The simulations were generated over a zenith angle range of 0^o to 45^o. The standard uncorrelated noise model was used. Gamma MC events were thrown with energies in the range 100 GeV to 100 TeV while proton MC events had an energy range of 50 GeV to 100 TeV. The data set used come from sub run 5953_108, taken on Oct. 19, 2004 at 06:28 UT. The number of events in this data sample is 355,475 , 57% of which landed on the pond. For gamma MC there were 95,909 events, 54% of these events landed on the pond. While for proton MC there were 1,005,920 events of which 76 % landed on the pond.

Tony Shoup's fitting routines were used. All three layers are used for the angular fitting of the shower. All distributions studied use the same event fit cut of **nFit > 80**[1] . This nFit value was selected to through poorly reconstructed events, later, its effects, should be discussed. For gamma MC, events must be reconstructed within 0.7^o of their true direction. This value represents the optimal bin radius for the nFit cut selected. For the MC distributions the Multiplicity trigger of 55 PMTs was used, while for data the risetime trigger was used. This may be part of the reason for differences between the monte carlo simulations and data.

[1]For this analysis nFit includes tubes from the outriggers, Airshower, and the Muon layers.

Many different distributions were studied. Only distributions that showed the greatest differences between gamma and proton MC and the best agreement between proton MC and data were studied in more details. Efficiency distributions and distributions of the Q-Factor were then generated and studied for the best cuts to be applied. These cuts were then tested on 416 days of Crab data to see if there are any improvements in the significance from the Crab[2].

All of the variables studied involved the number of hit or fit tubes in the numerator and include cxPE in the denominator. cxPE was first introduced by Sinnis in [2]. cxPE is expected to keep gamma showers that land on the pond. The naming convention chosen for these variables are: CX_2 for the analog of X_2, but using cxPE instead of mxPE in the denominator, and AX_n for the other variables, where A stands for Aous, and n gives the number of basic variables used. We will first tour these variables and their performance and then summarize in a table and compare results in sections 6 and 7.

[2]compared to the current compactness cut criteria.

2 $\mathbf{AX_4}$ & $\mathbf{CX_2}$

In this section the two discriminating variables , $\mathrm{AX_4}$ and $\mathrm{CX_2}$ are studied in more details, where:

$$\mathrm{AX_4} = \frac{\mathrm{nOut*nTop*nFit}}{\mathrm{cxPE}} \tag{1}$$

$$\mathrm{CX_2} = \frac{\mathrm{nb2}}{\mathrm{cxPE}} \tag{2}$$

Distributions of both variables are shown in figure 1 for gamma MC, proton MC, and data. The two upper plots show the distribution of $\mathrm{AX_4}$, while the lower ones show the distribution of $\mathrm{CX_2}$. Distributions are shown on the right with a log scale for the y-axis. The $\mathrm{AX_4}$ distribution of proton MC is indistinguishable form the data, and the $\mathrm{CX_2}$ distribution for proton MC is in reasonable agreement with that for the data.

Figure 2 shows two dimensional distributions of $\mathrm{AX_4}$ vs $\mathrm{CX_2}$ for gamma MC, proton MC, and data. Distributions are shown on the right with a log scale for the z-axis. From these distributions we see that the proton distribution is well concentrated in the region $\mathrm{AX_4} < 40000$ and $\mathrm{CX_2} < 6.00$ while the gamma distribution extends beyond this region. This suggest a two sided cut be made on these two variables, i.e. a cut where only events that passes a cut threshold on both variables are kept:

$$\mathrm{AX_4} \geq \alpha \tag{3}$$

$$\mathrm{CX_2} \geq \beta \tag{4}$$

where α and β are the cut thresholds for $\mathrm{AX_4}$ and $\mathrm{CX_2}$, respectively. The exact values of α and β that are expected to optimize the Q-Factor while keeping as much gamma, after the cuts, as possible, will be decided upon after studying the Q-Factor and the efficiency distributions. In figure 3 we zoom in the distributions for proton MC and data. The denser look of the distribution for data is due to the fact that the data sample has much more statistics than the proton MC sample [3]. Figure 4 shows the two dimensional efficiency plots for the distributions in figure 2. Each point in these two dimensional efficiency plots represents the efficiency of the cut made on the corresponding values of the x and y axes. The efficiency distribution for proton MC is in good agreement with the data. A main characteristic of the gamma efficiency distribution is that this distribution doesn't fall off as fast as those for proton and data.

The Q-Factor represents the relative increase in sensitivity expected for a given selection criterion and is defined as:

$$\mathrm{Q} = \frac{\epsilon_s}{\sqrt{\epsilon_b}} \tag{5}$$

where ϵ_s is the efficiency for retaining the signal events, and ϵ_b is the efficiency for retaining background events. Figure 5 shows the Q-Factor as a function of the cut values. The upper plot shows the Q-Factor obtained by comparing gamma (signal) to proton MC (background), while the lower one shows the Q-Factor obtained by comparing gamma MC(signal) to data (background)[4]. The Q-Factor for gamma MC vs. data does not plateau as it does for gamma MC vs. proton MC, We chose to operate at a moderate value of the Q-Factor so as to have higher photon efficiency.

[3]The number of triggered proton MC events is 24,904.

[4]The fact that most of the data collected with Milagro consists of events from hadron-initiated air showers allows the use of the data sample as a background.

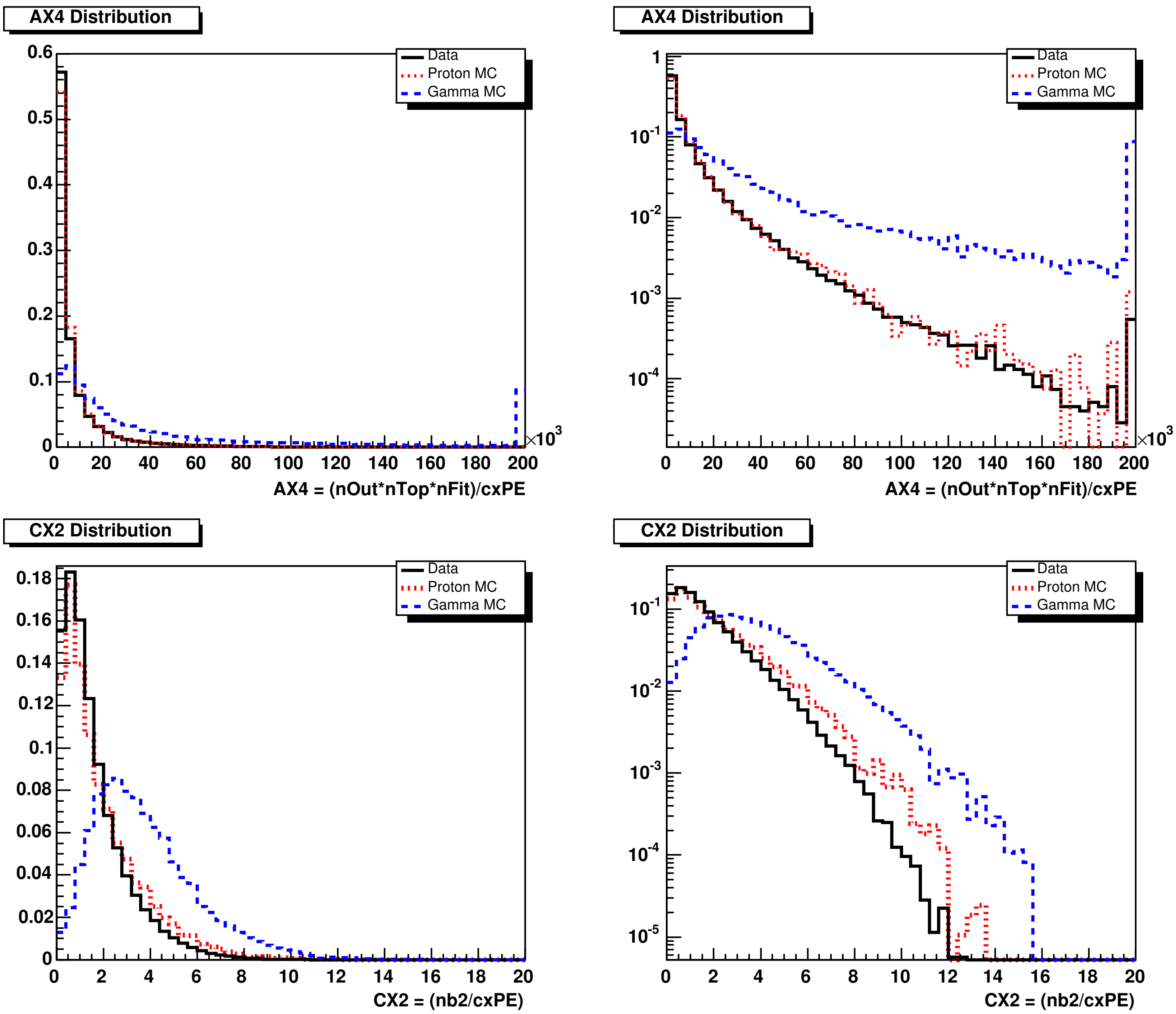


Figure 1: Distributions of AX_4 and CX_2 for gamma MC, proton MC, and data. The two upper plots show the distribution of AX_4, while the lower ones show the distribution of CX_2. Distributions are shown on the right with a log scale for the y-axis . All histograms have been normalized to have unit area.

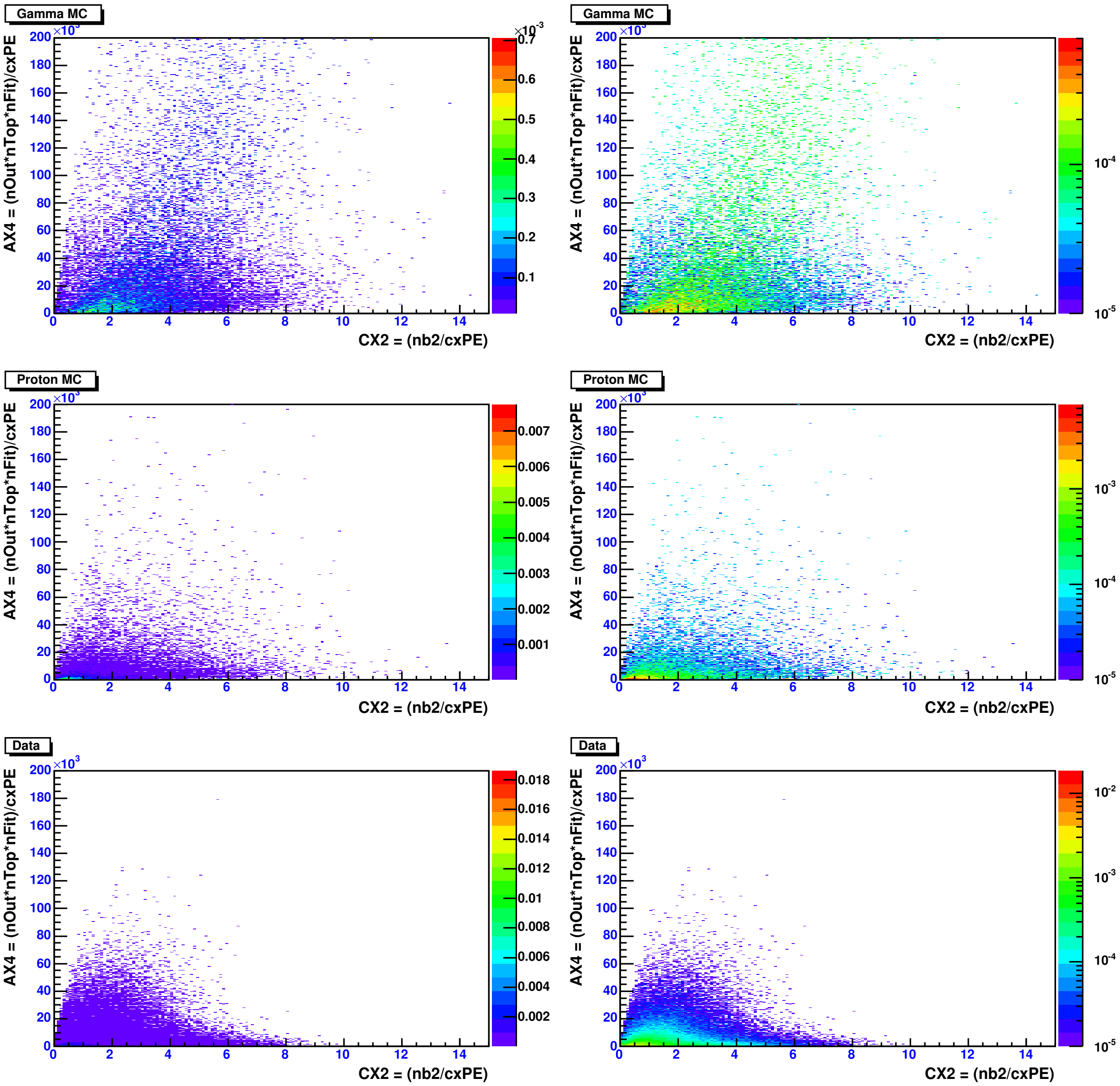


Figure 2: Distributions of AX_4 vs. CX_2 for gamma MC, proton MC, and data. Distributions are shown on the right with a log scale for the z-axis. All histograms have been normalized to have unit area.

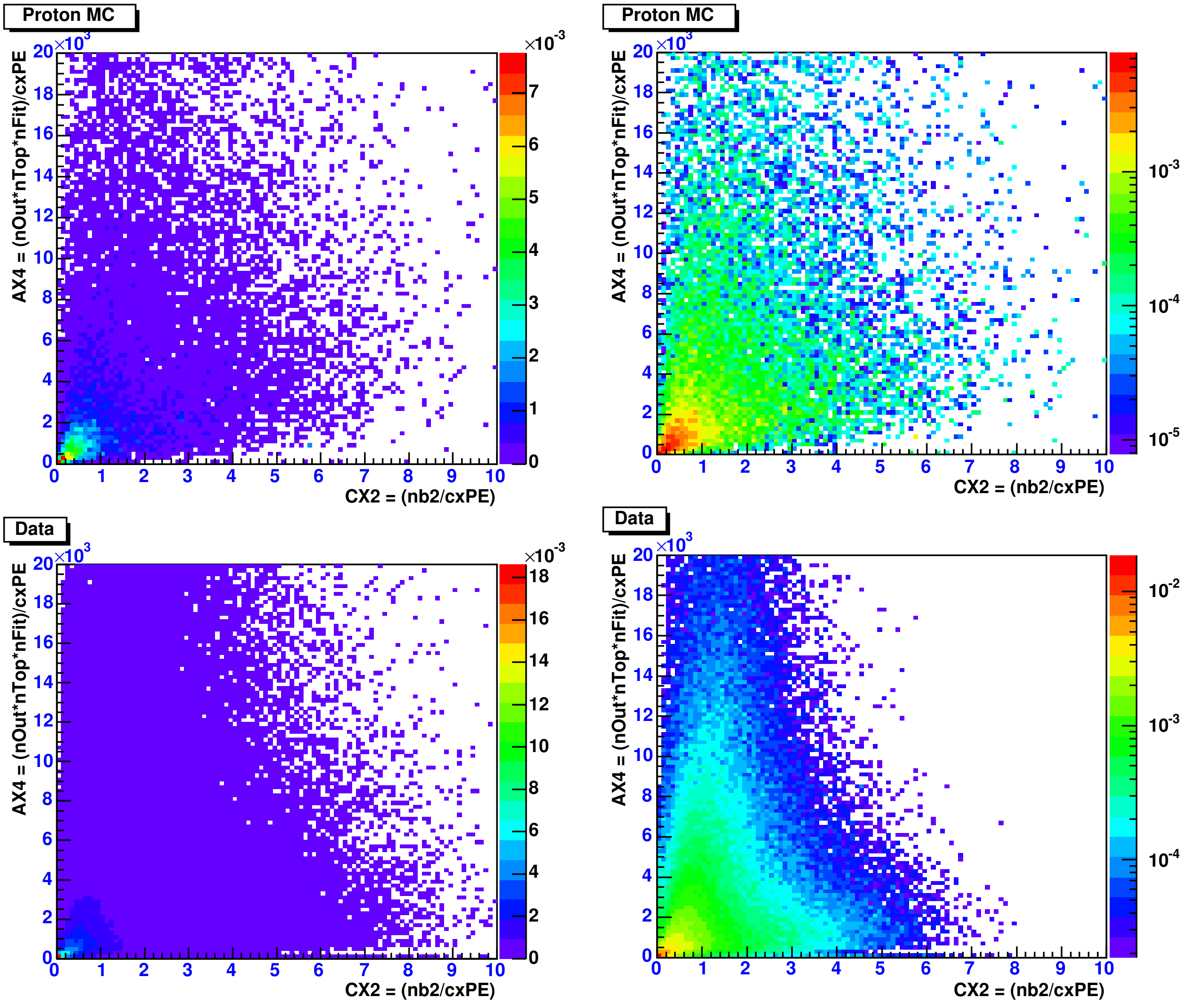


Figure 3: Distributions of AX_4 vs. CX_2 for proton MC and data for the highest density regions. Distributions are shown on the right with a log scale for the z-axis. All histograms have been normalized to have unit area.

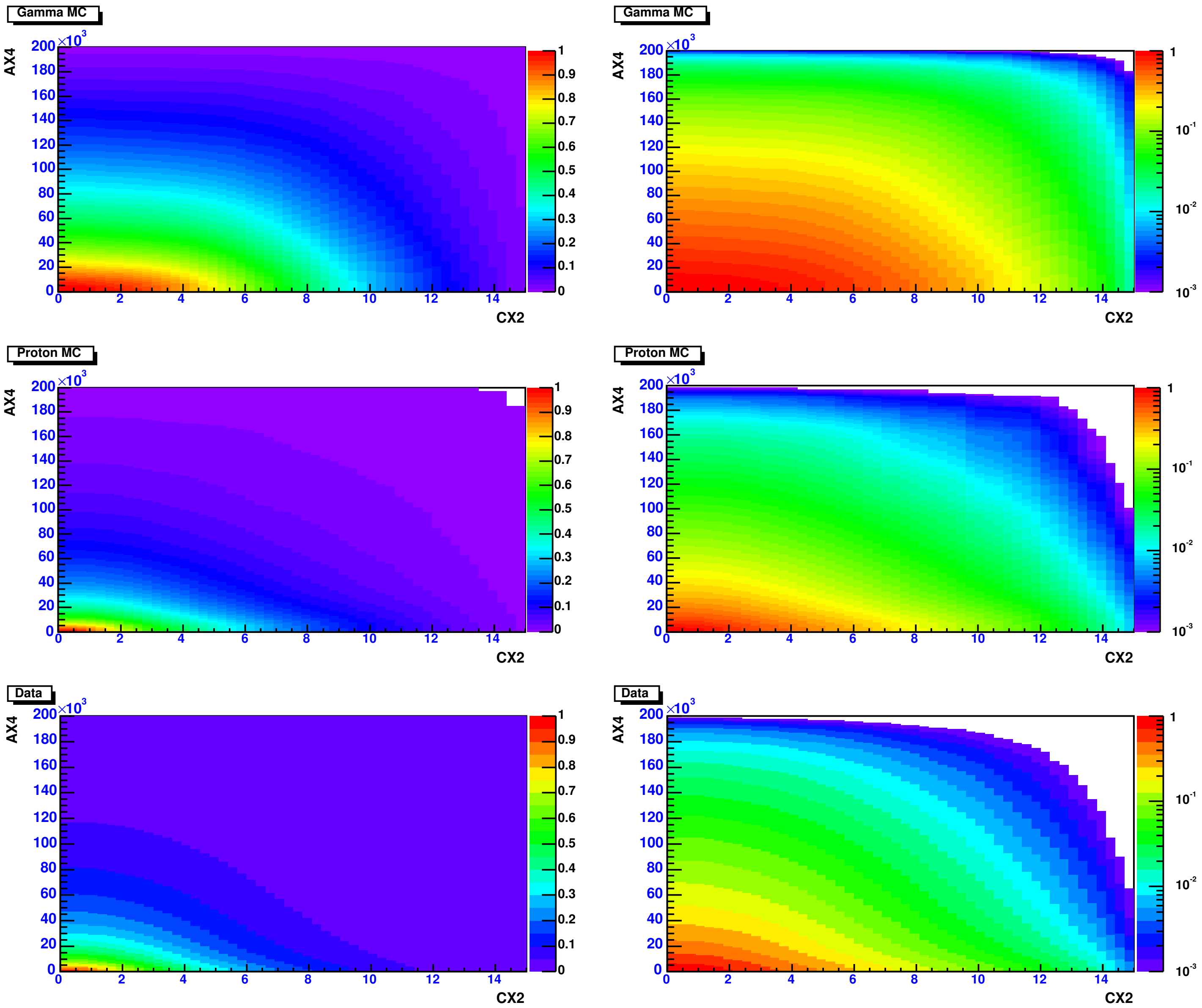


Figure 4: Efficiency plots for the distributions shown in figure 2. Distributions are shown on the right with a log scale for the z-axis.

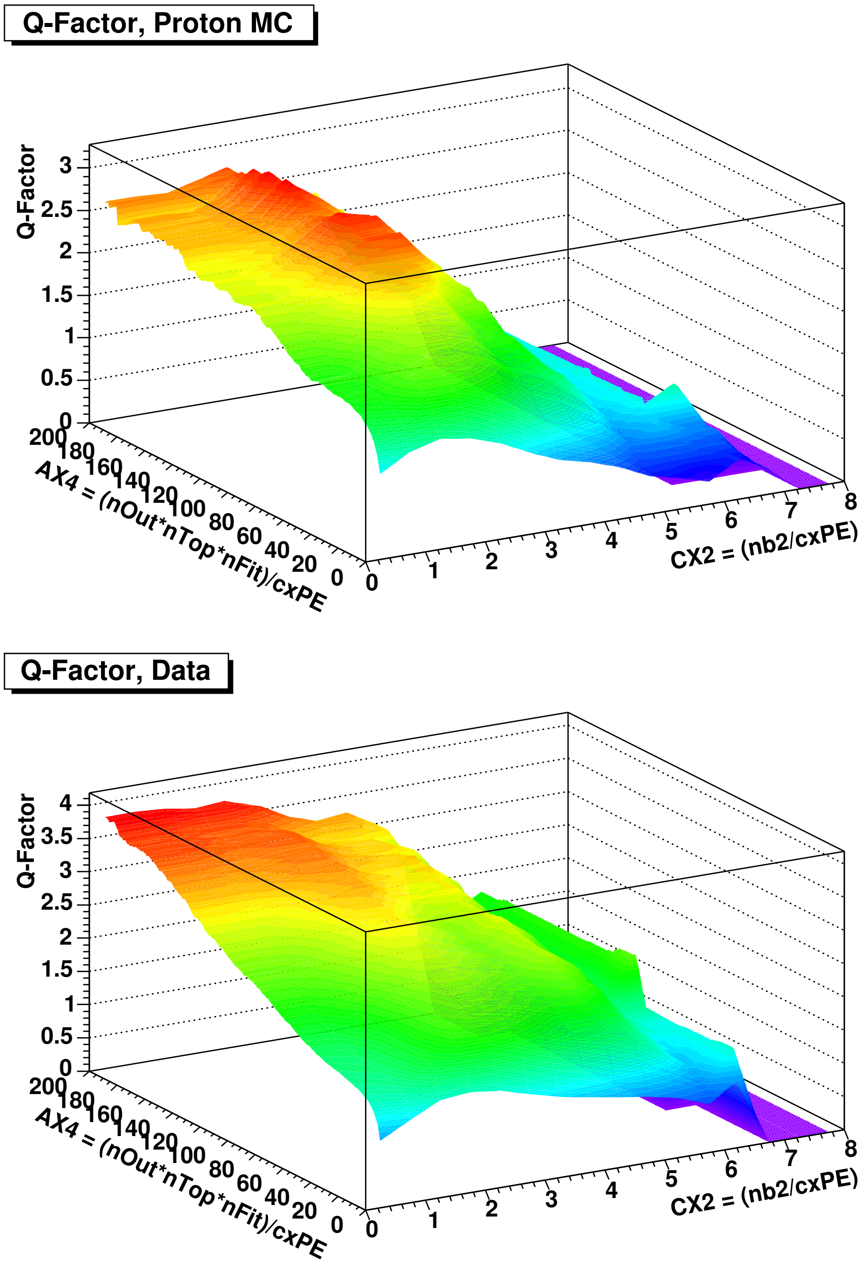


Figure 5: Q-Factor plots for AX_4 and CX_2, the upper plot is for gamma MC and proton MC, while the lower one is for gamma MC and data. The plots show the Q-Factor as a function of cut threshold in each variable. The AX_4 values are multiplied by 10^3.

2.1 Cuts

Table 1 shows a list of cuts that were considered for AX_4 and CX_2 along with the corresponding efficiencies and Q-Factors. For the analysis that follows the following cuts were chosen:[5]:

$$AX_4 \geq 10^5 \quad (6)$$

$$CX_2 \geq 4.28 \quad (7)$$

requiring events to pass these cuts rejects 99.73 % of proton MC and 99.79 % of tha data(for this data sample), and retains 13.88% of gamma MC. This results in a predicted Q-Factor of 2.65 comparing proton MC to gamma MC events, and 2.99 comparing data to gamma MC. Thus, the Monte Carlo predicts an improvement in significance over all triggered data of 2.65, or a factor of 2.65/1.62 = 1.64 over the present compactness cut.

Cuts Applied		**Eff_G**	**Eff_P**	**Eff_D**	**$Q_{MC:MC}$**	**$Q_{MC:D}$**	**$\frac{Q_{MC:MC}}{Q_{MC:D}}$**
AX_4 >	**CX_2 >**						
80000	2.00	0.2053	0.0089	0.0075	2.18	2.37	0.92
85000	2.50	0.1897	0.0068	0.0058	2.29	2.50	0.92
90000	3.00	0.1753	0.0048	0.0042	2.52	2.69	0.94
95000	3.50	0.1607	0.0037	0.0033	2.63	2.82	0.93
100000	4.28	0.1388	0.0027	0.0021	2.65	3.00	0.88
105000	4.50	0.1282	0.0022	0.0018	2.75	3.05	0.90
110000	5.00	0.1114	0.0016	0.0013	2.80	3.06	0.92
115000	5.50	0.0954	0.0013	0.0010	2.67	3.09	0.86

Table 1: List of cuts applied for AX_4 and CX_2 along with the corresponding efficiencies and Q-Factors. $Q_{MC:MC}$ is the Q-Factor obtained by comparing gamma MC to proton MC, while $Q_{MC:D}$ is the Q-Factor obtained by comparing gamma MC to data.

[5]Event selection cuts were applied for all MC and data files. Please see the last paragraph on page 1.

2.2 Energy dependence on the Cuts

In figure 6 three plots are shown. The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, The second plot shows the energy distribution of gamma and proton MC after the cuts have been applied, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy. From the second and third plots we see that the efficiency of gamma-ray-initiated events is dependent on the energy of the primary gamma ray, while for protons the efficiency is nearly independent of energy. We also notice that, with these cuts, we don't expect to see any gamma-ray-initiated events with primary energy less than 1 TeV, and that the median for gamma-ray-initiated events passing the cuts is around 13.4 TeV compared to 3.72 TeV for all triggered events[6]. It is somewhat difficult to tell the effect on cosmic rays since so few pass the cuts, but the median energy has moved up from 1 TeV to 7.6 TeV.

2.3 Test Of The Cuts On The Crab Nebula

Figure 7 shows the map of statistical significance in the region around the Crab nebula with the above cuts applied. For comparison a map of the same region with the compactness cut applied is shown in figure 8.[7] From these two figures we see that the new cuts improved the significance of the Crab from 3.64 (for the compactness cut) to 5.05 standard deviations (for the new criteria) a factor of 1.39. Although the Monte Carlo led us to hope for an improvement in the Q-Factor by a factor of 1.65, the critical thing is that the new selection criteria have gained us a significant improvement in sensitivity. The Q-Factor indicates the relative sensitivity of the detector to steady source,

[6]Table 5 shows a list of median energies for all cuts applied.

[7]The two maps shown are for the same time period of 416 days of data collection.

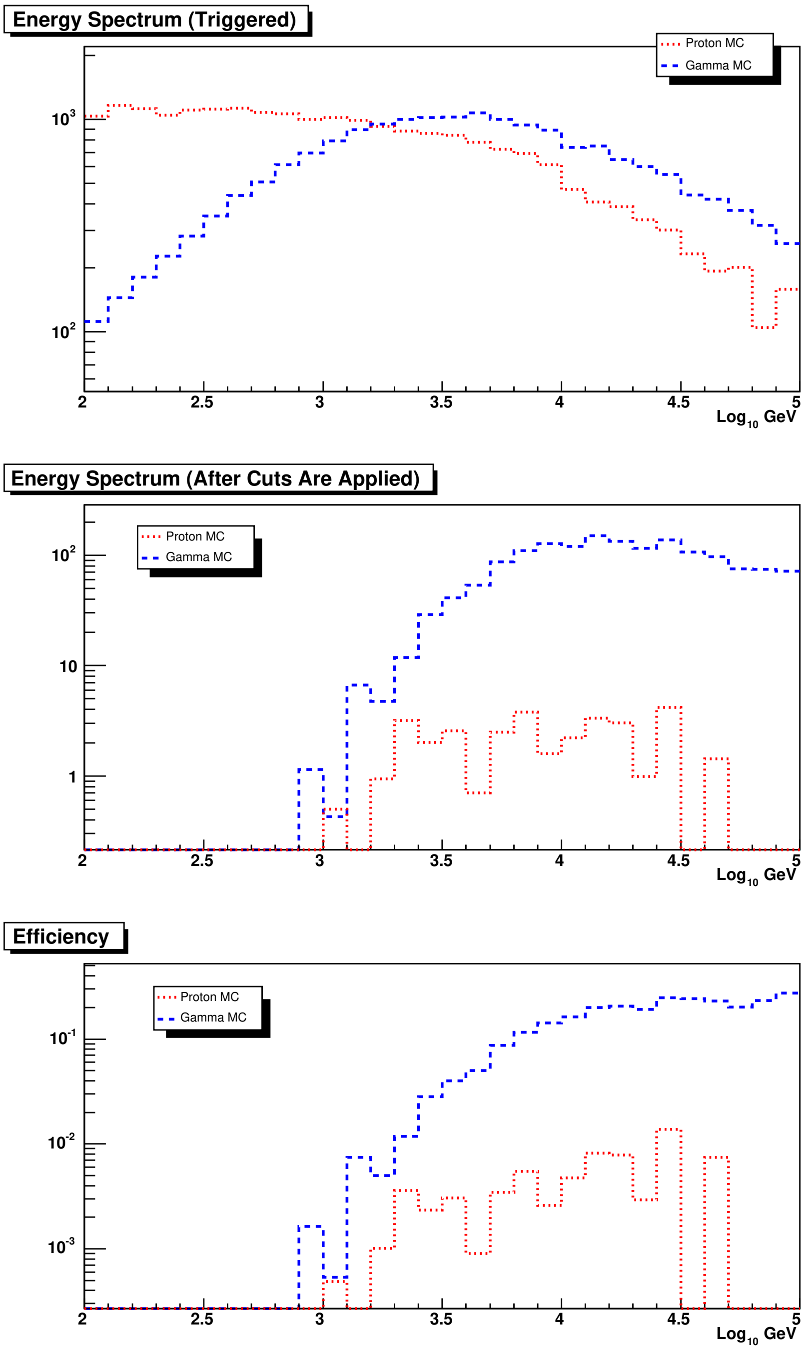


Figure 6: The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after the cuts were applied , and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy.

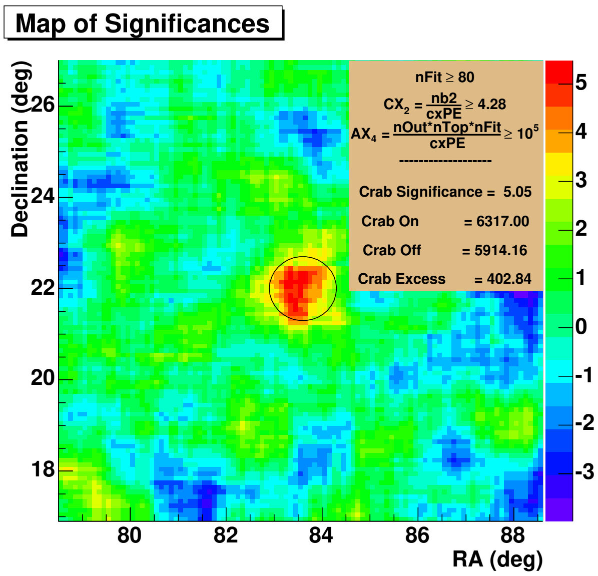


Figure 7: Map of the statistical significance in the region around the Crab Nebula after the $AX_4 > 10^5$ and $CX_2 > 4.28$ cuts have been applied.

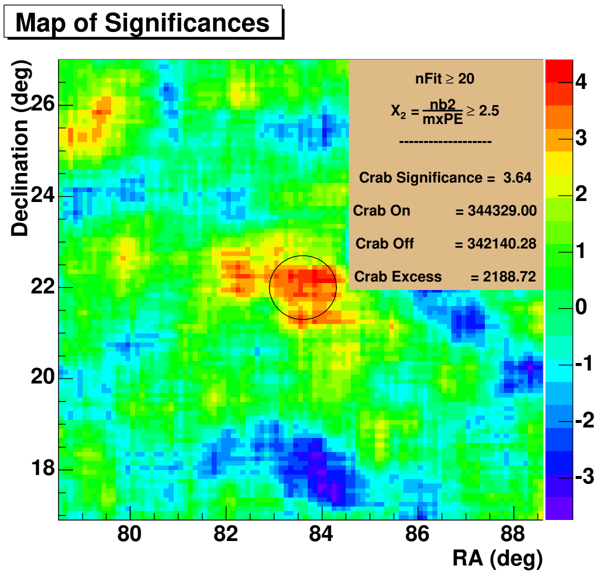


Figure 8: Map of the statistical significance in the region around the Crab Nebula after the compactness cut($X_2 > 2.5$) has been applied.

and the time to reach a given level of statistical significance for a steady source decreases[8] as $(1/\beta^2)$, where β is the improvement factor(Q(new Cuts)/Q(compactness cut)). A Crab-like source can be seen with roughly half as much data with the new cuts provided the source produces photons with energy > 1 TeV.

2.4 On-Pond Events

Figure 9 shows distributions of AX_4 and CX_2 for on the pond events[9]. The two upper plots show the distribution of AX_4, while the lower ones show the distribution of CX_2. The distributions are shown on the right with a log scale for the y-axis. The AX_4 distribution of proton MC is indistinguishable form the data, and the proton MC CX_2 distribution is in reasonable agreement with the data. In comparison with all events, on-pond events have smaller AX_4 and CX_2.

2.4.1 Cuts For On-Pond Events

For on-pond events two sets of cuts were tested. The first set of cuts is identical to the set of cuts applied for all events :

$$AX_4 \geq 10^5 \tag{8}$$

$$CX_2 \geq 4.28 \tag{9}$$

Requiring on-pond events to pass these cuts rejects 99.97 % of proton MC and 99.98 % of the data(for this data set), and retains 1.3 % of gamma MC. This results in a predicted Q-Factor of 0.74 comparing proton MC to gamma MC events, and 0.94 comparing data to gamma MC.

The second set of cuts is:

$$AX_4 \geq 3200 \tag{10}$$

$$CX_2 \geq 2.1 \tag{11}$$

[8] See Appendix.

[9] On-Pond events are events with their fit core on the pond.

Requiring on-pond events to pass these cuts rejects 71.24 % of proton MC and 82.77 % of the data(for this data set), and retains 65.9 % of gamma MC. This results in a predicted Q-Factor of 1.23 comparing proton MC to gamma MC events, and 1.59 comparing data to gamma MC.

2.4.2 Test of On-Pond Cuts On The Crab Nebula

When applying the above cuts for on-pond events from the Crab[10] statistical significances of 1.77 and 2.55 were obtained, respectively. The $X_2 \geq 2.5$ cut gave a statistical significance of 3.42 for on-pond events which is better than the results we got with the above cuts.

Although the cuts $AX_4 > 10^5$ and $CX_2 > 4.28$ gave good results for the case of all events, these cuts are not good for on-pond events and we may have to apply different sets of cuts for on and off the pond events.

[10] For the same time period of 416 days.

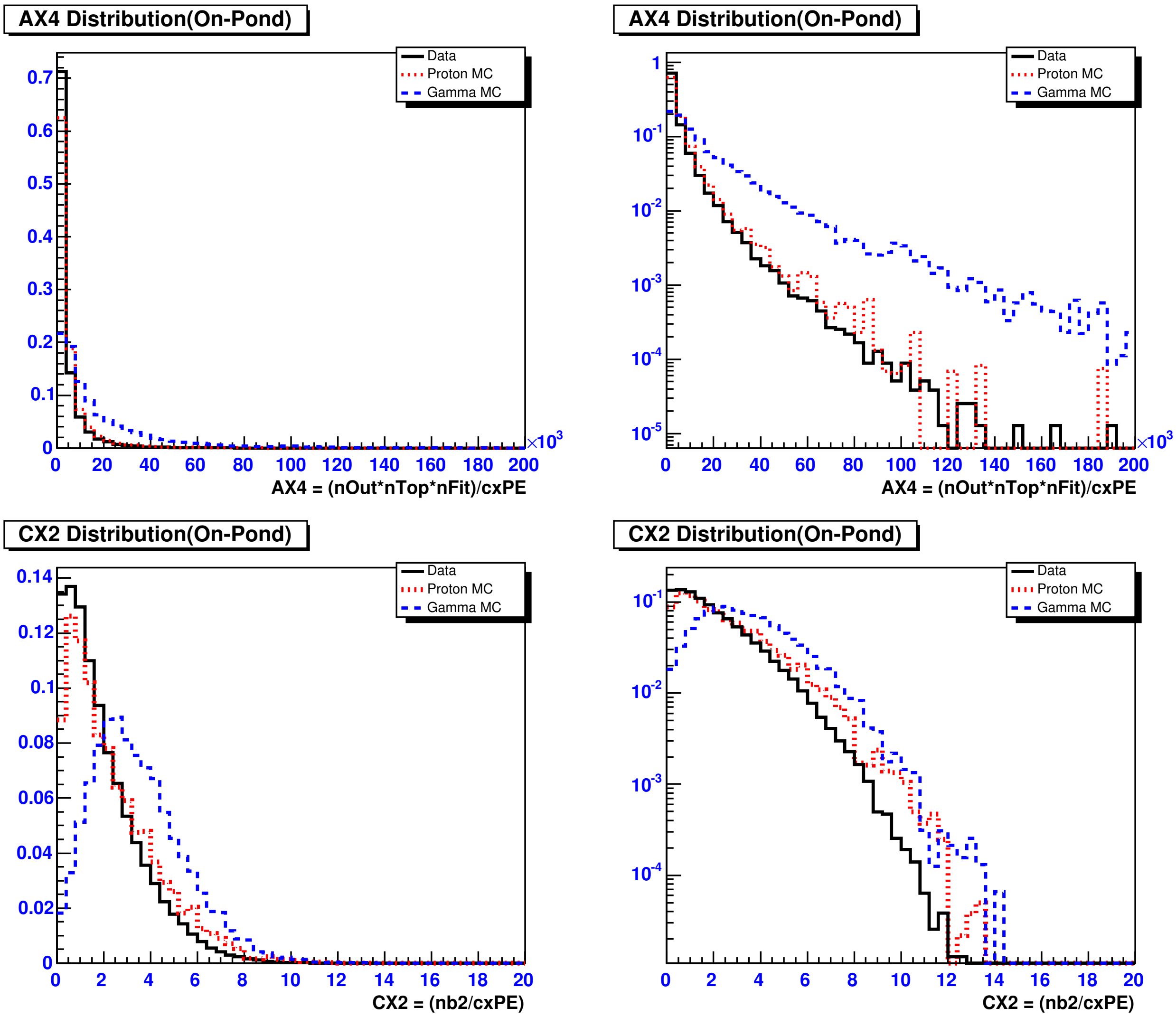


Figure 9: Distributions of AX_4 and CX_2 for gamma MC, proton MC, and data for On-Pond Events. The two upper plots show the distribution of AX_4, while the lower ones show the distribution of CX_2. Distributions are shown on the right with a log scale for the y-axis . All histograms have been normalized to have unit area.

3 AX_4P & CX_2

In this section the variables AX_4P and CX_2 are studied in more details. AX_4P is a modified version of AX_4 :

$$AX_4P \; = \frac{(\text{nOut}+\text{nTop})^*\text{nFit}}{\text{cxPE}} \tag{12}$$

and CX_2 is given by equation 2. The distribution of AX_4P is shown in figure 10 for gamma MC, proton MC, and data[11]. The distribution is shown on the right with a log scale for the y-axis. Although not as good as the AX_4 distribution, the proton MC distribution of AX_4P is in reasonable agreement with the data.

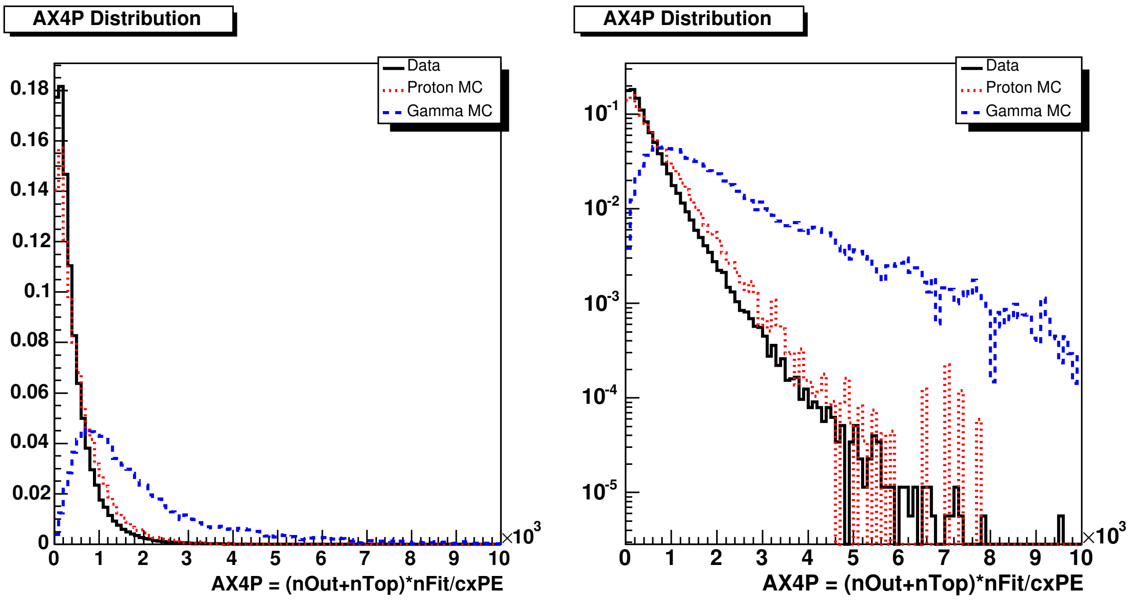


Figure 10: Distribution of AX_4P for gamma MC, proton MC, and data. The distribution is shown on the right with a log scale for the y-axis. All histograms have been normalized to have unit area.

Figure 11 shows two dimensional distributions of AX_4P vs. CX_2 for gamma MC, proton MC, and data. Distributions are

[11] CX_2 distribution is shown in figure 1.

shown on the right with a log scale for the z-axis. From this figure we see that AX_4P and CX_2 are more strongly correlated than AX_4 and CX_2. In figure 12 we zoom in the distributions for proton MC and data. The strong correlation between AX_4P and CX_2 is clear here. Figure 13 shows the two dimensional efficiency plots for the distributions in figure 10. Each point in these two dimensional efficiency plots represents the efficiency of the cut made on the corresponding values of the x and y axes. The efficiency distribution for proton MC is in good agreement with the data. A main characteristic of the gamma efficiency distribution is that this distribution doesn't fall off as fast as those for proton and data.

Figure 14 shows the Q-Factor plots, the upper plots are for gamma MC and proton MC, while the lower ones are for gamma MC and data. On the right are the top view of the corresponding plots on the left. Somewhat to our surprise, the correlation between AX_4P and CX_2 causes the efficiency of the AX_4P to be weakly related to CX_2 and the predicted Q Factor nearly independent of CX_2. More on this in section 4.

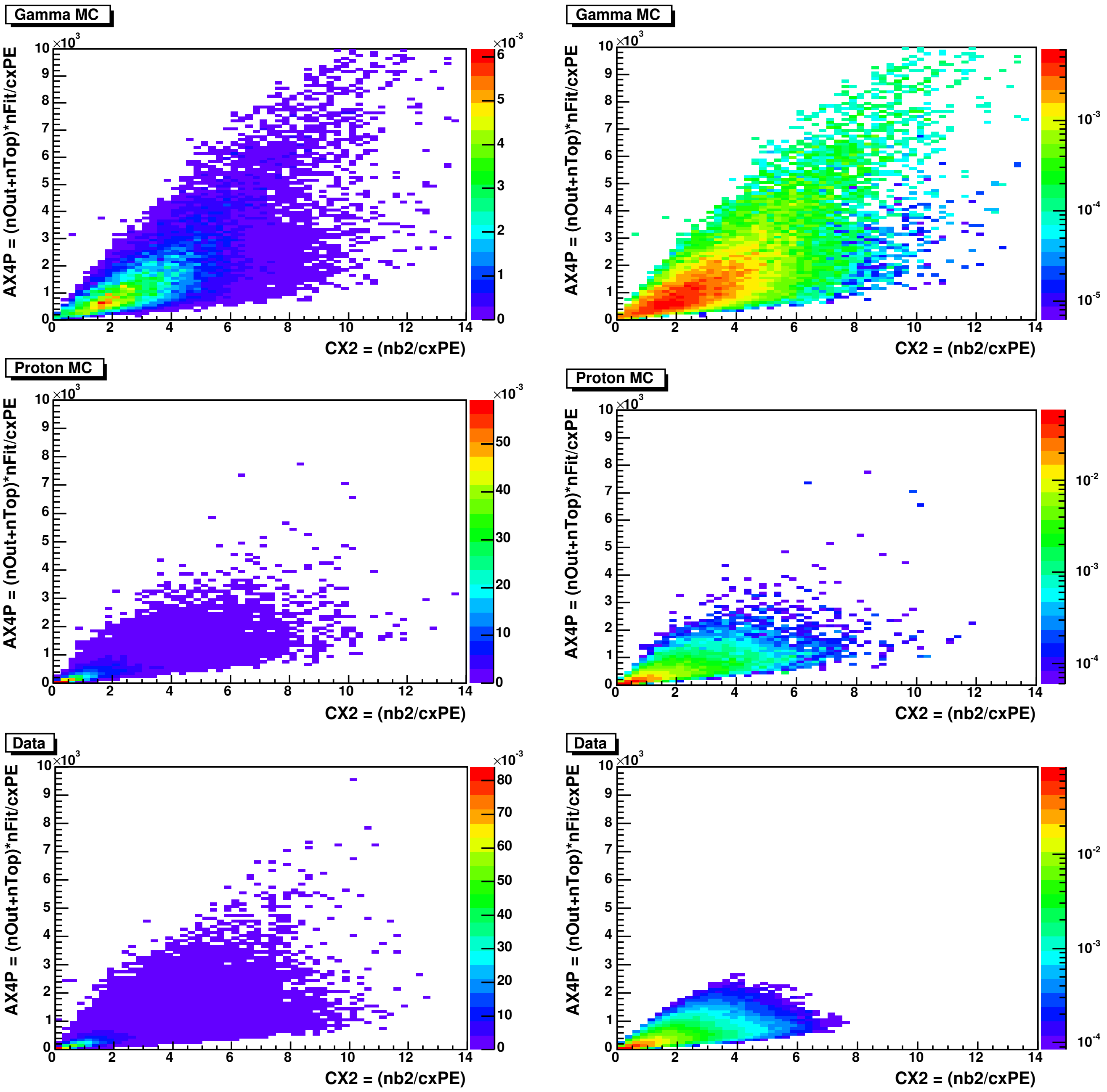


Figure 11: Distributions of AX_4P vs. CX_2 for gamma MC, proton MC, and data. Distributions are shown on the right with a log scale for the z-axis. All histograms have been normalized to have unit area.

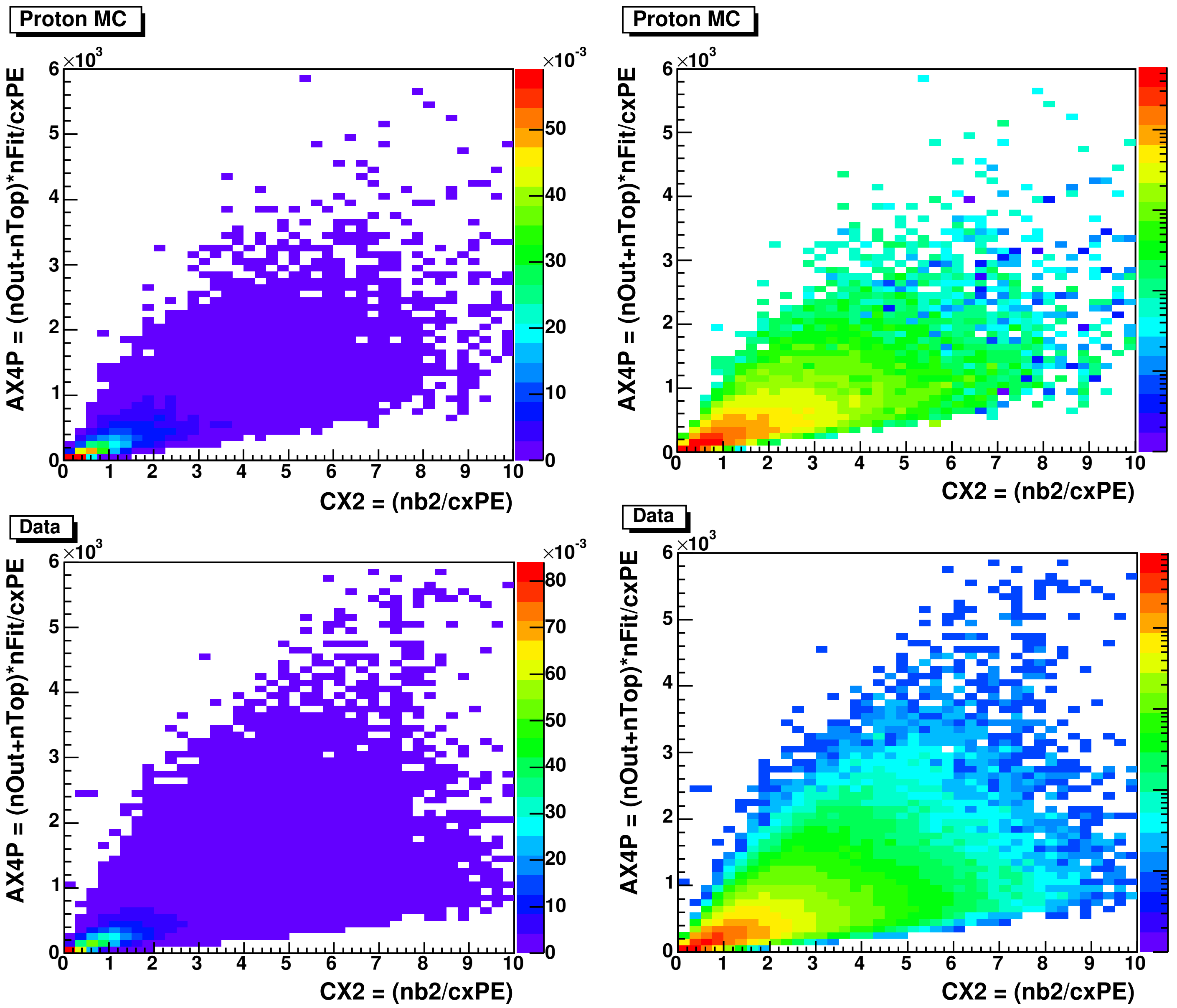


Figure 12: Distributions of AX_4P vs. CX_2 for proton MC and data for the highest density region. Distributions are shown on the right with a log scale for the z-axis. All histograms have been normalized to have unit area.

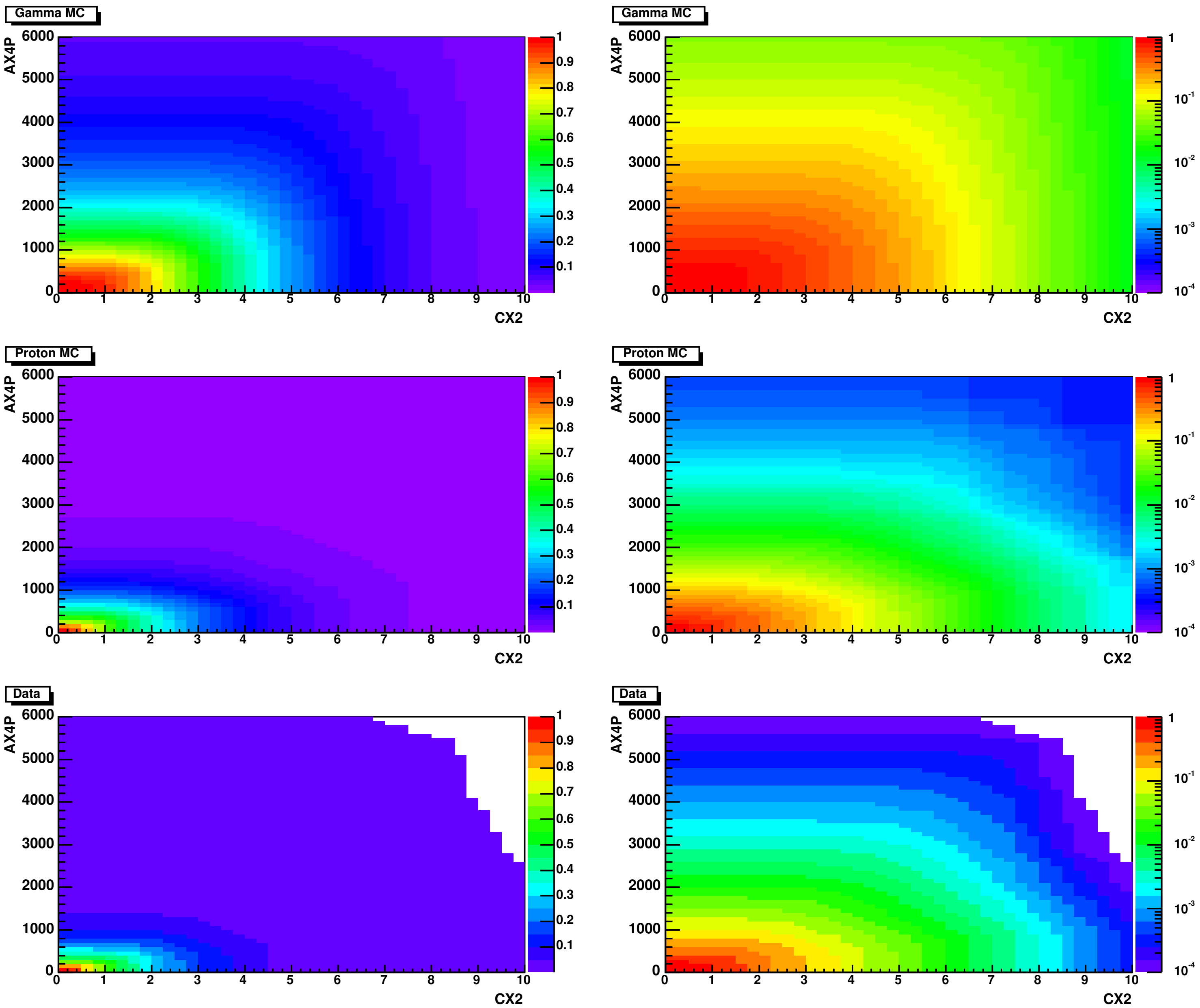


Figure 13: Efficiency plots for the distributions shown in figure 10. Distributions are shown on the right with a log scale for the z-axis.

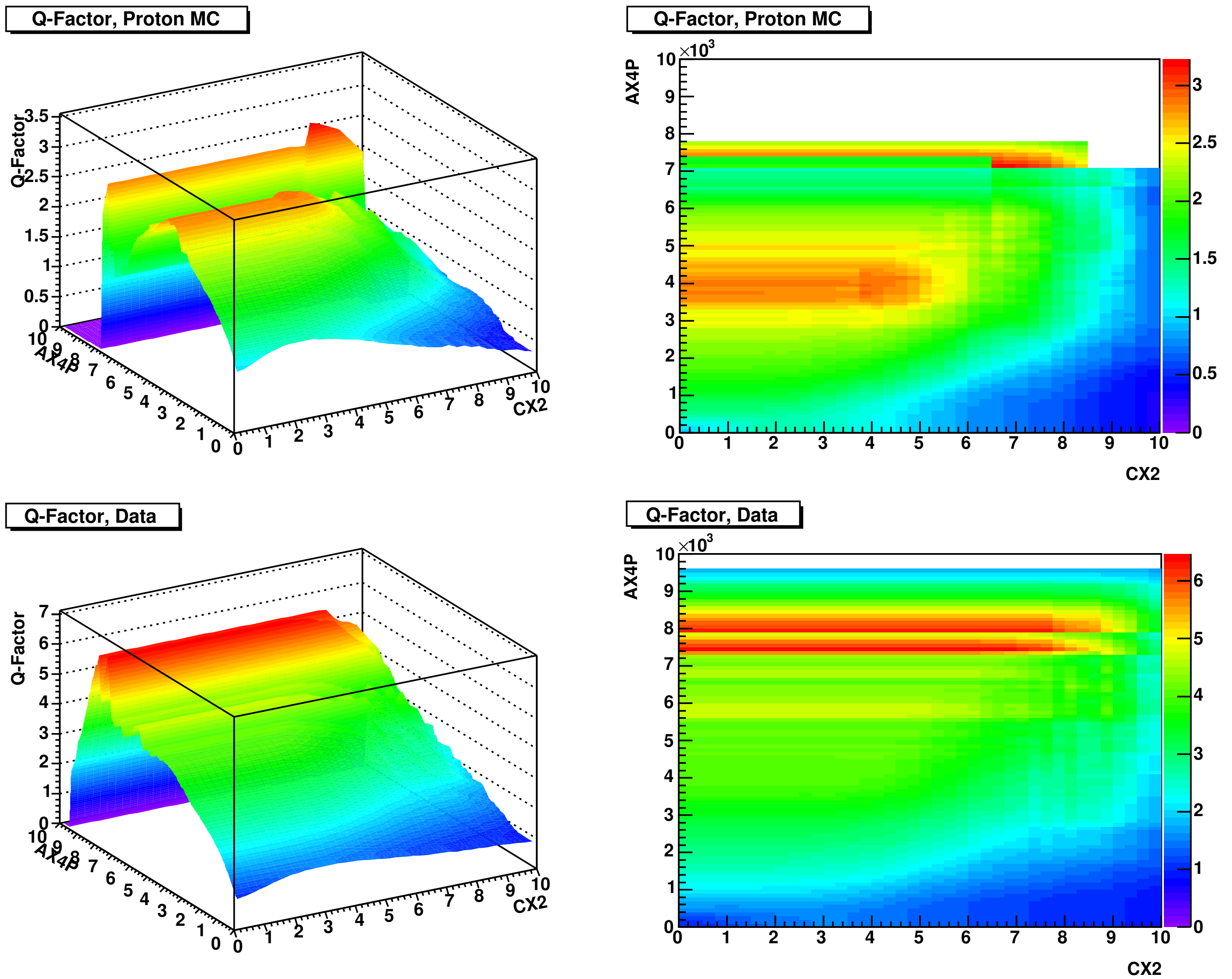


Figure 14: Q-Factor plots, the upper plots are for gamma MC and proton MC, while the lower ones are for gamma MC and data. On the right are the top view of the corresponding plots on the left. The plots show the Q-Factor as a function of cut threshold in each variable.

3.1 Cuts

Table 2 shows a list of cuts that were considered for AX_4P and CX_2 along with the corresponding efficiencies and Q-Factors. For the analysis that follows the following cut was choosen: [12]:

$$AX_4P \geq 3800 \tag{13}$$
$$CX_2 \geq 4.00 \tag{14}$$

requiring events to pass these cuts rejects 99.788 % of proton MC and 99.891 % of tha data(for this data sample), and retains 13.044% of gamma MC. This results in a predicted Q-Factor of 2.83 comparing proton MC to gamma MC events, and 3.95 comparing data to gamma MC. Thus, the Monte Carlo predicts an improvement in significance over all triggered data of 2.83, or a factor of 2.83/1.62 = 1.75 over the present compactness cut.

Cuts Applied		**Eff_G**	**Eff_P**	**Eff_D**	**$Q_{MC:MC}$**	**$Q_{MC:D}$**	**$Q_{MC:MC}/Q_{MC:D}$**
AX_4P >	**CX_2 >**						
3500	2.50	0.1535	0.0030	0.0016	2.78	3.85	0.72
3700	3.00	0.1403	0.0025	0.0013	2.83	3.93	0.72
3800	3.00	0.1333	0.0023	0.0011	2.77	4.00	0.69
3800	3.50	0.1326	0.0023	0.0011	2.75	4.00	0.69
3800	4.00	0.1304	0.0021	0.0011	2.83	3.95	0.72
3800	4.50	0.1244	0.0020	0.0011	2.79	3.83	0.73
3800	5.00	0.1159	0.0020	0.0010	2.62	3.68	0.71
3900	5.00	0.1124	0.0018	0.0009	2.63	3.71	0.71
4000	5.50	0.0992	0.0016	0.0008	2.46	3.59	0.68

Table 2: List of cuts applied for AX_4P and CX_2 along with the corresponding efficiencies and Q-Factors. $Q_{MC:MC}$ is the Q-Factor obtained by comparing gamma MC to proton MC, while $Q_{MC:D}$ is the Q-Factor obtained by comparing gamma MC to data.

[12]Event selection cuts were applied for all MC and data files. Please see second paragraph on page 2.

3.2 Energy dependence on the Cuts

In figure 15 three plots are shown. The first plot shows the energy distribution of gamma and proton MC that trigger Milagro. The second plot shows the energy distribution of gamma and proton MC after applying the cuts, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy. From the second and third plots we see that the efficiency of gamma-ray-initiated events is dependent on the energy of the primary gamma ray, while for protons the efficiency is nearly independent of energy. We also notice that, with these cuts, we don't expect to see any gamma-ray-initiated events with primary energy less than 500 GeV, and that the median for gamma-ray-initiated events passing the cuts is around 16 TeV.

3.3 Test Of The Cuts On The Crab Nebula

Figure 16 shows the map of statistical significance in the region around the Crab nebula with the above cuts applied. A map of the same region with the compactness cut applied is shown in figure 8.[13] from these two figures we see that the new cuts improved the significance of the Crab from 3.64 (for the compactness cut) to 3.94 standard deviations (for the new criteria), a rather meager factor of 1.08.

3.4 On-Pond Events

Figure 17 shows distributions of AX_4P for on the pond events[14]. The distribution is shown on the right with a log scale for the y-axis. As for AX_4P, on-pond events have smaller values.

[13] The two maps shown are for the same time period of 416 days of data collection.

[14] The distribution of CX_2 for on-pond events is shown in figure 9.

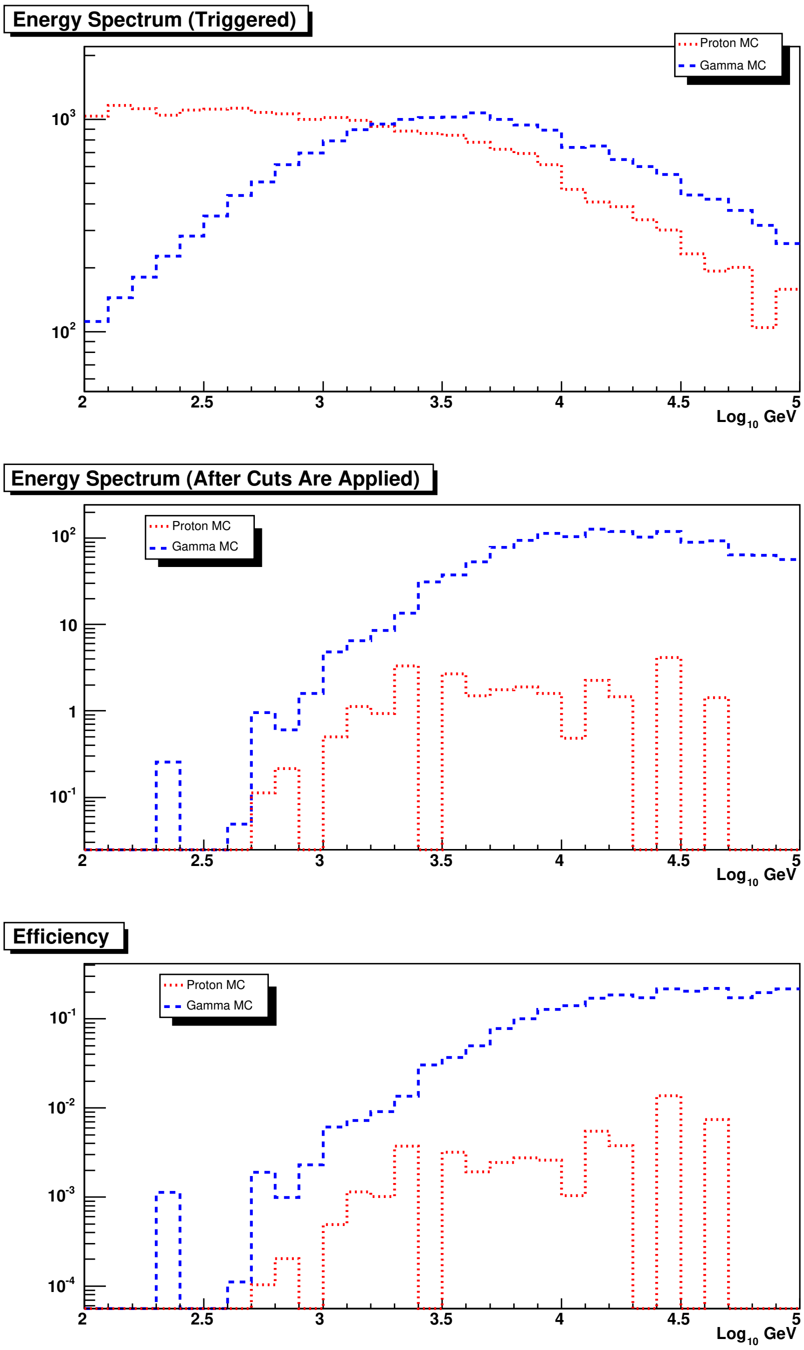


Figure 15: The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after applying the cuts, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy.

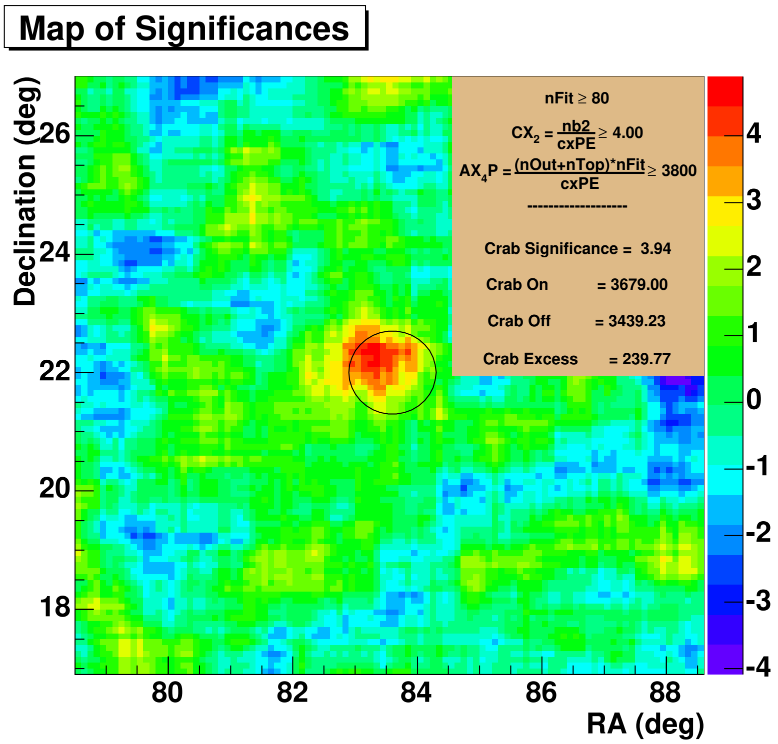


Figure 16: Map of the statistical significance in the region around the Crab Nebula after the $AX_4P > 3800$ and $CX_2 > 4.00$ cuts have been applied

3.4.1 Cuts For On-Pond Events

For on-pond events two sets of cuts were applied. The first set of cuts applied is identical to the set of cuts applied for all events :

$$AX_4P \geq 3800 \tag{15}$$
$$CX_2 \geq 4.00 \tag{16}$$

Requiring on-pond events to pass these cuts rejects 99.95 % of proton MC and 99.98 % of the data(for this data set), and retains 1.64 % of gamma MC. This results in a predicted Q-Factor of 0.72 comparing proton MC to gamma MC events, and 1.23 comparing data to gamma MC.

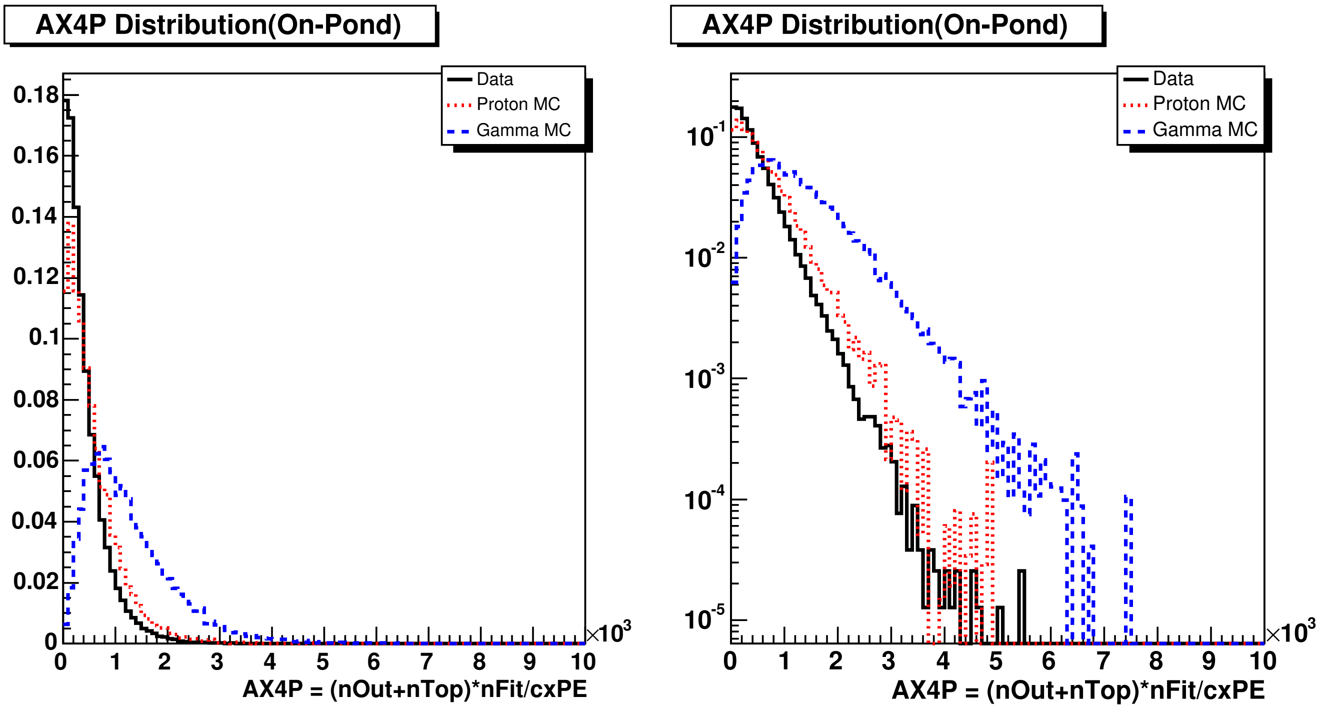


Figure 17: Distributions of AX_4P for gamma MC, proton MC, and data for On-Pond Events. The distribution is shown on the right with a log scale for the y-axis . All histograms have been normalized to have unit area.

The second cut applied is:

$$AX_4P \geq 800 \tag{17}$$

Requiring on-pond events to pass this cut rejects 75.91 % of proton MC and 86.18 % of the data(for this data set), and retains 66.14 % of gamma MC. This results in a predicted Q-Factor of 1.35 comparing proton MC to gamma MC events, and 1.78 comparing data to gamma MC.

3.4.2 Test of On-Pond Cuts On The Crab Nebula

When applying the above cuts for on-pond events from the Crab[15] we get statistical significances of 0.39 and 2.00, respectively. Still the results we get from the new cuts for on-pond events are worse than that with the $X_2 \geq 2.5$ cut.

From the above we see that although the cuts $AX_4P > 3800$ and $CX_2 > 4.00$ gave good results for the case of all-events, these cuts are not good for on-pond events and we may have to apply different sets of cuts for on and off the pond events.

[15] For the same time period of 416 days

4 AX_4P

In the previous section the distribution of AX_4P and CX_2 were studied. However, by looking at Figure 11 which shows a two dimensional distribution of these variables, a good correlation between these two variables is seen. Also, if by looking at Figure 14 which shows the Q-factor distributions, one can see that the Q-Factor distribution is almost independent of CX_2. This suggests that a cut only on AX_4P be applied.

Figure 18 shows efficiency and Q-Factor plots for AX_4P. The upper plot shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to proton MC. The plot on the bottom shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to the data.

4.1 Cuts

Two cuts for AX_4P were applied, the first cut is:

$$AX_4P \geq 3800 \tag{18}$$

this is the same cut applied for AX_4P in the previous section.

Requiring events to pass this cut rejects 99.77% of proton MC and 99.89% of tha data(for this data sample), and retains 13.33 % of gamma MC. This results in a predicted Q-Factor of 2.77 comparing proton MC to gamma MC events, and 4.00 comparing data to gamma MC. Thus, the Monte Carlo predicts an improvement in significance over all triggered data of 2.77, or a factor of $2.77/1.62 = 1.71$ over the present compactness cut.

The second cut applied is:

$$AX_4P \geq 3000 \tag{19}$$

Requiring events to pass this cut rejects 99.35 % of proton MC and 99.68 % of tha data(for this data sample), and retains

19.61 % of gamma MC. This results in a predicted Q-Factor of 2.44 comparing proton MC to gamma MC events, and 3.49 comparing data to gamma MC. Thus, the Monte Carlo predicts an improvement in significance over all triggered data of 2.44, or a factor of $2.44/1.62 = 1.51$ over the present compactness cut.

4.2 Energy dependence on the Cuts

Figures 19 and 20 show the energy dependence on these cuts. In each figure the first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after applying the cuts, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy. From the second and third plots for both cuts we see that the efficiency of gamma-ray-initiated events is dependent on the energy of the primary gamma ray, while for protons the efficiency is nearly independent of energy .We also notice that for the $AX_4P > 3800$ cut we don't expect to see any gamma-ray-initiated events with primary energy less than 500 GeV, and that the median for gamma-ray-initiated events passing this cuts is around 16 TeV. However with the $AX_4P > 3000$ cut we are still able to see gamma-ray-initiated events with primary energy as low as 100 GeV. and the median is around 13 TeV.

4.3 Test Of The Cut On The Crab Nebula

Figures 21 and 22 show the map of statistical significance in the region around the Crab nebula with the $AX_4P > 3800$ and $AX_4P > 3000$ cuts applied, respectively. A map of the same

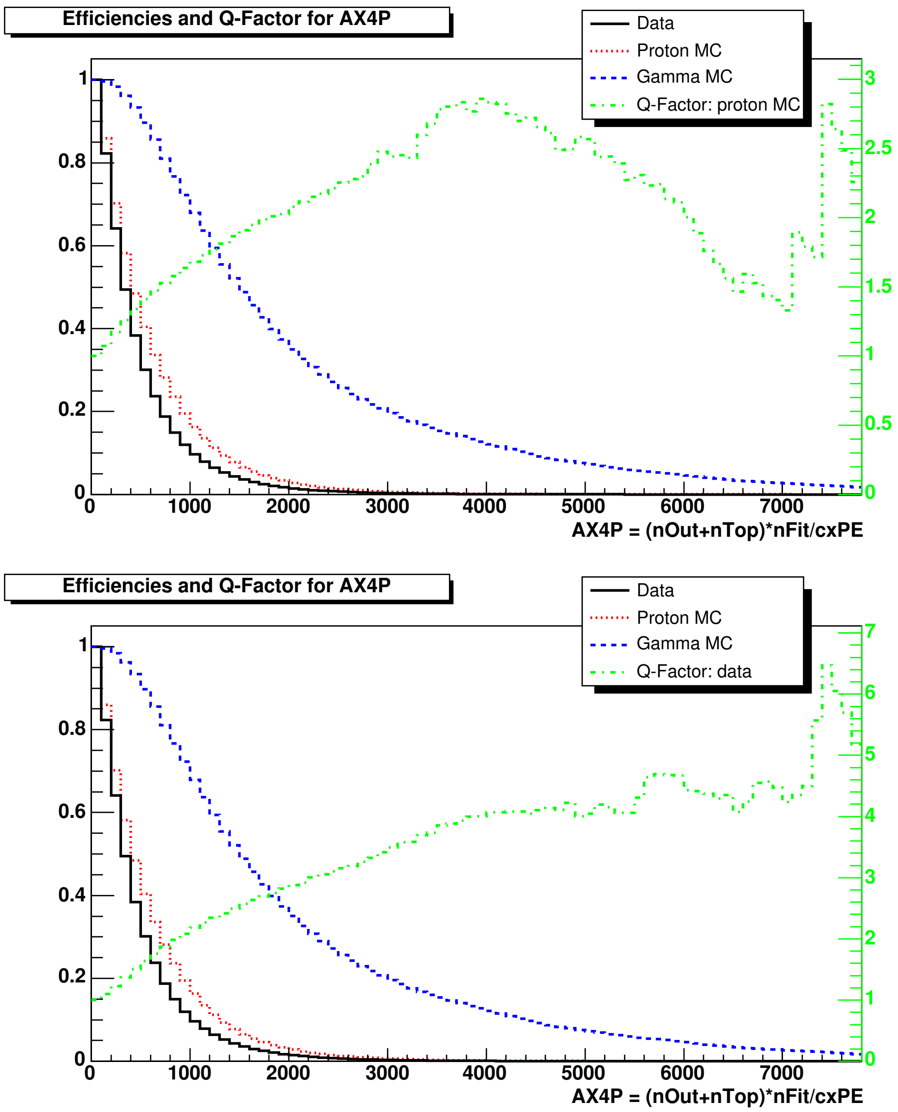


Figure 18: Efficiency and Q-Factor plots for AX_4P. The upper plot shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to proton MC. The plot on the bottom shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to the data.

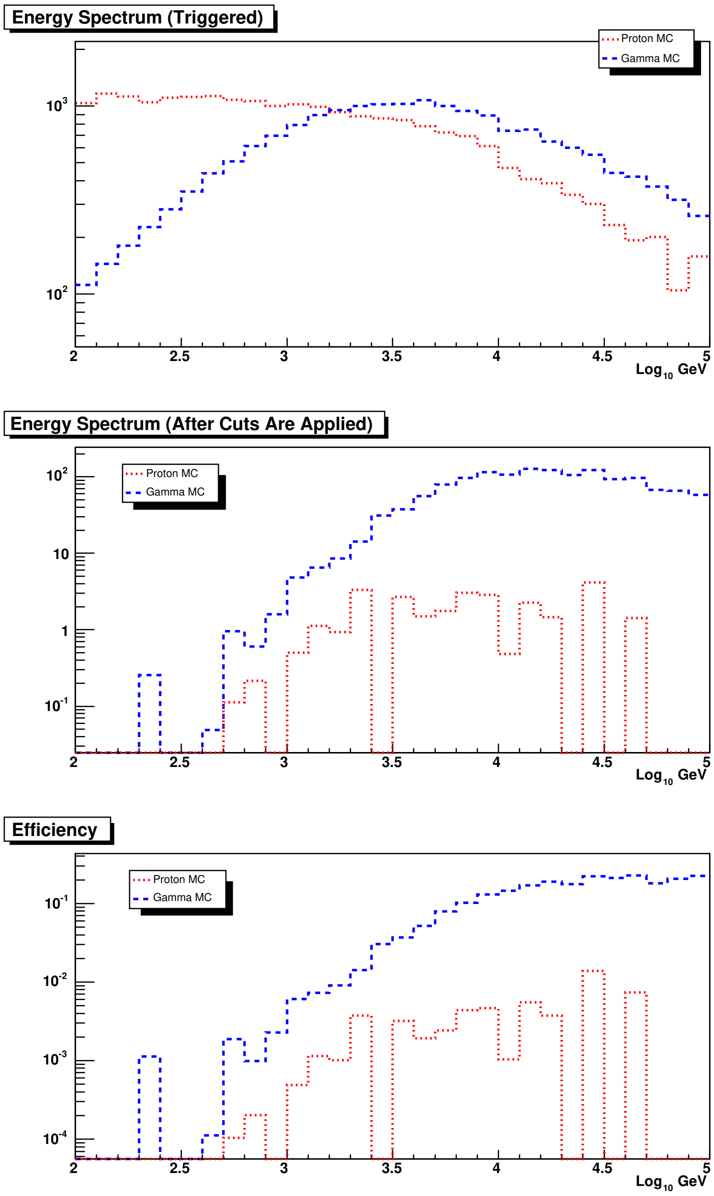


Figure 19: The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after applying the $AX_4P > 3800$ cut, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy.

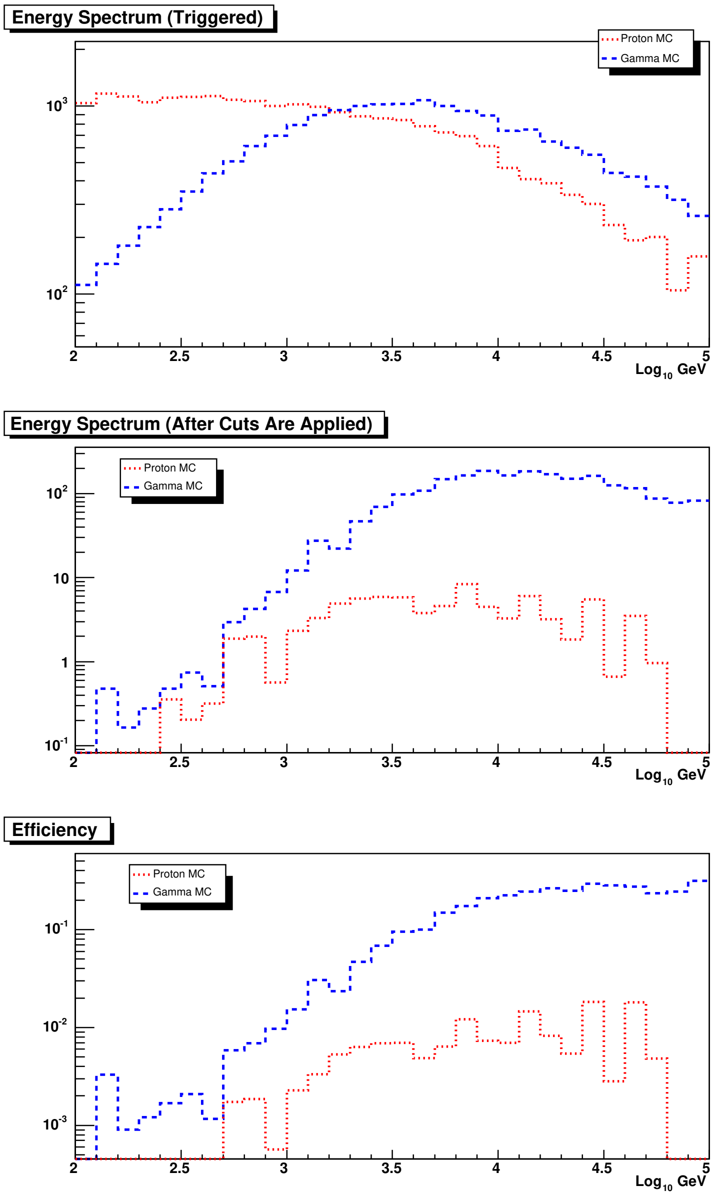


Figure 20: The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after applying the $AX_4P > 3000$ cut, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy.

region with the compactness cut applied is shown in figure 8.[16] From these figures we see that the $AX_4P > 3800$ cut improved the significance of the Crab from 3.64 (for the compactness cut) to 4.14 standard deviations(for the new criteria), a factor of 1.14, while the $AX_4P > 3000$ cut improved the significance of the Crab from 3.64 (for the compactness cut) to 4.74 standard deviations(for our new criteria) a factor of 1.30. Not only does the second cut increases the statistical significance on the Crab, it also increases the number of excess events by a factor of 2 compared to the $AX_4 \geq 3800$ cut .

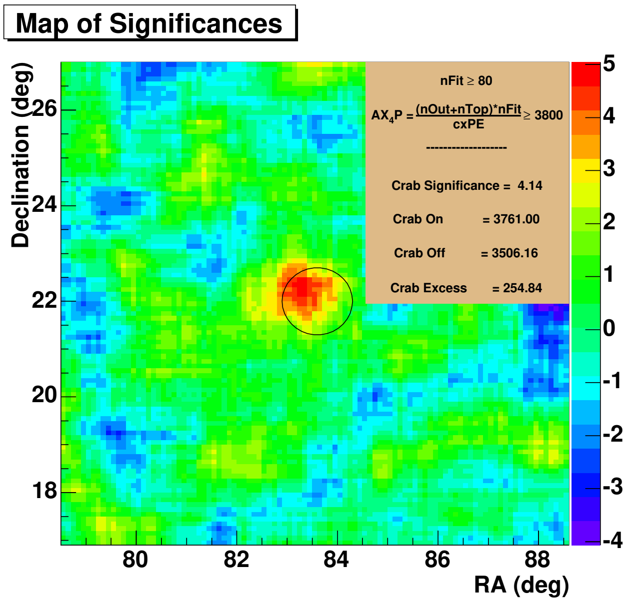


Figure 21: Map of the statistical significance in the region around the Crab Nebula after the $AX_4P > 3800$ cut has been applied.

[16]The maps shown are for the same time period of 416 days of data collection.

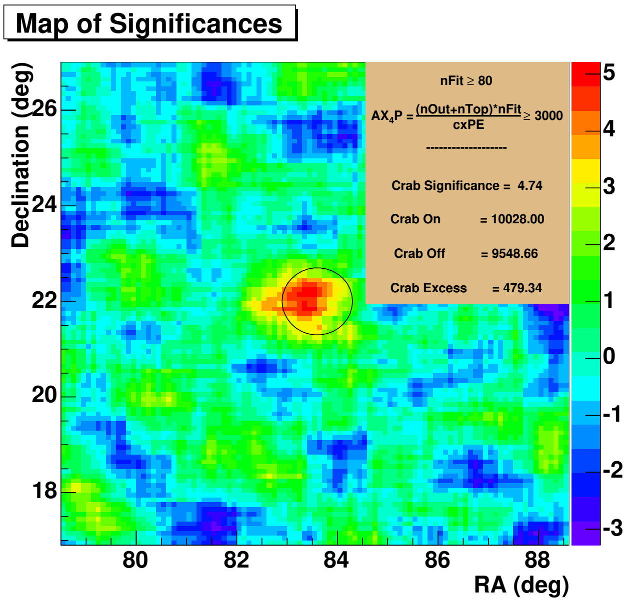


Figure 22: Map of the statistical significance in the region around the Crab Nebula after the $AX_4P > 3000$ cut has been applied.

5 AX_3

In this section the variable AX_3 is studied in more details, where:

$$AX_3 = \frac{nTop^*nFit}{cxPE} \tag{20}$$

The distributions of AX_3 is shown in figure 23 for gamma MC, proton MC, and data. The distribution is shown on the bottom with a log scale for the y-axis. The distribution of proton MC is in reasonable agreement with the data.

In figure 24 two plots are shwon. The upper plot shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to proton MC. The plot on the bottom shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to the data.

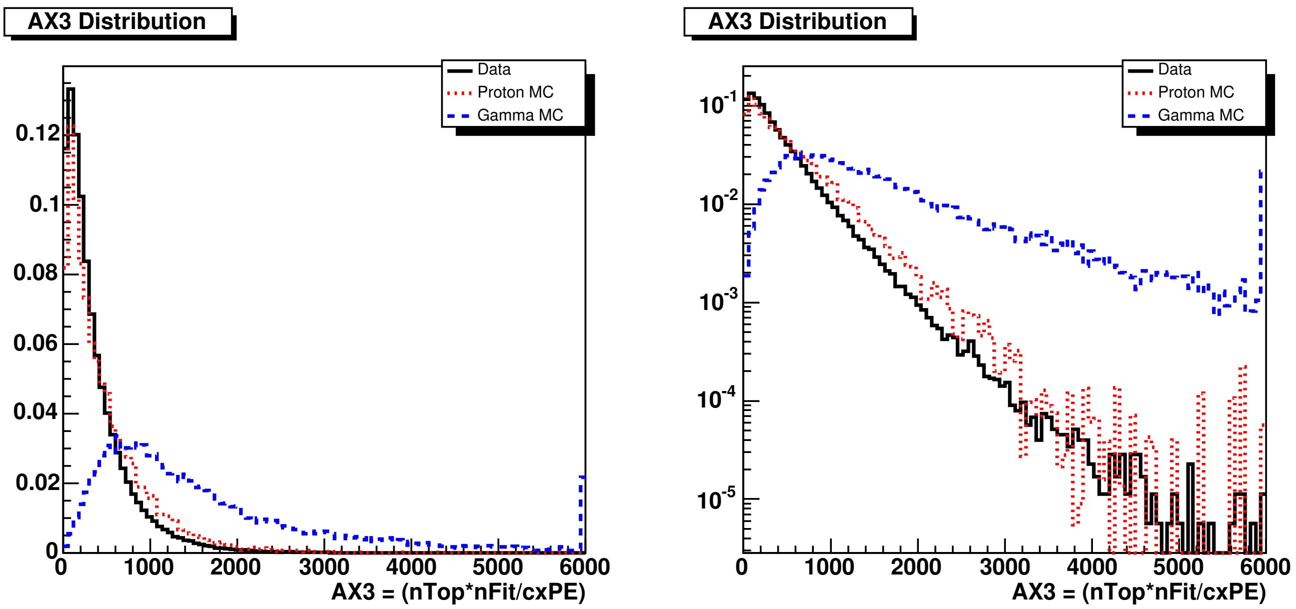


Figure 23: AX_3 Distribution of gamma MC, proton MC, and data. The distribution is shown on the bottom with a log scale for the y-axis. All histograms have been normalized to have unit area.

5.1 Cuts

Table 3 shows a list of cuts applied for AX_3 along with the corresponding efficiencies and Q-Factors. For this analysis two cuts were applied, the first cut is:[17]

$$AX_3 \geq 4000 \tag{21}$$

Requiring events to pass this cut rejects 99.80 % of proton MC and 99.89 % of tha data(for this data sample), and retains 13.2% of gamma MC. This results in a predicted Q-Factor

[17]Event selection cuts were applied for all MC and data files. Please see second paragraph on page 2.

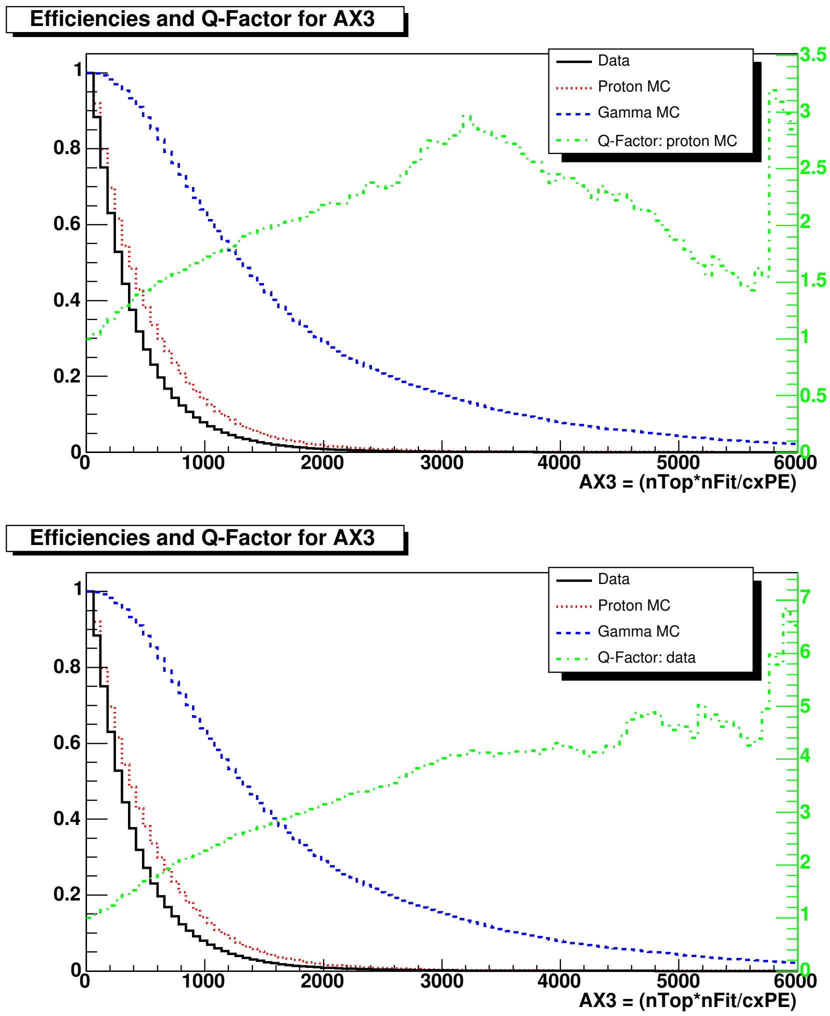


Figure 24: Efficiency and Q-Factor plots for the distributions shown in figure 16. The upper plot shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to proton MC. The plot on the bottom shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to the data.

of 2.96 comparing proton MC to gamma MC events, and 4.09 comparing data to gamma MC. Thus, the Monte Carlo predicts an improvement in significance over all triggered data of 2.96, or a factor of $2.96/1.62 = 1.83$ over the present compactness cut. The second cut applied is:

$$AX_3 \geq 3000 \quad (22)$$

Requiring events to pass this cut rejects 99.70 % of proton MC and 99.86 % of tha data(for this data sample), and retains 14.90 % of gamma MC. This results in a predicted Q-Factor of 2.72 comparing proton MC to gamma MC events, and 4.02 comparing data to gamma MC. Thus, the Monte Carlo predicts an improvement in significance over all triggered data of 2.72, or a factor of $2.72/1.62 = 1.68$ over the present compactness cut.

Cuts Applied	Eff_G	Eff_P	Eff_D	$Q_{MC:MC}$	$Q_{MC:D}$	$\frac{Q_{MC:MC}}{Q_{MC:D}}$
$AX_3 >$						
3000	0.1490	0.0030	0.0014	2.72	4.02	0.68
3200	0.1318	0.0020	0.0010	2.96	4.09	0.72
3700	0.0950	0.0013	0.0005	2.59	4.16	0.62
4000	0.0779	0.0010	0.0003	2.46	4.34	0.57
4200	0.0683	0.0009	0.0003	2.31	4.06	0.57
4500	0.0579	0.0006	0.0002	2.27	4.44	0.51

Table 3: List of cuts applied for AX_3 along with the corresponding efficiencies and Q-Factors. $Q_{MC:MC}$ is the Q-Factor obtained by comparing gamma MC to proton MC, while $Q_{MC:D}$ is the Q-Factor obtained by comparing gamma MC to data. Note that the data efficiency is smaller than the proton efficiency by a factor of 2-3 as the cut is tightened.

5.2 Energy dependence on the Cuts

Figures 25 and 26 show the energy dependence on these cuts. From the second and third plots for both cuts we see that the efficiency of gamma-ray-initiated events is dependent on the energy of the primary gamma ray, while for protons the efficiency is nearly independent of energy We also notice that for the AX_3

> 4000 cut we don't expect to see any gamma-ray-initiated events with primary energy less than 700 GeV, and that the median for gamma-ray-initiated events passing this cuts is around 19 TeV. However with the $AX_3 > 3000$ cut we are still able to see gamma-ray-initiated events with primary energy as low as 100 GeV and the median is around 14 TeV.

5.3 Test Of The Cuts On The Crab Nebula

Figures 27 and 28 show the map of statistical significance in the region around the Crab nebula with the $AX_3 > 4000$ and $AX_3 > 3000$ cuts applied, respectively. A map of the same region with the compactness cut applied is shown in figure 8.[18] From these figures we see that the $AX_3 > 4000$ cut improved the significance of the Crab from 3.64(for the compactness cut) to 4.89 standard deviations(for the new criteria) a factor of 1.34, while the $AX_3 > 3000$ cut improved the significance of the Crab from 3.64(for the compactness cut) to 4.58 standard deviations(for the new criteria) a factor of 1.26. Not only does the $AX_3 \geq 3000$ cut increases the statistical significance on the Crab, it also increases the number of events observed by a factor of 2 compared to the $AX_3 \geq 4000$ cut.

[18] The two maps shown are for the same time period of 416 days of data collection.

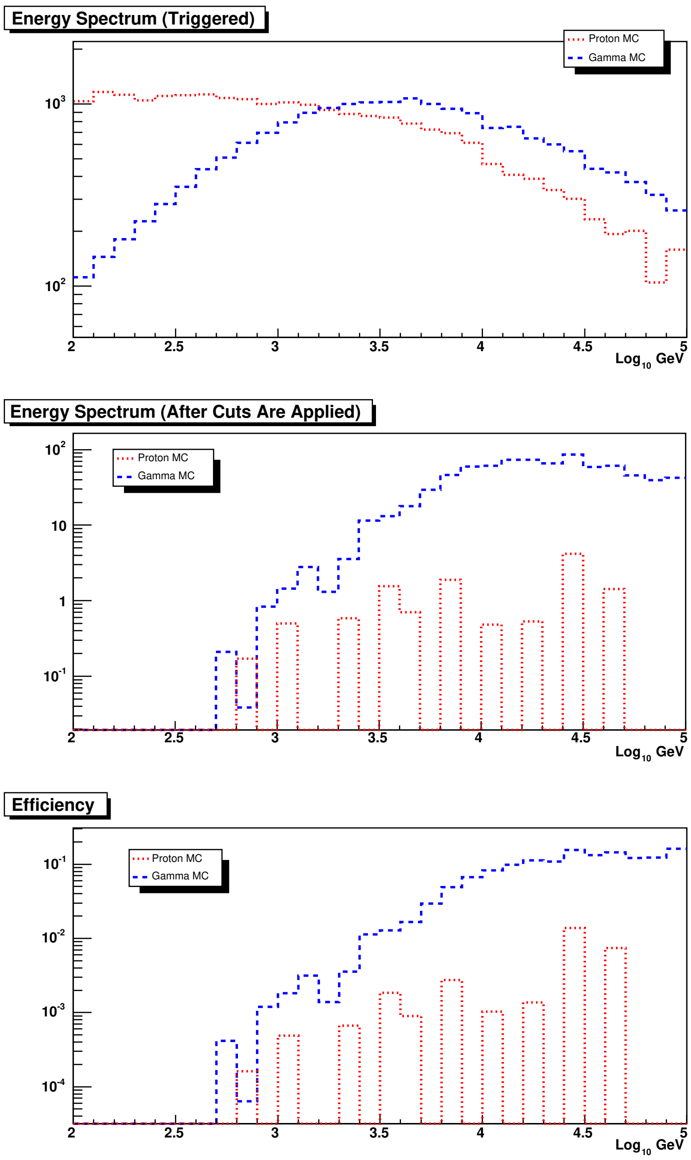


Figure 25: The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after applying the $AX_3 > 4000$ cut, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy.

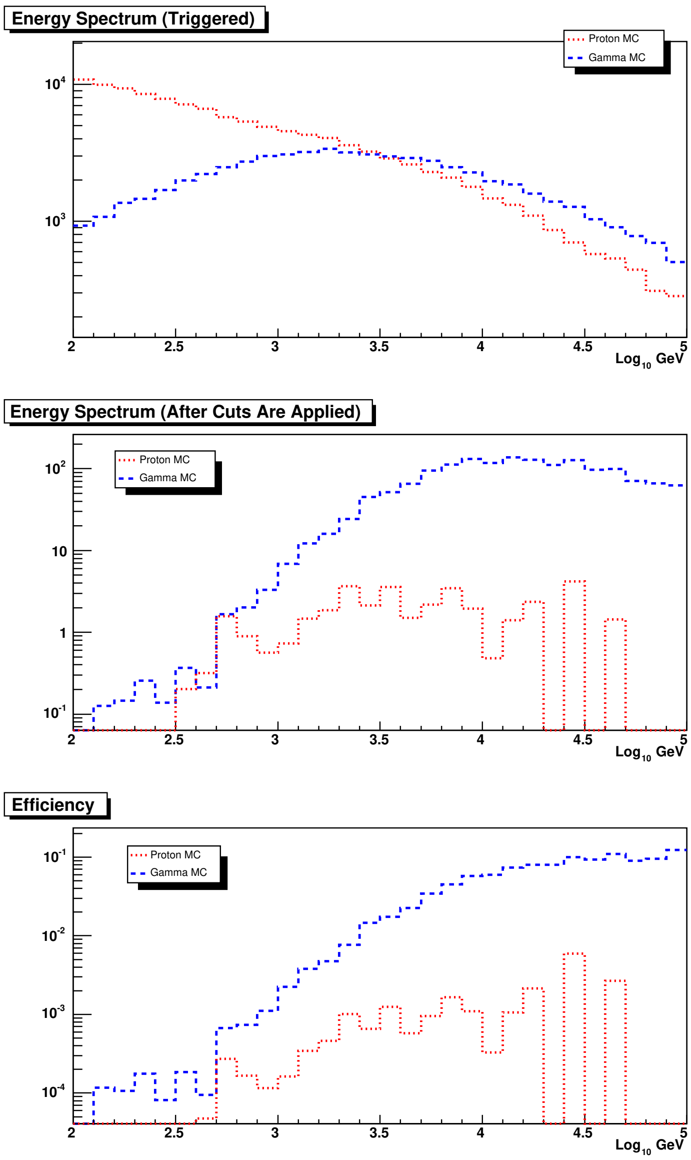


Figure 26: The first plot shows the energy distribution of gamma and proton MC that trigger Milagro, the second plot shows the energy distribution of gamma and proton MC after applying the $AX_3 > 3000$ cut, and the third plot shows the fraction of gamma and proton MC retained by these cuts as a function of primary energy.

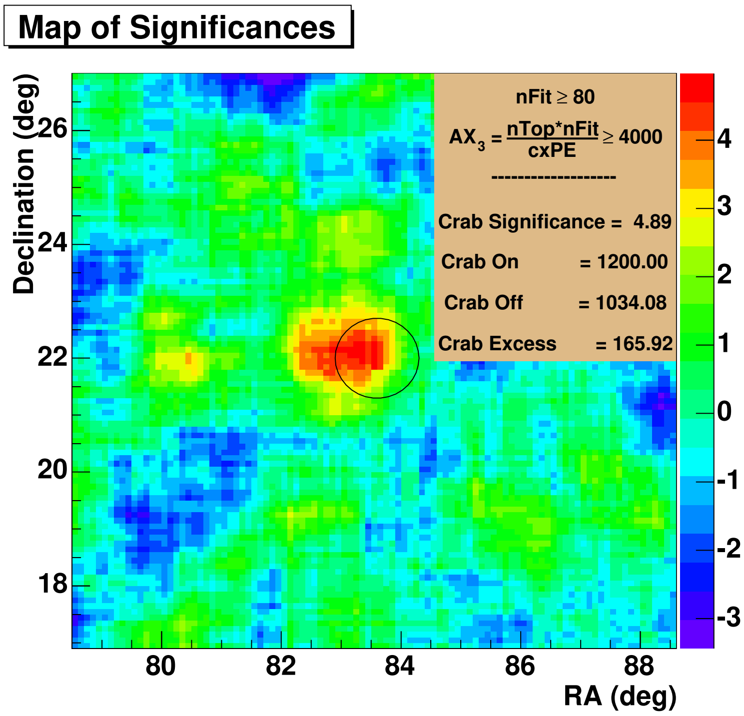


Figure 27: Map of the statistical significance in the region around the Crab Nebula after the $AX_3 > 4000$ cut has been applied.

5.4 On-Pond Events

Figure 29 shows distributions of AX_3 for on the pond events. The distribution is shown on the bottom with a log scale for the y-axis. Figure 30 we show two plots. The upper plot shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to proton MC. The plot on the bottom shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to the data.

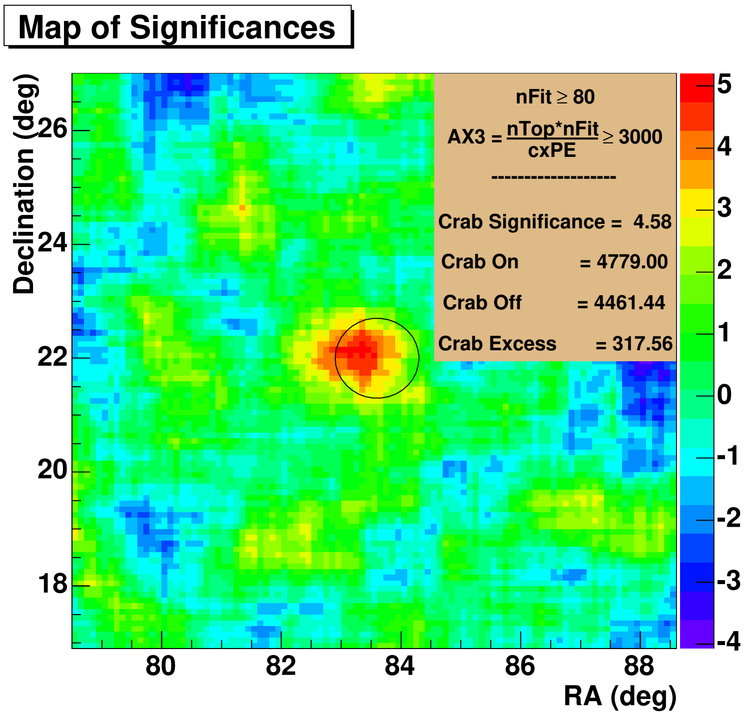


Figure 28: Map of the statistical significance in the region around the Crab Nebula after the $AX_3 > 3000$ cut has been applied.

5.4.1 Cuts For On-Pond Events

For on-pond events two sets of cuts were applied. The first set of cuts applied is identical to the set of cuts applied for all events :

$$AX_3 \geq 4000 \tag{23}$$

Requiring on-pond events to pass these cuts rejects 99.97 % of proton MC and 99.99% of the data (for this data set), and retains 0.66 % of gamma MC. This results in a predicted Q-Factor of 0.38 comparing proton MC to gamma MC events, and 0.75 comparing data to gamma MC.

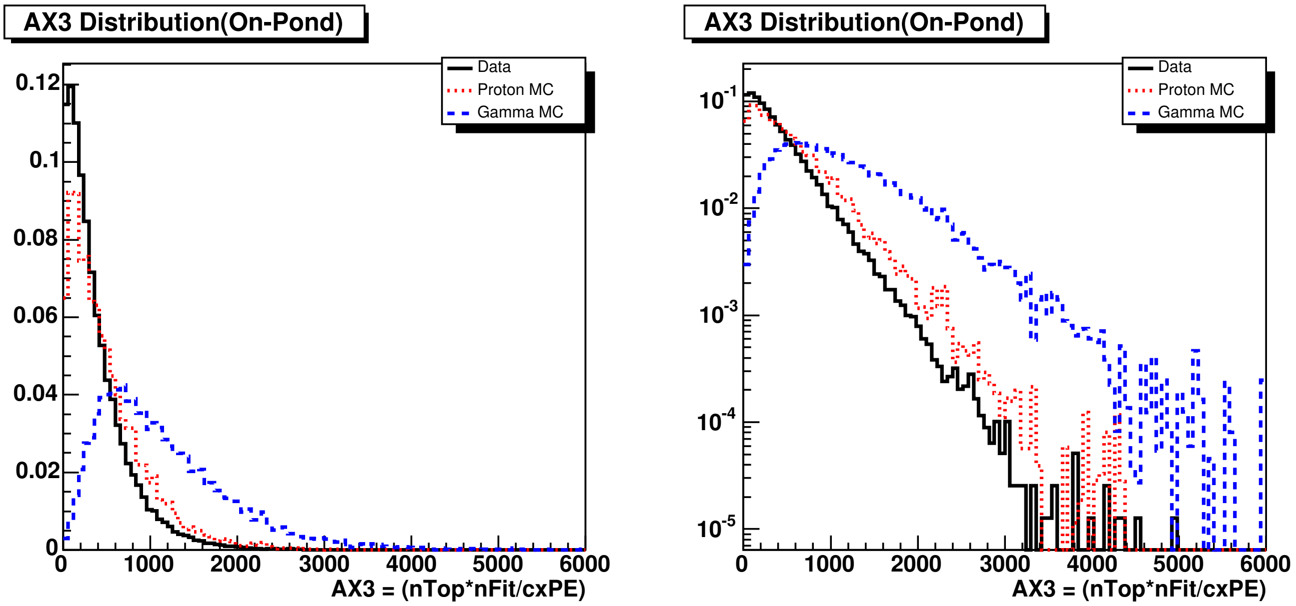


Figure 29: AX_3 Distribution of gamma MC, proton MC, and data for on-pond events. The distribution is shown on the bottom with a log scale for the y-axis.

The second cut applied was:

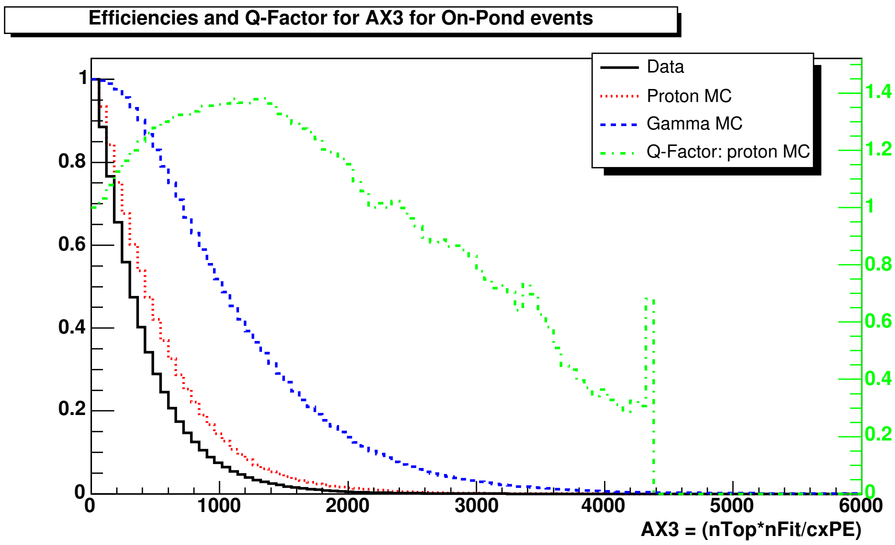
$$AX_3 \geq 800 \tag{24}$$

Requiring on-pond events to pass these cuts rejects 78.9 % of proton MC and 88.2 % of the data(for this data set), and retains 62.2 % of gamma MC. This results in a predicted Q-Factor of 1.35 comparing proton MC to gamma MC events, and 1.81 comparing data to gamma MC.

5.4.2 Test of On-Pond Cuts On The Crab Nebula

When applying the above cuts for on-pond events from the Crab[19] we get statistical significances of 0.86 and 2.65, respectively.

[19]For the same time period of 416 days.



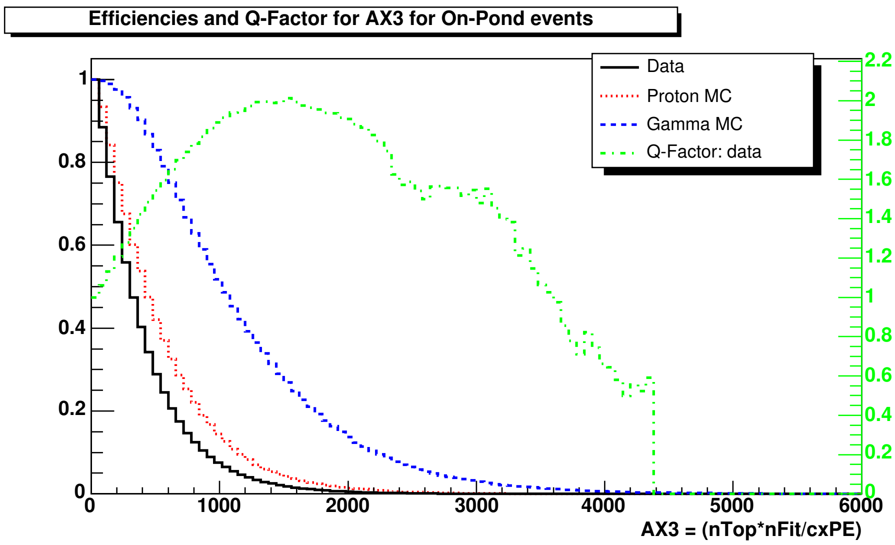


Figure 30: Efficiency and Q-Factor plots for the distributions shown in figure 29. The upper plot shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to proton MC. The plot on the bottom shows the efficiency distributions for gamma MC, proton MC, and data along with the Q-Factor obtained by comparing gamma MC to the data.

From the above we see that although the $AX_3 > 4000$ cut gave us good results for the case of all-events, this cut is not good for on-pond events and we may have to apply different sets of cuts for on and off the pond events.

6 Summary of Results

In this section we show a summary table and discuss some of the results shown in the table. For each cut applied, the table shows the following:

Eff_G	:	Efficiency of the cut for gamma MC.
Eff_P	:	Efficiency of the cut for proton MC.
Eff_D	:	Efficiency of the cut for data.
$\mathrm{Q}_{\mathrm{MC:MC}}$	:	Q-Factor predicted by comparing gamma MC to proton MC.
$\mathrm{Q}_{\mathrm{MC:D}}$	:	Q-Factor predicted by comparing gamma MC to data.
$\mathrm{Z}_{\mathrm{ACT}}$	:	Actual significance from the Crab data.
Z_{MC}	:	Significance predicted from MC and X_2 data.
$\mathrm{Z}_{\mathrm{scaled}}$	:	Adjusted significance prediction(see below)
Excess	:	Signal counts beyond background.
Flux	:	(Excess)/(Eff_G).
Flux Ratio	:	(Flux for the cut)/(Flux for X_2).

Z_{MC} is the expected significance based on the observed significance from the relevant X_2 cut multiplied by the ratio of predicted MC Q-Factors for the cut in question and X_2. i.e. :

$$\mathrm{Z}_{\mathrm{MC}} = \mathrm{Z}_{\mathrm{X}_2} \frac{\mathrm{Q}_{\mathrm{MC}}}{\mathrm{Q}_{\mathrm{X}_2}} \tag{25}$$

$\mathrm{Z}_{\mathrm{MC}}/\mathrm{Z}_{\mathrm{ACT}}$ examines how well the MC does at calculating the ratios of efficiencies.

The intention of the Flux Ratio column is to compare the excess for various cuts after applying the MC efficiency correction. This may give some insight into the relative accuracy of the MC efficiency calculations for various variables.

Marked in italics are the cases where $\mathrm{Z}_{\mathrm{ACT}}$ for a new cut is better than the corresponding X_2 cut for nFit = 80.

Z_{scaled} is an empirical attempt to adjust the predicted Z by basing the value not on the observed Z for X_2 but for the Z from the first cut of a given variable. Then the Z_{scaled}/Z_{act} column measures how well the MC does at predicting the efficiencies of the particular variable for different cut values, rather than comparing with the X_2 efficiencies.

A few of the on-pond efficiencies are not very well determined (30% or worse) and are marked in italics; these are all cases of un-optimized cuts from all events unsuccessfully applied to the on-pond case.

6.1 General Observations

For each of the variables introduced, a higher statistical significance than for X_2 was predicted, and actually found, in the Crab data. Hard cuts in the new variables look at high energy photons. The new variables do not perform well for data on the pond.

The 2-D variable (AX_4, CX_2) gives the best improvement over X_2: $5.05/4.01 = 1.26$ after an nFit=80 cut for both. The measured significance is not quite as large as the predicted: $2.65/1.67 = 1.59$ (a factor of 1.26 still higher than observed) but this is still quite a significant improvement. The 1D variables, AX_4P and AX3 reached measured Z = 4.74 and 4.89 respectively. Each improvement is a factor of 1.3 to 1.4 lower than MC would have predicted.

The flux (photon excess corrected for MC gamma efficiency) for the new variables is typically a factor of 2 lower than that for X_2. We used the flux after the X_2 cut as a reference [20] for all events, or for on-pond events as the case may be. The efficiency of a cut for data is typically 2-3 times smaller than for proton MC. This suggests that the efficiency for the new variables, at

[20] The excess before cuts have large statistical errors.

least when cutting hard, is lower than the MC predictions (or that the X_2 efficiency is larger than the MC predictions).

A better understanding of the MC and especially the gamma efficiencies is needed to use these variables for an absolute flux measurement. This could be in part due to the incorrect values of the Quantum Efficiency of the PMTs used in the simulation; a new simulation with the correct quantum efficiencies are being generated in Maryland now.

6.2 Crab paper vs. 416 days

As a reference we show the results from the Crab paper[3].

The Crab paper quoted a significance of 6.4 for an effective exposure of some 917 days; scaling to the effective exposure (courtesy of Andrew Smith) of 385 days leads us to expect about 4.15 . One might hope that this is the expected significance for the nFit=20 data, but the nFit used in this analysis contains tubes from both the muon and outriggers, so the signal to background ratio is worse and we should expect a smaller significance. In addition, the bin size was larger for the Crab paper analysis and is if anything inappropriately small for nFit=20 data as the choice was made for nFit=80. Applying 4.15 to the nFit=80 data is also complicated: the significance is expected to improve by the square root of the bin area ratio improvement due to outriggers (about 1.7), but the efficiency for the nFit=80 (this note) is expected to be lower than for the Crab paper nFit=20 (old definition of nFit). Also, the different trigger selection used in the Crab paper was different.

7 Comparisions for X_2 Results

One problem in comparing the significance before and after an X_2 cut is that the excess before cuts has a large error bar, making it an unsuitable reference value. As a striking example, notice the un-physical change in the excess before cuts between nFit=20 and nFit = 80 for the 416 day sample: the real excess of photons cannot have doubled by applying an nFit cut. Thus we will use the excess after the X_2 cut as a better reference value. As seen in the flux ratio column, the apparent flux is significantly larger after the X_2 cut in both the Crab paper and the independent 416 day sample (both for all events and the on-pond sample). Perhaps this is a coincidence; certainly the statistical uncertainty prevents any strong conclusions. The problem of apparent X_2 efficiency > 1 does not occur in the nFit > 80 data.

The on-pond X_2 flux is 90% of the total flux in both nFit = 20 and nFit = 80 data. This is surprising, as the efficiencies are defined with respect to the number of photons which pass the other cuts and are incident anywhere, and on the pond, respectively. So according to the Monte Carlo, we would have expected the on-pond flux (efficiency-corrected excess) to be roughly a third of the total flux, rather than 90%.

7.1 nFit = 20 vs. nFit = 80

For the 416 day data, raising the cut from nFit=20 to 80 raises the significance after the X_2 cut by a afctor of $4.01/3.64 = 1.1$ (and the same factor for on-pond data). However, the nFit=20 data was analyzed with the bin size optimized for nFit = 80, so it is at a somewhat unfair disadvantage.

The tighter nFit cut also doubles the apparent excess without cuts, on and off the pond, decreasing the discrepancy between the flux before and after cuts. Since the excesses without cuts are smaller than the excesses after cuts, this is un-physical and is completely dominated by the large (40 to 90%) statistical errors on the excesses before cuts.

7.2 On vs. Off the Pond

The situation with the on-pond data is not very satisfactory. None of the new variables produces a significance better than X_2. We are not sure whether this is because of the nature of the variables chosen (all involve cxPE), but it does appear that the new variables are better applied to the data off the pond. In all variables, it is harder to distinguish gammas from protons for on-pond data than for off-pond data. This is in some sense surprising as the data should have most information about shower shapes when the bulk of the shower is on the pond .

8 Followups

8.1 Analysis

- Rerun on reconstructed data with Airshower and Outrigger layers in the angle fitter only; will require revisiting nFit cuts (smaller than 80, probably) and perhaps choice of variables.
- Are any of these new variables of any use applied to old data without outriggers?
- Redo nFit = 20 vs. 80 comparison with appropriate bin size for both.

8.2 Understanding the Variables

- Plot correlations among AX_4, AX_4P, AX_3.
- Try separate analyses for off pond only data; use X_2 for on pond.
- Would mxPE instead of cxPE work better or worse with our numerators? Try on-pond and off-pond separately.

8.3 Monte Carlo

New MC samples should be used, with He, the new noise model, quantum efficiency, angular response. The new Geant4 MC angle optimization should be used. The data selection should use the rise time trigger on the MC if it has been certified. Possibly throw events beyond 45 degrees zenith, though it may not matter for cuts tuned to high energy.

Some things to check with the new MC: does it do a better job of predicting these variables? Does it do a better job of predicting the on-pond fraction after the data selection? The noise model might be a particularly important effect of the outrigger contribution to these variables.

8.4 Energy Studies

In a future memo we intend to study further the energy dependence properties of these variables.

- Try cutting directly on E for MC: what is best possible Q if had a perfect E variable.
- Plot correlation of variables vs E.
- What does the nFit cut do to energy: E distribution for 20 and 80? E distribution on and off-pond.

9 Questions for the Collaboration

- Is the off-centerness of some of the Crab plots significant?
- Any ideas on why most of the flux would show up as on-pond?

- Why are on-pond showers harder to distinguish? This is in some sense surprising as the data should have most information about shower shapes when the bulk of the shower is on the pond.

- Has the 416 day Crab Flux been compared to the value from the paper?

- Why is off-pond rejection so much better? Is it because the number of outriggers is smaller for on-pond events? because the angular resolution is worse, so that the bin size is actually too small for these events? Or is the energy systematically lower for on-pond events?

Cut(s) Applied	Eff_G	Eff_P	Eff_D	$Q_{MC:MC}$	$Q_{MC:D}$	$\frac{Q_{MC:MC}}{Q_{MC:D}}$	Z_{ACT}	Z_{MC}	$\frac{Z_{MC}}{Z_{ACT}}$	Z_{scaled}	$\frac{Z_{scaled}}{Z_{ACT}}$	Excess	Flux	Flux Ratio
Crab Paper														
No cuts	1.00	1.00	1.00	1.00	1.00		1.90	1.90				8342	8342	
$X_2 > 2.5$	0.51	0.10	0.09	1.61	1.75	0.92	6.40	3.06	0.48			**9717**	**19053**	**2.28**
416 days nFit >= 20														
No cuts	1.000	1.000	1.000	1.000	1.00		1.28	1.28				2442	2442	
No cuts (On-Pond)	1.000	1.000	1.000	1.000	1.00		1.14	1.14				1584	1584	
$X_2 > 2.5$	0.5016	0.0902	0.0791	1.67	1.78	0.94	3.64	2.14	0.59			2078	**4143**	**1.70**
$X_2 > 2.5$ (On-Pond)	0.3405	0.0779	0.0754	1.22	1.24	0.98	3.05	1.41	0.46			1330	**3907**	**2.47**
416 days nFit >= 80														
No cuts	1.000	1.000	1.000	1.000	1.00		2.65	2.65				4042	4042	
No cuts (On-Pond)	1.000	1.000	1.000	1.000	1.00		2.60	2.60				2727	2727	
$X_2 > 2.5$	0.5352	0.1287	0.1031	1.49	1.67	0.89	4.01	3.95	0.99	2.84	0.71	2078	**3883**	**0.96**
$X_2 > 2.5$ (On-Pond)	0.3559	0.1225	0.1008	1.02	1.12	0.91	3.42	2.64	0.77	1.94	0.57	1247	**3504**	**1.28**
AX_4 & CX_2														
$AX_4 > 10^5$ & $CX_2 > 4.28$	0.1388	0.0027	0.0021	2.65	3.00	0.88	5.05	7.02	1.39	5.05	1.00	403	2903	0.72
On-Pond														
$AX_4 > 10^5$ & $CX_2 > 4.28$	0.0130	0.0003	0.0002	0.74	0.94	0.78	1.77	1.92	1.08	1.41	0.80	42	**3229**	**1.18**
$AX_4 > 3200$ & $CX_2 > 2.1$	0.6590	0.2876	0.1723	1.23	1.59	0.77	2.55	3.20	1.25	2.34	**0.92**	1119	1698	0.62
AX_4P & CX_2														
$AX_4P > 3800$ & $CX_2 > 4.00$	0.1304	0.0021	0.0011	2.83	3.95	0.72	3.94	7.51	1.91	5.40	1.37	240	1840	0.46
$AX_4P > 3800$	0.1333	0.0023	0.0011	2.77	4.00	0.69	4.14	7.33	1.77	5.28	1.27	255	1912	0.47
$AX_4P > 3000$	0.1961	0.0065	0.0032	2.44	3.49	0.70	4.74	6.47	1.37	4.66	**0.98**	479	2444	0.60
On-Pond														
$AX_4P > 3800$ & $CX_2 > 4.00$	0.0164	0.0005	0.0002	0.72	1.23	0.59	0.39	1.88	4.83	1.38	3.54	8	487	0.18
$AX_4P > 800$	0.6614	0.2409	0.1382	1.35	1.78	0.76	2.00	3.50	1.75	2.57	1.28	794	1200	0.44
AX_3														
$AX_3 > 4000$	0.0779	0.0010	0.0003	2.46	4.34	0.57	4.89	6.52	1.33	4.69	**0.96**	166	2130	0.53
$AX_3 > 3000$	0.1490	0.0030	0.0014	2.72	4.02	0.68	4.58	7.21	1.57	5.19	1.13	318	2131	0.53
On-Pond														
$AX_3 > 4000$	0.0066	0.0003	0.0001	0.38	0.75	0.51	0.86	1.00	1.16	0.73	0.85	9	1422	0.52
$AX_3 > 1400$ & nb2 > 100	0.1994	0.0181	0.0085	1.48	2.16	0.69	2.80	3.86	1.38	2.83	**1.01**	307	1539	0.56

Table 4: Summary Table

Cuts Applied	Median (TeV) Gamma MC	Median (TeV) Proton MC
Triggered (55 Mult.)	3.72	1.00
Triggered & nFit >= 80	1.63	1.25
Trig. & nFit > 20 & X_2 > 2.5	5.12	1.37
Triggered & nFit >= 80 &		
$AX_4 > 10^5$ & $CX_2 > 4.28$	13.44	7.60
$AX_4P > 3800$ & $CX_2 > 4.00$	15.60	6.04
$AX_4P > 3800$	15.83	6.60
$AX_4P > 3000$	12.90	6.10
$AX_3 > 4000$	18.60	7.11
$AX_3 > 3000$	14.37	4.69

Table 5: Median energies.

References

[1] Abdo, A., New Gamma-Hadron Separation Technique and Variables in Milagro
Milagro memo.

[2] Sinnis, G., miniHAWC: A First Step Toward HAWC
Milagro memo.

[3] Atkins, R.,et al, 2003, APJ, 595, 803

APPENDIX

10 Proof of the equation $(t^i/t^j) = (Q^j/Q^i)^2$

The proof of the above argument is the following:
Suppose we have two sets of cuts i and j that give us Q-Factors of Q^i and Q^j on the same set of data(same time of data collection), respectively. And suppose that

$$\mathrm{Q}^i = \beta\ \mathrm{Q}^j. \tag{A-1}$$

where β is the ratio between the two Q-Factors. The statistical significance σ is defined as the ratio of the signal(S) to the square root of the background(B):

$$\sigma = \frac{\mathrm{S}}{\sqrt{\mathrm{B}}}. \tag{A-2}$$

The relation between the Q-Factor and the statistical significance σ is:

$$\mathrm{Q}^i = \frac{\sigma^i}{\sigma^o}. \tag{A-3}$$

where σ^i is the statistical significance after the cut i has been applied, and σ^o is the statistical significance before any cuts have been applied.

From equations A-1, A-2, and A-3 we get:

$$\frac{\sigma^i}{\sigma^o} = \beta\ \frac{\sigma^j}{\sigma^o} \tag{A-4}$$

$$\sigma^i = \beta\ \sigma^j \tag{A-5}$$

For a steady source the signal and background are linearly proportional to the time of observation(t):

$$S(t) = R_S t \tag{A-6}$$

$$B(t) = R_B t \tag{A-7}$$

where R_S (R_B) is the rate of the signal(background) with respect to time

$$R_S = \frac{d\mathrm{S}}{dt} \tag{A-8}$$

$$R_B = \frac{d\mathrm{B}}{dt} \tag{A-9}$$

From equations A-2, A-5, A-6, and A-7 we have:

$$\frac{S^i(t)}{\sqrt{B^i(t)}} = \beta \frac{S^j(t)}{\sqrt{B^j(t)}} \tag{A-10}$$

$$\frac{R_S^i t}{\sqrt{R_B^i t}} = \beta \frac{R_S^j t}{\sqrt{R_B^j t}} \tag{A-11}$$

$$\frac{R_S^i}{\sqrt{R_B^i}} = \beta \frac{R_S^j}{\sqrt{R_B^j}} \tag{A-12}$$

To reach the same level of statistical significance with the two sets of cuts, i and j, we should have:

$$\sigma^i(t^i) = \sigma^j(t^j) \tag{A-13}$$

where t^i and t^j are the observation times required to reach the desired σ^i and σ^j, respectively.

From equations A-2, A-6, A-7, and A-13 we have:

$$\frac{R_S^i t^i}{\sqrt{R_B^i t^i}} = \frac{R_S^j t^j}{\sqrt{R_B^j t^j}} \tag{A-14}$$

$$\frac{R_S^i}{\sqrt{R_B^i}}\sqrt{t^i} = \frac{R_S^j}{\sqrt{R_B^j}}\sqrt{t^j} \tag{A-15}$$

Finally, from equations A-12 and A-15 we get:

$$\frac{R_S^i}{\sqrt{R_B^i}}\sqrt{t^i} = \frac{1}{\beta}\frac{R_S^i}{\sqrt{R_B^i}}\sqrt{t^j} \tag{A-16}$$

$$\begin{aligned}
\sqrt{t^i} &= \frac{1}{\beta}\sqrt{t^j} && \text{(A-17)} \\
t^i &= \frac{1}{\beta^2}t^j && \text{(A-18)}
\end{aligned}$$

In the case of comparing the new cut to the compactness cut, β is equal to $5.05/3.64 = 1.39$ and the observation time ,with our cuts, that is required to achieve the same level of σ is $1/1.39^2 = 0.52$ that with the compactness cut.