University of California
Los Angeles

Composition Analysis of Ultrahigh Energy Cosmic Rays Using the Pierre Auger Observatory Surface Detector

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics

by

David Scott Barnhill

2005

The dissertation of David Scott Barnhill is approved.

Kuo-Nan Liou

David Saltzberg

Alexander Kusenko

Katsushi Arisaka, Committee Chair

University of California, Los Angeles

2005

To my lovely wife, Robyn

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I'd like to thank, first off, all the people who have made these last 5 years enjoyable. I am grateful to Tohru for discussions, tutoring, and advice as a comrade in arms here at UCLA and in the trenches in Malargüe. I owe Federico Suarez my life because he has taken charge of the PMT testing and all the administrative work that goes along with it. Without him and his crew (Andres, Augustin, Mariela) I would have been living in Malargüe and would have never met the woman who is now my wife.

Here at UCLA I am indebted to Arun Tripathi who has been working with me from day one. He has taught me a lot whether he is aware of it or not. I am also thankful for the work of Matt Healy, and Joong Lee who have contributed in a big wayto the analysis effort here at UCLA, of which I am a beneficiary.

In reality, I am grateful for every member of the Pierre Auger Collaboration for building a great detector. In particular I give a great big thanks to Paul Clark, Patrick Allison, Rishi Meyhandan, and Jim Beatty for all the assistance in finding various ways to entertain ourselves while at collaboration meetings. In addition, there are those that actually helped direclty in my analysis as well, namely Markus Risse, Cecile Roucelle, Jean-Cristophe Hamilton, and Alan Watson to name but a few.

I also must thank Katsushi Arisaka, my advisor, who has pushed me the whole time I have been here. Each time he asks me to do something I usually think there is no way I can do all that. But it is only because he has required so much of me that I have been able to grow as much as I have. He may also be the one to blame for my stubborn attitude in some respects.

My research has been funded through the U. S. Department of Energy.

Vita

1975	Born, Murray, Utah, USA.
2000	B.S. (Physics) Brigham Young University
2000	Del Sasso Fellowship, UCLA Physics Department
2001	M.S. (Physics), UCLA, Los Angeles, California.
2000	Teaching Assistant, Physics Department, UCLA. Taught sections of Physics 4B (beginning lab teaching electromagnetic principles)
2000-2005	Research Assistant, Physics Department, UCLA.

Publications and Presentations

D. Barnhill, C. Jillings, T. Ohnuki, A. Tripathi, K. Arisaka, F. Suarez, M. Videla, and B. García, Production Test System and Results on Large PMTs for Pierre Auger Surface Detectors in *Proceedings of the 28th ICRC*, Tsukuba Japan, 2003

D. Barnhill, Current Status of the Pierre Auger Observatory in *The Meeting of the Division of Particles and Fields of the American Physical Society*, Riverside, California, August 26-31, 2004

A. K. Tripathi, S. Akhanjee, K. Arisaka, D. Barnhill, C. D'Pasquale, C. Jillings, T. Ohnuki, and P. Ranin, A systematic study of large PMTs for the Pierre Auger Observatory, Nucl. Instr. and Meth. A **497**, 331 (2003)

Abstract of the Dissertation

Composition Analysis of Ultrahigh Energy Cosmic Rays Using the Pierre Auger Observatory Surface Detector

by

David Scott Barnhill

Doctor of Philosophy in Physics
University of California, Los Angeles, 2005
Professor Katsushi Arisaka, Chair

The origin and composition of ultra high energy cosmic rays has been and continues to be a topic of much study and debate. The Pierre Auger Observatory was designed to investigate the highest energy cosmic rays and resolve some of these problems. In this dissertation, I present a description of the Pierre Auger Observatory and a study of the performance of the surface array as well as work done on the photomultiplier tubes used in the surface array. I also present an analysis done on the composition of the events detected in the surface detector paying special attention to a photon primary assumption. Monte carlo simulations of extensive air showers are put through a simulation of the surface detector and observables are compared to real data. The mean behavior of the real data is compared to various baryonic primary assumptions. For photon primaries, a method is described to set an upper limit on the flux of photons based on comparing real events to expected distributions for photon initiated air showers. An upper limit on the photon flux is presented and compared with predictions from various exotic models of cosmic ray origins.

CHAPTER 1

Introduction

The field of cosmic ray physics is approaching its 100 year anniversary. The birth, then, of cosmic ray physics was in 1912 and Victor Hess is the father. The discovery of cosmic radiation by Victor Hess [1] came through studying ionization rates in a series of balloon flights. He observed that the rate of radiation increased with higher altitudes concluding that the source of radiation was extraterrestrial.

1.1 Cosmic Ray Discoveries

Early particle physics owes many discoveries to cosmic rays. Somewhere in the galaxy, and beyond, particles are being accelerated to energies unattainable by accelerators on earth. This is as true today as it was when many important advancements in particle physics were taking place. This free source of high energy particles hitting earth provided the circumstances for the creation of previously undiscovered particles. When Dirac proposed that there is a sea of electrons with negative energy states, and when an electron leaves this sea it leaves a hole, correctly later interpreted to be an anti-electron or positron, it was Anderson [2] in late 1931 who provided the proof of the existence of a positively charged particle with an identical mass as the electron using cosmic rays. Anderson and Hess shared the Nobel in 1936 for their work.

In fact, when Yukawa proposed the existence of a particle associated with the strong nuclear force, it was two groups studying cosmic rays that first thought they found it. Anderson and Neddermeyer, and at the same time Street and Stevenson [3], announced the discovery of a particle with about the right mass in 1937. It was not until 1947, however, that it was discovered that there were actually two particles with similar masses that abound in cosmic ray air showers, the muon (μ) and the pion (π), with the pion being the particle Yukawa predicted [4].

Finally, in December of 1947, a new type of particle was discovered that was different from the "normal" particles being studied. A new particle with the mass of at least twice that of pions was discovered [5], later called the kaon (K^0). This was the first particle discovered in a series of similar discoveries of particles that were produced on a short time scale (10^{-23} s) but decayed relatively slowly (10^{-10} s). The particles were created via the strong force, but decayed via the weak interaction. They were the first particles with "strangeness" discovered, or consisting of strange quarks, and they were discovered using cosmic rays.

These discoveries and the potential for studying high energy particles fueled the interest in the cosmic ray field. The difficulty was in detecting the cosmic rays. Up to energies of around 1-10 TeV, direct detection of cosmic rays is possible using high altitude detectors due to the high flux of radiation at these energies. At higher energies, the flux drops and a much larger collection area is necessary for reasonable statistics. In 1938, a key discovery was made by Pierre Auger that would enable the cosmic ray community to continue to grow. Pierre Auger and his colleagues [6] recorded coincidences in arrays of particle counters. They were able to record multiple counter coincidences at sea level as well as mountain altitudes using electronics with microsecond timing. From electromagnetic cascade theory, he deduced the energy of these showers to be around 10^{15} eV.

More advances were made by Bassi *et al.* at MIT in 1953 [7] when they used the timing information from arrays of scintillation detectors to reconstruct the original direction of the cosmic ray. In 1963, using the Volcano Ranch array, Linsley reported a cosmic ray with an energy of 10^{20} eV [8]. This corresponds to around 16 J of energy packed into one particle or nucleus! Another step forward was in 1962 when Suga and Chudakov proposed that the atmosphere could be used as a large scintillator for air shower detection [9], and in 1968 when Tanahashi actually detected an air shower with an energy of 10^{19} eV using fluorescence in the atmosphere [10]. And, in a final giant step forward, Volcano Ranch recorded a fluorescence event in conjunction with an event detected by the ground array; one cosmic ray shower detected by two different methods at the same location [11]. It is this hybrid technique that the Pierre Auger Observatory will implement, becoming the largest cosmic ray detector in the world designed to detect cosmic rays of the highest energies.

1.2 Cosmic Ray Physics

When discussing cosmic rays, two fundamental questions arise. What are they? And, where do they come from? Or, to formulate it in a more systematic way, there are three interrelated topics of discussion. There is the matter of the chemical composition and energy distribution of cosmic rays. There is a matter of the propagation of these cosmic rays from the location where they are accelerated or created to the earth. And finally, there is the matter of the arrival direction distribution on the sky and its relation to the distribution of sources of these cosmic rays, whether they are accelerated or created via some exotic phenomenon.

These three topics are not independent. The source of a cosmic ray is directly related to the chemical composition and energy distribution of cosmic rays from that source. For example, if the cosmic rays are heavy nuclei, they must come from a source that has access to heavy elements and that is capable of accelerating them to the observed energies. Also, the arrival direction distribution on the earth will affect the possibility of determining the source of the cosmic ray. Taking into account the propagation of charged particles in magnetic fields, if the sources are distant there may be no correlation of the arrival direction with the source location in the sky.

The energy spectrum contains valuable information that relates to the source of cosmic rays and their propagation to earth. Figure 1.1 is the energy spectrum over many orders of magnitude in energy and flux. The spectrum is an overall power law with a break in the spectrum at around 10^{15} eV, referred to as the "knee" of the spectrum, and another at around 10^{18} eV, referred to as the "ankle". This one plot makes very clear that the sources must generate a power law spectrum. The two features in the spectrum imply significant changes in the characteristics of cosmic rays, whether it be the chemical composition or the location of the sources or a combination of both. Calculating the energy density shows that it is 1 eV/cm^3 while the energy density of starlight and galactic magnetic fields are 0.6 eV/cm^3 and 0.2 eV/cm^3 respectively [12]. Calculating the energy density of the highest energy cosmic rays shows that these particles must be non-thermal due to the enormous energy that would be required. They must be either accelerated or created.

1.2.1 Composition

At lower energies (0.1 - 100 TeV), the composition of cosmic rays can be measured using direct detection techniques, such as spectrometers and calorimeters. Measurements of the composition of cosmic rays at these energies is approximately 50% protons, 25% α particles, 13% CNO, and 13% iron [14]. A more detailed

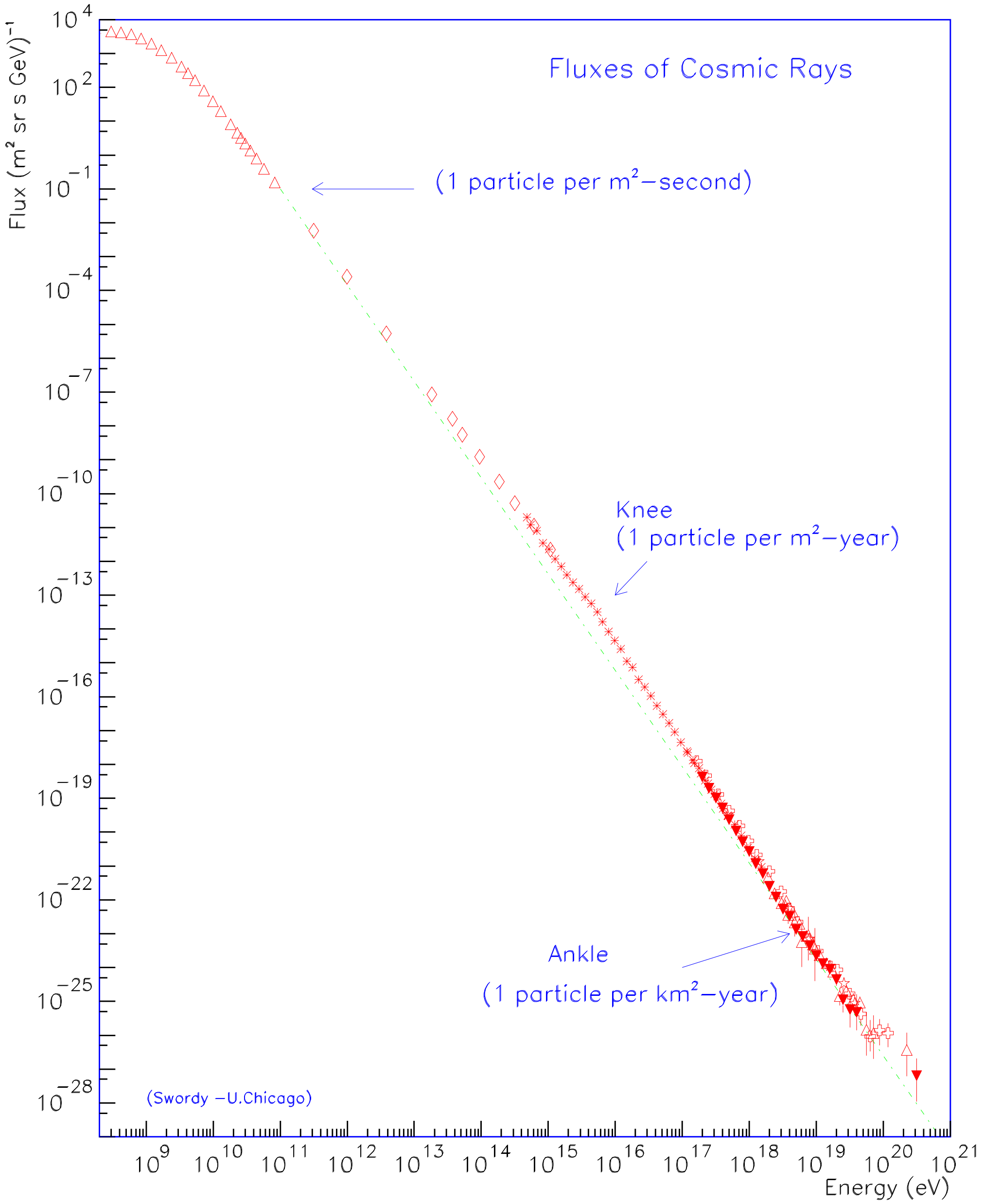


Figure 1.1: Energy spectrum of cosmic ray particles compiled using many different experiments. The dotted line is an E^{-3} power law. Plot is from [13].

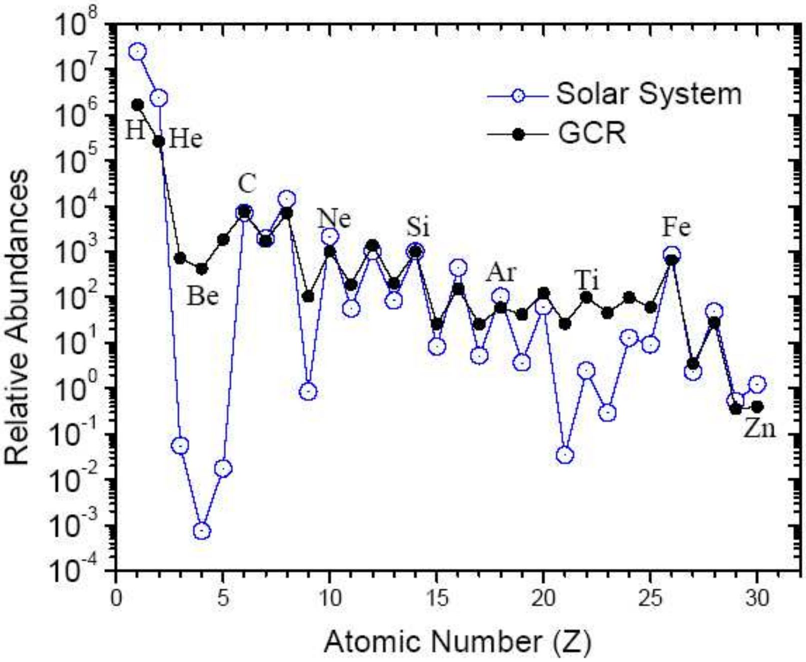


Figure 1.2: Composition of low energy cosmic rays compared to solar system abundances. Figure is from [14].

comparison with solar system abundances shows that cosmic rays are deficient in H and He, which is not fully understood. It may indicate that heavy elements are easier to ionize and accelerate or it is a direct reflection of the source composition. On the other hand Li, Be, and B are too abundant along with Sc, Ti, V, Cr, and Mn. This can be understood as spallation of C and O for the first group and Fe spallation for the second group, see Fig. 1.2.

At higher energies, the flux is too low to measure the composition directly, indirect measurements must suffice. As was explained in Section 1.1, two techiniques are used to measure high energy cosmic rays through the extensive air showers they cause, fluorescence detection and ground particle sampling. In this paper, a study of the composition of ultrahigh energy cosmic rays is discussed using only the array of surface detectors in the Pierre Auger Observatory. Another commonly used method for composition studies employs the depth of shower maximum (X_{max}) as measured by fluorescence detectors. The physics behind extensive air showers will be discussed below, but it is sufficient to state at this point that the depth of shower maximum is related to the primary composition of cosmic rays. Using the data from the HiRes detector, measured X_{max} data is compared to simulated X_{max} values for proton and iron using different hadronic models, see Fig. 1.3 [15]. The data seems to indicate a mixed composition, with a tendency for lighter nuclei at higher energies. Also, there seems to be a shift from heavy to lighter nuclei as energy increases, but not much more can be concluded.

1.2.2 Sources

Having discussed the chemical composition and energy spectrum, the next question to address is the source of the cosmic rays. One possibility is that cosmic rays are protons or other nuclei that are accelerated to the observed energies by the source, or the so-called "bottom-up" method. Another possibility is that these cosmic rays are created at these energies from decays of super heavy dark matter particles or from massive particles released by topological defects or some other exotic phenomena. These theories are called "top-down" models and will be discussed at greater length in Chapter 2. The focus in this chapter is on acceleration models.

Particle acceleration can occur in a direct fashion or in a statistical process. Direct acceleration requires a strong electromagnetic field, and the result is fast acceleration. Another possibility is that there is a strong rotating magnetic field which results in a large electromotive force, EMF. This can trap the particle while

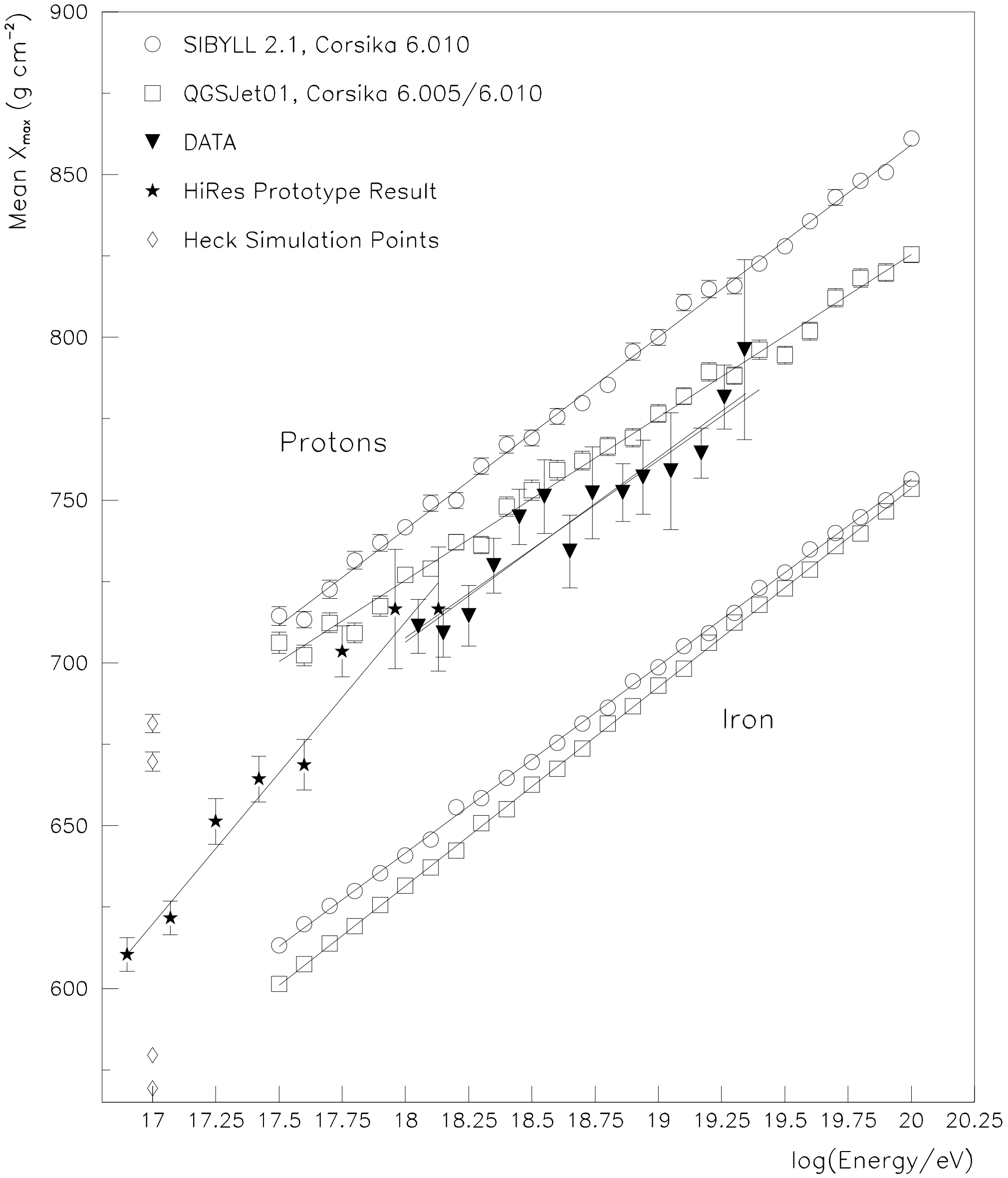


Figure 1.3: A plot of measured X_{max} values compared to simulated values for proton and iron as a function of energy. Figure is from [15].

accelerating it to high energies. There are some problems with this method, however. First, the power law spectrum is not immediately obvious in this scenario. Also, the acceleration occurs in a dense region of space where chances for energy loss are high. Optical photons are dense which leads to meson photoproduction, photonuclear fission, and pair creation. This affects both the energy spectrum and the composition of the resulting cosmic rays.

Another possibility, first proposed by Fermi [16], is a statistical acceleration process. In this process, the build up of energy is slow and takes place over a long period of time compared to direct acceleration. This statistical acceleration can take place in collisions with magnetic clouds or in shockwaves from supernovae, active galactic nuclei (AGNs), or even gamma-ray bursts (GRBs). A benefit of this approach is that the power law is a natural consequence of the acceleration mechanism.

Collisions with magnetic clouds, as first proposed by Fermi, is referred to as second-order Fermi acceleration. It can be understood as particles colliding with magnetic clouds and in these collisions, they can either gain or lose energy. There is a minimum energy that the particle can reach, so on average the particles gain energy in these collisions. It is referred to as second order Fermi acceleration because the acceleration goes as the square of the velocity of the magnetic cloud ($\Delta E/E \sim \beta^2$). It is a slow process and the energy loss from ionization is large for slow particles. Thus, it is difficult to efficiently accelerate particles to high energies.

Shockwave acceleration, however, is much more efficient. It is referred to as first order Fermi acceleration because it is linear with the speed of the shockwave ($\Delta E/E \sim \beta$), resulting in faster acceleration. A shockwave passes through a medium of gas or dust and creates a density gradient at the shock front. The shockwave creates kinetic energy in the medium and there is a resulting net motion as it passes. Particles diffuse and randomly travel in the medium and have a probability to hit the shock front and be accelerated and then scatter back downstream passing the shock front again, gaining more energy. The acceleration continues until energy losses match energy gains, which depends on ambient conditions, and results in a power law spectrum.

In a paper by Drury [17], it was shown that through diffusive shock acceleration the maximum energy attainable is:

$$E = kZeBR\beta c \tag{1.1}$$

where B is the magnetic field of the shockwave, R is the size of the shock region, βc is the shock speed and k is a number less than one, related to efficiency. For example, the case where the acceleration is limited by the age of the shock and not the escape of the particle from the shock region, $k = 3/20$. An easy relationship comes by assuming optimal acceleration, $k = 1$ and $\beta = 1$, which leads to the equation of the highest attainable energy given a region of space and the associated magnetic field:

$$E = 0.9ZBR \tag{1.2}$$

where E is in EeV, B is in μG, and R is in kpc. These equations give a rough estimate of the conditions necessary to accelerate particles to a certain energy. In Fig. 1.4, there is a plot of the magnetic field versus the size of the acceleration region, with certain astronomical objects placed for reference, to show what objects are capable of accelerating particles to the highest energies recorded.

From Fig. 1.4, it is apparent that there are few objects capable of accelerating cosmic rays to the observed highest energies. Another difficulty is that the objects that may be capable of accelerating cosmic rays to the energies of interest are

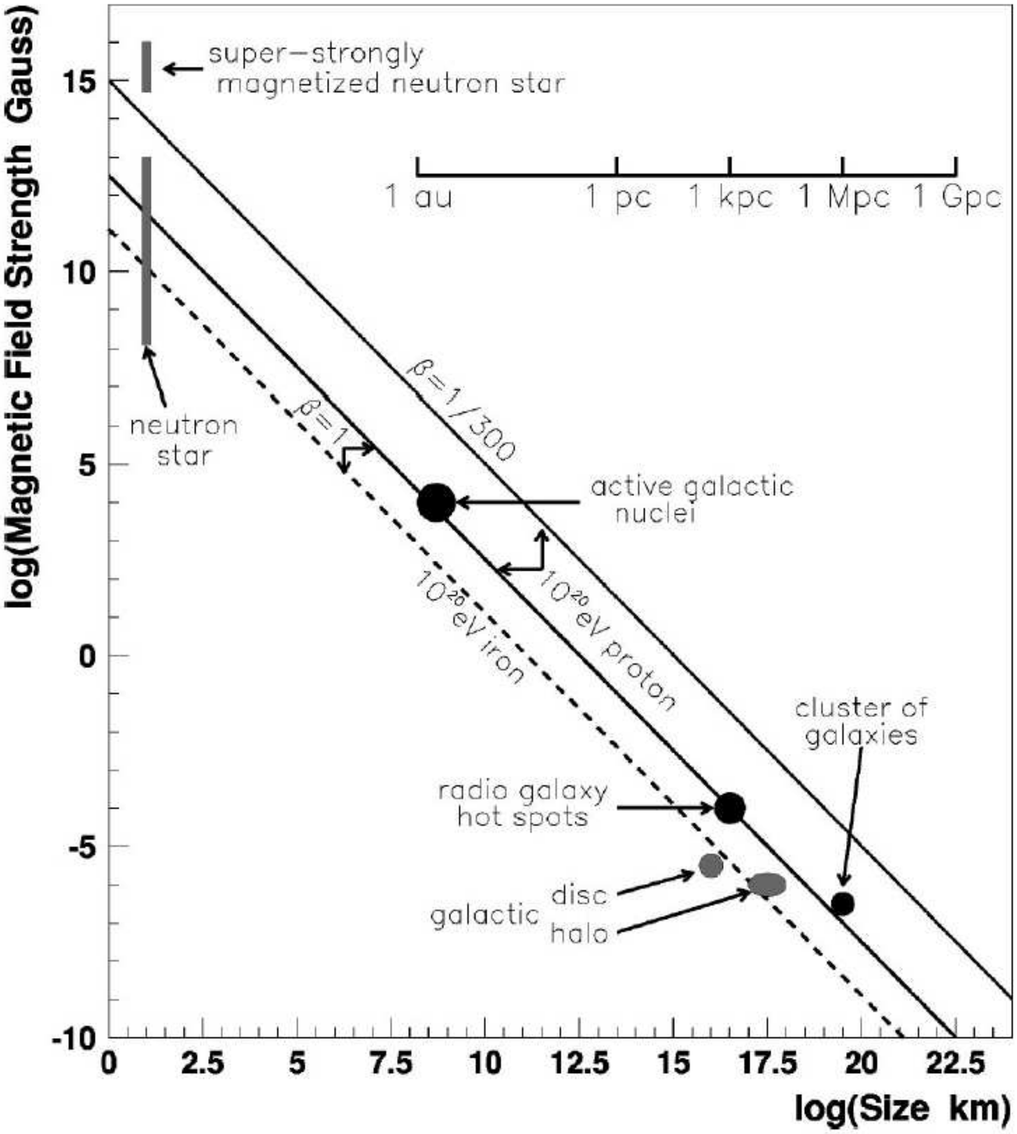


Figure 1.4: Modification of Hillas plot showing possible acceleration sites given the size of acceleration region and corresponding magnetic fields. Figure is from [9].

located at great distances from the earth. Thus, the particles must travel a long way to reach the earth and propagation effects change what is observed on earth from what is generated at the source. Charged particles are bent in intervening magnetic fields to add another level of difficulty in identifying the sources of cosmic rays.

1.2.3 Propagation

Particles traveling through space to reach the earth may interact with ambient radiation, dust, or gases to change both the composition and the energy spectrum of the generated cosmic rays. If the initial cosmic ray is a heavy nucleus, such as iron, there is a probability that the nucleus will photodisintegrate or pair create on the cosmic microwave background radiation (CMBR):

$$A + \gamma_{2.7K} \rightarrow (A-1) + N \tag{1.3}$$

$$A + \gamma_{2.7K} \rightarrow (A-2) + 2N \tag{1.4}$$

$$A + \gamma_{2.7K} \rightarrow A + e^{+} + e^{-} \tag{1.5}$$

All of these interactions result in a loss of energy. The nucleus may interact with the infrared photon background, which is only important below 5×10^{19} eV, while energy loss due to disintegration on the CMBR is important above 2×10^{20} eV, and energy lost in pair creation is dominant in the energy range $5 \times 10^{19} - 2 \times 10^{20}$ eV [18]. The typical attenuation length for Fe and Si in the energy range 40-100 EeV goes from $\sim 10^{3}$ Mpc to that of nucleons. Therfore, above 10^{20} eV the attenuation length is around 10 Mpc [19]. The result is that the observed energy spectrum differs from the energy spectrum at the source, containing features due to the interactions described above and containing fewer higher energy particles than were created. Another effect is that the composition has changed to contain more light nuclei than were present at the source.

For protons, there is the well known feature called the GZK cutoff, named for Greisen, Zatsepin, and Kuz'min [20, 21], who predicted a sharp cutoff in the spectrum due to protons interacting with the CMBR. A proton with an energy of 5×10^{19} eV sees a CMBR photon as a 300 MeV photon, which is the threshold for photopion production. Following the discussion in [22], with the temperature of the CMBR being 2.74 K (corresponding to an energy of 2.36×10^{-4} eV), the delta resonance energy becomes $\sim 10^{20}$ eV for protons ($p + \gamma_{2.7K} \rightarrow \Delta^{+} \rightarrow N + \pi$). Now, using the cross section for the delta resonance as $10^{-28}\mathrm{cm}^{-2}$ and the photon density as $420(1+z^{3})\mathrm{cm}^{-3}$, the mean free path for this interaction is $\sim$8 Mpc. In each interaction, the proton loses about 20% of its energy. After a certain distance, the energy of the proton will decrease to an energy below the delta resonance threshold no matter what energy it started with, see Fig. 1.5.

At lower energies, protons can also interact with the CMBR and pair create ($p + \gamma_{2.7K} \rightarrow p + e^{+} + e^{-}$). This effect is smaller because the energy loss in each interaction is much smaller for the proton. It may, however, contribute to the shape of the spectrum below the GZK cutoff if the primaries are protons from distant sources.

If the primary cosmic ray is a photon, then pair creation with background photons is the dominant form of energy loss. Pair creation with the CMBR is important above 4×10^{14} eV while attenuation from pair creation with the radio background dominates the energy loss above 2×10^{19} eV [24]. The attenuation length for photons with an energy around $\sim 10^{20}$ eV is 10-40 Mpc, depending on the radio background photon density [25].

Figure 1.6 is a compilation of all the previously discussed interactions with the CMBR. The GZK feature is apparent for sources of protons located outside

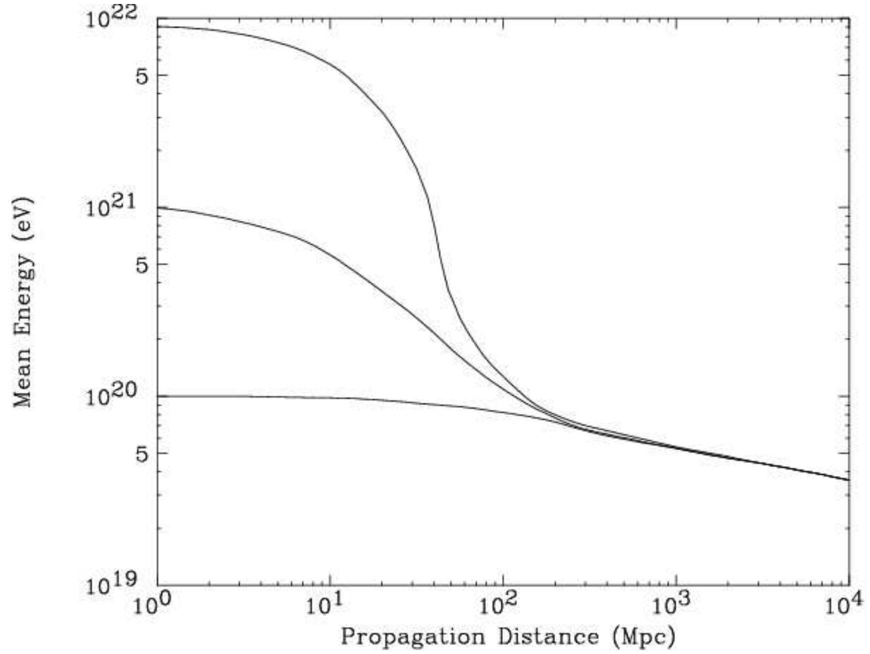


Figure 1.5: The mean energy of protons as a function of distance traveled through the CMBR. The three curves correspond to protons with initial energy of $10^{20}, 10^{21}, 10^{22}$ eV. Figure is from [23].

of the local cluster. Heavy nuclei will experience a more drastic cutoff in the same energy region.

1.3 Extensive Air Showers

Once the cosmic ray reaches the earth, the most efficient method of detection depends on the energy of the particle. For low energies, 0.1 - 100 TeV, direct detection methods are sufficient due to the large flux. Cosmic rays with higher energies, on the other hand, have a much lower flux and require the detection of the extensive air showers that result when cosmic rays interact with the molecules in the atmosphere. The resulting air shower can be detected by observing the

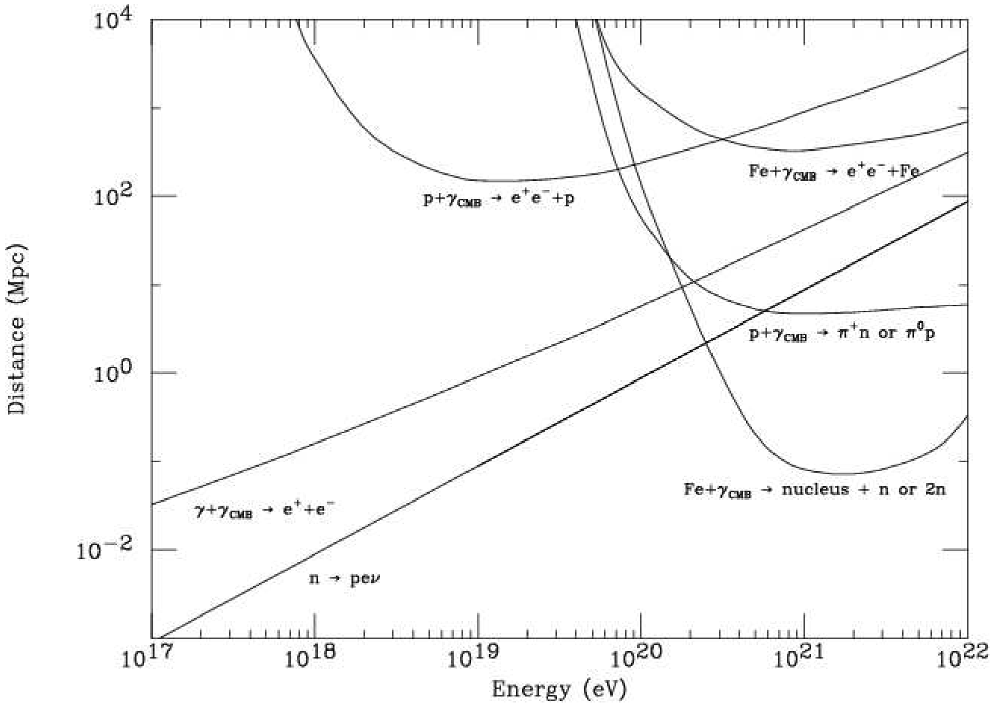


Figure 1.6: Various interactions with the CMBR. The curves labeled $p + \gamma_{CMB} \rightarrow e^+e^- + p$ and $Fe + \gamma_{CMB} \rightarrow e^+e^- + Fe$ are the distances for which the proton and the iron nucleus lose $1/e$ of their energy due to pair production. $p + \gamma_{CMB} \rightarrow N + \pi$ is the mean free path for photopion production. $Fe + \gamma_{CMB} \rightarrow \text{nucleus} + n$ or $2n$ is the mean free path for spallation. $\gamma + \gamma_{CMB} \rightarrow e^+e^-$ is the mean free path for pair creation for photons with the CMBR. $n \rightarrow pe\nu$ is the mean decay length for a neutron. Figure is from [23].

flourescence caused by electrons and positrons exciting the nitrogen or by measuring the shower particles that reach the ground.

1.3.1 Electromagnetic Cascade

When a cosmic ray enters the atmosphere, it collides with a nitrogen or oxygen nucleus. In this collision, pions are created as well as the original nucleus fragmenting. Among the pions are neutral pions which subsequently decay into two photons. This section focuses on the resulting electromagnetic cascade caused by the decay photons.

Assume that an initial photon has an energy E_0 and travels a distance R before creating an electron and positron pair. On average, each resulting particle will have energy $E_0/2$. The e^+e^- pair travel another distance R before they bremsstrahlung and generate one photon each with the photon taking half the initial energy of the electron or positron. After a distance nR, there will be 2^n particles, each with an energy of $E_0/2^n$. This process continues until the average energy of the particles is below a critical energy, E_c. For electrons and positrons, E_c is the energy where the cross section for bremsstrahlung is smaller than the cross section for ionization. For photons, E_c is the energy where Compton scattering is the dominant interaction over pair production.

For high energies, the length for pair production, ϵ_0, is approximately equal to the radiation length for bremsstrahlung. If R is the distance where the probability for pair production or bremsstrahlung is 1/2, then $R = \epsilon_0/\ln 2$ [12]. The number of distances, then, for the shower to reach the maximum number of particles is:

$$\frac{\ln(E_0/E_c)}{\ln 2} \tag{1.6}$$

The depth of the shower maximum (X_{max}), or the depth at which the number of charged particles reaches a maximum (N_{max}), is proportional to the log of the initial energy and N_{max} is proportional to the energy.

1.3.2 Hadronic Cascade

An extensive air shower initiated by a hadron is just a superposition of electromagnetic cascades from π^0 decays fed by a hadronic core. In addition to the electromagnetic cascades, charged pions decay to muons. The decays occur in the region where the probability to decay is higher than the probability to interact, or high in the atmosphere where it is less dense. Thus, the majority of the muons that arrive at the ground are created in the initial stages of the extensive air shower. Deep in the shower development, electrons and positrons are created via the decay of muons. Thus, the electromagnetic cascade is not fully attenuated deep in the shower development, it persists due to muon decays.

Since the neutral pions are fed by the hadronic core, X_{max} depends on the hadronic interaction model and composition of the cosmic ray. Protons have a longer mean free path in the atmosphere while an iron nucleus is much shorter. In addition, for a cosmic ray with energy E_0, the average energy per nucleon for iron is much lower (E_0/A) than if it were just a proton. The result is that X_{max} is shallower for iron nuclei and fluctuates less than for a proton initiated shower. Thus, the study of X_{max} is sensitive to the primary mass, and to a lesser extent, the hadronic model, see Fig. 1.3. Figure 1.7 is a representation of the interactions in an extensive air shower with the electromagnetic cascade on the left, the muons in the middle, and the hadronic core on the right.

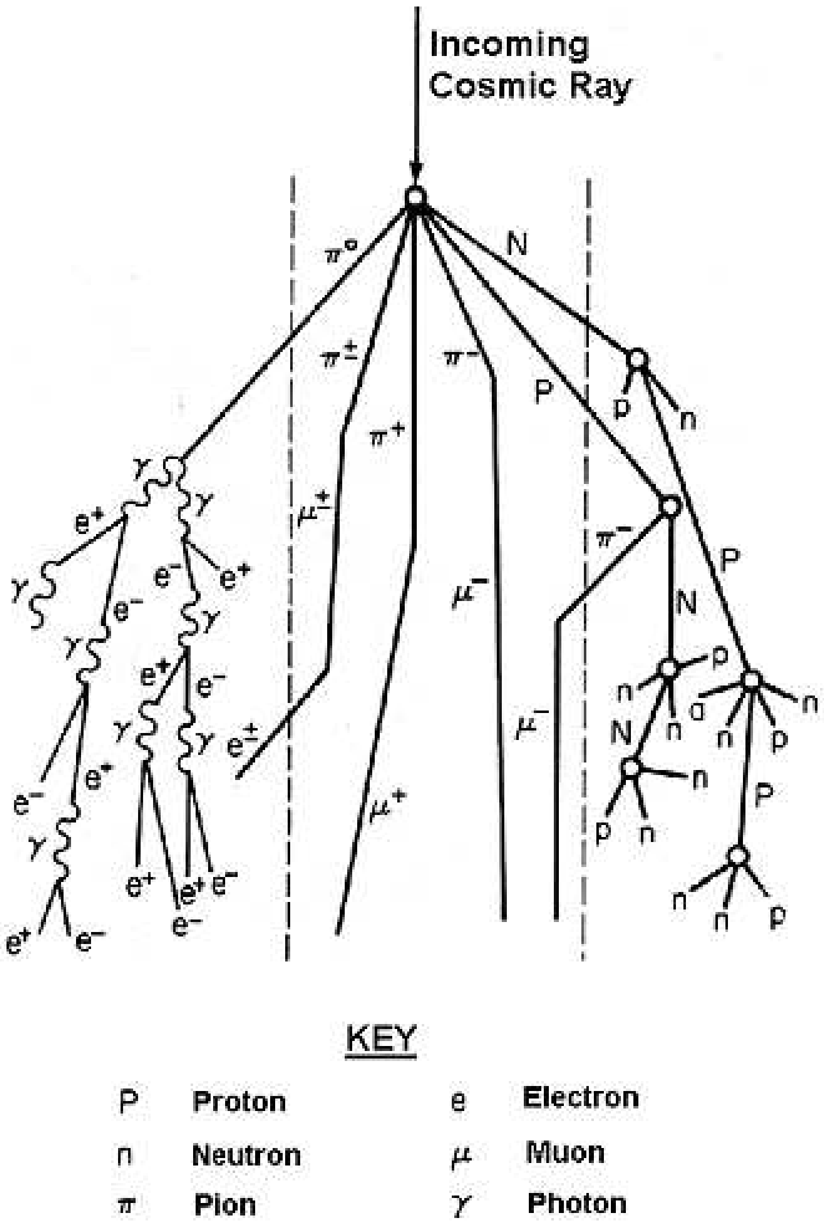


Figure 1.7: The incoming cosmic ray interacts with a nucleus creating pions and fragmenting the original nucleus. Neutral pions decay into photons which create an electromagnetic cascade. Charged pions decay into muons and neutrinos. There is also a hadronic core that generates more pions.

1.4 Conclusions

The field of cosmic ray physics has a distinguished past and is an ever evolving field. There are still questions to be answered by cosmic ray observatories; namely the nature of the cosmic rays, the composition and energy spectrum, and the possible sources of these cosmic rays. The end goal, really, is to discover the composition to determine the nature of the sources, whether they be “top-down” or “bottom-up” methods. The energy spectrum, namely whether or not there is a GZK feature, is indicative of the source distribution. If there is no GZK feature, a possible explanation is that there are “top-down” sources of cosmic rays located relatively close to the earth.

The method of detecting ultrahigh energy cosmic rays depends on the properties of the extensive air showers. Fluorescence detectors depend on the development of the electromagnetic cascade with the energy being related to N_{max}, and the composition being related to X_{max}. Detectors on the ground sample the shower at one particular depth, but contain information about the lateral distribution of particles. It will be shown later in this paper that there are observables on the ground that are related to the energy and composition.

CHAPTER 2

“Top-Down” Models

In the previous chapter, conventional models of accelerating cosmic rays to the highest observed energies were discussed. In this chapter, more exotic models of the origins of ultrahigh energy cosmic rays will be presented. These models are referred to as “top-down” models because the cosmic rays are created with the observed high energies, not accelerated.

Several events above the GZK cutoff have been reported by various experiments [8, 26, 27, 28]. In particular, the AGASA array in Japan published a high energy spectrum with no apparent cutoff at the highest energies as would be expected by the GZK interaction, see Fig. 2.1 [29]. This is immediately interesting because of the difficulty in accelerating particles to those energies and the added complication of traveling from the creation site to the earth and arriving with the observed energies. The seemingly improbable circumstances necessary to bring about such a spectrum prompted several alternative theories as to the origin of these highest energy events. The theories propose that these events were created at the observed energies at relatively close distances from the earth. This avoids the acceleration and propagation difficulties of the more conventional theories as to the origin of these super-GZK events.

In this chapter, three general theories that may explain the existence of the highest energy events are presented. These “top-down” models involve neutrino interactions with the relic neutrino background (Z-bursts), topological defects

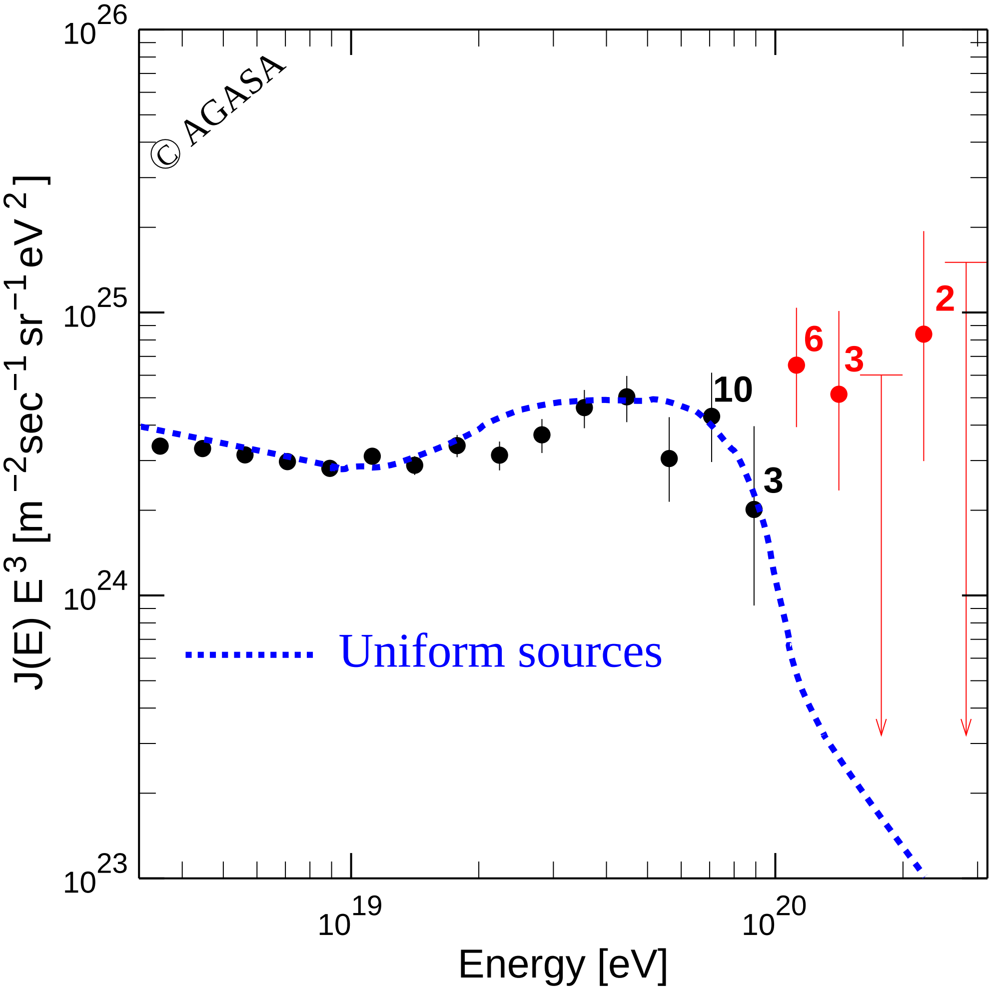


Figure 2.1: Cosmic ray spectrum as reported by the Akeno Giant Air Shower Array experiment (AGASA). Figure is from [29].

(TDs), and super heavy dark matter (SHDM). Each of these models is flexible enough to describe the AGASA energy spectrum and predict that a significant fraction of the highest energy events would be photons. Indeed, high energy photons are the signature of "top-down" models. The predicted high energy photon flux would produce events in the Pierre Auger Observatory that would have noticeably different characteristics from baryonic cosmic rays, see Chapter 7. Thus, the Pierre Auger Observatory may be sensitive enough to test the predicted photon fluxes from these theories.

2.1 Z-Bursts

The cosmic microwave background radiation originated when the universe became transparent to photons. Photon decoupling occurred at around 400,000 yrs after the big bang, when the mean energy was about 1/2 eV. In an analagous manner, neutrinos decoupled at an energy of around 1 MeV, or when the universe was about 1 s old [30]. These neutrinos persist today and are referred to as relic neutrinos. These relic neutrinos do not undergo reheating as the photons do, so the temperature and density are different. The density and the temperature of the relic neutrinos are directly related to the CMBR density and temperature. $T_\nu = (4/11)^{1/3} \times 2.73\text{K}$ with the factor of 4/11 coming from the reheating stage ($e^+e^- \rightarrow \gamma\gamma$) of the photons. The result is that $T_\nu \approx 1.95$ K or $\sim 10^{-4}$ eV. It is safe to assume that the relic neutrino background is non-relativistic, i.e. $m_\nu \gg 10^{-4}$ eV. The neutrino density (n_ν) is $(4/11) \times n_\gamma \sim 108\ \text{cm}^{-3}$ [31].

Since the relic neutrino background is related to the CMBR, the question can be asked if it affects cosmic ray propagation in a similar fashion, i.e. the GZK interaction. The mean free path of a particle through the relic neutrino background [30] is $1/n_\nu \sigma_W$ where:

$$\sigma_W \approx (G_F^2/\pi)[s/(1+s/M_W^2)] \leq (G_F^2/\pi)s \tag{2.1}$$

where G_F is the Fermi constant, M_W is the W boson mass, and $\sqrt{s}$ is the center of mass energy. Then:

$$\sigma_W \leq (G_F^2/\pi)E\langle\epsilon\rangle \tag{2.2}$$

where E is the energy of the cosmic ray and $\langle\epsilon\rangle$ is the average energy of the relic neutrino background. Plugging this last equation back into the formula for the mean free path, one gets that

$$\lambda > \pi/2G_F^2 E\rho_0 \tag{2.3}$$

where ρ_0 is the energy density of the relic neutrinos. The mean free path is long enough, i.e. it is longer than the size of the universe, that there is no energy loss due to interactions with the relic neutrino background for conventional cosmic rays unless the energy of the cosmic ray is greater than 10^{23} eV.

On the other hand, ultra high energy neutrinos are sensitive to the relic neutrino background. A high energy neutrino may annihilate with a relic neutrino to create a Z-boson ($\nu + \overline{\nu} \rightarrow Z$). The resonant energy for the incoming neutrino is:

$$E^R_{\nu_j} = \frac{M_Z^2}{2m_{\nu_j}} = 4(\mathrm{eV}/m_{\nu_j}) \times 10^{21}\mathrm{eV} \tag{2.4}$$

where the subscript j refers to the 3 neutrinos. The width of the energy resonance is:

$$\frac{\delta E^R_{\nu_j}}{E^R_{\nu_j}} \sim 2\frac{\delta M_Z}{M_Z} \sim 2\frac{\Gamma_Z}{M_Z} \sim 0.06 \tag{2.5}$$

The conclusion is that the annihilation process converts neutrinos with an energy within 3% of the peak resonant energy into highly energetic Z bosons. The resulting Z boson is boosted forward with a gamma factor of $E_\nu/M_Z \sim 10^{10}$

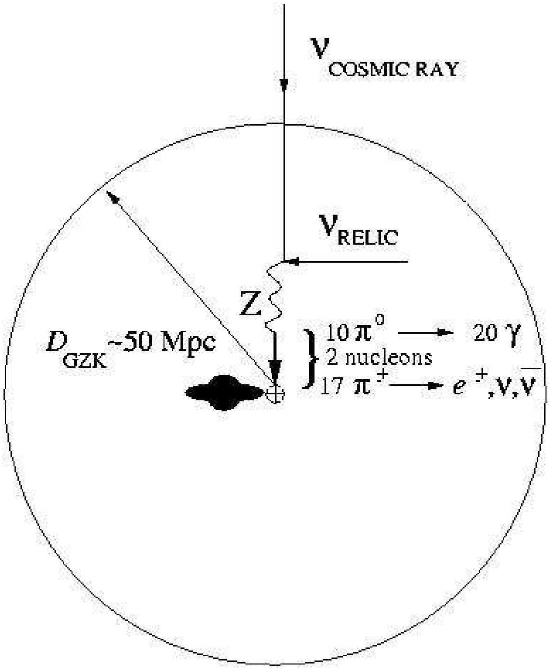


Figure 2.2: A representation of a Z-Burst. The ultra high energy neutrino enters the GZK sphere and interacts with a relic neutrino. The resulting Z-boson is boosted toward earth and decays into pions and nucleons. The π^0 decay into photons which reach the earth. Figure is from `http://www.hep.vanderbilt.edu/~weiler/weiler_research.html`

which means when the Z decays (decay time is 3×10^{-25} s in its rest frame) the decay products are beamed, $\theta \sim 1/\gamma_Z \sim 10^{-10}$. This beam from the decayed Z boson is referred to as a Z-burst. If the Z-burst is pointed towards the earth and is within the GZK sphere, super-GZK events can be detected, see Fig. 2.2. This resonance also indicates that a tell tale signature of Z-bursts would be a rather abrupt cutoff in the observed energy spectrum around 10^{21} eV.

When the Z decays, 70% of the time it is a hadronic decay. The products of the hadronic Z boson decay on average contain 15 π^0 and 2.7 baryons in addition to other particles [32]. The mean energy per hadron is about 40 times less than the resonant energy $E^R_{\nu_j}$ because the mean multiplicity is 40. In addition, the mean energy of the baryons is 10 times higher than the resulting mean energy per photon. It may be possible, then, that the highest energy events observed on earth are due to Z-bursts and a significant fraction of these events are photons.

The next concern is if the annihilation rate is sufficient to account for the observed high energy spectra. To do this, a lower bound is estimated by assuming a universal distribution of the relic neutrino background, no clustering. Realistically, the relic neutrinos are massive and non-relativistic, so there is no reason to assume that they do not cluster in gravity potential wells, and this scenario is discussed later. The annihilation cross section (σ_{ann}) is [33]:

$$\langle\sigma_{ann}\rangle = 4\pi G_F/\sqrt{2} = 4.2 \times 10^{-32}\text{cm}^2 \tag{2.6}$$

The annihilation rate is then related to the mean free path for neutrinos in the universally distributed relic neutrino background:

$$\lambda = (\sigma_{ann} n_{\nu_j})^{-1} = 4.4 \times 10^{29}\text{cm} \tag{2.7}$$

Comparing this distance with the Hubble distance:

$$D_H \equiv cH_0^{-1} = 0.9h_{100}^{-1} \times 10^{28}\text{cm} \tag{2.8}$$

where h_{100} is H_0 in units of 100 km/s/Mpc, the cosmic neutrino will travel D_H to the earth with a small probability to annihilate. The probability, then, for a neutrino with resonant energy to interact is $P = D_H/\lambda = 2h_{100}^{-1}\%$. To create super-GZK events, the Z-burst should occur within the GZK sphere. From (2.7), the probability to interact in each 50 Mpc traveled is approximately 3.6×10^{-4}. Including the probability that the Z will decay into hadrons (70%), 1/4000 of the resonant neutrino flux will interact within the GZK sphere and create hadrons and photons [33]. This is the absolute minimum. If there is clustering in the neighborhood of the earth, the rate will increase.

If the relic neutrinos cluster, the probability increases that Z-bursts will be seen on earth, especially if the clustering occurs in the galactic halo (GH). Fargion [34], using an adiabatic approximation of the GH neutrino density, claims that the density could be 10^5 to 10^7 times higher than for a universal distribution. Thus, the number density of relic neutrinos within the galactic halo could be $n_{\nu_r} = 10^{7-9}\text{cm}^{-3}$. The result is that the probability of Z-bursts within the galactic halo is close to unity, increasing the possibility that the highest energy events observed on earth are due to neutrino annihilation with the relic neutrino background.

There are several conditions that affect the probability of observing these Z-bursts. First is the existence of neutrinos with an energy greater than 10^{21} eV. The second is that the neutrino mass be in a range between 0.1 and 10 eV. Finally, as described in the paragraph above, if the relic neutrinos cluster locally. The first concern may be the most pressing, since recent neutrino oscillation measurements indicate that the mass scale required is reasonable, and neutrino clustering seems to be a viable scenario.

Ultra high energy neutrinos may be created at sites where particle acceleration occurs. If protons or other nuclei get accelerated to extremely high energies but then immediately interact with surrounding gases or dust, neutrinos may be created. These neutrinos may travel without interacting for long distances, being able to reach the earth. Escaping the local relic neutrino clusters at the acceleration site may be difficult for the neutrinos as pointed out above. Another solution is that the protons escape the acceleration site and interact with the CMBR, creating neutrinos which then annihilate and make a Z-burst. The difficulty is that the original proton or nuclei must be accelerated to enormous energies, around 10^{22} eV, which is very difficult in the framework of the current understanding of astrophysical phenomena, see Fig. 1.4. As an example, to create a photon with the highest observed energy (320 EeV by Fly's Eye) via neutrino annihilation into a Z boson, the following chain may be followed [31]:

$$p + \gamma \rightarrow (p, n) + 12\pi$$
$$\pi \rightarrow \mu + \nu$$
$$\mu \rightarrow e + \nu_e + \nu_\mu$$
$$\nu + \overline{\nu} \rightarrow Z^* \rightarrow \pi^0 + X$$
$$\pi^0 \rightarrow \gamma + \gamma$$

where the resulting photons have an average energy of $10^{-4}E_p$. Thus, the initial energy of the proton must be $\sim 3\times10^{24}$ eV!

2.2 Topological Defects

The next two sections on "top-down" models focus on the decays of super massive particles, which will be referred to as X particles. There are some common constraints on these models that will be beneficial to explain here. First, the decay of the X particles must happen in the recent cosmological epoch or at non-cosmological distances (less than 100 Mpc). The X particle must be sufficiently massive to create particles at the highest observed energies. Lastly, the number density and decay rate must be such that the observed flux can be a result of these decays.

The characteristics that make this X particle scenario attractive are that the decay products are quarks and leptons, with the quarks hadronizing creating light mesons, namely neutral pions, which decay into gammas. The exact nature of the decay products depends on QCD, but approximations do not differ significantly in their predictions. According to these approximations, the resulting spectrum is hard and is the same for all daughter particles (photons, neutrinos, etc.), following an inverse power law with a slope of ∼1.3-1.5 [35, 36]. The number of photons and neutrinos created dominates over the resulting number of nucleons by a factor of ∼10 [37]. Of course, the ability to detect the resulting photon dominance depends on the diffuse radio background, the extragalactic magnetic field, and the source distribution.

The absolute value of the normalization of the flux depends on $\dot{n}_X$. For super heavy dark matter this is the lifetime of the X particle while for topological defects this is the production rate as it decays almost immediately. Using $X \rightarrow lq$ and that the resulting quarks hadronize producing pions which in turn produce photons, it is possible to estimate $\dot{n}_X$. Assuming a uniform distribution, normalizing to the observed high energy spectrum, using $m_X = 10^{16}$ GeV, the slope of the energy spectrum 1.5, and the fraction of the initial energy carried by the pions being 0.9, the necessary decay rate is ∼13 $\mathrm{AU}^{-3}\mathrm{yr}^{-1}$ [37]. This estimate of around 10 decays per solar system volume per year can be used as a benchmark. If the sources are clustered, this number would change.

Cosmic topological defects may be the sources of these massive X particles. Topological defects may be objects such as magnetic monopoles, cosmic strings, domain walls, superconducting cosmic strings, etc. They are a result of symmetry breaking phase transitions in the early universe from GUTs. There are related phenomena in condensed matter, such as vortex lines in superfluid He [38]. These topological defects are made of trapped quanta of massive gauge and Higgs fields of underlying spontaneously broken gauge theory. The X particles are then massive fermions trapped within these TDs due to their coupling with these fields. More specifically, topological defects are defects because at their core (lines for strings and points for monopoles), the Higgs field is zero and symmetry is unbroken while this region is surrounded by a non-zero field and broken symmetry. The mass of these defects is the energy density which is trapped inside these regions by the non-trivial "winding" of the fields around the core.

2.2.1 Cosmic Strings

Cosmic strings are topological defects where the symmetric, zero Higgs field is a line. These lines may be open ended or loops. There is a finite width associated with cosmic strings, $w \sim \eta^{-1}$, where η is the vacuum expectation value of the relevant Higgs field. The energy per unit length, then, is $\mu \sim \eta^2$.

The way energy is released through X particles is possible via string intersection, loop shrinking, or cusps. String intersection is when two portions of the string overlap and there is a discontinuity in the Higgs field in that region. As a result, a topology removal event occurs, releasing an X particle with energy $w\mu \sim \eta$ and $\eta \sim m_X$, and the strings then join at that point to maintain continuity. In this way, loops can be formed from self-intersecting strings. Loops then radiate energy in the form of gravitational radiation, which in turn makes them shrink. Once the loop gets small enough, i.e. the radius is of comparable size to the width of the string w, one X particle is released and the loop vanishes. Cusps are kinks or other deformations in the cosmic string and create X particles in much the same ways described above.

2.2.2 Monopoles

Magnetic monopoles are topological defects predicted by most realistic GUTs. As was pointed out in [39, 40], if monopoles were created in the early stages of the universe, then the formation of monopole anti-monopole metastable bound states, “monopolonium”, would occur.

At a temperature T, the monopolonia would be created with a binding energy $E_b \geq T$ and initial radius $r_i \sim g_m^2/(2E_b)$, where g_m is the magnetic charge related to the electric charge through the Dirac condition $eg_m = N/2$ with N being the Higgs field winding number making the monopole a topological defect [37]. The monopolonium may exist in quantized energy states with the associated radii being $r = n^2 a_m^B$ with n being a positive integer and $a_m^B = 8\alpha_e/m_M$ the magnetic Bohr radius. Since the magnetic Bohr radius is much smaller than the Compton wavelength of the monopole ($a_m^B \ll m_M^{-1}$), the monopolonium does not exist in the ground state. It is created in an excited state ($n \gg 1$) and radiates gammas, then gluons, then Z bosons, then X particles until reaching a state where the cores of the monopole and anti-monopole overlap, annihilating releasing ~80 X particles (phenomenologically, the mass of a monopole is typically $\sim 40 m_X$) [37].

2.2.3 Cosmic Necklaces

Cosmic necklaces are just closed loop cosmic strings with monopole “beads” on them [41]. These objects may form if there is a two stage symmetry breaking where monopoles are formed in the first stage and cosmic strings are formed in the second. In this scenario, each monopole attaches to two strings with the monopole magnetic flux channeled along the string.

X particles are produced because the monopoles make the motion of the closed strings irregular, resulting in self-intersections. At these intersection points, monopole and anti-monopole may meet and annhilate, creating ~80 X particles [42]. This scenario is attractive because of the capability to produce many ultra high energy cosmic rays, be clustered together, and to be located within our GZK sphere.

2.3 Super Heavy Dark Matter (SHDM)

There are two outstanding problems in physics that may be correlated. The universe is filled with matter of an unknown composition called dark matter. Also, the origins of ultra high energy cosmic rays are unknown. Super heavy dark matter is a possible solution to both of these problems. Super heavy dark matter (SHDM) is what the name implies, particles with a large mass that do not direclty interact with particles (except for maybe gravitationally or via the weak interaction) and are therefore "dark". The SHDM may decay or self-interact and annihilate to create the highest energy cosmic rays.

SHDM is expected to be a fraction of the total dark matter in the universe. It will be clustered around galaxies and in the galactic halo of the Milky Way. The signatures, then, of SHDM as the source for ultra high energy cosmic rays will be no GZK cutoff due to the proximity of the source, which also conforms to the requirements that the cascade radiation be suppressed [43]; and that the arrival directions will be nearly isotropic as SHDM will be evenly distributed in the galactic halo. The decay of SHDM produces cosmic rays in the method described in the previous section.

The origins of SHDM is not easily explained. The X particles may be created via topological defect necklaces, explained in the previous section, or they are created thermally. If they are created thermally, they must be created after inflation or the density would be too small to account for the effects attributed to dark matter. The particles must have been created in the reheating phase after inflation. However, the reheating temperature must then be on the same order as the mass of the X particles, $m_X \sim 10^{12-15}$ GeV, which is not allowed in many models. However, in models with dynamically broken supersymmetry, the lightest super partner is the gravitino whose mass is $m_{3/2} \leq 1$ keV and decouples from the thermal bath relatively late. This may be the cold dark matter particle. In this scenario, all constraints on the reheating temperature disappear and it can reach the necessary levels [44].

2.3.1 SHDM Decays

Difficulties arise when considering the requisite long lifetime of SHDM. The lifetime must be longer than the age of the universe $\tau > t_0$, yet it must decay. There must be a way to exponentially suppress the interaction that causes the decay of the X particle. Several methods are proposed such as using "the hidden sector of supersymmetry breaking" (cryptons) [45]. Two methods will be briefly mentioned here.

The first way is to assume that the X particle is a neutral fermion belonging to the $SU(2) \times U(1)$ group [46]. The stability is protected by a discrete symmetry, associated with the quantum number R', respected by all interactions except quantum gravity through wormhole effects. The X particle then decays via a dimension 5 operator that is inversely proportional to the Planck mass and suppressed by a factor of e^{-S} where S is the action of a wormhole absorbing one R' charge. An example where $X \to \nu q\bar{q}$, the lifetime is [46]:

$$\tau_X \sim \frac{192(2\pi)^3}{(G_F v_{EW}^2)^2} \frac{m_{PL}^2}{m_X^3} e^{2S} \tag{2.9}$$

For $m_X > 10^{13}$ GeV and $\tau_X > t_0$, S must be greater than 44, which is an acceptable value [47].

A second method is to say that instantons are responsible for X decays [48]:

$$\tau_X \sim m_X^{-1} \exp\left(4\pi/\alpha_X\right) \tag{2.10}$$

where α_X is the coupling constant of the spontaneously broken gauge symmetry involved. Following the discussion presented in [48], a toy model is put forward where $SU(2)_X$ gauge interactions are added to the standard model. Also, assume that there is a broken gauge symmetry at high energy for $SU(2)_X$ resulting in 2 left-handed fermionic doublets (X and Y) and 4 right-handed singlets, all of which are singlets under the standard model $SU(2)_L \times SU(3)_c$. When $SU(2)_X$ breaks down, X and Y acquire mass and are not allowed to mix. The lightest of X and Y may interact via instantons which violate X and Y quantum numbers, so that if X is heavier than Y:

$$X \rightarrow Y + q + l \tag{2.11}$$

with τ_X from (2.10). The benefit of this method is that the decay products can produce the observed cosmic ray spectrum, dark matter may be the Y particle with a small mixture of X particles, and dark matter does not interact strongly or via the electroweak forces.

2.3.2 SHDM Annihilation

To avoid the difficulty of requiring an extremely long lifetime by invoking wormhole effects, instantons, or cryptons, it is suggested by Blasi, Dick, and Kolb in [49] that SHDM does not decay but annihilates. These particles are referred to as WIMPZILLAS, because they are heavier than the proposed WIMPS that are thermal relics with a maximum mass ∼100 TeV. WIMPS have a small mass due to annihilation cross section arguments, if the mass is too large, the cross section is small resulting in too many WIMPS. WIMPZILLAS, on the other hand, are not thermal relics because they were never in chemical equilibrium during the early stages of the universe [50], and can be massive ($m_X \sim 10^{12-19}$ GeV).

When the WIMPZILLAS (referred to as X particles from now on) annihilate, two jets are produced each with an energy equivalent to the mass of the X particle. These then fragment into many particles with the leading particles carrying the majority of the original energy. This is very similar to the results of the decaying X particles and in a similar fashion many high energy photons are created.

A problem arises, however, if the dark matter is assumed to be smoothly distributed in the galactic halo. To get the observed spectrum, the annihilation cross section violates unitarity [49], i.e. $\sigma_{ann} < m_X^{-2}$. This may be avoided by introducing non-standard physics or by assuming that the distribution within the galactic halo is clumpy. If the distribution is clumpy, there should be a strong anisotropy towards the galactic center. The consequence of the clumpy solution is that the massive X particles can no longer be a large fraction of the dark matter and another particle must be introduced to solve the problem.

2.4 Predicted Photon Fluxes

In a paper by Gelmini, Kalashev, and Semikoz [51], various top-down models were used to fit the observed cosmic ray spectra from AGASA and HiRes. In particular, the predicted photon fluxes from the top-down models were estimated based on these spectra. Several plots are presented here to show the range of predicted photon fluxes based on different models and resulting spectra. In Chapter 7, an upper limit on the flux of photons as determined from data taken with the Pierre Auger Observatory will be presented and compared to the values estimated here.

In Fig. 2.3, six plots are shown of fits from top-down models to AGASA and HiRes spectra. The top plots are the predictions according to the Z-burst model where the simulation assumes a relic neutrino mass of 0.4 eV. The middle plots are the predictions according to SHDM where $m_X = 2 \times 10^{12}$ GeV. The bottom row are plots from the assumption that the source of the ultra high energy cosmic rays are due to topological defects, namely necklaces. The mass of the

X particle in this case is $m_X = 2 \times 10^{13}$ GeV and the fragmentation from the QCD spectrum assumes no supersymmetry and predicts roughly 3 photons per nucleon in the decay products. For all the plots, a certain low energy component (LEC) was assumed from astrophysical sources to better fit the low energy end of the spectrum. The protons, separate from the LEC, also arise from the top-down models.

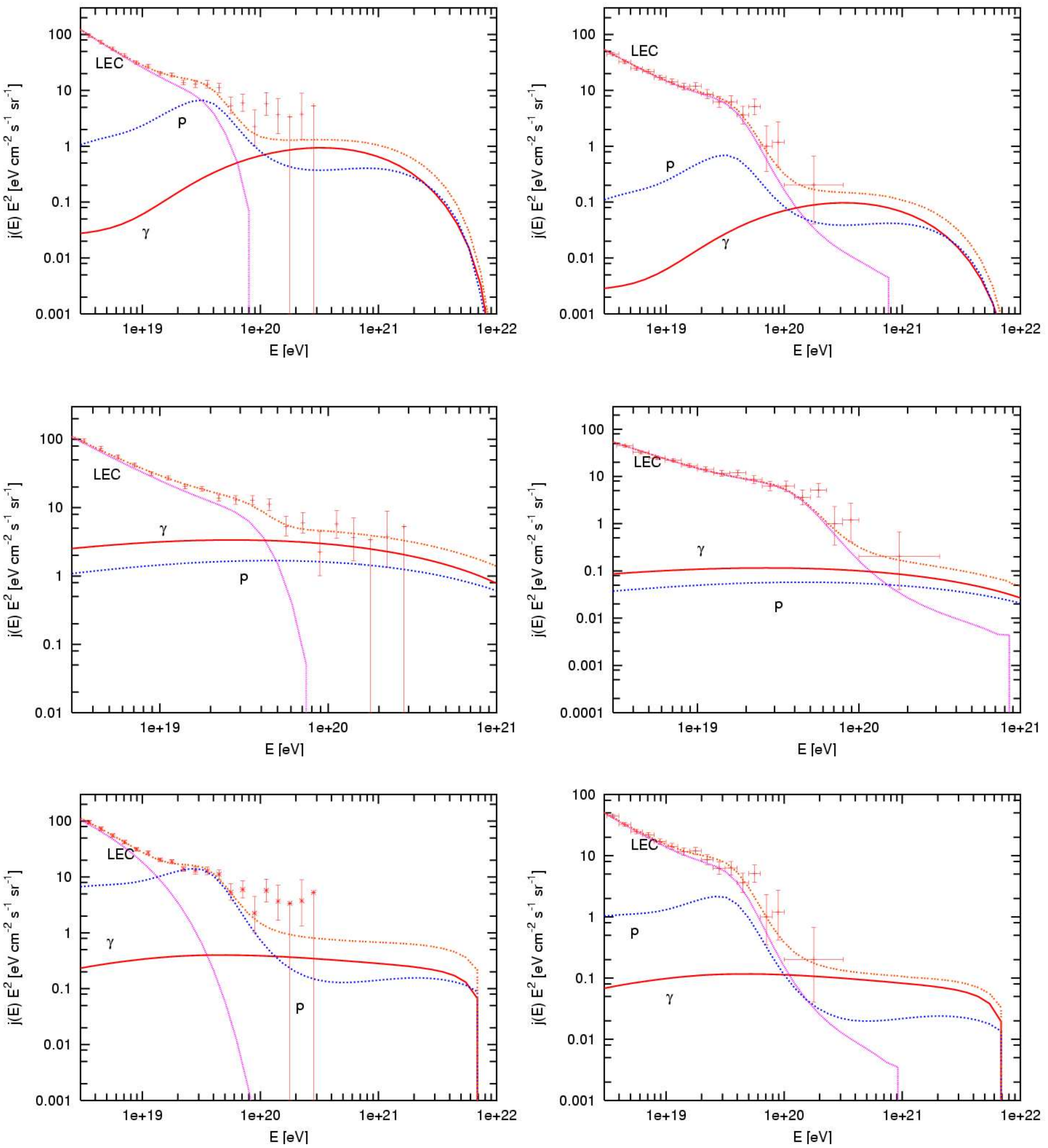


Figure 2.3: Predictions of various top-down models fit to the AGASA spectrum (left) and the HiRes spectrum (right). Top Row: Z-burst. Middle Row: SHDM. Bottom Row: TD. The solid red line is the photon flux, the dotted blue line is the proton flux, the dotted pink line is a low energy component (LEC), and the dotted red line is the sum.

CHAPTER 3

The Pierre Auger Observatory

The Pierre Auger Observatory (PAO) was designed to investigate the outstanding puzzles in cosmic ray physics, namely determining the origin and composition of the highest energy cosmic rays. The design incorporated two measurement techniques used with success in the past: detecting the nitrogen fluorescence in the atmosphere caused by an extensive air shower and measuring the lateral distributions of particles that reach the ground. This hybrid technique of detecting extensive air showers is unprecedented for an observatory the size of the PAO.

The PAO will have an array of water Cherenkov detectors that will cover 3000 km^2 using 1600 detectors spaced 1.5 km apart in a triangular grid. On the edges of the surface detector array there will be 4 fluorescence telescopes that will view up to 30 degrees in elevation and 180 degrees in azimuth coinciding with the surface detectors (SD). Thus, for a small subset ($\sim$10%) of cosmic rays, the air showers will be recorded with both techniques which will allow for energy and arrival direction cross-checks. These "hybrid" events will be valuable in determining systematic errors inherent in both techniques as well as providing more information to determine particle kind and check hadronic interaction models. A map of the site of the PAO in the Mendoza province in Argentina is shown in Fig. 3.1.

Both techniques to measure the energy of cosmic rays mentioned above have different systematic errors associated with them. The fluorescence detector en-

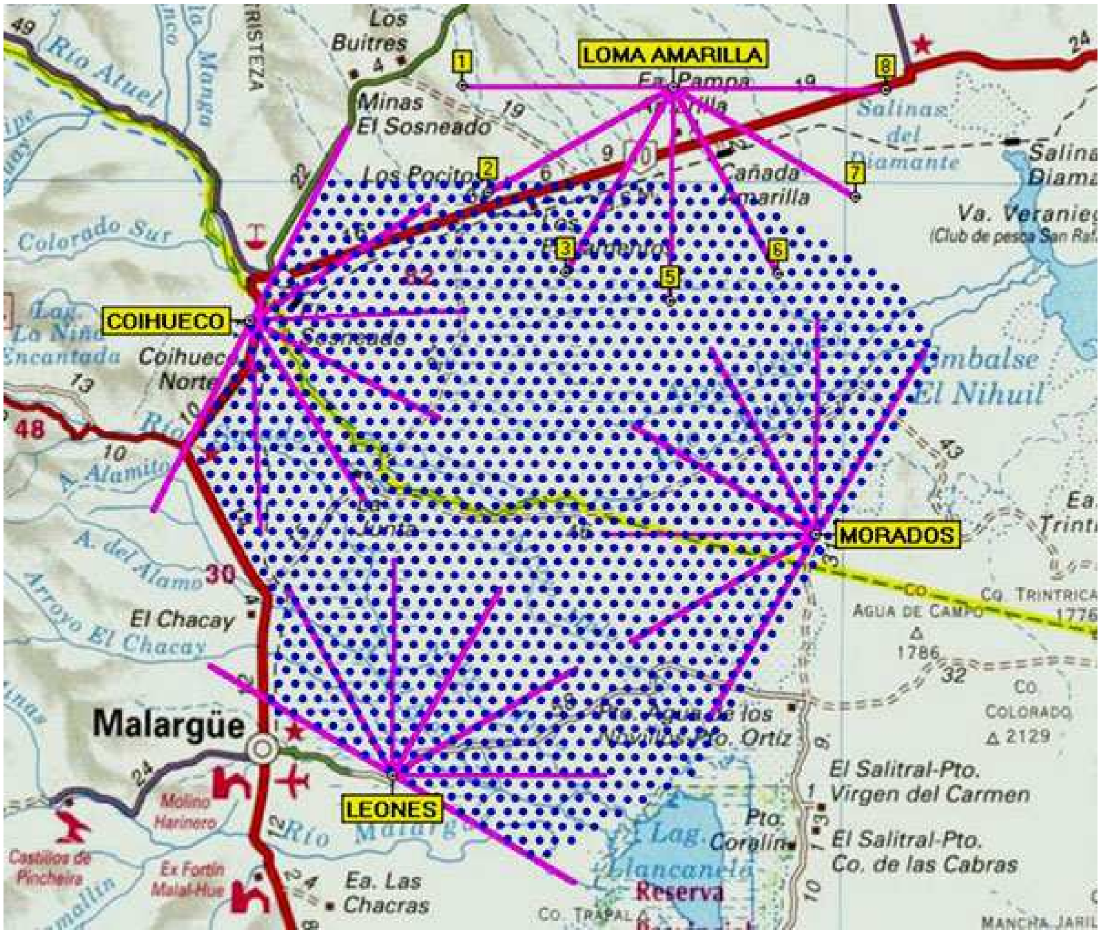


Figure 3.1: A map of the Pierre Auger Observatory with 1600 water tanks (blue dots) and 4 fluorescence detector sites, labeled in yellow, located next to Malargüe, Mendoza in Argentina. For scale, the distance from Malargüe to Coihueco is 40 km.

ergy measurement relies on the photon yield, or the number of photons fluoresced per unit length for an electron. Any systematic error in this measurement will then propagate to the energy estimate made using fluorescence data. Another possible source of systematic error using fluorescence arises in determining the atmospheric conditions at the time of a given cosmic ray shower. This is important because the light that reaches the detector from nitrogen fluorescence must travel through kilometers of atmosphere which will attenuate the intensity of the light. The attenuation length must be known to calculate the number of photons created at a given location in the air shower. This brings into focus another possible source of systematic error, calculating the absolute number of photons at the detector. The signal measured from the readout electronics must be converted into an absolute number of photons at the detector. All of these possible errors are being addressed and in this chapter the atmospheric monitoring and absolute calibration of the detector will be discussed.

As for the array of surface detectors, the absolute calibration point is given by atmospheric muons. Thus, the possible errors described above do not apply to the ground array, but there are systematic uncertainties in the energy estimate from the surface detectors. In Chapter 6 these uncertainties will be discussed. It is sufficient to state at this point that the systematic uncertainties arise mainly from the unknown composition of the cosmic ray and hadronic models used in the monte carlo simulations. The main difference, then, between the uncertainties listed for fluorescence and those for the ground array are that the fluorescence errors may be reduced through careful measurements while the uncertainty in the composition and hadronic models in the simulations remains regardless of the care taken in calibrating and monitoring the detector. Thus, the two techniques in the PAO have different systematics and combining the data from both will hopefully constrain the problem such that the uncertainties will be minimal.

3.1 The Fluorescence Detector

As electrons and positrons pass through the air, they excite the nitrogen in the air which in turn fluoresces. By measuring the amount of fluorescent light at different atmospheric depths, the shower development can be studied. From this measurement, the depth of shower maximum (X_{max}) and the number of charged particles at shower maximum (N_{max}) can be calculated. The PAO will have 4 fluorescence detectors (FD) overlooking the SD that will measure these parameters.

The FD was designed to achieve certain physics objectives. Since the principal reason for the FD is to measure the longitudinal profile of the shower development (i.e. X_{max} and N_{max}), there is a certain minimum resolution in atmospheric depth necessary for any useful results to be derived. A resolution of 20 g/cm^2 is desirable to distinguish between iron and proton primaries which have a mean X_{max} that differs by $\sim$100 g/cm^2. An energy resolution of 10% is achieved by certain signal to noise measurements which also lead to a 20 g/cm^2 resolution [52].

The base design of the FD is such that these objectives are met. Each FD building, or "eye", has six telescopes, or "cameras", made of 440 pixels each. Each pixel covers a 1.5° area of the sky. The pixels are arranged in a 22×20 matrix so that the resulting coverage is 30 degrees in azimuth and 28.6 degrees in elevation, see Fig. 3.2. The light detector for each pixel is a hexagonal PMT that is sampled by a 12-bit ADC every 100 ns. There is a data acquisition system at each eye that records all the data from the six cameras and checks if trigger levels are met in the raw data. This data is then transferred to a central data acquisition system (CDAS) for the entire observatory that checks for coincidence with the SD or any other FD eye and builds the events from the trigger data from all the detectors.

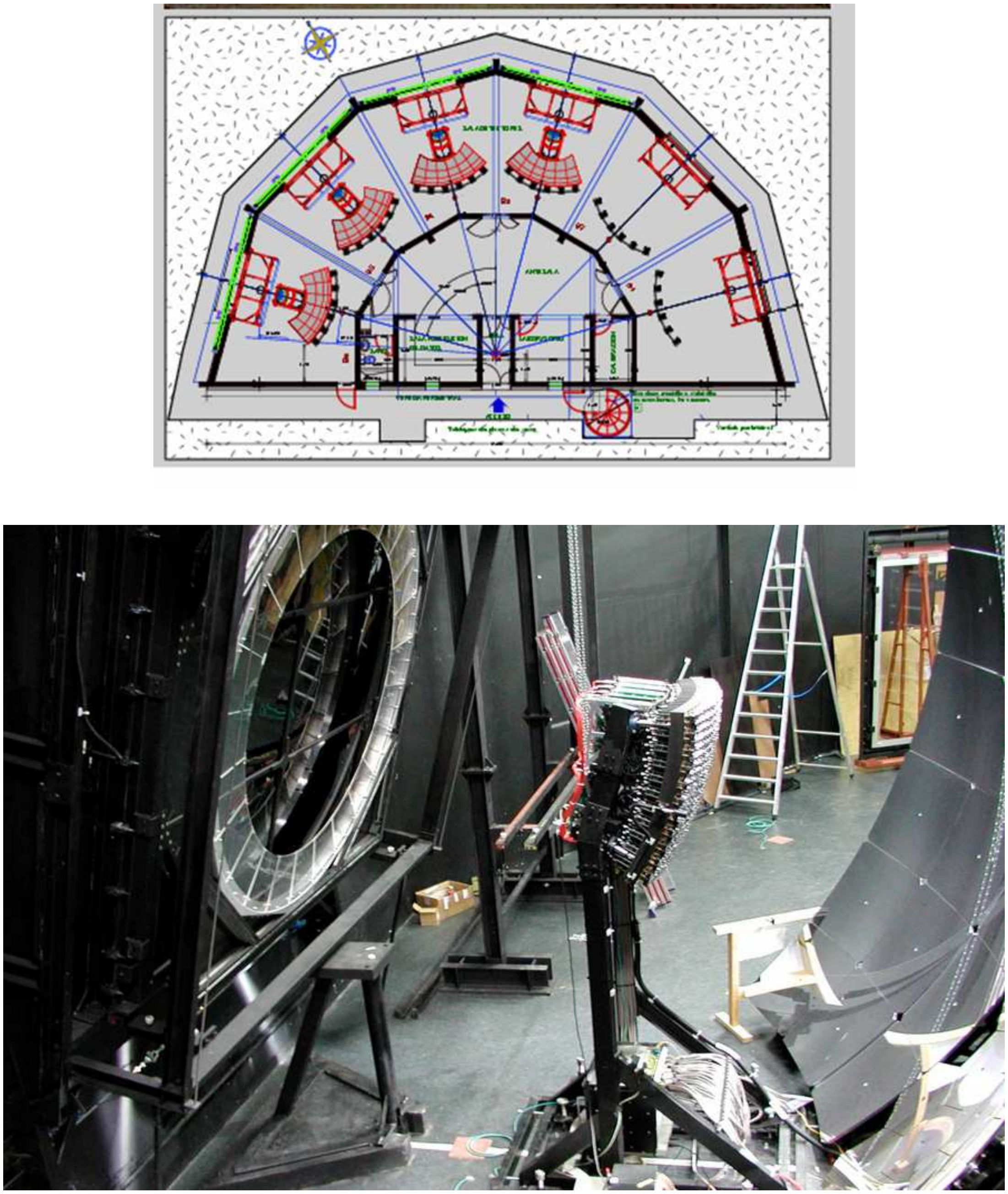


Figure 3.2: Top: Schematic layout of a FD building with six telescopes. Bottom: Picture of a telescope in the FD building. On the left is the optics (filter, diaphragm, corrector ring), in the middle are the 440 PMTs, and on the right are the mirrors.

The design of the telescope is driven by the desire to increase the signal to noise ratio while maintaining good angular resolution. The relationship between the pixel size and the signal to noise ratio is:

$$S/N \sim \frac{\text{mirror diameter}}{\text{pixel angular diameter}} \tag{3.1}$$

However, to maintain the ability for a good angular reconstruction, simulations show that the pixel size should not exceed 1.5° [52]. A good angular reconstruction is necessary to determine correctly the longitudinal profile which in turn determines X_{max} accuracy.

Fluorescent light enters the telescope through a 1.1 m radius diaphragm and is collected using a 3.5 m×3.5 m spherical mirror. Schmidt optics are used to eliminate coma aberration which is a problem in spherical mirrors covering a large solid angle [53]. Each telescope diaphragm has a UV transparent filter that restricts the incoming light to the range of wavelengths in fluorescent light ($300 < \lambda < 420$ nm) which also reduces night-sky noise [54].

To correctly determine the size of the shower at a given depth of development, there are several factors that must be accounted for. First, the number of photons emitted via nitrogen fluorescence for an electron that travels through a certain distance in the atmosphere, or the photon yield, must be known. Next, the attenuation of this fluorescent light through scattering in the atmosphere must be measured and corrected for. Finally, the calibration of the detector must be such that for a given pixel the integrated signal can be converted into an absolute number of photons. The photon yield has been measured by independent experiments and those values are used in data analysis. The atmosphere must be monitored during data taking to parameterize attenuation lengths and scattering due to aerosols. The absolute calibration is done 3 or 4 times a year, but there is a relative calibration that is run nightly to monitor any changes in the system.

The absolute calibration of the FD is an end-to-end calibration in that it accounts for all the components of the system from mirrors to PMTs to the readout electronics. The calibration is done with a diffuse light source that is 2.5 m in diameter that is placed in the aperture of the telescope. The light intensity and uniformity are measured in the lab. The light intensity is measured using NIST-calibrated (National Institute of Standards and Technology) photodiodes while the uniformity is measured using a CCD (charge coupled device) camera. The diffuse light source is able to uniformly illuminate all the pixels in the camera with a known light intensity. Knowing the light intensity at the pixel makes it possible to calculate the conversion from integrated signal to number of photons incident on the pixel. This calibration currently has 12% uncertainties [55].

It is not possible to perform this calibration every night as it is labor and time intensive. The calibration must be monitored on a nightly basis to track any changes in the performance of the telescopes and correct for these changes in the data analysis. Calibration monitoring is accomplished using an array of LEDs and a diffuser located at the collector mirrors. Light pulses are fed into the pixels and at the same time a portion of the light is directed into a calibrated photodiode to monitor the light source. Correcting for the stability of the LED array, any changes in the response of the optical system (mirror, PMT, and readout) can be monitored on a nightly basis [56].

From the absolute calibration of the optical system in the telescope, there is still a need to correct for the distance the light travels to the telescope to calculate the absolute number of photons emitted at the shower axis. Atmospheric conditions must be monitored closely, specifically the aerosols in the air and atmospheric depth and temperature profiles.

Aerosols can strongly affect the propagation of fluorescence light in the at-

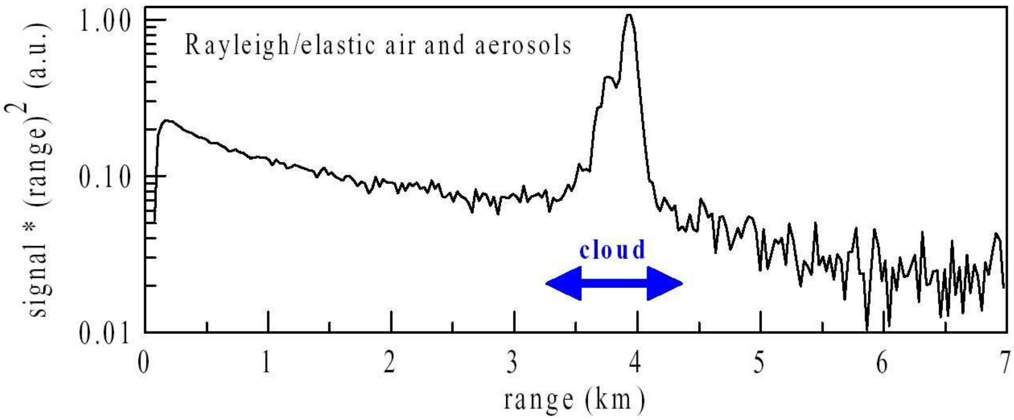


Figure 3.3: The x-axis is the height from which the backscattered light arrives and the y-axis is a parameter proportional to the signal from backscattered light. A clear night exhibits a linear behavior while any aerosols or clouds will cause anomalies as indicated. Figure is from [58].

mosphere. Several methods are used to characterize the aerosols present in the air at any given time during data taking. For instance, backscatter LIDARs are steerable UV lasers located at each FD eye. Each LIDAR has a PMT that detects the backscattered light from the UV laser pulses. The timing information from the PMT gives information about the aerosol content of the air at any given spot along the path of the laser, see Fig. 3.3. This system is able to check various locations and directions in the sky and eventually will be able to "shoot the shower". Every time a large event is recorded by a FD, the LIDAR will shoot laser pulses along the reconstructed track to measure the backscattered light and calculate the attenuation length along the path of the shower [57]. There are also cloud monitors and star monitors to detect clouds and track the stars and any changes in their intensity due to changing atmospheric conditions.

In addition to the aerosol content of the atmosphere, it is important to know the atmospheric depth and temperature profiles. Photon yield has both a pressure

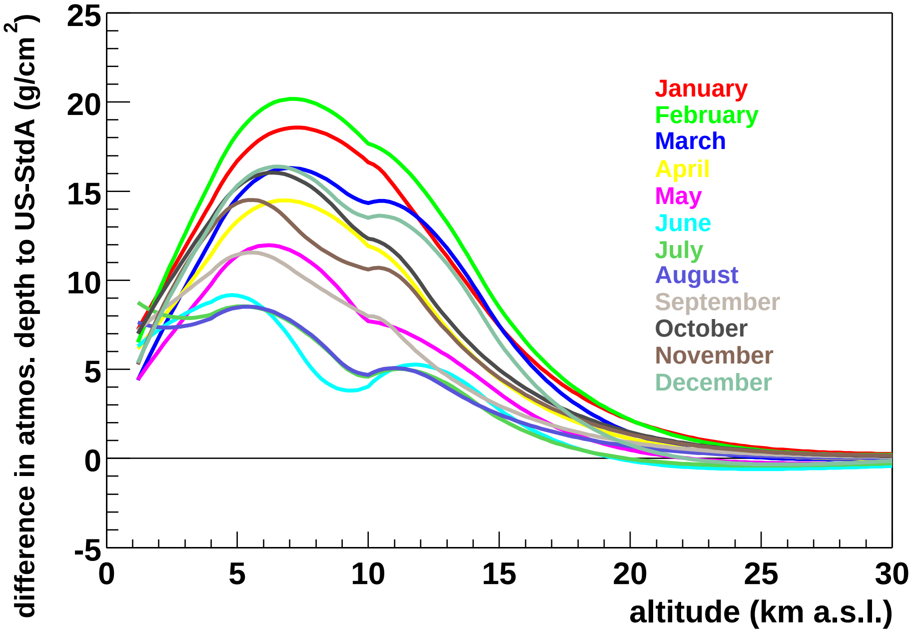


Figure 3.4: The difference from the U.S. Standard Atmosphere 1976 (US-StdA) for the Malargüe atmosphere at different times of the year is plotted here as a function of altitude. Figure is from [59].

and temperature dependence, which change with altitude. In the past, a parameterized atmosphere was assumed in analyzing data based on the assumption that atmospheric conditions were relatively stable. In Malargüe, meteorological radio soundings have been performed as well as monitoring ground based weather stations to understand the atmosphere. Radiosondes are launched with helium balloons above the PAO and data is taken every 20 m in altitude until reaching 25 km above sea level. The profiles are recorded and then compared to the parameterization used previously. If the parameterized values are used instead of the measured profiles, X_{max} values change on average 15 g/cm^2 while energy changes less than 1% [59]. The deviations from the parameterized values are shown in Fig. 3.4.

A useful tool to cross-check the calibration and atmospheric monitoring is the central laser facility (CLF). The CLF, as indicated by the name, is located in the center of the array and has a steerable UV laser with an optical fiber that injects a portion of the calibrated, pulsed laser light into a surface detector. The CLF is used to check angular reconstruction, atmospheric conditions, the relative timing between the SD and FD, and the calibration of the telescopes [60]. The pulsed laser light is scattered by the air and is detected in the FD providing a "test beam" to cross-check all important quantities in determining the properties of an extensive air shower.

3.2 The Surface Detector

The surface detector is made up of 1600 individual particle detectors [61]. Each particle detector is a water Cherenkov detector that is a cylindrical tank with the top an area of 10 m^2 and a height of 1.5 m. Each detector is completely independent from all other detectors and is driven by two 12 V batteries that are recharged by solar panels and communicates with the CDAS via a wireless communications system. Thus, when a detector is deployed in the array, it can begin data taking immediately regardless of the status of other detectors.

The water Cherenkov detector, also called a tank or station, is filled with purified water to a height of 1.2 m, or 12 cubic meters of purified water for each tank. The water is contained within a bag that has a diffuse reflective white interior with three windows on top. In these windows sit three 9" PMTs used to detect the Cherenkov light when particles pass through the detector. The signal from each PMT is split into two channels, a high gain and a low gain channel. The high gain channel comes from tapping the last dynode in the PMT dynode chain and amplifying it by a factor of 40. The low gain channel is the signal straight from the anode. These signals are then passed through filters and read out by a flash analog to digital converter (FADC) that samples at a rate of 40 MHz. Thus, each station has six signals associated with it, two from each PMT. The timing information for each station comes from a GPS system located on each tank and the timing resolution is better than 20 ns [62]. A picture of a station is shown in Fig. 3.5.

The design of the station allows detection of a wide range of signal sizes, from a few to around 10^5 photoelectrons. It is this dynamic range that allows the calibration of the station to occur using single particles while still detecting, without readout saturation, large extensive air showers that deposit thousands of particles in a station. The dynode to anode ratio of a tank is the ratio of the high gain to low gain signals. It is necessary to measure and monitor this ratio to be able to extend the calibration values, small signals typically, to the detected signals of large air showers. The monitoring of this value and other important values is discussed in Chapter 5.

Calibrating the stations is accomplished via atmospheric muons. Nature provides these particles at a rate which allows frequent calibration of the stations. Each physics signal is measured in units of vertical equivalent muon (VEM or Q_{VEM}) which is the charge deposited by a vertical, through-going (one that does not stop inside the tank) muon in a station. In nature, the muons (and other particles) come from many different directions so the calibration is done by setting a low threshold, 3 PMT coincidence trigger for the station and making a histogram of the charge deposited [63], see Fig. 3.6. The relationship between the peak in the charge histogram (Q_{VEM}^{peak}) in Fig. 3.6 and the charge deposited by a vertical through-going muon (Q_{VEM}) can be studied in a reference tank and is expected to vary little between detectors because it is a geometrical factor. Arneodo *et al.*

Figure 3.5: A picture of a water Cherenkov detector deployed in the Pampa Amarilla in Argentina. The beige dome contains all the electronics of the station and the panel mounted on the top of the station is the solar panel. The antenna for the wireless communications system can also be seen as well as the GPS antenna mounted on top of the antenna mast. Three "propellers" can be seen, one with the beige dome on top, where the PMTs are housed.

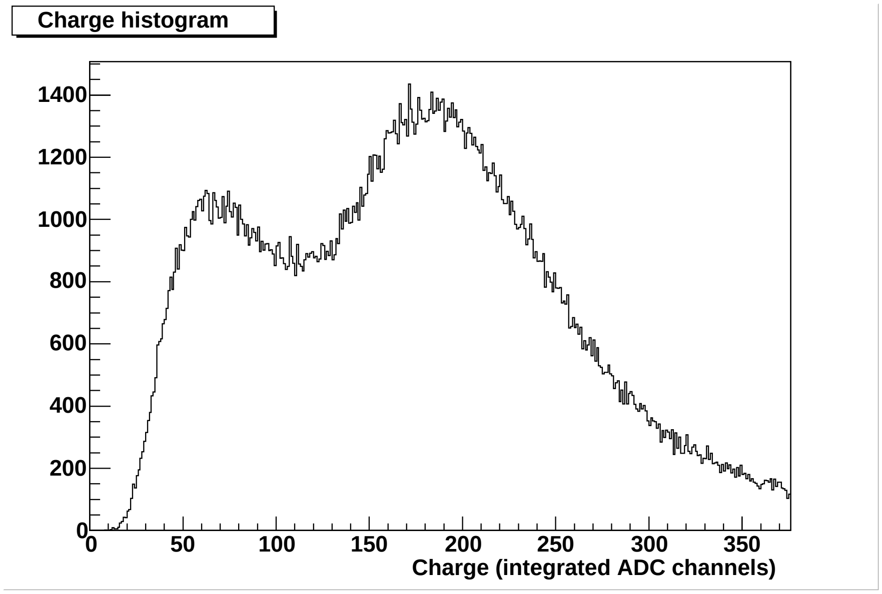


Figure 3.6: Charge histogram in integrated FADC channels where the trigger is set requiring a peak signal that is greater than 5 channels above baseline, seen in all 3 PMTs in the station. The first peak is a triggering effect and the second peak is Q_{VEM}^{peak}. Plot is from [63].

[64] performed this study by using a muon telescope as a trigger for the reference tank and they determined the factor to be $Q_{VEM}^{peak} = 1.05\ Q_{VEM}$.

Local station triggers are not set using integrated FADC traces because of strict timing requirements. To calculate the sum of a trace would require too many clock cycles on the CPU and would retard the data taking of a station. Thus, local triggers are set using the maximum pulse height in FADC channels. Channels in the FADC are a measure of the photocurrent from the PMTs. A vertical through-going muon induces a certain maximum photocurrent in the PMTs (I_{VEM}), and the trigger is set relative to that value. The procedure is the same as determining Q_{VEM}. A histogram is made of maximum pulse heights using a low threshold 3-fold coincidence trigger and that is compared to a histogram of maximum pulse heights for only vertical through-going muons. The relationship between the peak of the pulse height histogram (I_{VEM}^{peak}) and the average maximum pulse height for a VEM (I_{VEM}) follows the relationship between Q_{VEM}^{peak} and Q_{VEM} [63].

The trigger is set relative to I_{VEM}, but it is desirable that this trigger is similar, if not the same, for all stations in the array in units of FADC channels. This is desirable to achieve a similar dynamic range in all the stations and uniform triggering. A similar trigger level is achieved using a rate-based calibration method. For a given station, the voltages on the PMTs are adjusted so that the rates of events with a peak signal above a threshold are identical. The target rate is 100 Hz and the threshold is 150 channels above baseline. From a reference tank, a 100 Hz rate at 150 channels above baseline corresponds to I_{VEM} being 50 channels above baseline. Requiring that all the PMTs satisfy the above condition, the end to end gains of the PMTs are roughly identical, where the end to end gain means that for an identical energy deposit in a tank, the electronic signal (i.e. number of FADC channels) is identical.

The trigger thresholds are dynamically changed when the station is operating. Changing temperatures, PMT and electronics drifts, or any other effects may change the triggering rate of a station. Assuming, to first order, that the rate of particles hitting the detector is constant, any significant change in the triggering rate can be attributed to changes in the station. To maintain a constant trigger rate, the threshold value for the trigger is changed dynamically, where the changes are on the order of a single FADC channel [63]. This does not affect calibration, however, since absolute physics calibration is done using charge histograms as explained in the previous paragraphs. The changing trigger threshold assures uniformity in the triggering across the array in that it compensates for any possible changes in the electronics or PMTs that would affect where the Q_{VEM} is determined to be. As long as the individual station triggers are set such that the entire array is triggering on the same physical quantity (related to Q_{VEM}), uniform behavior can be expected which simplifies the calculation of important quantities such as the aperture and trigger efficiency.

Since the trigger thresholds may change with time, it is necessary to monitor the calibration quantities. Calibration quantities are taken every 3 minutes and are sent back to CDAS every 6 minutes [62]. These values can be monitored for each station in the array. In addition, high statistics (∼150,000 events) charge histograms, like Fig. 3.6, are sent back to CDAS with every physics event providing a method of checking the calibration when doing physics analysis.

3.3 Conclusions

The baseline design of the PAO is a hybrid detector for the highest energy cosmic rays. The hybrid technique will provide invaluable cross-checks between the fluorescence detection method and the ground particle detection method for extensive air showers. Each method has inherent systematic uncertainties but also adds another dimension of information. Using both detectors together should provide an illuminating look into ultra high energy cosmic rays. The bulk of the data will be taken using the array of surface detectors. A small subset (∼10%) will be events detected with both the FD and SD. This small subset will be used for determining systematic uncertainties in energy measurements, composition studies, and comparing hadronic interaction models. With these detectors, the puzzles behind the composition and origin of ultra high energy cosmic rays will begin to be resolved.

CHAPTER 4

Testing of Photomultiplier Tubes

The surface detector array of the Pierre Auger Observatory will have 1600 water Cherenkov detectors when completed. In each of these detectors, there are three 9" photomultiplier tubes (PMTs). These PMTs act as the "eyes" of the surface detector. It is important, then, to understand their behavior and characteristics.

A PMT is a device for detecting small amounts of light. They consist of a photocathode which is what converts light into electrons, or photoelectrons. This is due to the photoelectric effect where light can eject electrons from the surface of a material. These photoelectrons are then accelerated in an electric field which also focuses them onto a dynode. The dynode acts as a multiplier because the accelerated electrons hit the dynode, and more electrons are ejected than hit the dynode. These ejected electrons are accelerated in another electric field and focused onto another dynode. This process continues until there is a final amplification at the last dynode which is then collected by an anode where the signal can be read out. The number of electrons at the anode is controlled by the voltages supplied to the dynode chain, the number of stages in the dynode chain, the number of photons that are converted into photoelectrons (called the quantum efficiency), and the number of photoelectrons that reach the first dynode (called the collection efficiency). The gain of a PMT is the number of electrons at the anode given one photoelectron and a typical gain for the PMTs used in the surface detectors is 10^6.

4.1 Initial Testing of Photomultiplier Tubes

During the engineering phase of the PAO, three companies that manufacture large photomultiplier tubes were requested to give samples of their phototubes for testing and characterization. These tubes were tested at UCLA [65] and requests were made for changes in their original design to better suit the data taking conditions required in the surface detector. The three companies then sent redesigned PMTs that complied with the new design specifications. These PMTs were again tested to characterize them. After testing, a recommendation was given to the collaboration as to which PMT was most suitable for our needs.

The specifications sent to the companies [66] consisted of requirements on cathode sensitivity and quantum efficiency (QE), dark current and dark pulse rate behavior, after pulse percentage, pulse linearity, single photoelectron resolution, a standard operating gain, and a minimum gain at 2000 V. As an example of the measurements made, the QE measurements along with the measurements of dark current and gain are described.

4.1.1 Quantum Efficiency Measurements

Quantum efficiency is the fraction of photoelectrons produced at the photocathode from the incident photons. The quantum efficiency system at UCLA is designed to measure the photocathode current of two PMTs simultaneously at a given wavelength of light. One of these PMTs has a known QE which is used as a reference to measure the QE of the other PMT.

The setup can be seen in Fig. 4.1. The key features are a Xe arc lamp which gives a roughly uniform light output in the light frequencies of interest (between 250 and 700 nm), a monochromator to select specific wavelengths of light, and an integrating sphere to uniformly illuminate both the reference PMT and the test PMT. The QE of the test PMT is then:

$$QE_{test} = QE_{ref} \times \frac{I_{test}}{I_{ref}} \tag{4.1}$$

where I refers to the photocathode current and ref and $test$ refer to the reference PMT and test PMT respectively.

To ensure accuracy in the measurements, several issues need to be addressed. The reference PMT and the tested PMTs have different photocathode areas. The reference tube is a 2" hexagonal PMT from Hamamatsu (R6234) while the tested PMTs were 8" and 9" PMTs. The illumination of the photocathodes needs to be identical to make sure that the number of photons hitting each PMT is equivalent. This was achieved by using a 1.25" aperture at the interface of both PMTs to the integrating sphere.

In addition, a voltage is applied to the first dynode while the photocathode is held at ground to collect photoelectrons from the emission of the photocathode. With a constant light intensity, cathode current varies with the voltage applied to the first dynode because the efficiency of the photoelectrons to escape the photocathode increases with higher voltage. After a certain applied voltage, the cathode current levels off because the escape efficiency reaches a maximum. It is necessary to operate in this region.

Finally, at a constant applied voltage, the cathode current varies linearly with the light intensity. If the light intensity is too high, the relationship ceases to be linear due to the resistivity of the photocathode. Again, one must operate in the linear region. Results of these measurements can be seen in Fig. 4.2.

Once these potential error causing issues are resolved, data can be taken with confidence. The process followed was to apply voltage to both PMTs and measure

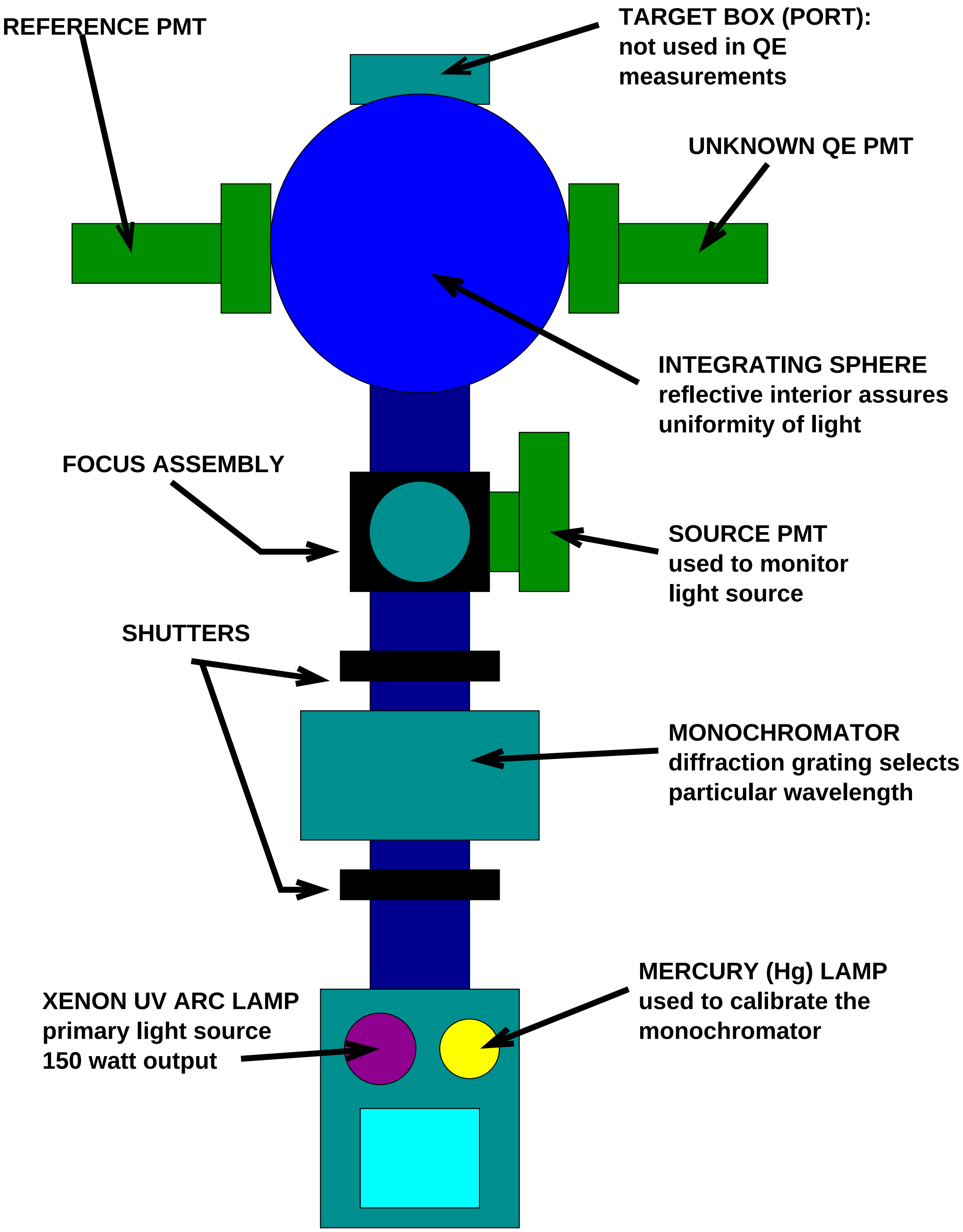


Figure 4.1: Layout of the quantum efficiency system.

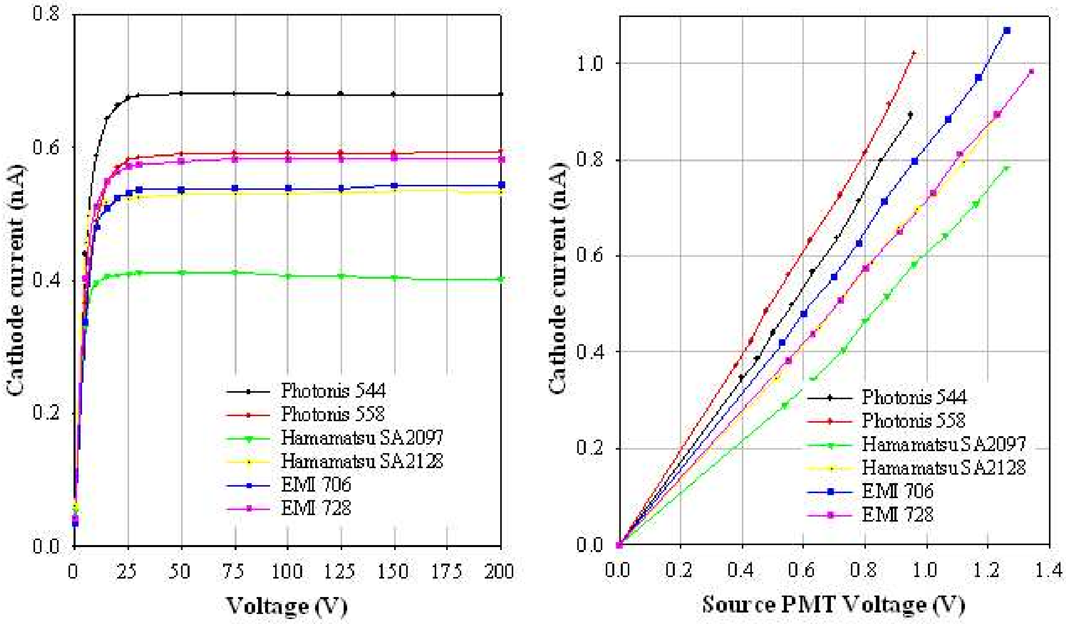


Figure 4.2: Left: Cathode current versus dynode voltage for 6 PMTs, 2 PMTs each from ETL, Hamamatsu, and Photonis. Right: Cathode current versus source PMT voltage (proportional to light intensity).

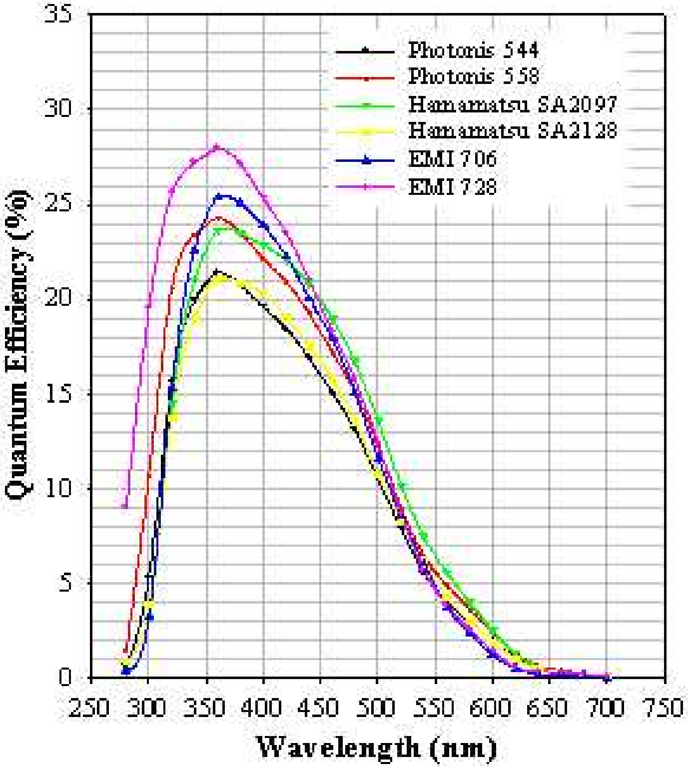


Figure 4.3: Quantum efficiency measurement of 6 PMTs, 2 PMTs each from ETL, Hamamatsu, and Photonis.

the photocathode current with no light: this dark current is used as a baseline for the other measurements. Then, light was applied to the PMTs at defined wavelengths and the cathode currents were measured with the dark current being subtracted. The QE was then calculated as specified in (4.1). Typical results can be seen in Fig. 4.3.

4.1.2 Dark Current and Gain Measurements

Both dark current and gain are measured using a very simple system, identical for both. It consists of a lamp (like the lamp used in the QE measurement) that passes through a shutter to a dark box. Using this setup, one can measure the current at the anode as a function of applied voltage when there is no light (the dark current) as well as with a constant light source (relative gain). To obtain the gain as a function of voltage for a given PMT, then, one can measure the single photoelectron spectrum at a specific voltage (which gives the absolute gain) and then use that value to normalize the relative gain curve.

Dark current is the current that flows between the photocathode and the anode when there is no light incident on the photocathode. It is caused mainly by the leakage current, thermionic emission, field emission and background radiation [67]. The dark current is also affected by the amount of time that it has spent in the dark. Therefore, one must wait a sufficient amount of time after the PMT has been exposed to light before doing any dark current and gain measurements. The dark current as a function of time in the dark box was studied to know how long these PMTs required before reaching a steady dark current. The results can be seen in Fig. 4.4.

The relative gain was obtained by measuring the anode current as a function of voltage for a given constant flux of light incident on the PMT. For accurate gain measurements, the intensity of the light incident on the PMT was adjusted such that the anode current for each tube was greater than 100 times the dark current and less than 1% of the bleeder current on the base. This is to ensure that the light intensity is high enough to avoid small photon statistics and low enough to be in a range where the PMT base can support the amount of current necessary. The voltage to the PMT was changed in 200 V increments, and the corresponding anode current and dark current were measured.

The gain obtained by this method is only relative, because the absolute intensity of the incident light is not known. The normalization of the gain curve

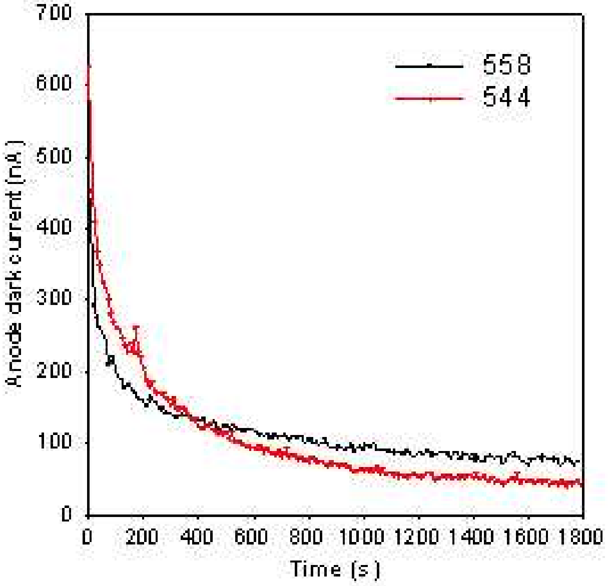


Figure 4.4: Dark current versus storage time in dark for two Photonis XP1805 PMTs, serial numbers 544 and 558.

is fixed by the absolute gain obtained from the single photoelectron method at a given voltage. Results for dark current and gain versus voltage are in Fig. 4.5. At lower voltages, the dark current levels off indicating that it is dominated by electronic noise. The behavior then is parallel to the gain curve, and then at higher voltages it increases dramatically indicating that this is an unstable region.

4.1.3 Remarks

The measurements discussed in this section are DC measurements that were only done on these prototype PMTs. In the following section, production testing of the phototubes is discussed which are all pulse measurements. This is necessary because the bases for the PMTs used in the surface detector are designed for pulse (charge) measurements, not necessarily direct current measurements. There were other tests performed on these prototype tubes and these tests became a part of

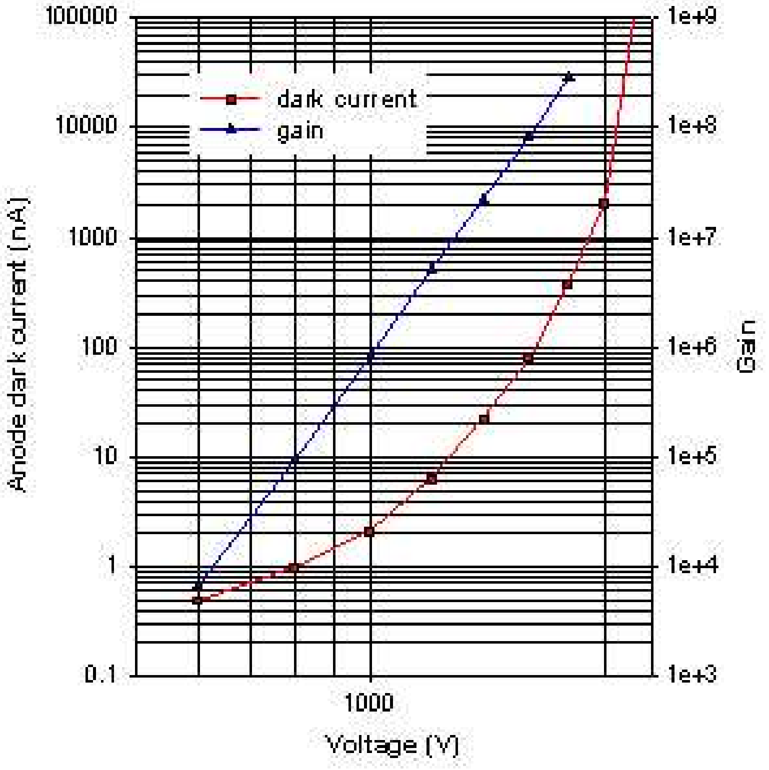


Figure 4.5: Dark current (red circles) and gain (blue squares) versus voltage. Dark current values are given on the left y-axis while gain values are given on the right y-axis. The dark current shows the PMT becomes unstable at around 2000 V.

the production test system. The tests will be described in conjunction with the production test system in the next section.

After the comparisons of the prototype tubes from three companies, Photonis was chosen to supply the more than 5000 PMTs needed for the surface array.

4.2 Production Testing of Photomultiplier Tubes

Once the decision was made as to which company was going to produce the PMTs for the surface array, it was necessary to build a system that could test the PMTs in large quantities accurately. A system was designed at UCLA, as documented in [68], and 171 PMTs were tested using that system. During this initial phase, it became apparent that it would be more time and cost effective if the system were in Malargüe, Argentina, where the construction of the observatory is taking place. The system was then duplicated and shipped to Argentina where it was assembled and commissioned in a prefabricated house that was given to the observatory for the purpose of PMT testing [69].

The system is designed to test 16 PMTs in a single run (see Fig. 4.6 and Fig. 4.7). However, to monitor the stability of the system, we have 4 permanent PMTs located at the corners of the test stand. These PMTs monitor the stability of the light source as well as the readout electronics and the performance of the system overall. Each test run lasts about 4 hours and is completely automated which makes it possible to test 24 PMTs per day.

The testing system is controlled with the DAQ computer. This computer controls the voltage delivered to the 16 PMTs via custom designed electronics. It also controls the intensity and the firing of 4 LEDs, 3 blue (Nichia NSPB520) and 1 UV (Nichia NSHU550E), through other custom designed electronics that also

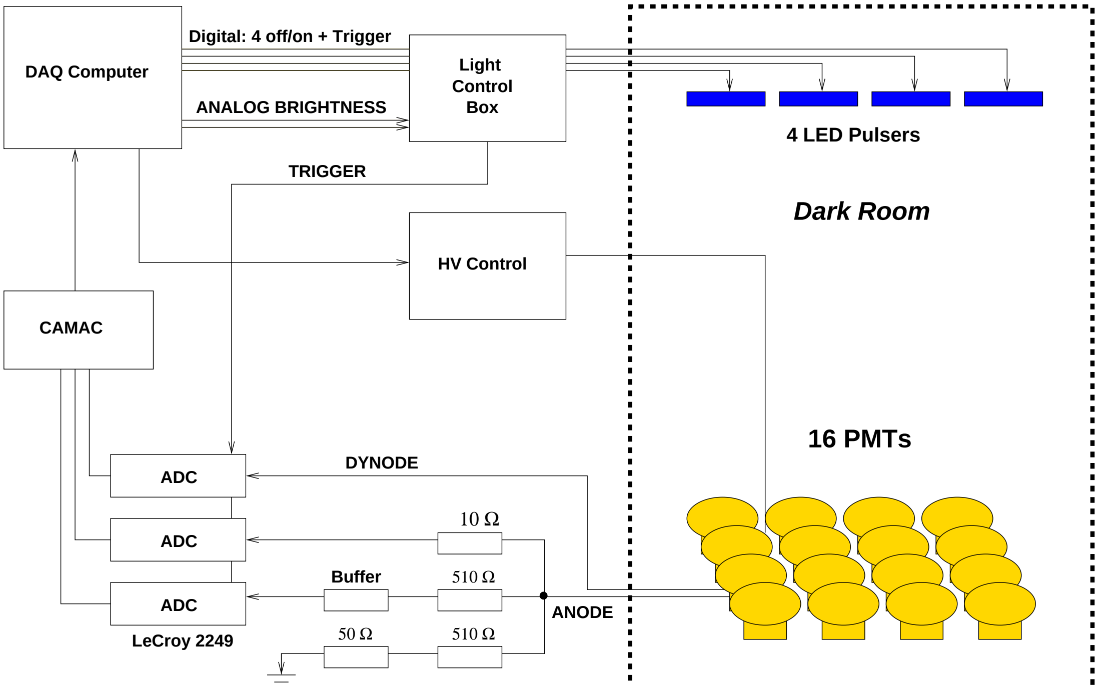


Figure 4.6: Layout of PMT test DAQ.

output a trigger signal to the data acquisition system. The signals of the PMTs comes from two places, the anode as is customary, but also from an amplified tap from the last dynode in the amplification chain of the PMT. This is done to extend the dynamic range of the PMTs when they are in the detectors taking data. To further extend the dynamic range of the system for testing purposes, the signal from the anode is broken into three components, one that is terminated to ground with the proper impedance to eliminate reflectance, one that goes through a buffer and a 510 Ω resistor, called the attenuated anode, and one that goes through a 10 Ω resistor, called the unattenuated anode, see Fig. 4.6. Again, this is done to extend the range over which the PMTs can be tested. These signals are then put into an ADC in a CAMAC crate where they are recorded and then passed to the computer. All the analysis is then done on the DAQ computer.

Figure 4.7: Top: PMTs in the dark room. Bottom: The DAQ setup with the CAMAC crate and custom electronics in the rack on the left.

Test	Specification
SPE Peak to Valley	$>$1.2
Gain versus Voltage	10^6 gain with V$<$ 2000
Dark Pulse Rate	$<$ 10 kHz at 1/4 pe threshold
Non-linearity	less than 6% non-linearity below 50 mA
Dynode to Anode Ratio	greater than 25 and less than 40
Afterpulse Ratio	less than 5%

Table 4.1: Specifications to determine if a PMT passed or failed a given test.

4.2.1 Test Results

Using this test system, 2272 PMTs have been tested out of the 5000 tubes needed for the surface detector. The tests run by the system are single photoelectron (spe) spectrum, gain versus voltage curve, dark pulse rate at 1/4 photoelectron (pe) threshold, non-linearity, dynode to anode ratio, excess noise factor, and afterpulse ratio. Each test will be described in greater detail in the following sections. In Table 4.1 the specifications are listed to determine which PMTs will be used in the surface detector. The specifications were set to ensure quality, longevity, and reliable operation of the PMTs over the lifetime of the experiment.

4.2.1.1 Single Photoelectron Spectrum

To obtain the absolute gain of the phototube at a certain voltage, a single photoelectron spectrum is measured. This is done by setting the PMT to a gain of around 2×10^6, according to measurements done at Photonis, and flashing the LED at an intensity such that 90% of the time there are no photoelectrons (pe) at the first dynode in the PMT. From Poisson statistics:

$$P(n) = \frac{e^{-\mu}\mu^n}{n!} \tag{4.2}$$

where $P(n)$ is the probability to see exactly n pe, if there are 0 pe 90% of the time, then $P(0) = e^{-\mu} = 0.9$, so that $\mu = 0.105$. Then, the probability of seeing 1 pe is: $P(1) = \mu e^{-\mu} = 0.095$, and the probability of seeing more than 1 is 0.005, or there is a ∼0.5% contamination of events caused by 2 or more pe in any given single photoelectron spectrum. The signal then is dominated by single photoelectron events, $P(1)/P(n > 1) = 21$.

A typical single photoelectron spectrum can be seen in Fig. 4.8. To calculate the gain from this spectrum, the pedestal (or the signal deposited with no light) and the standard deviation of the pedestal must be established in a separate measurement. Once this is known, data is taken with the light source such that 90% of the events are 0 pe events. The resulting spectrum contains events from 0, 1, 2, etc. pe. To calculate the gain, it is necessary to find the mean of the single pe distribution, correcting for 2 pe contamination and compensating for events that are under the pedestal.

To quantify the resolution of the single photoelectron peak, the peak to valley ratio is used. The peak to valley ratio is the ratio of the maximum of the single photoelectron peak to the minimum between the pedestal and the spe peak. To calculate this number, smoothing has to be done on the distribution, which is done by taking the average of the current bin with the 2 bins before it and the 3 bins that follow it. That value is then used as the value in that bin. One then just steps through the bins and finds the minimum value (between the pedestal and the mean of the single photoelectron distribution) and the maximum value of the histogram. In Fig. 4.8, the peak to valley ratio is 1.47. In Fig. 4.9, the distribution of peak to valley ratios for PMTs tested in the system is shown.

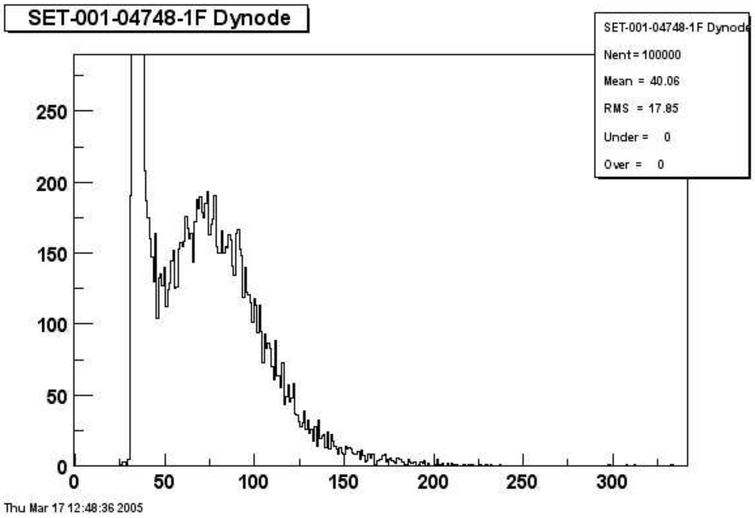


Figure 4.8: Typical single photoelectron spectrum. The first tall peak is the pedestal and second peak is due to 1 pe events. The peak to valley ratio in this plot is 1.47.

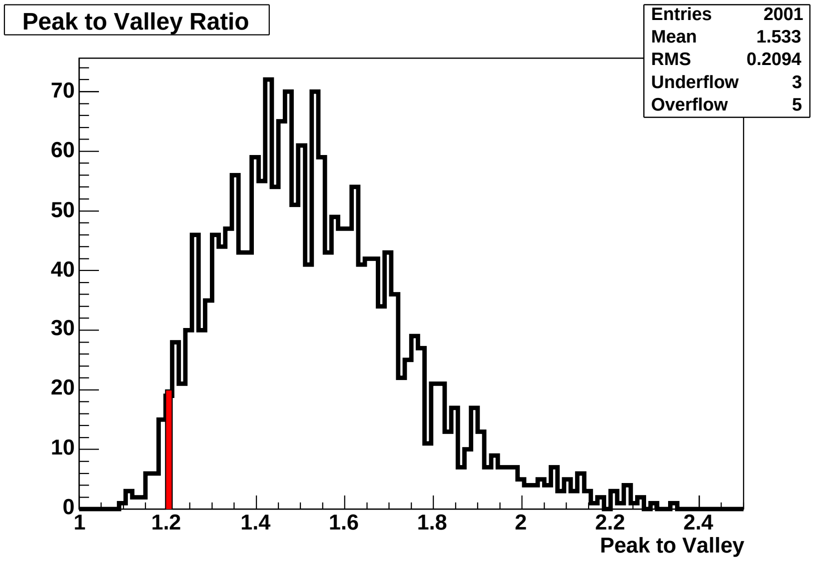


Figure 4.9: Distribution of peak to valley ratios for 2001 PMTs tested in the system at Malargüe. The red line represents the specification given to Photonis for the minimum desired peak to valley ratio. 98% of the PMTs pass this spec.

4.2.1.2 Gain versus Voltage

Once the gain is calculated from the spe spectrum, the absolute gain at that voltage is known. However, one needs several points to determine the equation for the gain as a function of the input voltage. It is known that the relationship between gain and input voltage is a power law [67]:

$$G = \alpha V^{\beta} \tag{4.3}$$

$$\ln G = \gamma + \beta \ln V \qquad \text{with} \qquad \gamma = \ln \alpha \tag{4.4}$$

with γ and β being parameters to be determined from measurements. The LED is pulsed with a constant intensity and the PMT is set at different voltages. This determines the slope of the relationship between gain and voltage. The value of the gain at the voltage used by the spe measurement is then used as a point that this line must pass through. This uniquely determines the two unknown parameters and one can calculate the gain of the phototube at any input voltage. An example of the gain versus voltage curve is given in Fig. 4.10, and for this particular PMT, γ = -12.190 and β = 5.7562.

In Fig. 4.10, the result of the measurements at Photonis is shown as well to illustrate the difference between two methods of measuring the gain. At Photonis, the measurements are made using a constant light source and measuring the current of the cathode and the anode. The gain is then the ratio of the anode current to the cathode current. In this method, the collection efficiency of the first dynode is included, whereas, in the spe method it is not. As a result, one can take the ratio of the gain as calculated by the spe method and the gain calculated using the method just described and obtain an estimate of the collection efficiency at a given voltage. Calculating the collection efficiency is not necessary in any PAO physics, but it is a nice bonus feature of the different

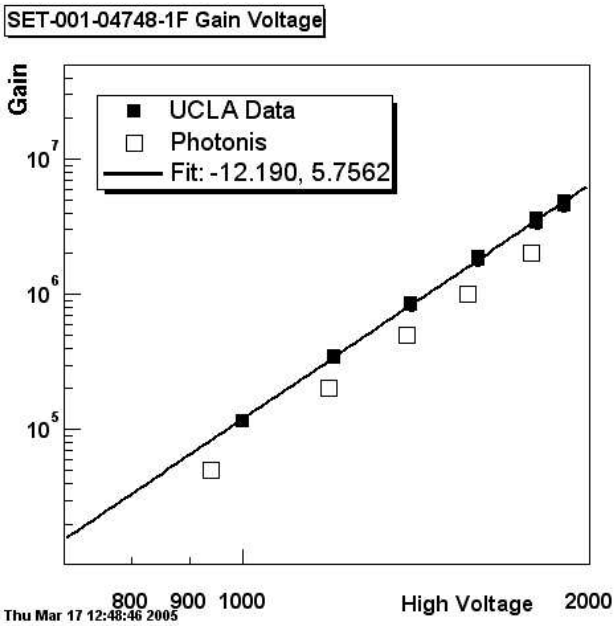


Figure 4.10: A typical gain versus voltage curve for a PMT tested in the system at Malargüe, with the curve obtained by Photonis plotted as well.

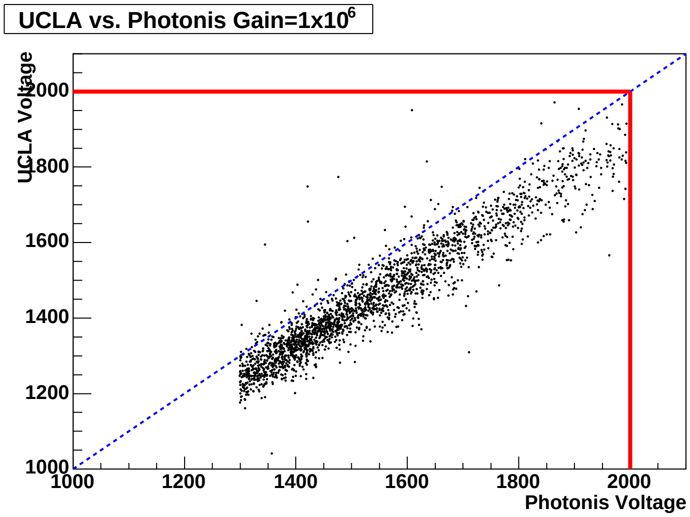


Figure 4.11: Required voltage necessary for 10^6 gain as measured by UCLA and Photonis. The red line is the spec. given to Photonis and the dashed line is y=x.

measurements. The difficulty arises in understanding the systematic errors in comparing two measurements from different systems. Figures 4.11 is a plot of the voltage necessary to get a gain of 10^6 as determined by the spe method (axis labeled UCLA Voltage) and using the current ratio method (axis labeled Photonis Voltage).

4.2.1.3 Dark Pulse Rate

When a PMT is exposed to a lot of light, like when it is exposed to the light in a room or daylight, it needs a chance to sit in the dark to reduce the dark pulse rate. The dark pulse rate is the rate at which signals above a certain threshold are observed in a given PMT. It is analogous to the dark current discussed earlier. In the test system, before other tests are run, the PMTs are allowed to cool down for a period of 2 hours, during which time the dark pulse rate above 1/4 pe and 1/2 pe threshold is monitored. Figure 4.12 is a plot of a typical dark pulse rate versus time.

Of course, to know the magnitude of the 1/4 pe threshold for a PMT, the absolute gain of the PMT must be known. Therefore, after the two hour cool down period, the gain of the PMT is calculated and the dark pulse rate measurement is taken again. The distribution of dark pulse rates for the measured PMTs is in Fig. 4.13. The specification given to Photonis was that after two hours in the dark the dark pulse rate above 1/4 pe would be below 10 kHz. This limit is set because dark pulse rate is related to the lifetime of a PMT, and the less noise in the PMT, the longer it will last.

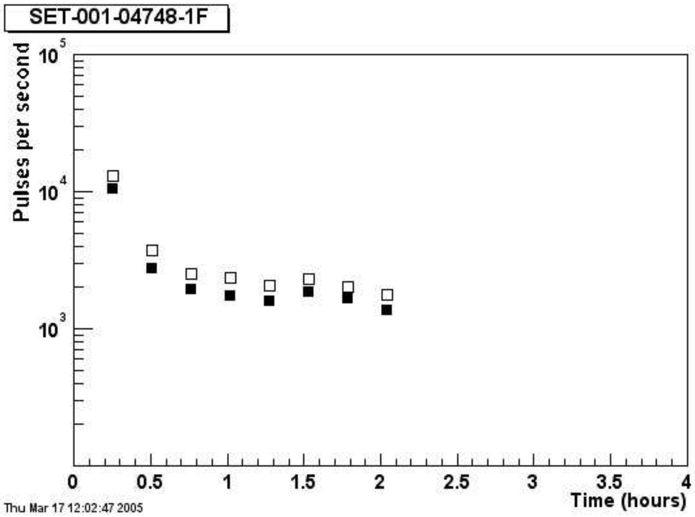


Figure 4.12: Dark pulse rate versus time spent in the dark room. Open boxes are 1/4 pe threshold while solid boxes are 1/2 pe threshold.

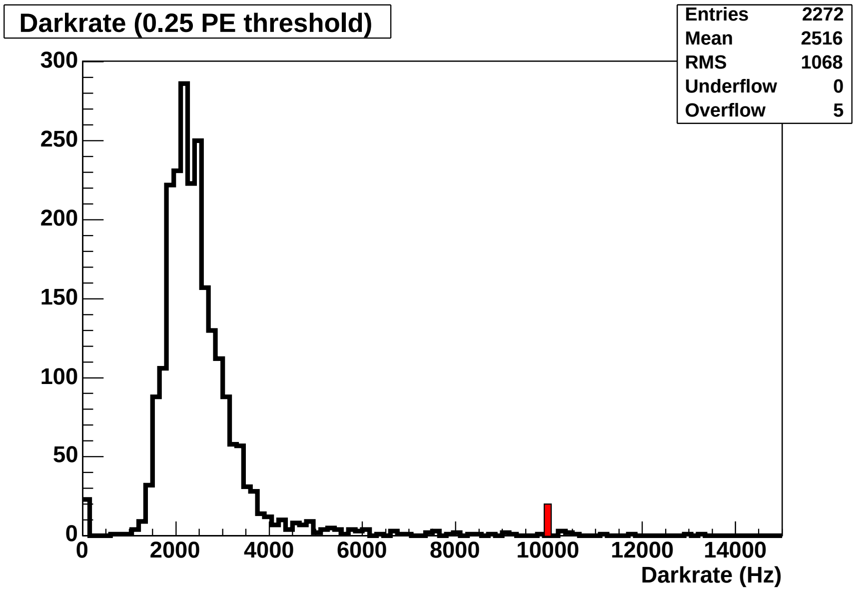


Figure 4.13: Distribution of dark pulse rates (1/4 pe threshold) after 2 hours in the dark room. The red line represents the spec. given to Photonis. More than 99% of the PMTs pass this spec.

4.2.1.4 Non-linearity

The linearity of the PMTs is important for the overall performance of the surface detector. The calibration of the PMTs is done with single muons passing through the tank whereas large showers can have thousands of particles passing through the tank. To extend the calibration from single muons to thousands of particles, the non-linearity of the PMT is a concern. Non-linearity normally occurs when the current gets so high that there is a space charge effect around the last dynode. This space charge effect is caused by an excessive amount of electrons which changes the electric field in that region and consequently causes the normal trajectory of the electrons to be skewed. Thus, the amount of electrons arriving at the last dynode, and hence the anode, is smaller than expected. This causes a negative non-linearity. In many of the PMTs from Photonis, there is a positive non-linearity. This is due to the design of the dynode chain. It is designed to more efficiently collect electrons at higher currents which causes signal to be lost at lower currents. This appears as a positive non-linearity due to the method used to calculate non-linearity.

The method to measure non-linearity uses two LEDs. LED A is fired, LED B is fired, LED A and B are fired simultaneoulsy, then no LED is fired (to obtain the baseline or pedestal). The non-linearity is then defined as:

$$NL(\%) = 100 \times \frac{Q_{AB} - (Q_A + Q_B)}{Q_A + Q_B} \tag{4.5}$$

where Q_A is the signal from firing LED A alone, Q_B is the signal from LED B alone, and Q_{AB} is the signal from firing LED A and B simultaneously (all are baseline subtracted). This sequence is repeated at several light intensities to map out the non-linearity as a function of peak anode current. A typical non-linearity curve is shown in Fig. 4.14. In this figure, the positive non-linearity feature is

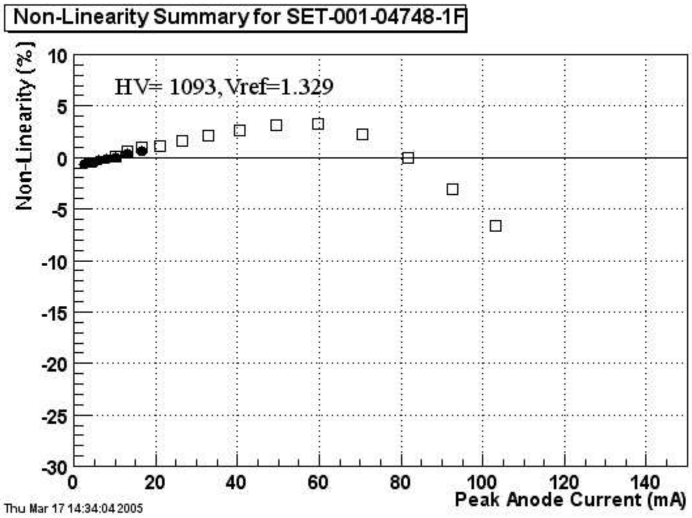


Figure 4.14: Non-linearity versus peak anode current for a typical PMT.

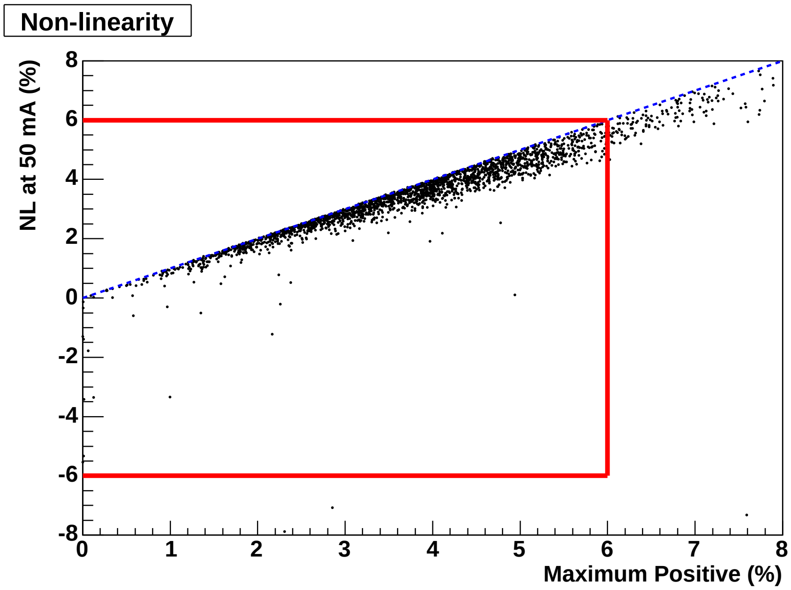


Figure 4.15: Non-linearity at 50 mA versus maximum positive non-linearity. The red lines represent the 6% limit given to Photons and the blue dotted line is y=x. 98% of the PMTs pass this spec.

evident as well as the following negative non-linearity due to the space charge effect.

To illustrate the properties of the non-linearity in the Photonis tubes, a plot of the maximum non-linearity versus the non-linearity at 50 mA is shown in Fig. 4.15. The reason 50 mA is chosen is because the specification given to Photonis was that the non-linearity be less than ±6% with a peak anode current of less than 50 mA. In the original design of the Pierre Auger Observatory, this was estimated to be the peak current at 1000 m from the core of an air shower initiated by a cosmic ray with an energy of 10^{21} eV.

In Fig. 4.15, it is evident that almost all the PMTs have a positive non-linearity, and the maximum positive non-linearity often occurs around 50 mA of peak anode current. It should be noted that the maximum positive non-linearity is defined as the maximum non-linearity with a peak anode current less than 50 mA. If non-linearity continues to increase after a current of 50 mA, that is not considered in the definition of maximum positive non-linearity, which is why there are no points above the blue dotted line in Fig. 4.15.

4.2.1.5 Dynode to Anode Ratio

To enable detection of small signals in the water Cherenkov detectors and extend the dynamic range of the detector, the signal from the last dynode is extracted and amplified. The amplification is fixed via the electronics on the board to be a factor of 40. There are then two signals from the PMT, the amplified dynode and the signal from the anode. As stated before, the calibration is done with single muons which give a small signal and are recorded using the dynode signal. Once a large number of particles pass through a detector from an air shower, the dynode reaches the maximum dynamic range of the FADC and the signal from the anode is used in the analysis. To be able to extend the muon calibration from the dynode to the air shower signal from the anode, the ratio of the dynode signal to the anode signal must be known. This ratio depends on the gain of the last dynode since the amplifier is fixed to a gain of 40.

$$D/A = \alpha = 40\frac{\delta - 1}{\delta} \tag{4.6}$$

Equation (4.6) gives the relationship between the dynode to anode ratio (α) and the gain of the last dynode (δ). The factor of 40 is the value of the gain of the amplifier. The factor of $(\delta - 1)/\delta$ represents that for every electron that hits the last dynode, δ electrons leave. This gives a signal of $1 - \delta$ on the last dynode and δ on the anode (in arbitrary units). The amplifier inverts and amplifies the signal from the last dynode to make it the same polarity as the anode.

To measure the dynode to anode ratio, the PMTs are set to a fixed gain and the light source is flashed at varying intensities. The signal of the dynode versus the signal of the anode is plotted and the slope of the line gives the dynode to anode ratio. Of course, the dynode to anode ratio depends on the gain of the last dynode, and the gain of the last dynode is dependent on the voltage to the PMT. A plot of the gain of the last dynode as a function of input voltage is in Fig. 4.16, where offset and slope refer to γ and β from (4.4). The distribution of dynode to anode ratios at a gain of 10^6 is in Fig. 4.17. The nominal value of the dynode to anode ratio is 32 so accepted values range from 25 to 40.

4.2.1.6 Excess Noise Factor

The amplification process in a PMT is a convolution of a number of statistical processes. From the photocathode to each dynode stage, the number of electrons that are ejected is a Poisson process. As a consequence, the number of initial

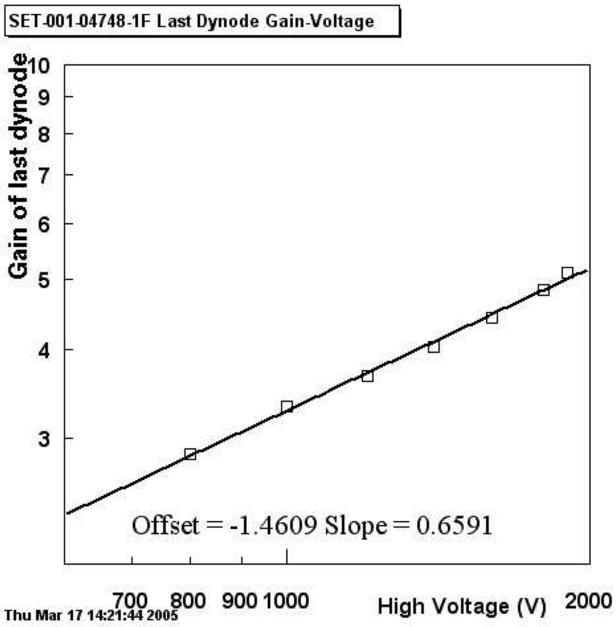


Figure 4.16: Gain of the last dynode versus input voltage for a typical PMT.

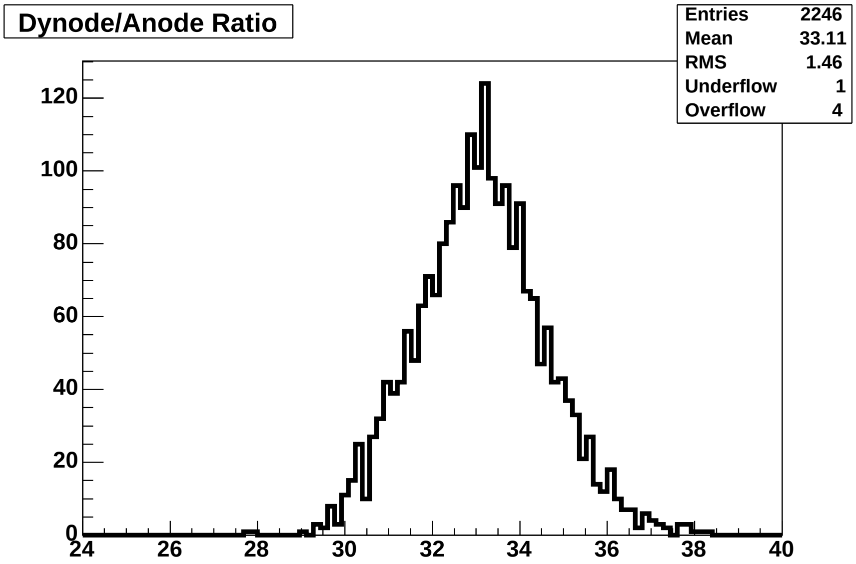


Figure 4.17: Distribution of dynode to anode ratios at a gain of 10^6 for 2246 PMTs. Only 5 PMTs fall outside the accepted 25 to 40 region (0.2%).

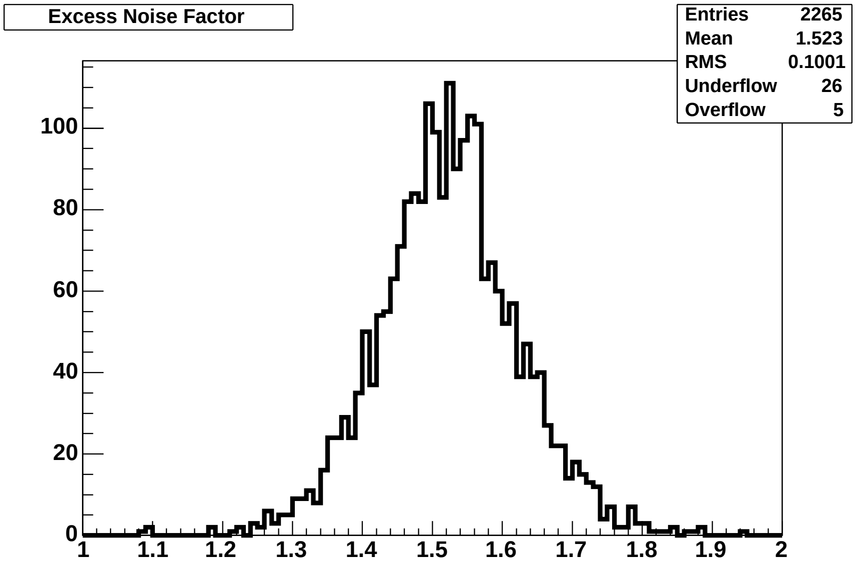


Figure 4.18: Distribution of excess noise factors at a gain of 10^6 for 2265 PMTs.

photons can not be calculated from the spread in measured values because the amplification introduces a widening of the distribution. The excess noise factor (ENF) is defined as:

$$\sigma^2_{input} = ENF \cdot \sigma^2_{output} \tag{4.7}$$

and is this “widening” factor. The distribution of excess noise factors at a gain of 10^6 is presented in Fig. 4.18.

The excess noise factor is related to the peak to valley ratio of the single photoelectron spectrum. The larger the ENF, the broader the distribution of output signals will be for the same input. Therefore, we expect an anti-correlation between ENF and peak to valley ratio. There is no specification for the excess noise factor of a PMT, but because it is related to the peak to valley ratio any excessive noise will cause a failure in the peak to valley requirement. This is

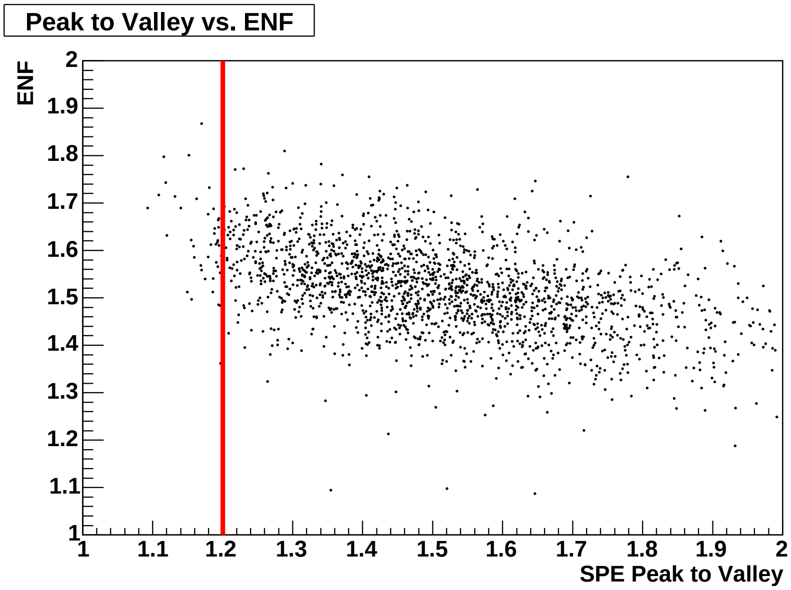


Figure 4.19: ENF versus peak to valley ratio showing anti-correlation. Red line represents the peak to valley specification given to Photonis (>1.2).

plotted in Fig. 4.19.

4.2.1.7 Afterpulse

One concern with PMTs is contamination of gases. The PMT is made of a glass envelope around a dynode structure with a vacuum inside the glass tube. If there are molecules of gas inside the glass envelope, as the photoelectrons pass through the gas the molecules will ionizes and these ions will travel back to the glass where they will make more electrons. This will cause a pulse proportional to the initial pulse delayed in time anywhere from hundreds of nanoseconds to microseconds, depending on the gas. Ultimately, this could cause a miscalculation of the energy deposited in a surface detector and thus in the energy of the primary particle.

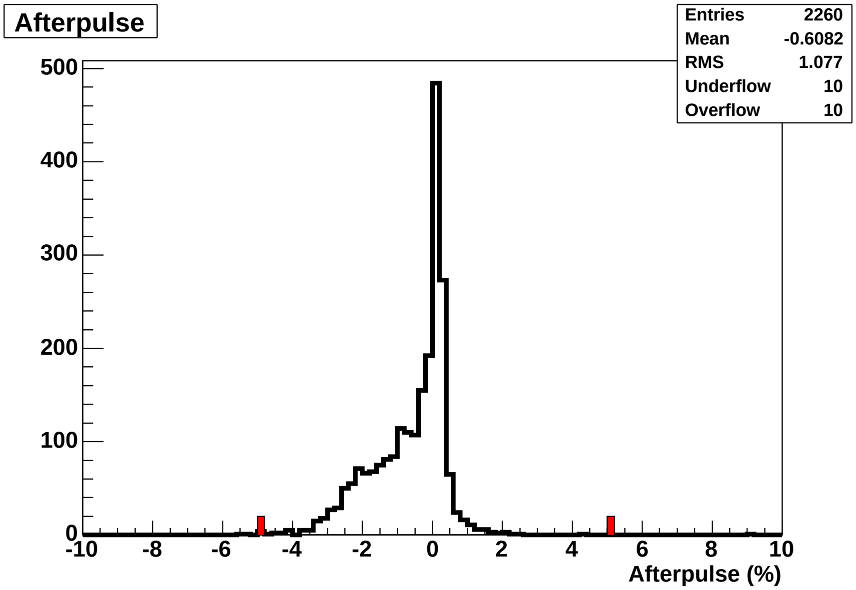


Figure 4.20: Afterpulse distribution for 2260 PMTs. 99% of the PMTs pass this spec.

In the PMTs tested, there is no significant afterpulsing measured indicating that the vacuum is free from gases. There is a systematic negative afterpulsing (see Fig. 4.20) due to the system used to measure the PMTs. This negative signal is due to the position of the baseline being lower for a period of time after a significant signal. It is only significant when integrating over long times ($\sim 5\mu$s) which is done in the afterpulse measurements and hence not an issue in the previous tests discussed which only integrate over a couple hundred of nanoseconds.

4.3 Test System Performance

The test system, as it is operated, has four PMTs that are left permanently in the test stand. These PMTs monitor the performance of the system as a whole,

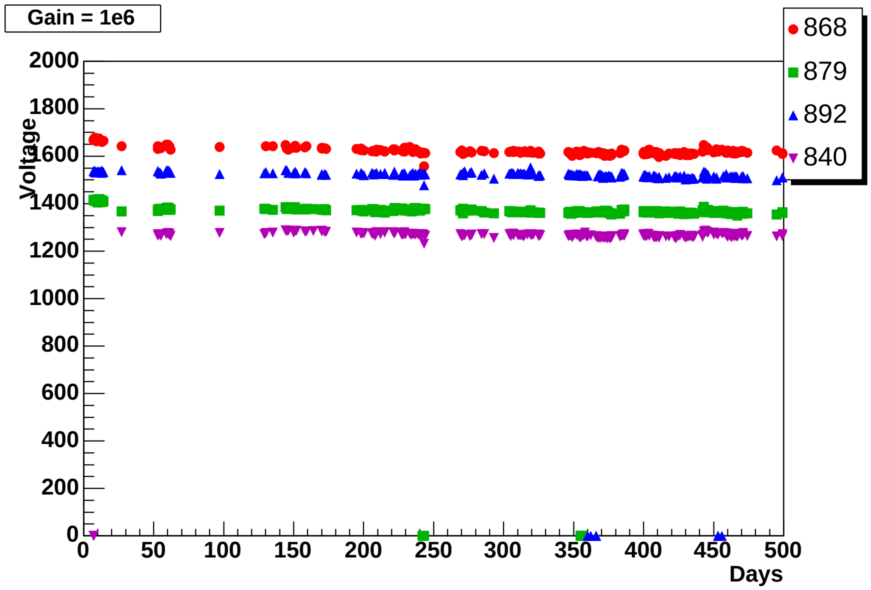


Figure 4.21: Voltage to get a gain of 10^6 for the four permanent PMTs over a 500 day period. The consistency of the value for a given PMT is what is monitored.

from the LEDs to the DAQ. The permanent PMTs provide a comparison for tests run currently versus tests run when the system was first commissioned, to see any systematic shifts or anomalous behavior. In addition, checking the spread of the measurements of a given parameter for the four permanent PMTs indicates the resolution of the test system, assuming that time and temperature effects are negligible.

Figure 4.21 is a plot of the voltage necessary for a gain of 10^6 for the four permanent PMTs during a period of 500 days. There is no noticeable drift of this value with time, and any temperature effects are lost in the spread of the voltages. One PMT is taken as an example, PMT 892, to show the spread of the measurements over this same time period (see Fig. 4.22). For this PMT, the

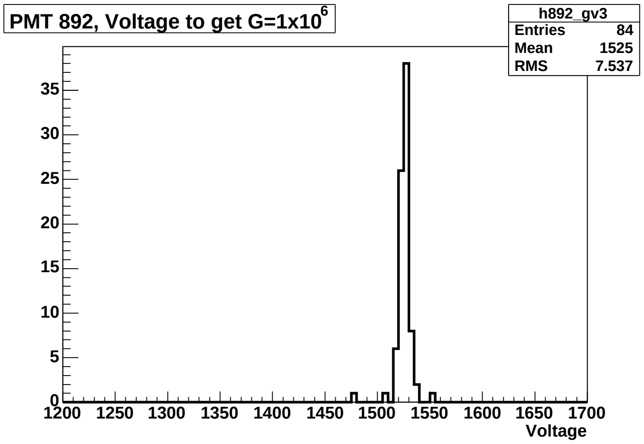


Figure 4.22: Distribution of the voltage for a gain of 10^6 for PMT 892 over a 500 day period.

standard deviation is 7.5 V with a mean of 1525 ± 1 V. This corresponds to a standard deviation of less than 3% on the gain for the given voltage. Repeating this process for all the tests, the resolution of the system is determined for each test and is reported in Table 4.2.

For an example of the monitoring capabilities of the four permanent PMTs, there was a drift detected in the system in the non-linearity measurements starting around day 350, see Fig. 4.23. Since all 4 PMTs experienced the same change in behavior, the drift can be attributed to something that is happening in the test system itself and the results for the tested PMTs can be corrected for this behavior. The cause of this drift is unknown, but it can be monitored and corrected for in the results.

To determine any systematic error associated with location in the test stand, test results are also plotted as a function of position in the test stand. Each position in the test stand is in a fixed location, meaning the orientation with

Test	868	879	892	840	Average
SPE Peak to Valley ratio [%]	4.5	4.7	5.5	5.7	5.1
G=2×10^5 Voltage [%]	0.64	0.48	0.62	0.57	0.58
G=1×10^6 Voltage [%]	0.66	0.46	0.49	0.61	0.56
G=2×10^6 Voltage [%]	0.72	0.48	0.48	0.66	0.59
Dark Pulse Rate [Hz]	897	963	643	764	817
Non-linearity at 50 mA [%]*	0.35	0.39	0.49	0.38	0.40
Max. positive non-linearity [%]*	0.31	0.39	0.48	0.38	0.39
Dynode to Anode Ratio	0.34	0.40	0.41	0.31	0.37
Excess Noise Factor	0.07	0.06	0.06	0.07	0.07
Afterpulse [%]*	1.5	2.0	2.0	1.2	1.7

Table 4.2: Resolution of the PMT Test stand based on the standard deviation of the measurements of the specified tests for the permanent 4 PMTs. For the peak to valley and the gain voltage values, the percentage is given of the standard deviation to the mean value. For all other tests, the value given is the raw value of the standard deviation of the measurements because the spread is independent of the mean value of the observable. Values with an asterisk (*) denote that the value is the raw spread in the variable which is a measurement of a percentage and are not to be confused with the percentage of the standard deviation to the mean value.

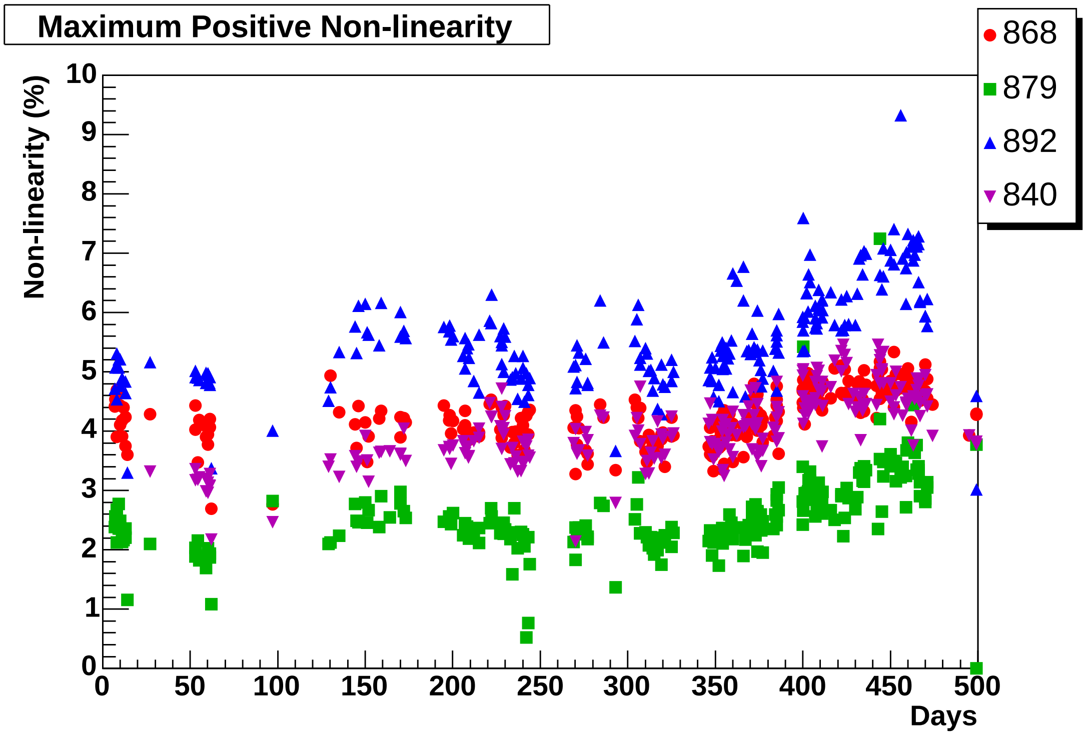


Figure 4.23: Maximum positive non-linearity for the four permanent PMTs. A system wide drift is noticeable starting around day 350.

respect to the LEDs is constant, as well as the data acquisition electronics is constant for the given location in the test stand. For examples of these plots see Fig. 4.24. In these plots, each data point for the given location in the test stand has anywhere from 100 to 200 PMTs to compute the average so any systematic difference with location is not due to poor statistics. In Fig. 4.24, the mean dark pulse rate does not depend on the location in the test stand, whereas non-linearity measurements vary with the location in the test stand by around 1%. Any systematic error due to the location in the test stand can be corrected.

4.4 PMT Testing Conclusions

The PMTs being used in the surface detector array for the Pierre Auger Observatory are being thoroughly tested. During the initial test phase, they were tested to make sure their performance would suit the needs of the experiment. From these first tests, requests were made to alter the original design of the PMTs. These modified PMTs were again tested and one company was selected to produce all the PMTs to be used in the surface detector array.

For all the PMTs used in the surface detectors, their characteristics and behavior are quantified in the PMT test system. These results are catalogued in a database for use in the analysis and monitoring of the surface detector. The tests run on the PMTs ensure uniformity and quality in the performance of the surface detectors and provide valuable feedback to the company producing them.

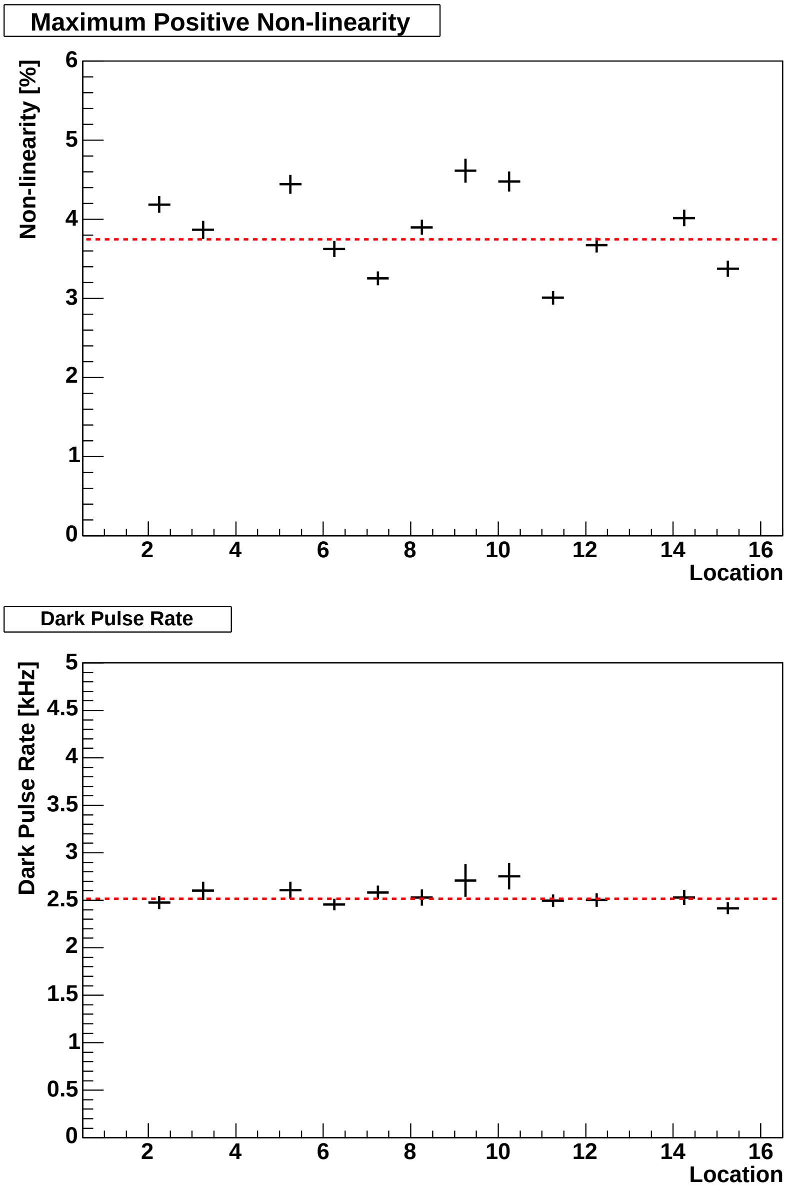


Figure 4.24: Top: Maximum positive non-linearity as a function of location in the test stand. Bottom: Dark pulse rate as a function of location in the test stand. In each, the red dashed line is the mean value of that parameter for all tested PMTs, and vertical lines on the points are the uncertainty on the mean.

CHAPTER 5

Monitoring the Performance of the Array

In Chapter 3, the calibration of the surface detector was discussed. In this chapter, the monitoring of certain calibration quantities is presented. The monitoring of calibration quantities is essentially monitoring the performance of the PMTs along with the associated readout electronics. For each surface detector, monitoring data is sent to the central data acquisition system (CDAS) in regular time intervals [62]. It is by using this data that the performance and long term behavior of the surface detector array can be monitored.

5.1 Monitoring Detector Performance

The monitoring data sent to CDAS by each station contains parameters that are used in physics analyses. These parameters are associated with the calibration of the station as well as the stability of the DAQ electronics. The particular parameters of interest that are monitored are the baselines of the anode and dynode channels in the FADC, Q_{VEM}, the ratio Q_{VEM}/I_{VEM}, and the dynode to anode ratio. Each parameter will be explained in further detail in the following sections. For each parameter, the behavior over time and as a function of temperature was investigated. To study the array as a whole, data was broken down into two month time periods to see any change in the quality of data or any particular parameter over long time periods.

5.1.1 Anode and Dynode Baselines

Each detector of the surface array has six channels of input, the dynode and the anode signal from each of the three PMTs. These signals are read out with a FADC. For each of these six channels, there is a baseline current in the FADC, usually between 20 to 100 channels (where the full range of the FADC is 1024 channels). An example of the baselines for a particular station is shown in Fig. 5.1. It is necessary to know the baselines, obviously, to be able to correctly determine the integrated signal for each PMT.

Any instability in the baselines indicates a problem with the FADC. It is important to track the baselines over time and check for any erratic behavior. A plot of the distribution of the standard deviation of the dynode baselines from July to August, 2004 and March to April, 2005 are in Fig. 5.2.

5.1.2 Q_{VEM} and Q_{VEM}/I_{VEM}

The calibration for the surface detectors is done using atmospheric muons. These muons are passing through the tanks at a high rate and can be used to calibrate the detectors with a high statistics "standard candle". The calibration point is the vertical equivalent muon (VEM), which is the energy deposited in a detector by a muon that transverses the entire height of the tank vertically and the point of entrance of the muon is located in the center of the detector. All signals in the detectors are calibrated in terms of this unit, Q_{VEM} [63].

Another related quantity is the peak current of a VEM. This quantity is necessary for triggering since the FADC measures current and the charge is only known after integrating the FADC trace. Therefore, the local trigger on the detector is a threshold trigger, or peak current trigger, rather than a charge

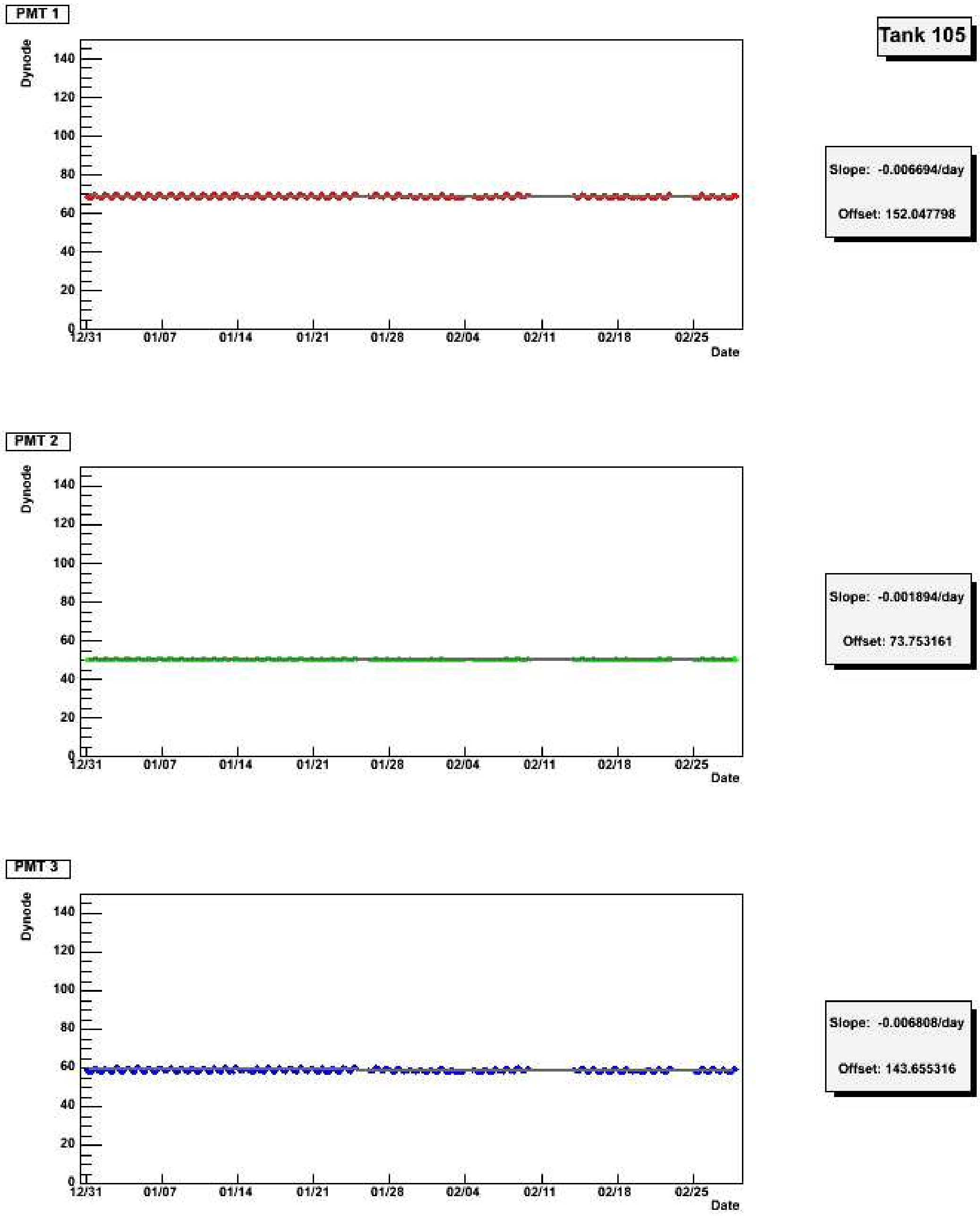


Figure 5.1: Plot of the baselines for a particular detector versus time. From top to bottom, they are for the dynode channel in PMT 1, 2, and 3.

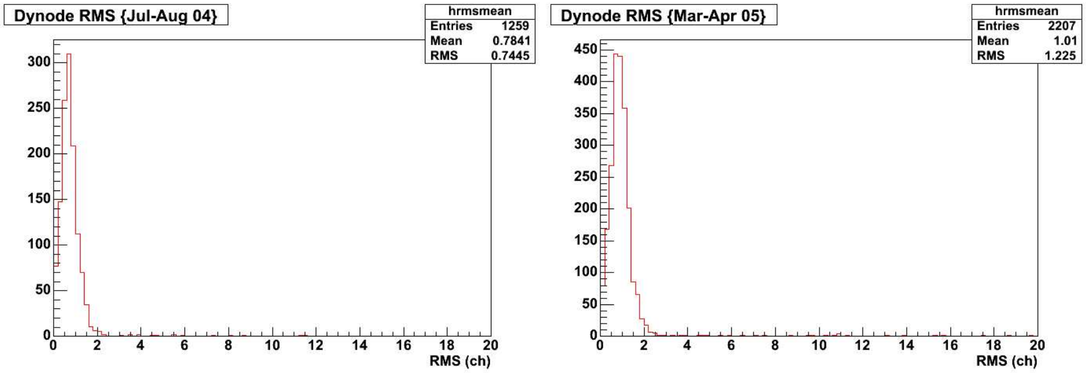


Figure 5.2: Dynode pedestal fluctuations (standard deviation) for two two-month periods.

dependent trigger. The peak current deposited by a VEM is called I_{VEM}. The ratio of Q_{VEM} to I_{VEM} is an indicator of the pulse shape, and thus related to the properties of the tank such as water absorption and reflectivity of the inside of the detector. The larger this ratio, called the area to peak ratio, the longer the pulse from a muon takes to attenuate in the detector, indicating a longer water absorption length and/or a more reflective inner surface. Conversely, dirty water and/or poor reflectivity result in a short pulse from a muon because the light is attenuated in a shorter time, giving a smaller area to peak ratio.

Q_{VEM}, then, is a useful parameter to use in monitoring the stability of the detector calibration and electronics. If any problem arises in a given PMT or DAQ channel, it will be noticed in Q_{VEM} parameter. As an example, Fig. 5.6 shows the distribution of the fluctuations of Q_{VEM} over two two-month periods. Any PMT channel that exhibits unstable behavior can then be investigated as to the cause of the large fluctuations. As can be seen from the plots in Fig. 5.3, there is a temperature dependence for Q_{VEM} which can be understood as a convolution of temperature effects on the gain of the PMT as well as effects in the electronics. Fitting Q_{VEM} parameter as a function of temperature with a line (see Fig. 5.4), the temperature coefficient can be obtained. In Fig. 5.5, a plot of all the temperature coefficients is shown indicating an average increase of 0.029±0.006% in the charge deposited by a VEM per degree.

The area to peak ratio is monitored to check the quality of the water and inner liner of the detector. Since any instability of the electronics that is manifest in Q_{VEM} affects I_{VEM} as well, the area to peak ratio is not as sensitive to these instablities, such as temperature. Therefore, it is useful to monitor the area to peak ratio as an indicator of the quality of the detector itself. Again, the fluctuation of the area to peak ratio is plotted for two two-month periods to monitor the stability of this parameter, see Fig. 5.6.

5.1.3 Dynode to Anode Ratio

As explained in Section 4.2.1.5, the dynode to anode ratio is essential in analyzing physics data. The method to calculate the dynode to anode ratio in the detector itself is more complicated than in the PMT test system. In the test system, the intensity of the light source can be increased uniformly over the dynamic range of the PMT. In the detector, due to the particle flux, there are a multitude of low energy signals in the tanks, with very few large energy signals. Thus, the same method to compute the dynode to anode ratio in the PMT test system fails due to poor statistics in the high energy signal area. Thus, a novel technique was developed to obtain the dynode to anode ratio using trigger rates or pulse shape fitting [70].

The dynode to anode ratio is monitored like the other variables discussed in the preceding sections, and the distribution of the standard deviation normalized by the mean for two two-month periods is shown in Fig. 5.7.

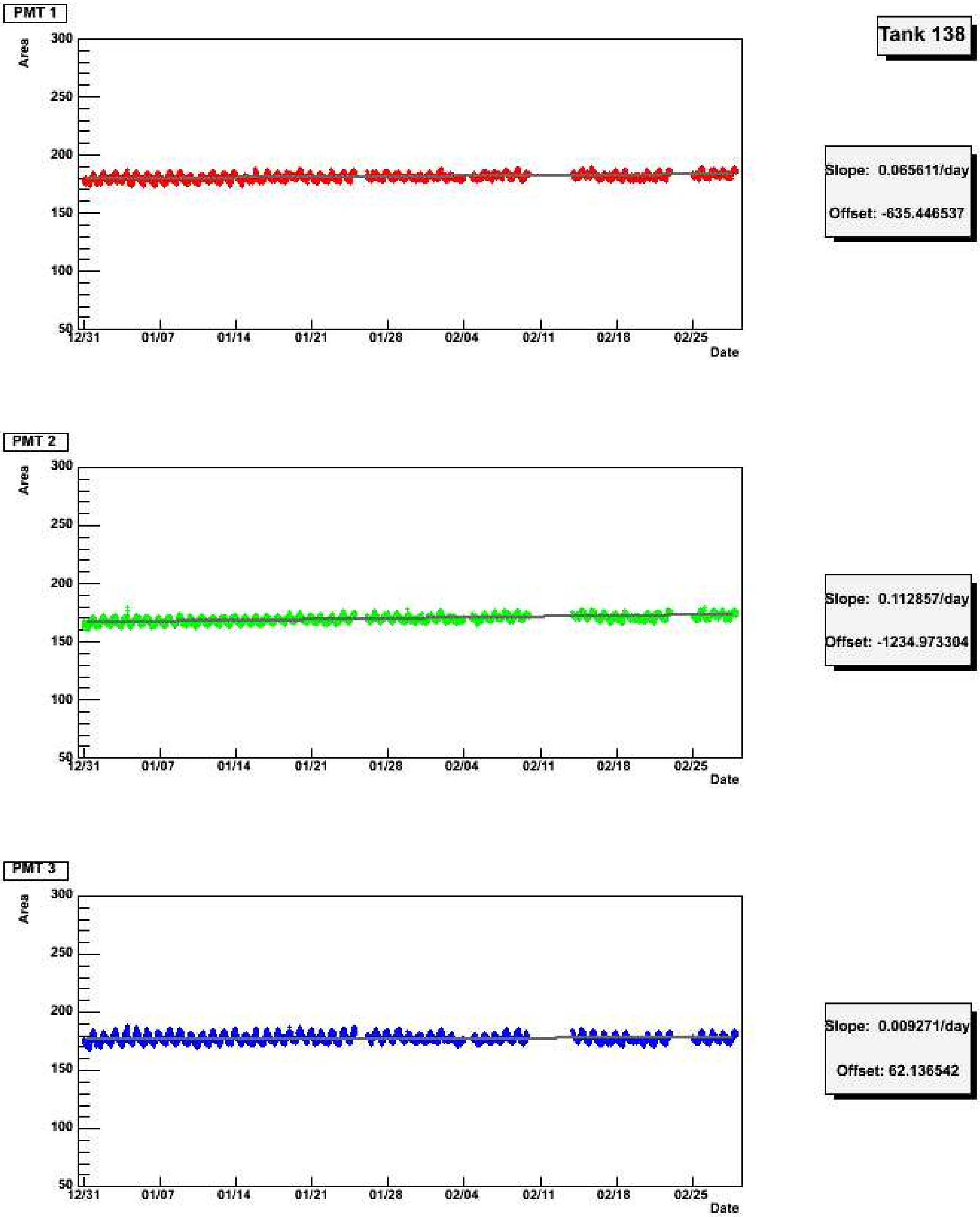


Figure 5.3: Plot of Q_{VEM} for a particular detector versus time. From top to bottom, they are for PMT 1, 2, and 3.

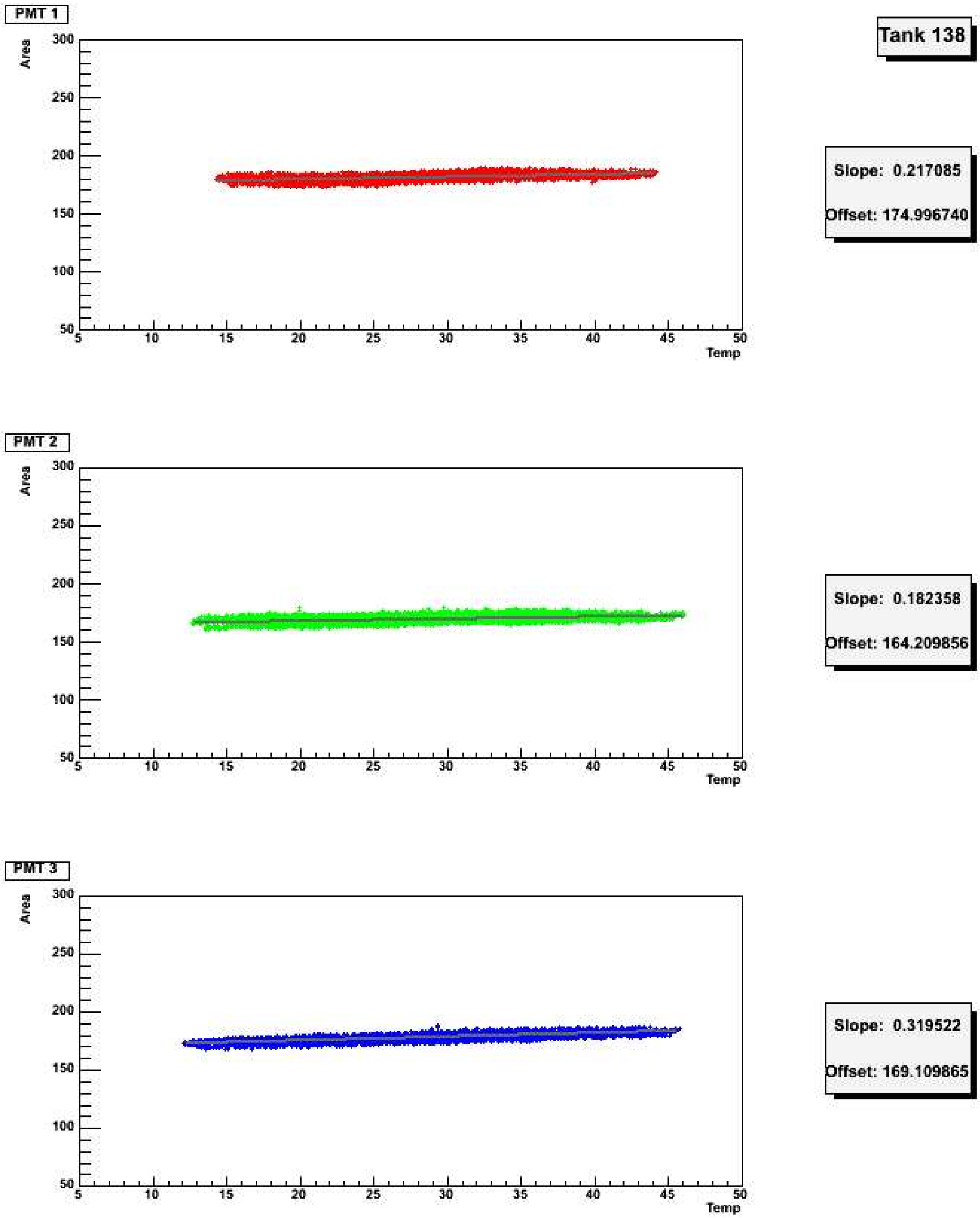


Figure 5.4: Plot of Q_{VEM} for a particular detector versus temperature. From top to bottom, they are for PMT 1, 2, and 3.

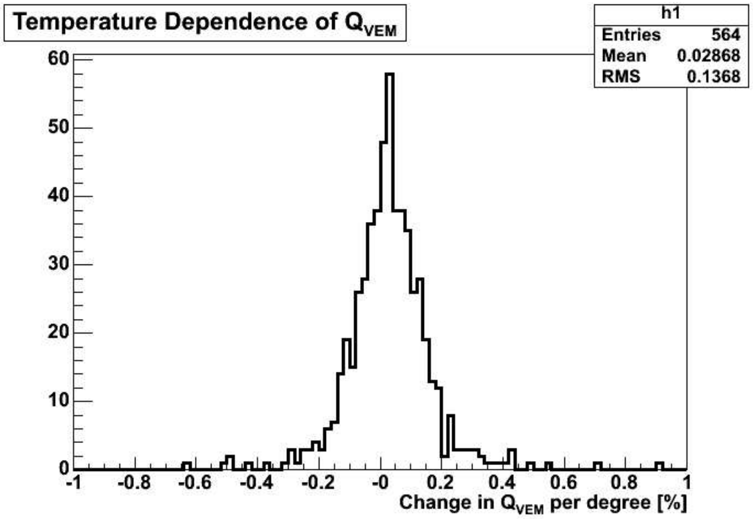


Figure 5.5: Q_{VEM} temperature dependence.

5.1.4 Number of photoelectrons per VEM

Combining the information from this and the previous chapter, the number of photoelectrons per VEM for a given PMT can be calculated [71]. The calculation involves data in the PMT testing database (gain as a function of voltage), PMT base database (input voltage to PMT voltage conversion), and from the monitoring database the dynode to anode ratio, input voltage supplied, and Q_{VEM}.

To operate the PMTs at a gain of $\sim 10^6$, 1000-2000 V would have to be supplied to the PMTs. It is not practical to supply the PMTs with these high voltages directly from the station electronics. A digital to analog converter (DAC) is used to supply a voltage between 0 and 2.5 V to the base of the PMTs. There is a module on the PMT base that then converts this DAC voltage to the required high voltages to achieve the desired gain. The conversion factor that relates the input DAC voltage to the output high voltage is a property of the high voltage

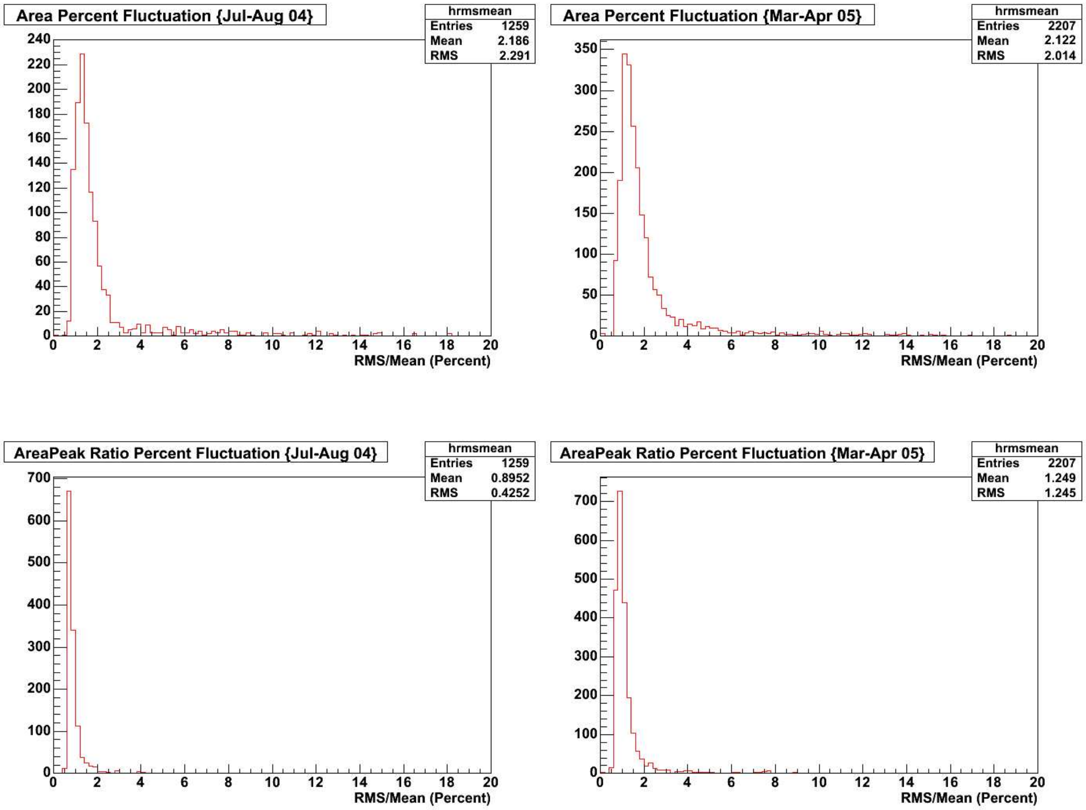


Figure 5.6: Q_{VEM} (top) and Q_{VEM}/I_{VEM} (bottom) fluctuations (standard deviation/Mean) for two two-month periods.

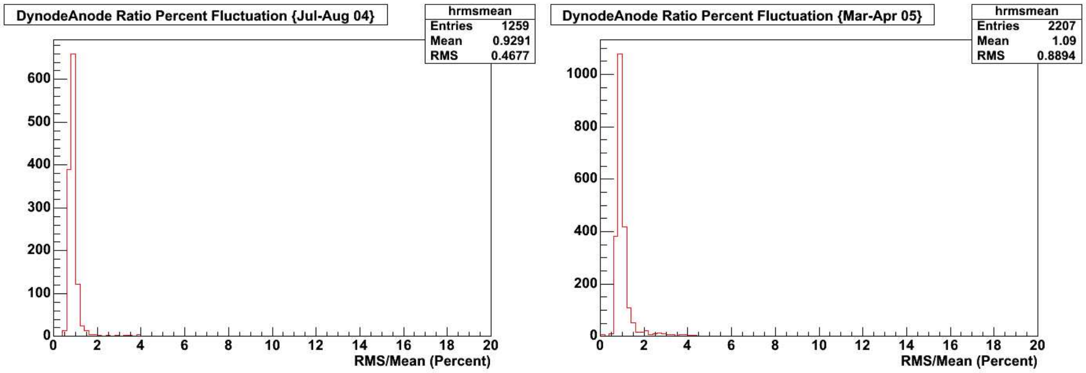


Figure 5.7: Dynode to anode ratio fluctuations (standard deviation/Mean) for two two-month periods.

module located on the base, and, therefore, is base specific. The input DAC value is an integer between 0 and 4095, a 10 bit DAC, which corresponds to 0 to 2.5 V DAC output, which is then amplified a factor of R_{base} by the module on the base. The voltage supplied to the PMT is then:

$$HV = \frac{2500\text{mV}}{4095\text{ch}} \times \text{DAC} \times R_{base} \tag{5.1}$$

where R_{base} is the base specific conversion from input to output voltage. Once the voltage of the PMT is known, the gain can be calculated from the PMT testing data, see (4.3). Combining these two equations and evaluating the constants, the gain of the PMT in the detector is:

$$G = 10^{\gamma}(0.61 R_{base}\text{DAC})^{\beta} \tag{5.2}$$

Once the gain is known, the number of photoelectrons at the first dynode can be calculated for any signal. It should be noted that the following equations are to obtain the number of photoelectrons at the *first dynode*. This is due to the fact that the gain as a function of voltage is obtained via the single photoelectron method. This, of course, only gives a result when a photon converts to a photoelectron and that photoelectron subsequently reaches the first dynode. To know the number of *photons* produced per VEM, the quantum efficiency (QE) and the collection efficiency (CE) must be known. These quantities are not reliably measured for each PMT, therefore the number of photoelectrons at the first dynode suffices.

For small signals, the dynode is used to calculate the integrated signal, or the charge deposited. Since the gain of the PMT refers to the number of electrons at the anode for 1 electron at the first dynode, the dynode to anode ratio (R_{DA}) must be known. The integrated charge at the dynode is then:

$$Q_{dynode} = e n_{pe} G R_{DA} \tag{5.3}$$

where e is the charge of an electron, and n_{pe} is the number of photoelectrons at the first dynode. This is the charge at the base, but when the signal enters the FADC, it passes through an amplifier with a gain of 0.5 (G_{FE}). This is to match the full scale of the base (0 - 2 V) to the full scale of the FADC (0 - 1 V). The charge deposited in the FADC is:

$$Q_{FADC} = e n_{pe} G R_{DA} G_{FE} \tag{5.4}$$

The 10 bit (1024 channels) FADC samples at a rate of 40 MHz ($t_s = 25$ ns) and goes through a 50 Ω load (Z_{load}). Thus, the charge as measured by the detector is:

$$Q_{meas} = \frac{G_D t_s \sum_i x_i}{Z_{load}} \tag{5.5}$$

where G_D is the range of the FADC ($\frac{1\text{V}}{1023\text{ch}}$), x_i is the FADC value of the ith sample, and the sum runs over the width of the pulse.

Setting Q_{FADC} and Q_{meas} equal to each other and solving for n_{pe}, the equation is obtained for the number of photoelectrons at the first dynode for a given integrated signal:

$$n_{pe} = \frac{G_D t_s (\sum_i x_i)}{e Z_{load} R_{DA} G_{FE} G} \tag{5.6}$$

Substituting in the values of the constants, the equation simplifies to:

$$n_{pe} = 6 \times 10^6 \frac{(\sum_i x_i)}{R_{DA} G} \tag{5.7}$$

To calculate the number of photoelectrons per VEM, the charge deposited per VEM (Q_{VEM}) must be known. This was discussed in Section 5.1.2 and is constantly monitored by the detector. With (5.2) and (5.7) and exchanging the sum over the pulse with Q_{VEM}, the number of photoelectrons at the first dynode per VEM is:

$$n_{pe} = 6 \times 10^6 \frac{1}{R_D A} \frac{10^{-\gamma}}{(0.61 R_{base} \mathrm{DAC})^{\beta}} Q_{VEM} \tag{5.8}$$

During a two day period, Q_{VEM} and R_{DA} were averaged for the PMTs in all the detectors. Only PMTs that had a constant input voltage supplied over this same two day period were used in the analysis. Using the values obtained in that two day period and combining that with the PMT test data, the number of photoelectrons at the first dynode per VEM was calculated. The histogram of these values is shown in Fig. 5.8. In Fig. 5.9, a plot of n_{pe} per VEM versus gain is shown. Due to the output matching calibration method referred to in Section 3.2, an anti-correlation is expected and seen. The anti-correlation arises from the fact that a PMT within a station that receives fewer photoelectrons per VEM must be set to a higher gain to match the output of a PMT with a higher number of photoelectrons per VEM. That is the goal of the output matching calibration, to have each PMT produce the same signal for the same energy deposited in the tank. The uncertainty in n_{pe} is dominated by the 10% uncertainty in the gain of the front end (G_{FE}).

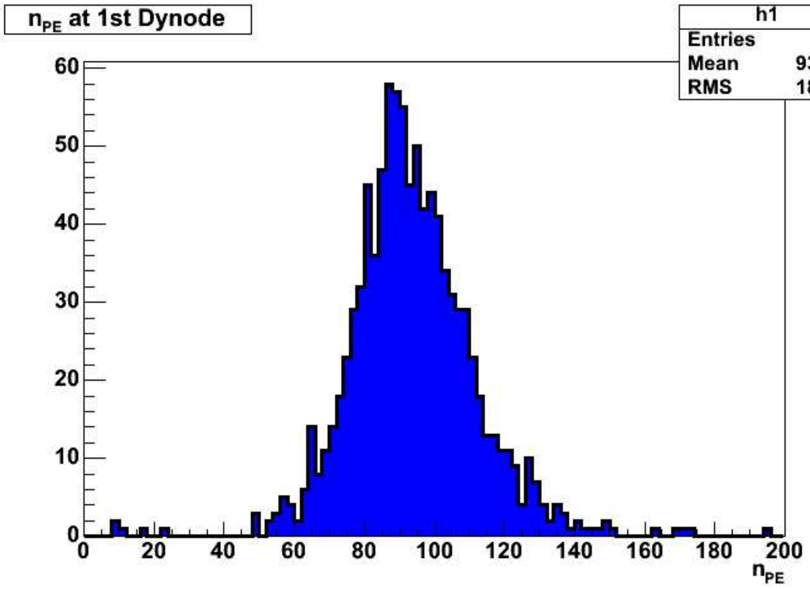


Figure 5.8: Histogram of the number of photoelectrons at the first dynode.

5.2 Conclusions

Monitoring the calibration and performance of the detectors in the surface array is necessary to instill confidence that the detector is stable, the calibration is stable and the fluctuations are known and below a desired level. Once this is confirmed, more credence can be given to the physics analysis. The monitoring capabilities have been demonstrated in this chapter and by using the techniques outlined here, the quantities of interest are stable with fluctuations less than 5% [72]. Combining the data from the PMTs and the detector monitoring, the number of photoelectrons at the first dynode per VEM can be calculated. For the array, the average n_{pe} per VEM is 93.7 ± 0.6 with a standard deviation of 18.2, and the average operating gain is $3.47 \pm 0.02 \times 10^5$ with a standard deviation of 0.56×10^5.

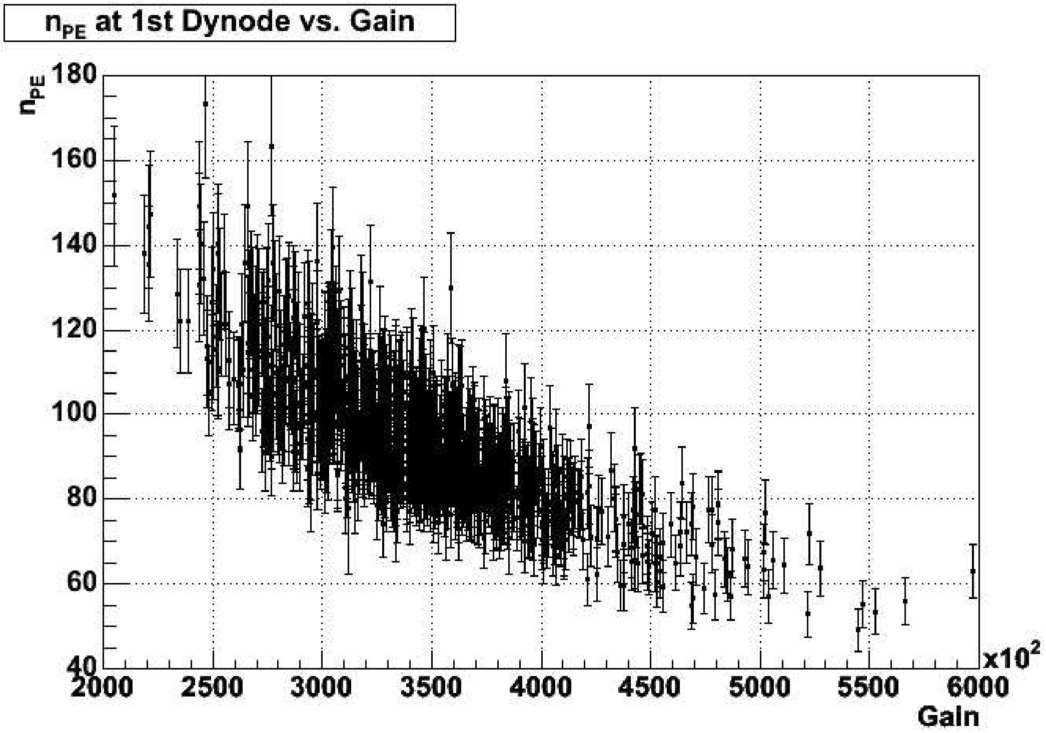


Figure 5.9: Number of photoelectrons per VEM versus gain. An anti-correlation is seen indicating that the PMT output for a VEM is the same (within uncertainties) across the array. The uncertainty in the nPE is dominated by the uncertainty in the gain of the front end (G_{FE}).

CHAPTER 6

Uncertainties in Energy Determination

In the preceding discussion it was shown that the calibration of the detector is stable, with less than 5% fluctuation in the parameters of interest. It is then advantageous to investigate any larger sources of uncertainty in the primary energy measurement of cosmic rays using the array of surface detectors (SD). Using the SD alone, the largest uncertainty arises from the lack of knowledge of the composition of the cosmic rays and uncertainties in the hadronic interaction models used in the simulations.

In this chapter, the physics behind these uncertainties will be discussed in conjunction with the Pierre Auger Observatory. Different air shower monte carlo programs will be discussed as well as different hadronic interaction models and their role in the uncertainty. From these studies, an estimate of the systematic and statistical uncertainties in the energy determination will be presented.

6.1 Monte Carlo Extensive Air Showers

In Chapter 1, the physics of extensive air showers was discussed. The shower grows in number of particles until the shower maximum, X_{max}, and then the number of particles decreases. The composition of the initial cosmic ray affects the depth at which the shower maximum occurs, and in a smaller degree so does the hadronic interaction model, see Fig. 6.1. The energy of the primary cosmic

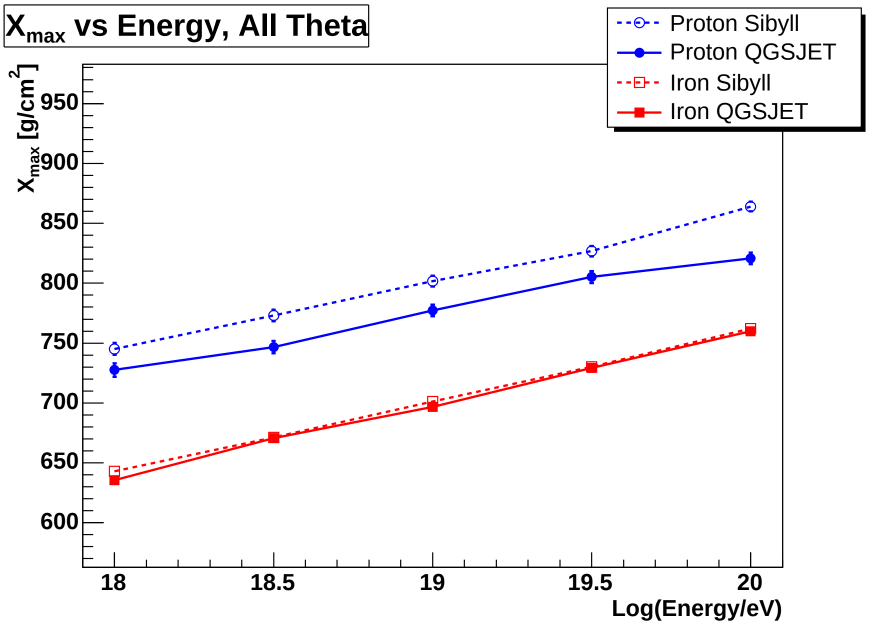


Figure 6.1: Mean X_{max} as a function of primary energy. Aires was used in conjunction with QGSJET and Sibyll for the hadronic interaction models with proton and iron as the primary cosmic rays.

ray is associated with the maximum number of charged particles, N_{max}, with little dependence on primary composition and hadronic interaction model, see Fig. 6.2.

The initial interactions between the cosmic ray and the atmosphere have a center of mass energy higher than any man made accelerator can attain. The monte carlo programs used to simulate these interactions are extrapolations of cross sections and particle multiplicities measured at lower energies. Thus, there are different programs used to simulate these high energy interactions based on different models and extrapolations. The effect of these different hadronic interaction models is that X_{max} is slightly affected while N_{max} remains virtually

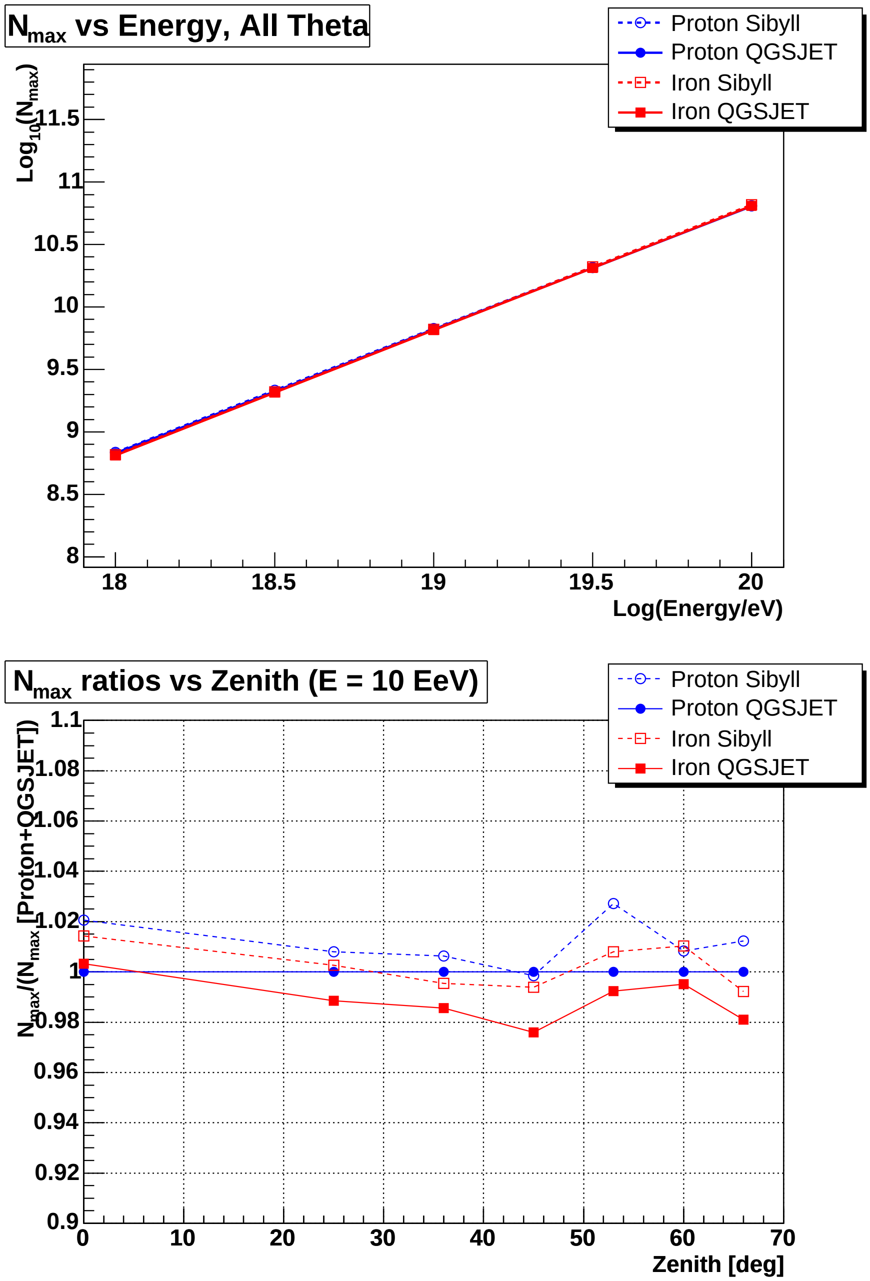


Figure 6.2: Top: Mean N_{max} Energy dependence. Showers were simulated at fixed energies using Aires and 4 primary/model combinations are plotted, but are on top of each other. Bottom: The ratio of different monte carlo predictions to the proton+QGSJET combination for the mean N_{max} of 10 EeV showers.

unchanged. The major difference is that the number of muons, important in the Pierre Auger Observatory, is very different for the different models, even for the same primary cosmic ray with identical primary energy. Thus, another parameter called muon richness [73] has been defined to characterize this difference. Muon richness (μ) is defined as:

$$\mu \equiv 100\frac{N_{max,\mu}}{N_{max}} \tag{6.1}$$

where $N_{max,\mu}$ is the maximum number of muons at the peak of the muon development, which develops slower than the total number of charged particles. The depth of the muon maximum, $X_{max,\mu}$, is deeper than X_{max}. In Fig. 6.3, a plot of X_{max} vs. $X_{max,\mu}$ is shown for simulations done using Aires [74] and proton and iron as initial cosmic ray primaries with Sibyll 2.1 [75] and QGSJET01 [76] as high energy hadronic interaction models. There is a clear correlation between X_{max} and $X_{max,\mu}$ which can be parameterized as:

$$X_{max,\mu} = 287 + 0.716X_{max} \tag{6.2}$$

A plot of μ is shown in Fig. 6.4 for Aires simulations with a primary energy of 10 EeV and different primaries (proton and iron) and different hadronic interaction models (Sibyll 2.1 and QGSJET01).

Once the shower develops and approaches X_{max}, the average particle energy is in a regime described by models that are well tested against accelerator data and the interactions are well understood. Consequently, the monte carlo simulations do not vary as much once this portion of the shower development is reached. There are still differences between low energy models and different monte carlo programs, such as Aires and Corsika [77]. These programs handle low energy interactions differently. For a given monte carlo program, like Aires, once the extensive air shower has neared shower maximum, the development is smooth

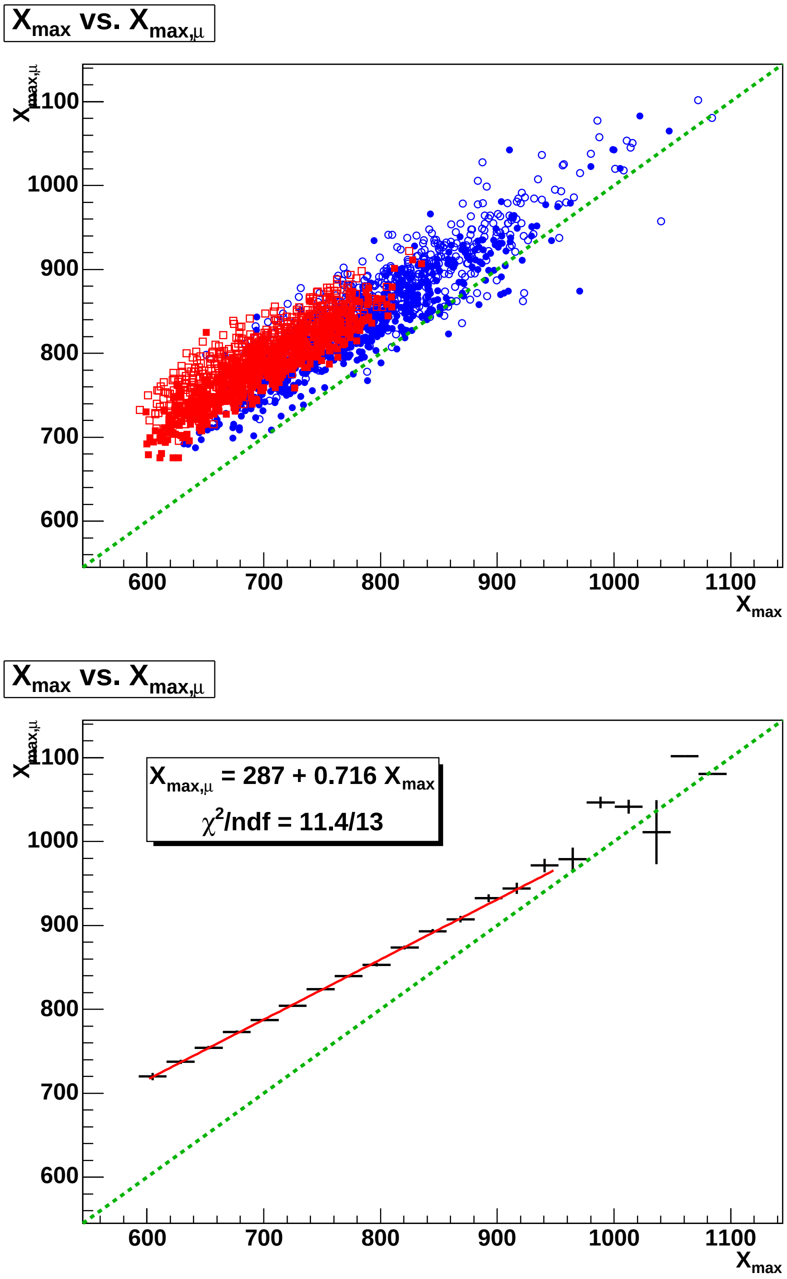


Figure 6.3: Top: X_{max} vs. $X_{max,\mu}$ for Aires monte carlo showers including 3 energies (10, 31, 100 EeV) and proton (blue circles) and iron (red squares) with Sibyll 2.1 (open markers) and QGSJET01 (closed markers) hadronic models. Bottom: The same plot as the top but with the data binned and a fit with a line.

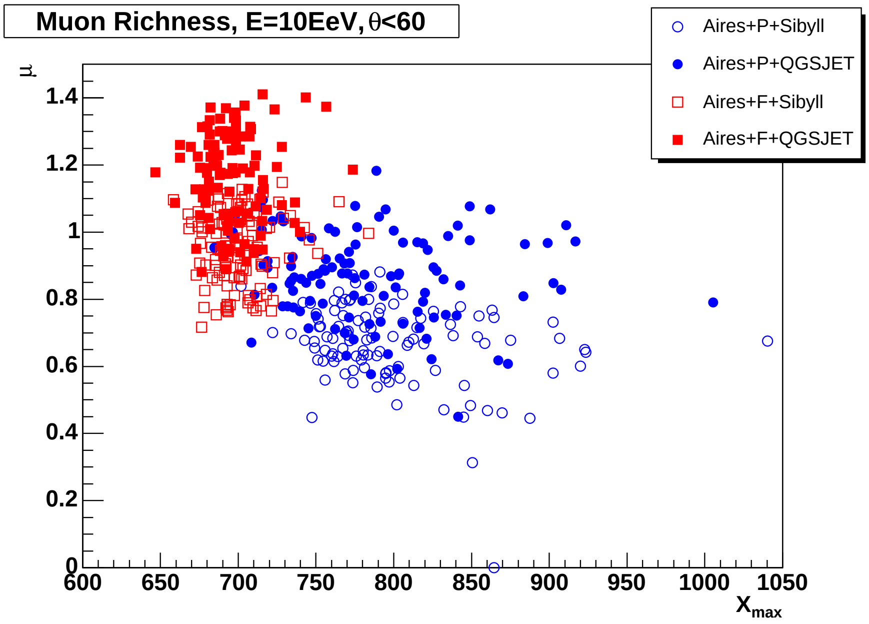


Figure 6.4: A plot of the muon richness for Aires 10 EeV showers and zeniths less than 60 degrees as a function of depth of shower maximum.

and predictable. Particle densities on the ground can then be predicted given only the three parameters N_{max} (or Energy), X_{max}, and muon richness (μ).

A plot of the density of photons on the ground normalized by N_{max} is shown in Fig. 6.5. This plot includes three different primary cosmic ray energies (10, 31, and 100 EeV), six different primary zenith angles (0, 25, 36, 45, 53, and 60 degrees), two different cosmic ray primaries (proton and iron nucleus), two different high energy hadronic models (Sibyll 2.1 and QGSJET01), and one simulation package (Aires). The density is defined as the particle density averaged over the area where the distance from the core in the plane perpendicular to the shower axis (called the shower plane) is greater than 575 m and less than 625 m and the internal azimuth angle (ξ) is between 80 and 100 degrees. The internal azimuth angle is defined such that 0 degrees corresponds to the intersection of

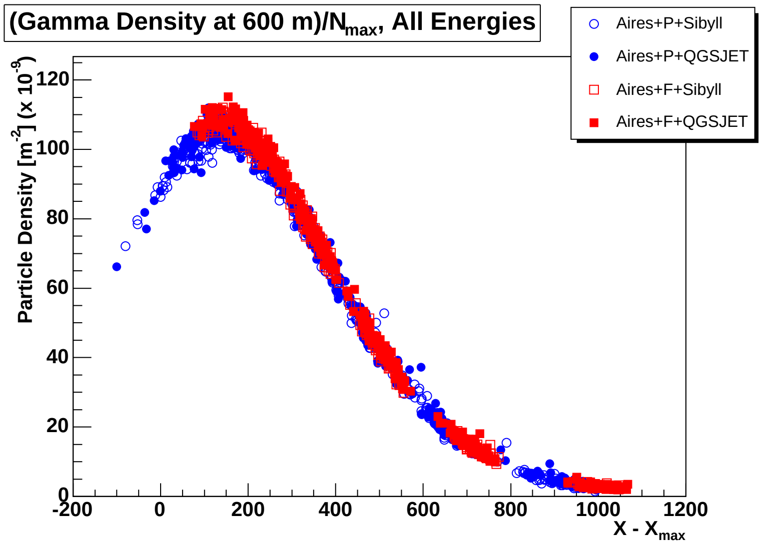


Figure 6.5: A plot of the photon density at 600 m from the core normalized by the maximum number of charged particles (N_{max}) vs. $X - X_{max}$

the line defined by the shortest distance between the shower axis and the ground and the shower plane. The x-axis is defined as the distance to the ground from the shower maximum ($X - X_{max}$), in units of g/cm^2. Figure 6.6 is a plot of the electron density normalized by N_{max}, where the density is defined according to the same qualifications as the photon density just described. From these plots it is apparent that the distribution of densities at a given distance from the core is smooth and can be described by a Gaisser-Hillas function. Given only N_{max} and X_{max} of an extensive air shower, the ground particle densities can be predicted for the electromagnetic component of the air shower for the given monte carlo program.

For the case of muons, it is more difficult. The muon development depends on the number of charged pions ($\pi^{\pm}$) created in the early part of the air shower which depends heavily on the hadronic interaction model used in the simulations.

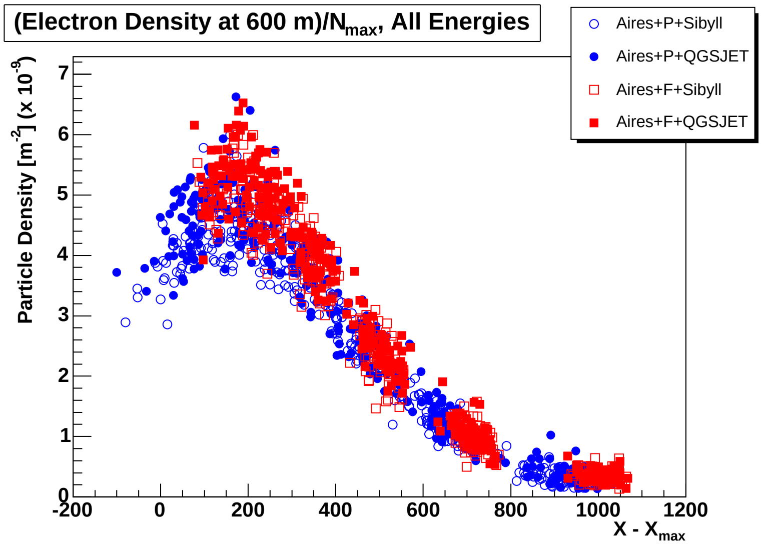


Figure 6.6: A plot of the electron ($e^{\pm}$) density at 600 m from the core normalized by N_{max} vs. $X - X_{max}$.

This can be seen in the top plot in Fig. 6.7 where the plot of muon densities at 600 m from the core for a fixed energy (10 EeV) are normalized by N_{max}. There is a clear separation between different primaries and hadronic interaction models. This is the major source of systematic uncertainties in determining the energy of a primary cosmic ray when using the SD and monte carlo methods. If, however, the N_{max} is replaced by $N_{max,\mu}$ as the normalization and $X - X_{max}$ is replaced with $X - X_{max,\mu}$, the distribution is smooth and predictable (bottom plot in Fig. 6.7). This shows that the longitudinal development of muons is smooth and predictable, similar to the electromagnetic development.

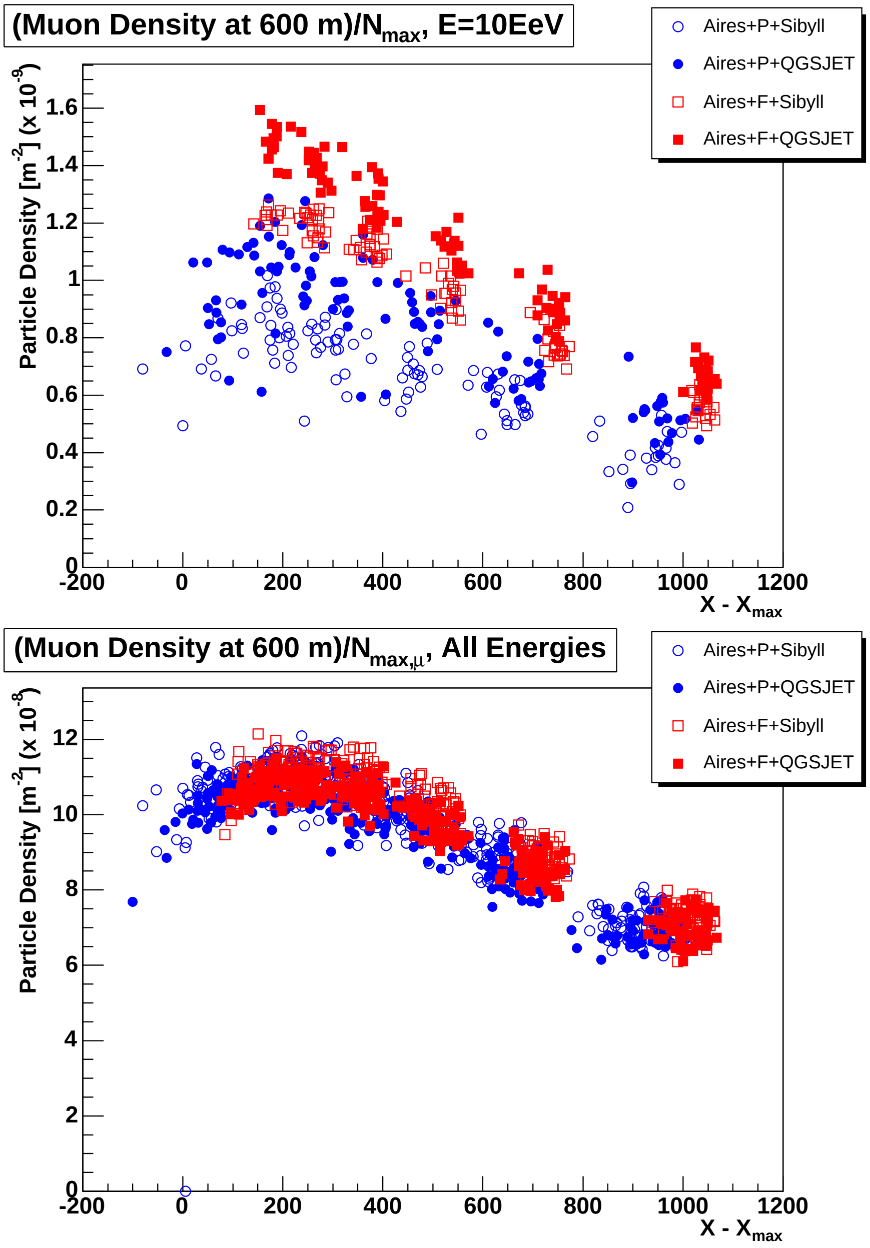


Figure 6.7: Top: A plot of the muon density at 600 m from the core normalized by N_{max} vs. $X - X_{max}$. A plot of the muon density at 600 m from the core normalized by the maximum number of muons in the shower development ($N_{max,\mu}$) vs. $X - X_{max}$.

6.2 Systematic and Statistical Uncertainties

The preceding plots and paragraphs motivate the discussion of systematic uncertainties in the primary energy measurement at the Pierre Auger Observatory using the SD. Figure 6.7 is a plot that shows the different muon densities on the ground due to different primaries and hadronic interaction models. Since the Pierre Auger Observatory employs the use of water Cherenkov detectors, the contribution from muons to the total signal is significant and leads to large systematic uncertainties ($\sim$10-20%) when using the SD alone to determine the primary energy of a cosmic ray.

The reason the muon signal is enhanced is due to a combination of factors. The average energy of a muon that reaches the ground in an extensive air shower has an energy greater than 1 GeV while the average energy of an electron (or positron) or photon is in the 1-10 MeV range, see figure 15 in [23]. Convolve this with a water Cherenkov detector that is 1.2 m high, and assuming that the particles enter vertically, each muon will deposit 240 MeV (2 MeV/cm) in the tank while the electromagnetic component will deposit all of its energy, 1-10 MeV. Thus, the signal in the detector enhances the muonic component of the shower. If the particles are inclined, it is even more pronounced. Assuming an average entry angle of 60°, the track inside the tank is now two times longer and the muon deposits twice the energy, while the electromagnetic particles still deposit the same amount of energy. It is necessary to then consider the absolute number of particles entering the detectors. The number of electromagnetic particles is $\sim$100 times larger than the number of muons for vertical showers while for 60° showers that ratio is $\sim$2, refer to Fig. 6.5 and Fig. 6.7. The changing ratio is due to the attenuation of the electromagnetic component of an air shower caused by an effectively deeper atmosphere for inclined showers.

In Fig. 6.8, a simple parameterized tank response for each particle in a shower is assumed and the signal in a detector is calculated at 1000 m. This is done to illustrate the competing effects of the number of particles versus the energy deposited by each particle type for a given shower as explained in the previous paragraph. For vertical showers, the signal ratio is about 1 to 1 for the electromagnetic and muonic contributions which is due to the large number of photons in a shower. As the showers become more inclined, the muonic signal clearly dominates. This is a result of the combination of tank geometry and the electromagnetic component of the shower attenuating for inclined showers. The end result is a systematic uncertainty in the energy determination due to the unknown mass of the primary cosmic ray and different hadronic models.

To quantify the systematic uncertainty due to unknown primary and mass, the ground parameter S(1000) was calculated from reconstructed monte carlo events. These events were generated by putting the ground particle output files from Aires and Corsika simulations through a resampling algorithm, then through a detector simulation based on Geant4. These simulated events were then reconstructed using the official reconstruction [78] of the Data Processing and Analysis (DPA) framework [79].

The ground parameter S(1000) is the energy deposited on the ground at 1000 meters from the core of the extensive air shower, where the distance from the core is defined in a plane perpendicular to the shower axis. This ground parameter is closely related to the energy of the primary cosmic ray and minimizes shower to shower fluctuations as will be discussed later in this chapter; thus it is used to estimate the energy. As was discussed above, S(1000) varies from primary to primary and with different hadronic interaction models, see Fig. 6.9. Thus, an energy estimate based on monte carlo data will either assume one mass and

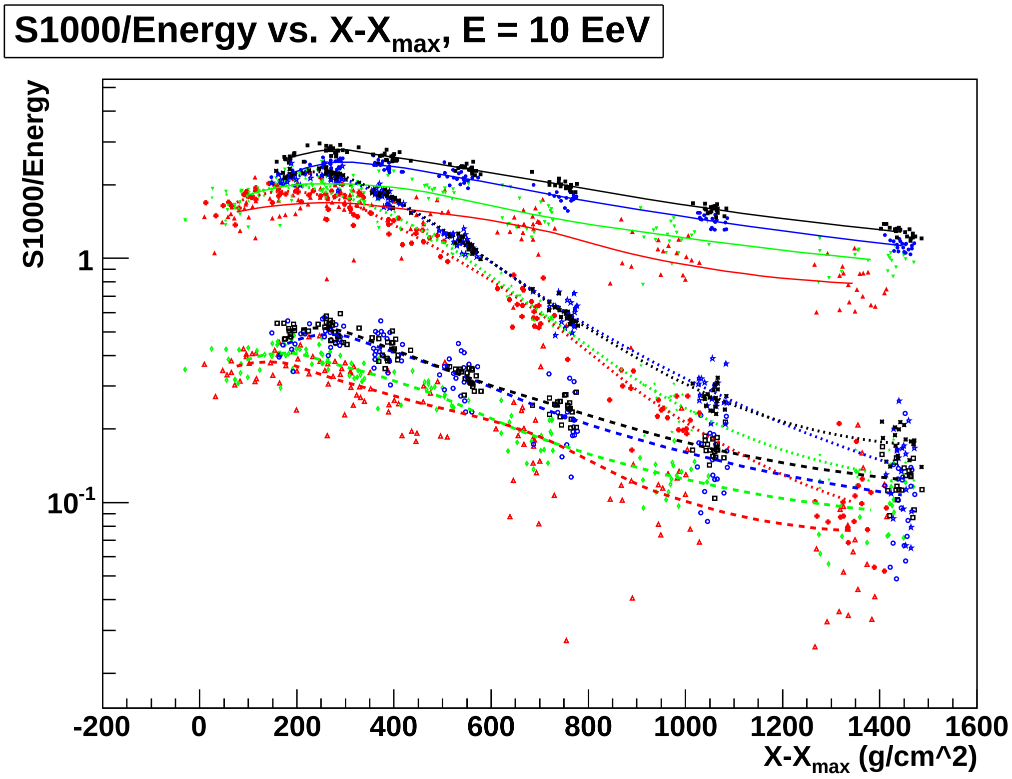


Figure 6.8: Tank response at 1000 m from the core in units of vertical equivalent muon (VEM) charge normalized by the energy for different particle types. There are 4 combinations of primary and hadronic model, RED: Proton+Sibyll, GREEN: Proton+QGSJET, BLUE: Iron+Sibyll, BLACK: Iron+QGSJET. The lines follow the average behavior where SOLID lines are for muon response, DOTTED lines are for photon response, and DASHED lines are for $e^{\pm}$ response.

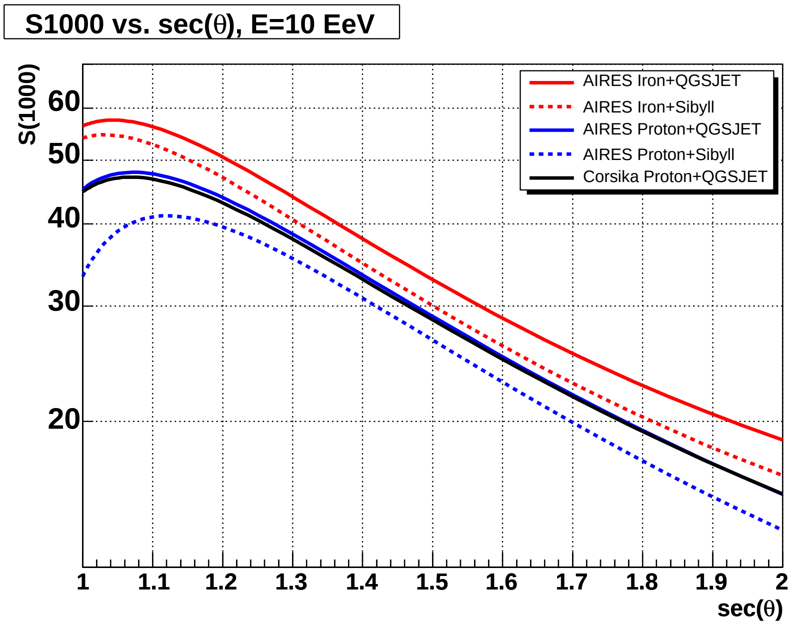


Figure 6.9: Average energy deposited in units of VEM for 10 EeV showers as a function of sec(θ).

model or will use an average expected S(1000) for a given primary energy. For the following calculation of systematic uncertainty, the latter approach is assumed. The systematic uncertainty is then defined as:

$$\text{uncertainty}[\%] = 100\frac{S(1000)_{max} - S(1000)_{min}}{S(1000)_{max} + S(1000)_{min}} \tag{6.3}$$

Equation (6.3) is the maximum deviation from the average S(1000) at any given input energy and zenith. A plot of the uncertainty as defined above as a function of zenith angle for four different energies is shown in Fig. 6.10. The uncertainty ranges from ∼20% for vertical and inclined showers to a minimum of ∼10% at 45°.

It is interesting to note that the systematic uncertainty is independent of energy. The differences between hadronic models and primary masses persists with changing energy. Also, the systematic uncertainty reaches a minimum around

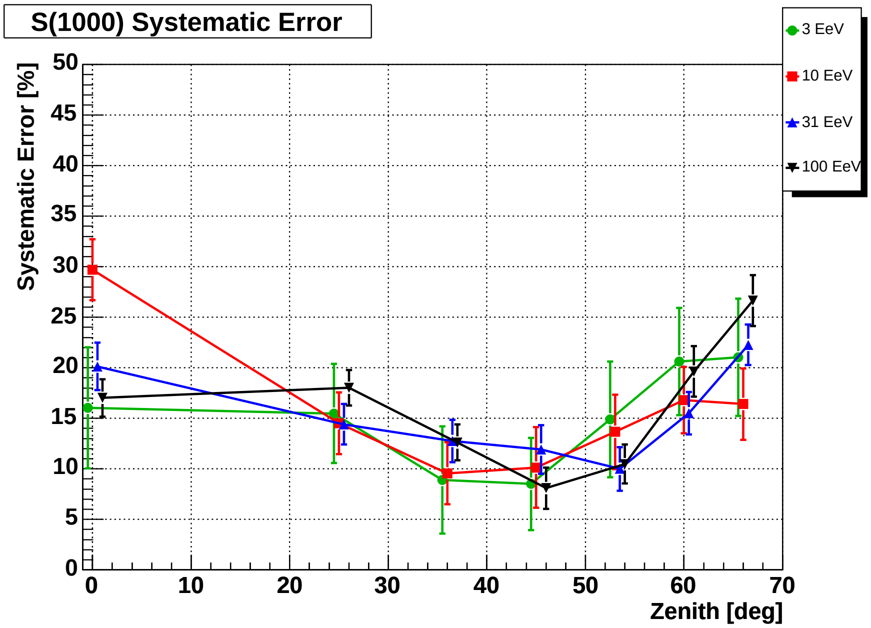


Figure 6.10: The systematic uncertainty from mass and model uncertainty as a function of zenith angle for 4 different energies.

35-45 degrees. This is due to the changing ratio of electromagnetic to muonic signal in the tanks. With increasing zenith angle, the muonic fraction of the total signal increases while the electromagnetic fraction decreases. At the range of zeniths where the systematic uncertainty is a minimum, the combination of muonic and electromagnetic signals for the different models and masses is such that the total signal deposited is nearly constant regardless of primary mass and hadronic model assumption. It is caused by the interplay of the depth of shower maximum with the muon richness of the shower. A shower initiated by a proton is deeper than that of an iron. For a vertical shower and the detector at the altitude of the site at Malargüe, the proton showers have a mean depth of shower maximum below the ground, while iron is above. For showers in the zenith range in question, at 1000 meters from the core, the shower development is around the maximum for protons, while iron has peaked and is attenuating. Thus, the electromagnetic signal is larger for proton showers than for iron showers. To counterbalance this, the iron showers produce more muons and the muonic signal is larger for iron showers than for proton showers. This compensation leads to minimal systematic uncertainties.

With the monte carlo data used to study systematic uncertainties, it is also possible to study statistical fluctuations and calculate statistical uncertainties arising from shower to shower fluctuations for different masses and hadronic models [80]. In Fig. 6.11, a plot of the statistical uncertainties for proton and iron are shown at 36 degrees for different energies. The influence of different hadronic models on the statistical fluctuation is negligible, what matters is the primary composition. Iron has a smaller statistical fluctuation in the depth of shower maximum, stemming from a shorter mean free path in air, than proton. This effect propagates to the ground where the energy deposited on the ground fluctuates less than proton. From Fig. 6.11, the iron statistical fluctuation is ∼16% for 3 EeV showers and steadily decreases to ∼5% at 100 EeV while proton goes from ∼14% at 3 EeV to ∼8% at 100 EeV.

As expected, the statistical uncertainties reduce with higher energy due to the larger sampling of the shower on the ground (higher detector multiplicity). The statistical uncertainty is influenced by the number of detectors on the ground that are used in the reconstruction as well as the geometry of the detectors in the shower plane. For example, if a nearly vertical shower lands near a detector, then that central detector has a signal that may or may not be usable due to it saturating the electronics while the next ring of triggered detectors all sit at roughly the same distance from the shower axis and, thus, give redundant information. This would lead to a larger uncertainty in reconstructed parameters

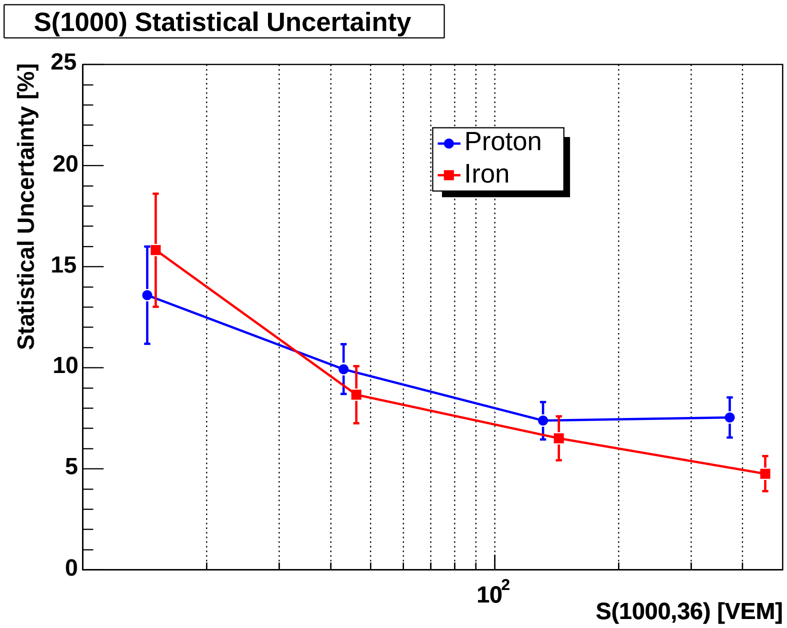


Figure 6.11: Statistical uncertainty for proton and iron as a function of mean S(1000) deposited on the ground for a shower with an inclination of 36 degrees. These mean S(1000) values correspond to energies of 3, 10, 31, and 100 EeV.

than for an inclined shower with multiple detectors being triggered at many varying distances from the shower axis.

6.2.1 Uncertainties and the Ground Parameter S(r)

As was stated earlier in this chapter, the ground parameter used to determine the energy of a cosmic ray is S(1000), or the energy deposited on the ground 1000 m from the core in the plane perpendicular to the shower axis. The distance of 1000 m is purely a geometrical factor. The 1.5 km spacing of the detectors in the surface array has the effect that when showers are to be reconstructed by fitting the signal in a detector as a function of distance from the core, on average, the fluctuations of the value of the fitted function at 1000 m from the core is minimized, see Fig. 6.12, with 10-15% fluctuation. Again, this is purely from the geometry of the detector [81], which explains why S(600) fluctuates from 20-25%. To determine S(600) accurately, it is necessary to have detectors triggered inside of 600 m, which happens less frequently due to the detector spacing.

The argument that the statistical fluctuations are reduced at this distance becomes less important as the number of events increases. The systematic uncertainty due to hadronic model and mass remain, however. It is desirable to evaluate whether there is a more optimal choice for the ground parameter used to estimate the primary energy that reduces systematic uncertainty.

In an extensive air shower, the electromagnetic component dominates close to the core. As the distance from the core increases, so does the contribution from muons. The muons are the major source of systematic uncertainty in the energy determination. The difference in N_{max} from one mass and hadronic model to another varies on the order of 5%, see Fig. 6.2. Therefore, if the electromagnetic component of the shower could be measured with minimal muonic contamination, the energy determination would have minimal systematic uncertainties due to mass and hadronic model uncertainty. This is shown in Fig. 6.13 where the systematic uncertainties are minimal closer to the core. For the ground parameter S(600), the systematic uncertainty is minimized at 36° with a value of ∼4% compared to ∼10% for S(1000).

6.3 Conclusions

In this chapter, it has been shown that the intrinsic systematic uncertainty from unknown primary composition and uncertainties in hadronic interaction models is on the order of ∼10-20% when determining the primary energy of a cosmic ray

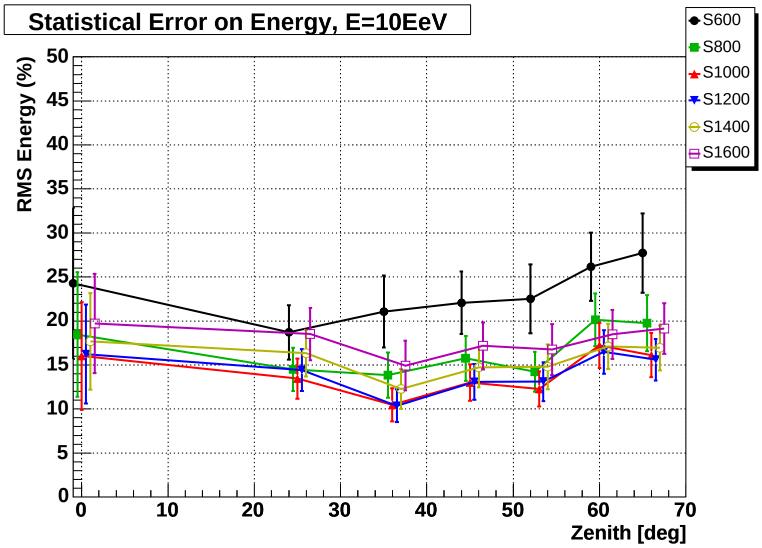


Figure 6.12: Statistical fluctuations for different ground parameters, S(r), where r is 600, 800, 1000, 1200, 1400, and 1600 m from the core. At 1000 m the fluctuations are minimized

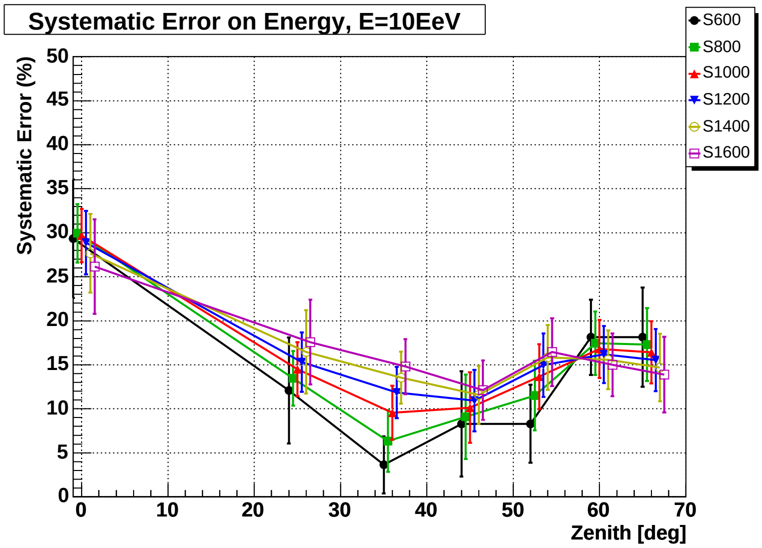


Figure 6.13: Systematic uncertainties for different ground parameters, S(r), where r is 600, 800, 1000, 1200, 1400, and 1600 m from the core. Closer to the core, the uncertainty is minimized.

using the SD. The main cause of this systematic uncertainty is the enhanced contribution of the muons by the water Cherenkov detectors in conjunction with the different predictions of the muon component of extensive air showers by different interaction models. Minimizing the muonic fraction of the total signal minimizes the systematic uncertainty discussed here and at 600 m from the core and for cosmic rays with a zenith angle of 36° that uncertainty is 4%.

Statistical fluctuations caused by shower to shower variation and detector geometry can not be avoided. For the ground parameter S(1000), the fluctuation is minimized to 10-15% and is the best choice for estimating the primary energy.

These are the uncertainties in the primary energy determination of the Pierre Auger Observatory, when using the SD and monte carlo to estimate the energies. Future studies may reveal a parameter that is related to energy that does not suffer from the same model and composition uncertainties.

CHAPTER 7

Composition Study

One of the key unanswered questions in the field of cosmic rays is the composition of the highest energy events. Once the composition is determined, energy determination for the events is less affected by the systematic errors described in Chapter 6. A particular composition and hadronic interaction model may more closely approximate the behavior of the real data, which would then lead to a particular S(1000) to energy conversion based on that composition and model combination. At this point, an event by event determination of the composition may not be possible, but the mean behavior can be studied. Also, monte carlo predictions of observables based on a photon primary assumption differ greatly from those of baryonic primaries. Due to this difference, a limit may be set on the flux of photon primaries.

7.1 Ground Observables

Traditionally, composition studies have been based on X_{max} information. The Pierre Auger Observatory will also present analysis based on X_{max} data from the fluorescence detector, but it is desirable to exploit the superior statistics of the ground array. What is necessary, then, is to establish ground observables that have strong composition dependence. In this section, a study is presented comparing various reconstructed ground observables from real data to the same reconstructed observables from monte carlo events.

7.1.1 Rise Time

There are many observables measured using the surface detector: slope of the lateral distribution function, shower front curvature, signal density at a certain distance from the core, rise and fall time of the signal, etc. Rise time, in particular, holds primary composition information due to the correlation with the muonic component of the shower as well as the depth of shower maximum, X_{max}, where the rise time of a given event is described in detail below. In brief, the rise time is related to the time an event takes to increase from 10% to 50% of the total energy deposited on the ground.

The relationship with shower maximum comes from the creation depth of shower particles. Showers with deeper X_{max} create shower particles along a longer path and deeper in the atmosphere, or closer to the detectors. This results in a longer path length difference from the creation point to the detector. This difference leads to a larger spread in arrival times, and thus a longer rise time. Conversely, for a shallow shower the different path lengths from creation point to the detector is smaller, and the particles arrive bunched resulting in a shorter rise time.

Another factor is the fraction of the signal from muons. Electromagnetic particles experience multiple scattering on their way to the ground and arrive spread out in time. Muons, on the other hand, travel in straight lines, and arrive tightly correlated in time on the ground. Thus, a shower with a higher muon fraction will have a shorter rise time than a muon poor shower. The combination of these effects is enhanced in the water Cherenkov detectors because they enhance the muon signal, as explained in Chapter 6. Thus, the muon number density is enhanced and is a big indicator of the behavior of the rise time.

It is apparent, then, that the rise time for a given station in an event varies with distance from the core due to the changing muon to electromagnetic ratio and path length arguments. Closer to the core, the particles arriving at these stations have a small distribution of path lengths. In other words, all the particles are tightly grouped together resulting in a fast rise time. Far from the core, the particles arriving at these stations have been created at varying depths in the shower's development, with varying paths to travel to arrive at the station. This results in a wider spread of arrival times and a longer rise time.

To define a simple parameter for the rise time for a given shower, we fit a simple function to the rise time vs. core distance for each station in a given event. From this fit, we can then estimate the rise time at 1000 meters from the core, or $\tau(1000)$. The function used is a two parameter, quadratic fit, which has been determined as an appropriate function to describe the behavior of the rise time as a function of distance from the core[82]:

$$40 + ar + br^2 \tag{7.1}$$

where the rise time is measured in nanoseconds and the y-intercept has been determined to be 40 ns from studying the rise times of extremely horizontal showers and is essentially the width of the shower at the core. The fit is restricted to tanks with a signal greater than 10 VEM, that have a non-saturated anode signal, and are closer than 2000 m from the core. This is done to restrict the fit to a range where the particle density is high enough to avoid any fluctuations due to small number statistics. The uncertainty of the rise time as a function of primary zenith angle and distance from the core has been studied [82] and the results were used in the fitting routine to give weights to the detectors used in the fit.

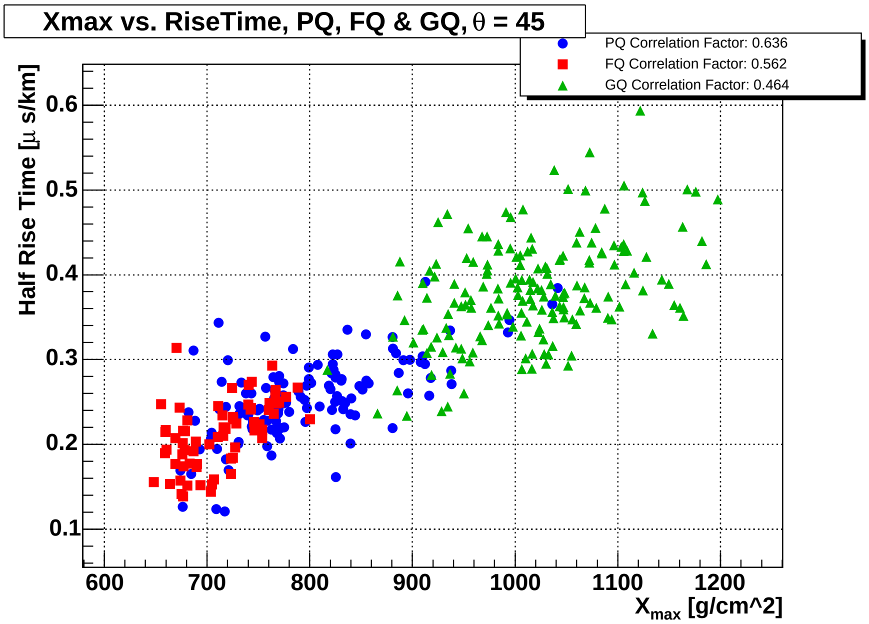


Figure 7.1: Monte carlo X_{max} vs. $\tau(1000)$ at 45 degree zenith angle. Iron (red squares), Proton (blue circles), and Photon (green triangles). Note: two energies (10 and 31 EeV) are included in this plot, so X_{max} for different primaries at a single energy is more distinct than is apparent in this plot which includes different energies.

The relationship between $\tau(1000)$ and shower maximum depth and muon richness (μ) are in Fig. 7.1 and Fig. 7.2. The parameter termed μ, defined in (6.1), is the ratio of the maximum number of muons in the shower development divided by the number of charged particles at shower maximum (multiplied by 100). This parameter holds composition information along with X_{max}, measured by the fluorescence detector. Being able to measure these parameters is vital in composition studies. The surface detector is not able to measure either of these two parameters directly, but the rise time observable is directly related to them.

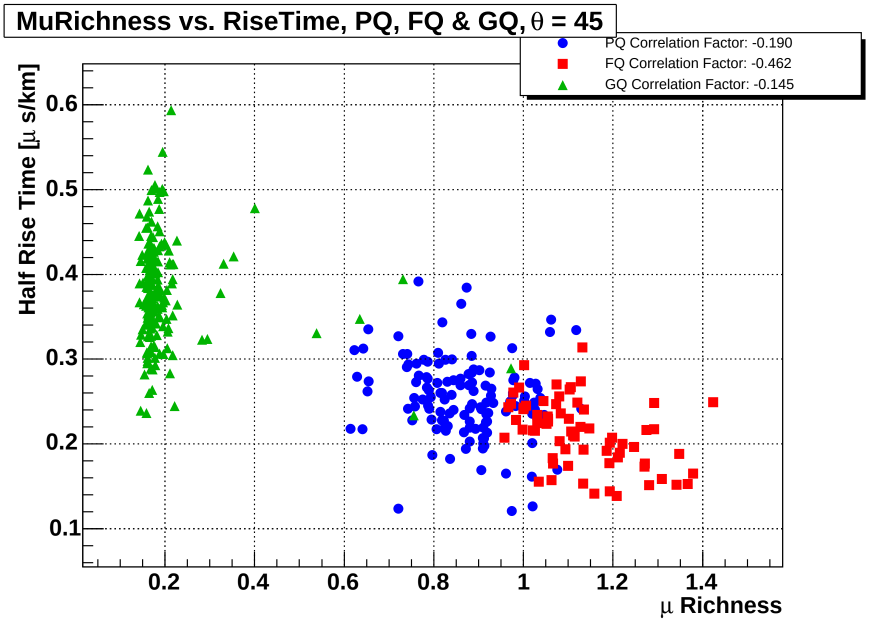


Figure 7.2: Monte carlo muon richness (μ) vs. $\tau(1000)$ at 45 degree zenith angle. μ is defined as $100 \cdot N_{\mu,max}/N_{max}$. Iron (red squares), Proton (blue circles), and Photon (green triangles). Note: two energies (10 and 31 EeV) are included in this plot, so μ for different primaries at a single energy is more distinct than is apparent in this plot which includes different energies.

7.1.2 Shower Curvature

Another observable that is related to the shower composition through the relationship with X_{max} and μ is the curvature of the shower front. The curvature of the shower front is the radius of curvature near the impact point of the shower. The shower can be approximated as beginning at a certain point in time and space, and the development expanding as a spherical front. The radius, R, is the distance to that initial point. In the reconstruction, the curved shower front fitting is done to better approximate the original zenith and azimuth of the cosmic ray [78].

Since the shower front curvature is related to the initial beginning point of the shower, it is related to the depth of shower maximum. For example, an iron primary has a shorter mean free path in air than proton or photon primaries. The initial point of the shower is higher in the atmosphere, resulting in a shallower X_{max} and a longer radius of the shower front curvature. Photon showers develop closer to the ground (deeper X_{max}) and have a smaller radius of curvature.

The muon content is also related to the shower front curvature through the arrival times at ground of the shower particles. Electromagnetic particles are scattered multiple times, so they must travel farther through the atmosphere and it takes longer to reach the ground when compared with muons which travel in straight lines. Thus, farther from the core, if the shower is more electromagnetic, the relative delay is larger than for muon rich showers, resulting in a larger reconstructed shower front curvature. The relationship between shower front curvature and X_{max} and μ is shown in Fig. 7.3 and Fig. 7.4.

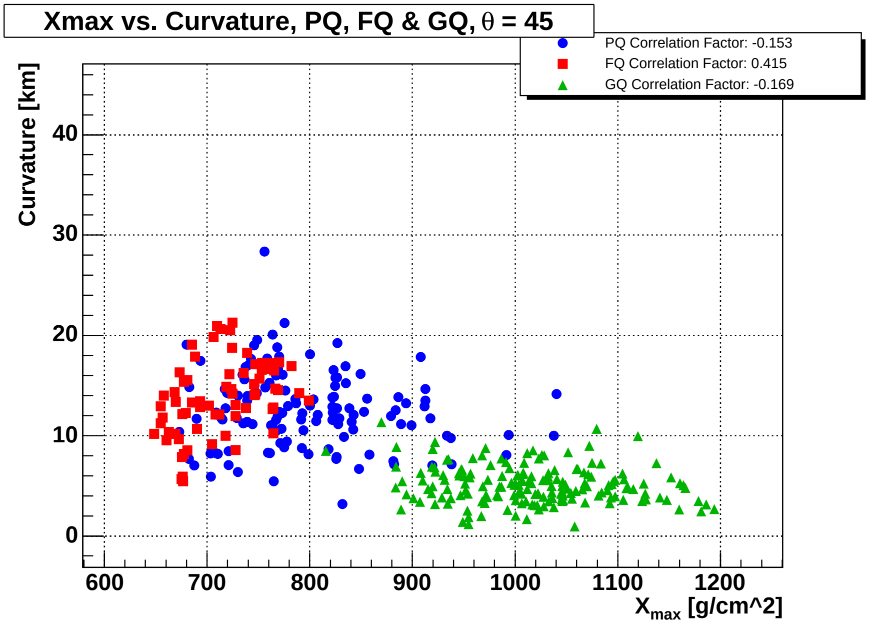


Figure 7.3: Monte carlo X_{max} vs. Curvature at 45 degree zenith angle. Iron (red squares), Proton (blue circles), and Photon (green triangles). Note: two energies (10 and 31 EeV) are included in this plot, so X_{max} for different primaries at a single energy is more distinct than is apparent in this plot which includes different energies.

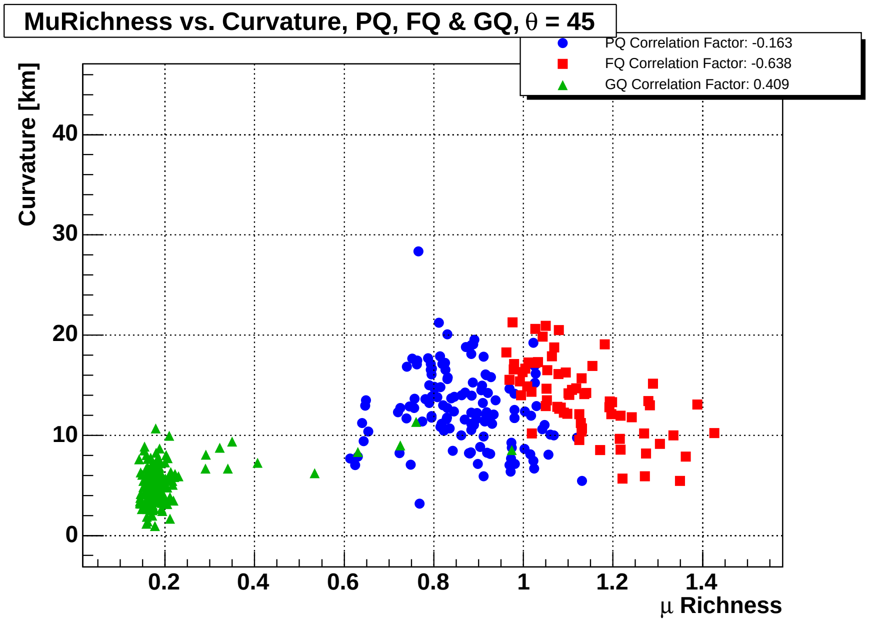


Figure 7.4: Monte carlo muon richness (μ) vs. Curvature at 45 degree zenith angle. μ is defined as $100 \cdot N_{\mu,max}/N_{max}$. Iron (red squares), Proton (blue circles), and Photon (green triangles). Note: two energies (10 and 31 EeV) are included in this plot, so μ for different primaries at a single energy is more distinct than is apparent in this plot which includes different energies.

7.2 Monte Carlo Events

The monte carlo events used in this study were generated using Aires 2.6 [74] and Corsika 6.2 [77]. For the Aires showers, proton and iron nuclei were used as the primary cosmic ray and Sibyll 2.1 [75] and QGSJET01 [76] were used for the high energy hadronic interactions. The showers were thinned using a relative thinning factor of 10^{-6} and statistical weight factor of 0.15, a factor used to limit the maximum weight of shower particles. The cosmic rays were simulated with four different primary energies (3, 10, 31, and 100 EeV), six different zeniths (0, 25, 36, 45, 53, and 60 degrees), and 10 different azimuths, from 0 to 324 degrees in 36° steps, with 2 showers per primary/model/energy/zenith/azimuth combination.

The Corsika showers were generated using proton as the primary, QGSJET01 as the high energy interaction model, and Fluka [83] as the low energy interaction model below 200 GeV. The showers were thinned using "optimal" 10^{-6} relative thinning, which again deals with the maximum weight factor and the relative maximum weight factors between hadronic and electromagnetic particles. The same energies, zeniths, and azimuths were used for the Corsika showers as the Aires showers.

Once the showers were generated, they were put through a detector simulation. A random location inside the surface array was chosen as the core location for the shower, and the ground particles were put through a resampling algorithm [84]. The resampled particles were then injected into a specific surface detector where Geant4 was used to simulate the physics inside the detector. The electronics are also simulated so that the simulated signal approximates what is really seen in the detectors. What is obtained is an event that is nearly identical to the real data taken by the observatory but containing the monte carlo input values in the event structure. Once the monte carlo events are in this format, they are treated identically to real data and the reconstruction programs are blind to the source of the input.

7.2.1 Thinning and Unthinning Effects

Due to the large number of particles created in an ultrahigh energy cosmic ray shower, it is computationally prohibitive to follow all generated particles in a monte carlo simulation. To reduce the time and memory needed to simulate these showers, thinning is introduced into monte carlo showers. Thinning is a method where all particles are tracked explicitly in the program until they reach a certain energy. When an interaction produces particles below this defined energy, a statistical method is used to determine which resulting particles to follow and the weight of each of these particles. In this manner, only a fraction of the particles must be followed and the time of the simulation is reduced and the memory required to store the resulting particles that reach the ground is smaller.

Of course, inherent in this method is the problem that not all particles are followed and may be too coarsely sampled resulting in artificial fluctuations of ground particle distributions. There is an optimization that must be determined where artificial fluctuations are reduced but the time and memory needed for the simulations is manageable.

For the simulations used in the studies presented in this paper, a relative thinning of 10^{-6} was used. Relative thinning defines the minimum energy for unthinned particles relative to the primary energy of the cosmic ray. As an example, a shower initiated with an energy of 10^{20} eV and relative thinning level of 10^{-6} will follow all particles without thinning until they reach an energy below 10^{14} eV.

Another problem when using thinning is that the weight factors for particles may become extremely large. To counteract this, a maximum weight can be defined such that any particle that has a weight greater than this factor, will not be thinned again. It will be followed explicitly until it exists no more or reaches the ground. In Aires, the maximum weight factor is set via the variable called the statistical weight factor (W_f). The maximum weight factor is defined as:

$$W_{max} = A_0 E_{th} W_f \tag{7.2}$$

where A_0 is a constant defined as 14 GeV^{-1}, E_{th} is the thinning factor times the primary energy, and W_f is the statistical weight factor. All the simulations were done using $W_f = 0.15$. For a primary energy of 10^{20} eV and relative thinning of 10^{-6}, the maximum weight factor would be 2.1×10^5. There is also a mechanism to set different maximum weight factors for electromagnetic and hadronic particles. This is done by setting a ratio for the maximum weight factors. In Aires the default ratio (W^{EM}_{max}/W^{H}_{max}) is 88 [74]. Continuing the preceding example, that would mean the maximum weight factor for a hadronic particle would be $\sim$2386.

Undoubtedly, thinning has an effect on simulated showers. This effect needs to be determined when doing systematic comparisons between monte carlo showers and real showers. Any effects in reconstructed observables introduced in the thinning procedure need to be understood to avoid any false conclusions based on monte carlo predictions. For the ground observables introduced above, a study has been done using Aires with relative thinning levels of 5×10^{-5}, 10^{-5}, 5×10^{-6}, 10^{-6}, 5×10^{-7}, and 10^{-7}, for showers initiated with a proton with an energy of 10^{20} eV and zenith angles of 25 and 60 degrees. For each data point, 20 showers were simulated. The results are presented in Fig. 7.5.

The differences in the plots are relative to the thinning levels used in the simulations to analyze the data in this chapter. For the observables of interest,

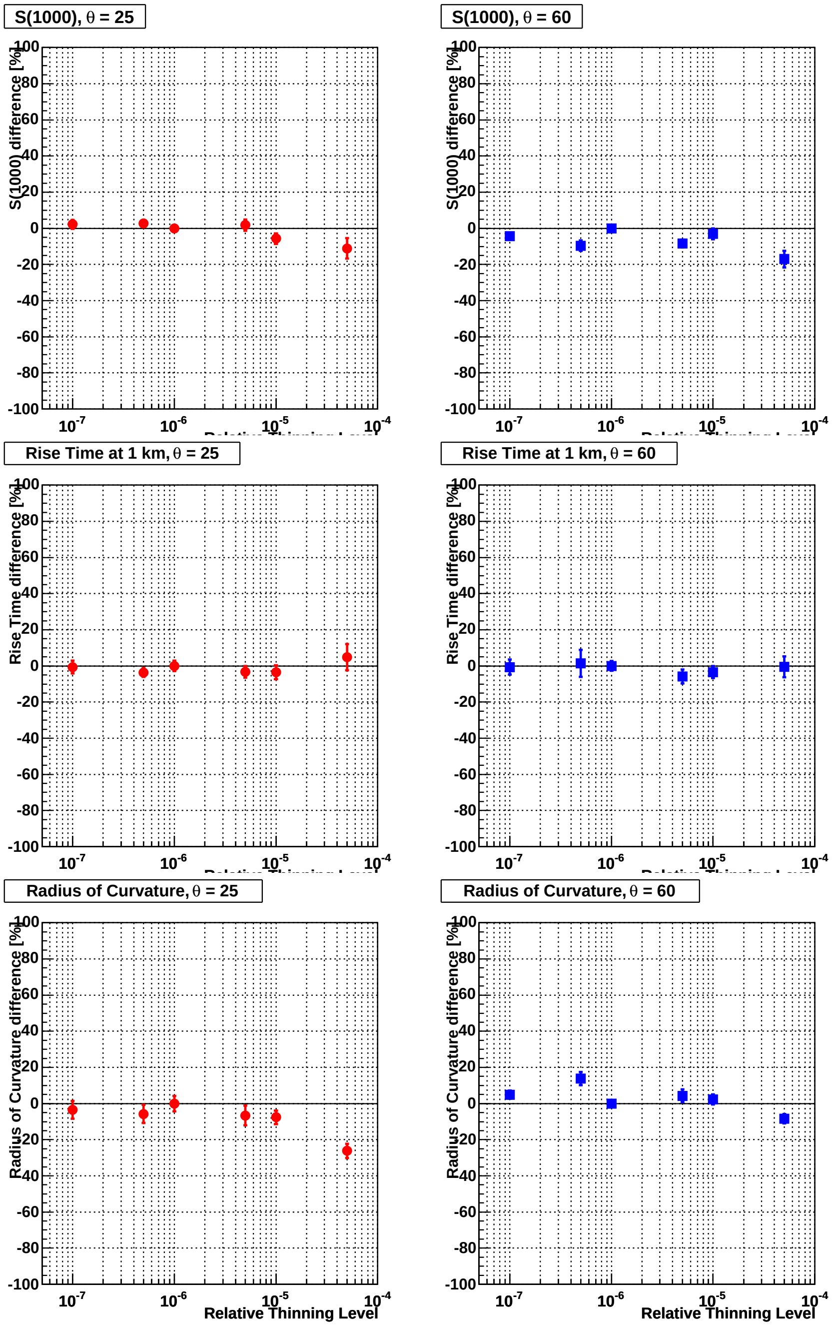


Figure 7.5: Aires: Systematic difference in S(1000) (top), rise time (middle), and curvature (bottom) as a function of thinning level for two zeniths: 25 (left) and 60 degrees (right). Difference is relative to 10^{-6} relative thinning.

namely rise time and curvature, there is no significant effect using Aires. For the time being, no corrections are made to compensate for thinning effects.

When a monte carlo shower is to be simulated in the surface detector, the ground particles need to be resampled in an effort to recover an accurate description of the ground particle distributions had there been no thinning. This unthinning procedure chooses an area around a detector and resamples the particles located within that area [84]. The procedure is to regenerate a number of particles from one thinned, weighted particle by generating a random number from a Poisson distribution with mean:

$$\mu = w\frac{A_{det}}{A_{sampling}} \tag{7.3}$$

where w is the weight of the given particle, A_{det} is the area of the detector, and $A_{sampling}$ is the sampling area. It is desirable to have this number less than one so that the fluctuations in the shower are due to intrinsic fluctuations and not the resampling procedure. Once the number of regenerated particles is determined, they are randomly placed on the detector with their time being shifted so as to maintain their relative timing within the shower front. The energy and direction are retained from the original thinned particle, as well as the particle type, obviously. These particles are then inserted into the detector and the detector simulation is run.

The default shower regenerator in the data processing and analysis (DPA) framework for the Pierre Auger Observatory uses an increasing $A_{sampling}$ with distance from the core. Also, it uses the same $A_{sampling}$ for hadronic and electromagnetic particles. This may be a problem because as was discussed earlier in this section, electromagnetic and hadronic particles have different maximum weight factors assigned to them resulting in different average weight factors, with the weights of hadronic particles being smaller. There is some flexibility in choosing

$A_{sampling}$, but it must be large enough to not introduce any artificial fluctuations that arise from too few particles to be resampled and it must be small enough that the average density in the sampling region remains roughly constant. If it is too large, the particle density may be artificially reduced by extending into regions where the particles are sparse.

To investigate this effect, the resampling program was altered in two ways. First, the resampling was done with $A_{sampling}$ being different for electromagnetic and hadronic particles, in the same ratio as their maximum weight factors, i.e. the sampling area for hadronic particles was 88 times smaller than for electromagnetic particles for Aires showers. Second, the sampling area was a fixed value, not increasing with distance from the core, and there were two sampling areas just as in the first modification described. The reasoning behind the second modification is that the average weight of particles actually decreases with distance from the core, see Fig. 7.6. Therefore, a sampling area that increases with core distance seems undesirable.

Twenty showers with an energy of 100 EeV and zenith angle 53 degrees were unthinned using the three resampling programs and the effect on ground observables was studied. In Fig. 7.7, the results are shown as difference from the default resampling program in the DPA framework. There seems to be no significant effect on the ground observables of interest.

7.2.2 Differences Between Aires and Corsika

In Chapter 6, ground particle distributions were discussed in conjunction with different monte carlo simulations. It is not immediately apparent the effect the small differences in particle densities will have on the ground observables of interest. It is necessary to simulate the entire detector chain and then reconstruct

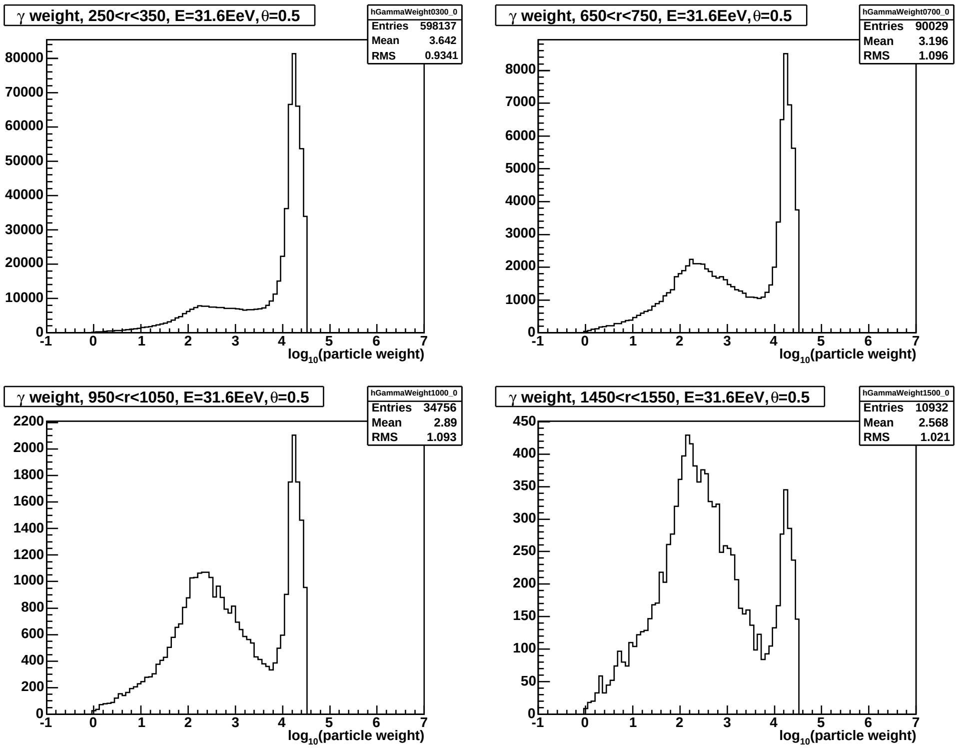


Figure 7.6: Distribution of particle weight factors at 4 distances for photons. The shower was initiated by a proton with an energy of 31.6 EeV and a zenith angle of 0.5 degrees. The four distances are 300, 700, 1000, and 1500 m.

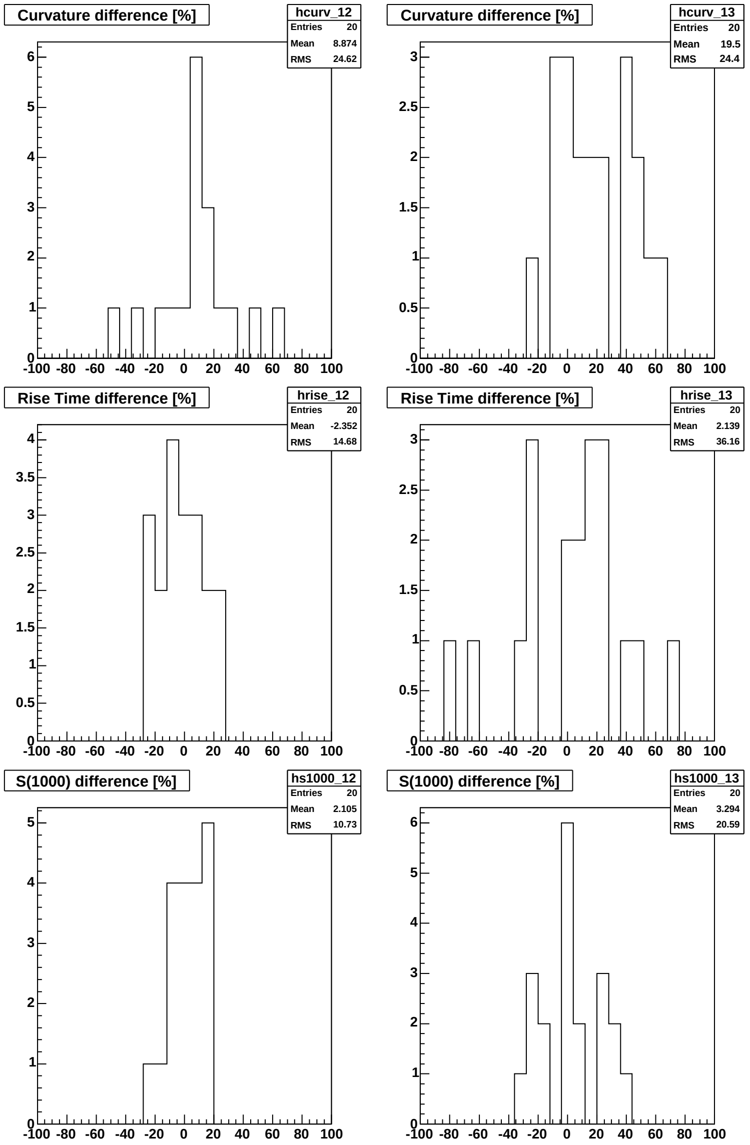


Figure 7.7: The difference between the default resampler and the two modifications, as explained in the text. Left: Comparison of modification 1 to default. Right: Comparison of modification 2 to default. Top: Curvature. Middle: Rise Time. Bottom: S(1000). Difference is calculated as $100 \times (X_1 - X_i)/X_1$, where 1 is the default resampler and i refers to one of the two modifications.

the showers to see any systematic difference between monte carlo programs.

For proton showers with the QGSJET01 hadronic interaction model, air showers were simulated using Aires and Corsika and then put through the detector simulation and reconstructed. The showers were simulated with common input parameters, at distinct energies and zenith angles. In Fig. 7.8 the values of rise time, curvature, and S(1000) are plotted for Aires and Corsika on the same plot to show any systematic difference in these observables at different angles and energies. The result is that there is no significant difference between Aires and Corsika monte carlo showers when simulating with the conditions discussed above.

7.3 Real Data

The data used in this study is the official data sample used in obtaining the spectrum presented in August 2005 [85]. It contains data from January 1, 2004 to June 5, 2005. The event selection criteria was that the detector with the largest signal had to have at least 5 neighbors properly functioning at the time of the event [86]. This eliminates events that fell on the border of the surface array or next to any holes in the detector configuration. The reconstruction of both the real data and the monte carlo events discussed above were done using the official reconstruction package of the DPA framework [78]. The reconstructed parameters used are the shower front curvature and the rise time at 1 km.

For each detector, the total signal and the distance from the core in the shower plane is fit using an NKG-like function:

$$S(r) = S_{1000} \left(\frac{r}{1000}\right)^{-\beta} \left(\frac{r+700}{1000+700}\right)^{-\beta} \tag{7.4}$$

This is the lateral distribution function (LDF). A maximum likelihood fit is made using the signals from the tanks, adjusting the slope and reconstructed core

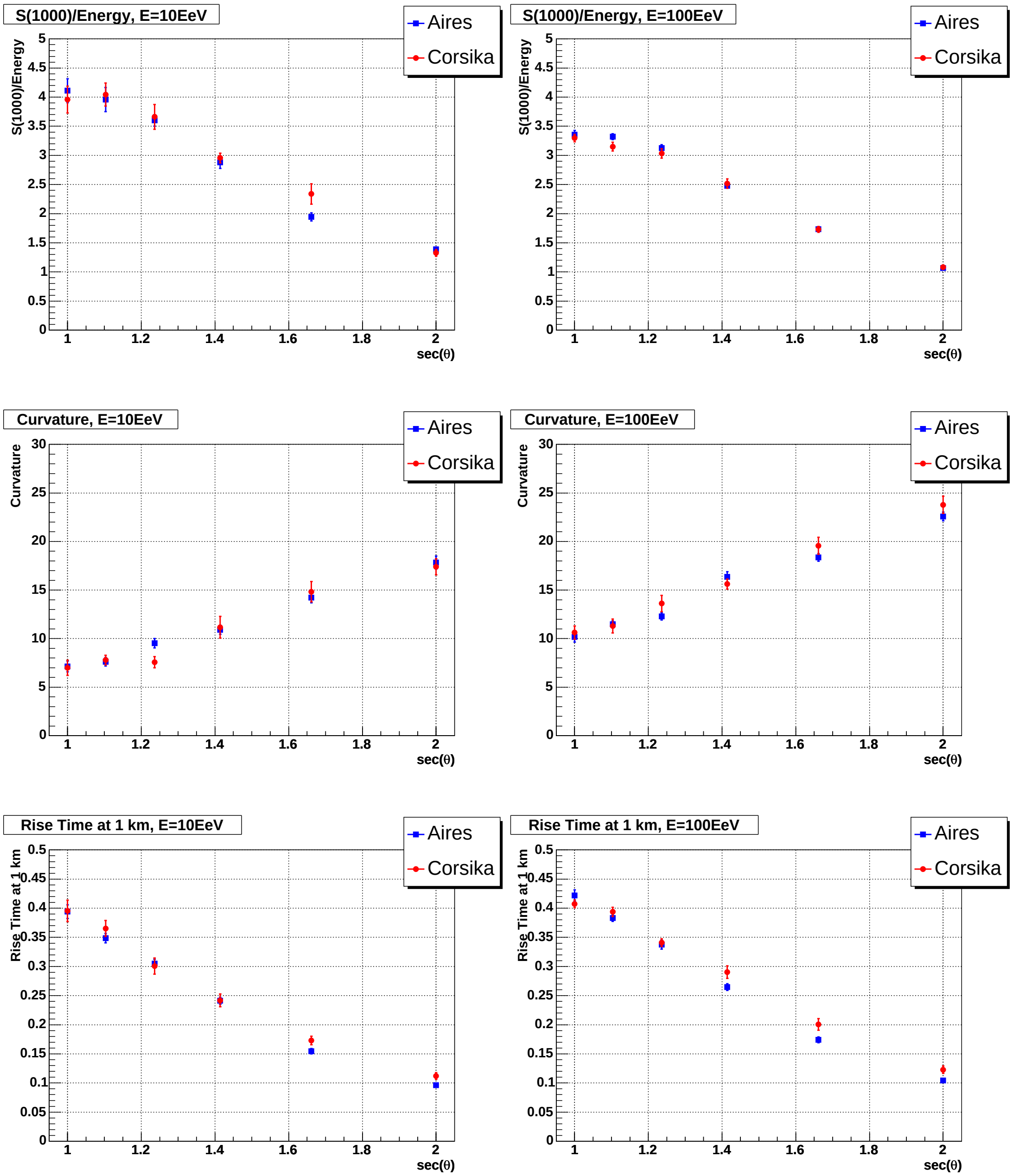


Figure 7.8: Aires and Corsika monte carlo showers at 10 EeV and 100 EeV. Top: S(1000). Middle: Shower front curvature. Bottom: Rise time at 1 km.

location to maximize the likelihood. With the reconstructed core location and slope, the signal at 1 km from the core, S_{1000}, is calculated and the energy is determined using a S_{1000} to energy conversion. For an example of an event with these parameters reconstructed, see Fig. 7.9 for the event display for a real event and a monte carlo event with similar input values for energy, zenith, and azimuth. The reconstructed rise and fall times for the same events are presented in Fig. 7.10 where fall times are analogous to rise times except the fall time is the time it takes to go from 50% to 90% of the total energy deposited.

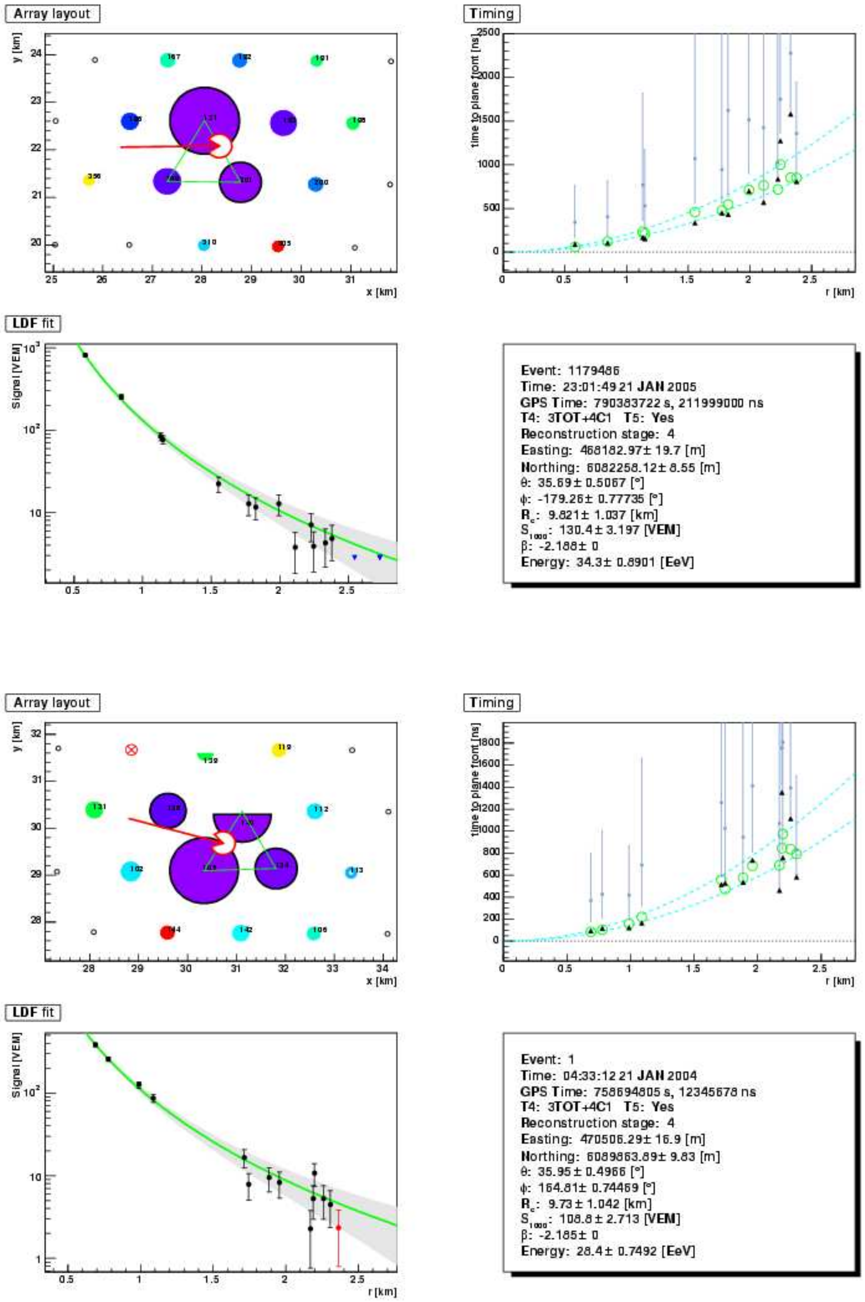
7.4 Analysis Method

To be able to compare monte carlo predictions with real data, it would be impossible to have dedicated simulations for each event in the surface array. The approach is to parameterize the predictions of certain monte carlo showers. More explicitly, the values of the ground observables for the given monte carlo showers at specific energies and angles are parameterized to enable interpolation of the predicted value of the observables at all energies and zenith angles to compare with real events. For example, if a real event has a reconstructed zenith angle of 43° and S(1000) of 53 VEM, what would be the mean and standard deviation of the shower front curvature and rise time at 1 km for monte carlo proton showers with the QGSJET01 hadronic model with the same S(1000) and zenith angle? This can be repeated for various primaries and models.

To obtain the parameterization, the dependence on log(energy) is assumed to be linear at all zenith angles. This leads to a function of the form:

$$f(S(1000), \theta) = m(\theta) \log S(1000) + b(\theta) \tag{7.5}$$

The first step is then to find the energy dependence of the observable at each



Event: 1179486
Time: 23:01:49 21 JAN 2005
GPS Time: 790383722 s, 211999000 ns
T4: 3TOT+4C1 T5: Yes
Reconstruction stage: 4
Easting: 468182.97± 19.7 [m]
Northing: 6082258.12± 8.55 [m]
θ: 35.69± 0.5067 [°]
ϕ: -179.26± 0.77735 [°]
R_c: 9.821± 1.037 [km]
S_{1000}: 130.4± 3.197 [VEM]
β: -2.188± 0
Energy: 34.3± 0.8901 [EeV]

Event: 1
Time: 04:33:12 21 JAN 2004
GPS Time: 758694805 s, 12345678 ns
T4: 3TOT+4C1 T5: Yes
Reconstruction stage: 4
Easting: 470506.29± 16.9 [m]
Northing: 6089863.89± 9.83 [m]
θ: 35.95± 0.4966 [°]
ϕ: 164.81± 0.74469 [°]
R_c: 9.73± 1.042 [km]
S_{1000}: 108.8± 2.713 [VEM]
β: -2.185± 0
Energy: 28.4± 0.7492 [EeV]

Figure 7.9: Event display for a real event (top) and a monte carlo event (bottom).

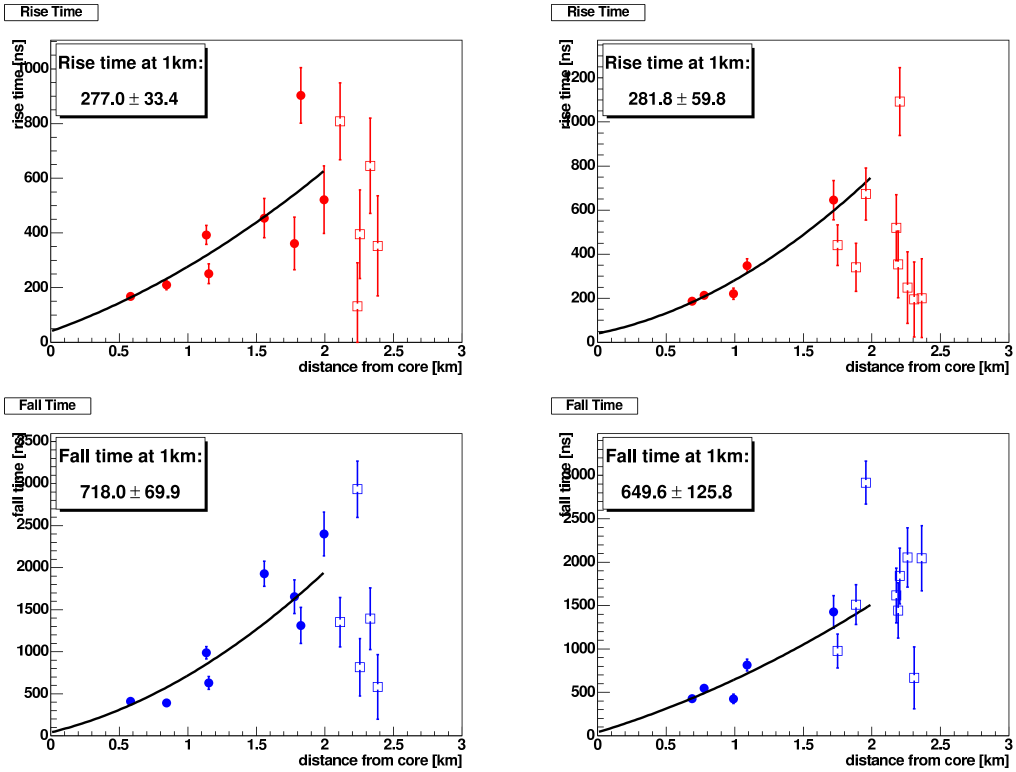


Figure 7.10: Event display for the rise and fall times for a real event (left) and a monte carlo iron initiated shower (right), same events as in Fig. 7.9. Rise times are in red, fall times are in blue and closed markers are the stations used to make the fit and open markers are the stations that failed the cuts.

zenith angle. This is achieved by finding the average S(1000) and rise time (or curvature) for each energy at the fixed zenith angle and fitting a straight line. This is seen in Fig. 7.11. The next step is to determine the zenith dependence of the 2 fit parameters, as seen in Fig. 7.12. With these fit values, the dependence on S(1000) and zenith angle is determined for the particular observable. This same procedure is repeated but for the standard deviation of the variable as a function of S(1000) and zenith, with the functional form:

$$\sigma(S(1000), \theta) = \frac{p(\theta)}{1 + \log S(1000)} \tag{7.6}$$

.

The next step is to investigate the consistency of the parameterization. Using the same monte carlo data that was used to determine the parameterization, the deviation from the parameterized prediction can be investigated. For each event, the predicted value for the particular observable, given S(1000) and the zenith angle, is compared with the reconstructed value of the observable. The deviation from the predicted mean in terms of the parameterized standard deviation is put into a histogram. The expected result is that the mean of the histogram is 0 and the standard deviation is 1. A plot of this histogram for Aires proton showers using QGSJET01 for the rise time at 1 km is shown in Fig. 7.13. A benefit of this approach is that if there are non-gaussian tails, it is not necessary to include that in the original parameterization because the final distribution, as shown in Fig. 7.13, will reveal any non-gaussian tails in the data. It is from this distribution that probabilities can be calculated, including non-gaussian tails. That is why the actual computed mean and standard deviation of the histogram differs from the gaussian fit, the fitting with a gaussian just reflects how well the parameterization approximates the data, ignoring non-gaussian behavior.

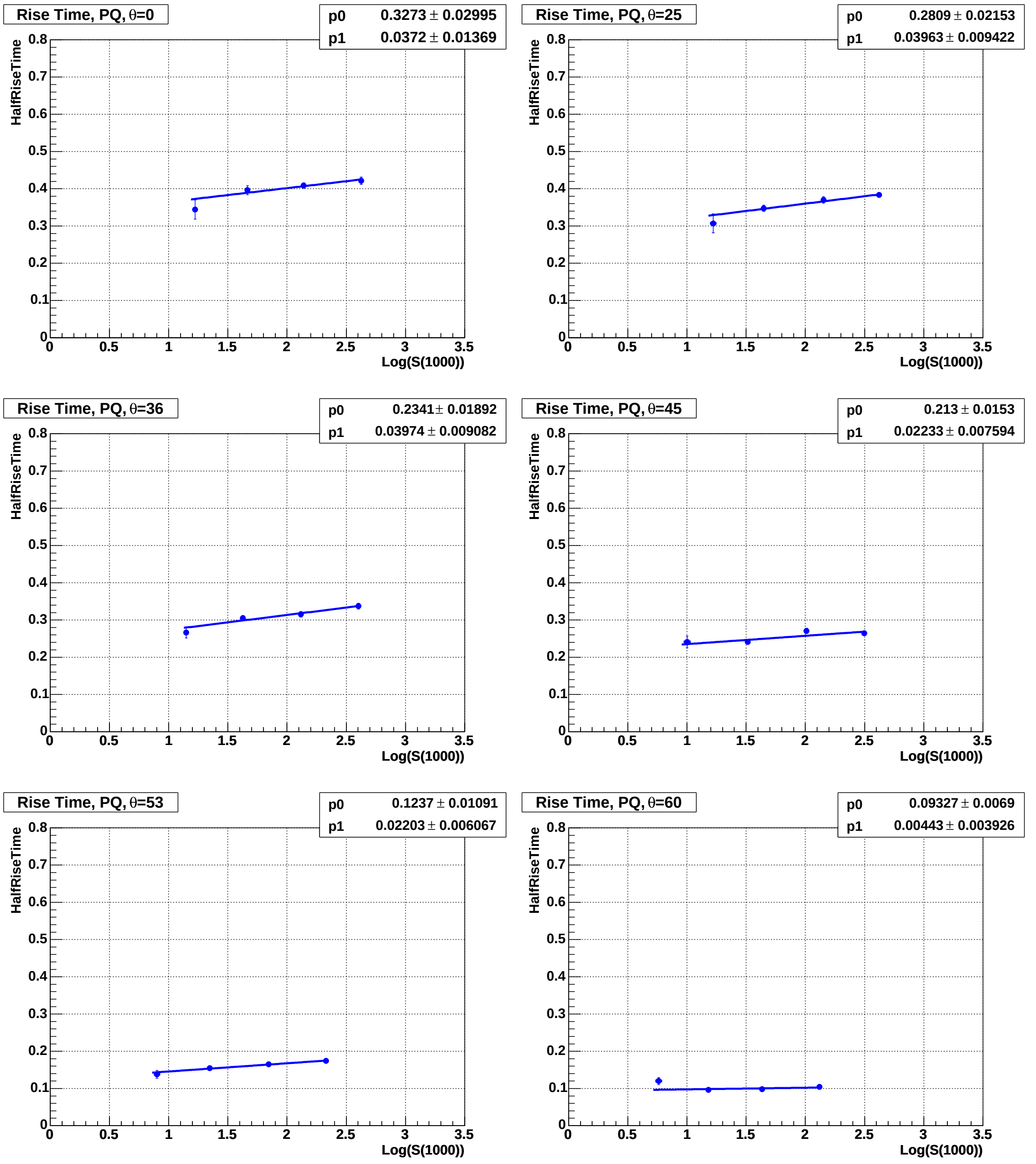


Figure 7.11: Plots of the parameterization vs. log(S(1000)) for the rise time at 1 km for Aires proton simulations using QGSJET01 hadronic model at 6 fixed zeniths (0, 25, 36, 45, 53, and 60°). The three points correspond to three energies (10, 31, and 100 EeV) the average S(1000) and rise time for each energy.

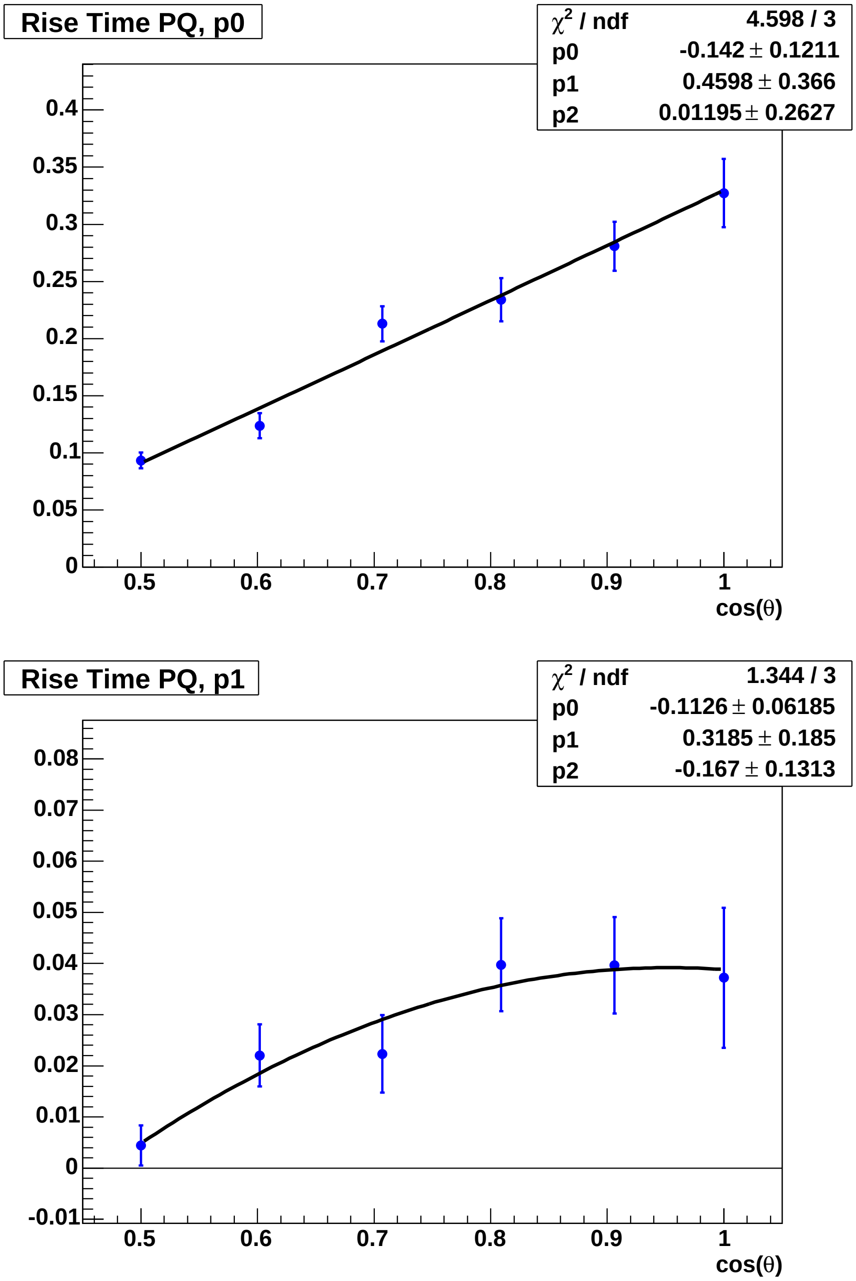


Figure 7.12: Plots of the fit parameters from the log(S(1000)) dependence vs. $\cos\theta$. The upper plot is $b(\theta)$ from (7.5) and the lower plot is $m(\theta)$. The zenith dependence is fit with a quadratic function in $\cos\theta$. This is again for the rise time of Aires proton simulations using QGSJET01 hadronic model.

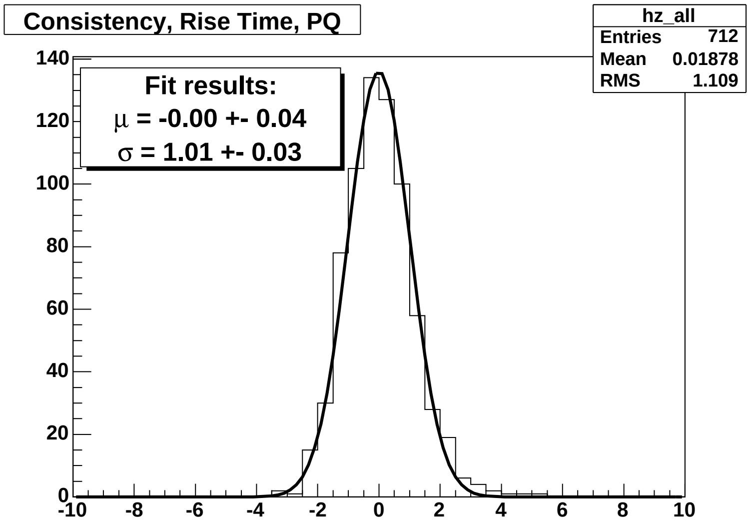


Figure 7.13: Gaussian fit to the resulting distribution of monte carlo events when compared to the parameterized prediction. If monte carlo distribution is gaussian and parameterization is good, mean should be 0 and rms should be 1, which is the case here.

7.5 Analysis Results: Baryon Primaries

Seeing that the resulting parameterization is satisfactory, real data can now be analyzed. As with the consistency check explained in the previous paragraph, for each real event, the predicted mean value of a given observable is compared with the actual reconstructed observable using the parameterized function and the reconstructed S(1000) and zenith of the real event. The deviation from the predicted mean is defined as the distance from the mean along the axis of the observable in units of the parameterized standard deviation. Real data was restricted to zeniths less than 60° and energy greater than 7 EeV as reconstructed using a constant intensity cut S(1000) to energy converter [85]. The same cuts are applied to the monte carlo data as well so that the same minimum S(1000) is required as for the real data. The results of the comparisons are shown in Fig. 7.14 and Fig. 7.15.

It is possible to use the monte carlo predicted values to assign a probability on an event by event basis according to where it falls in the monte carlo distributions. For example, event x has a curvature close to the predicted mean of the Aires proton QGSJET01 curvature distribution, therefore it has a high probability that it is an “Aires proton QGSJET01” event. On the other hand, the distribution of real data can be compared as a whole to the monte carlo predicted distribution. Thus, the probability that the real data, on average, is like “Aires Iron Sibyll 2.1” can be calculated using the mean of the real data distribution and the uncertainty on the mean compared with the mean of the simulated data and that uncertainty:

$$\chi^2 = \frac{(\mu_{real} - \mu_{mc})^2}{\Delta\mu_{real}^2 + \Delta\mu_{mc}^2} \tag{7.7}$$

It is this latter approach that will be used here, the goal being to see what model best describes the mean behavior of the real data and any changes with energy.

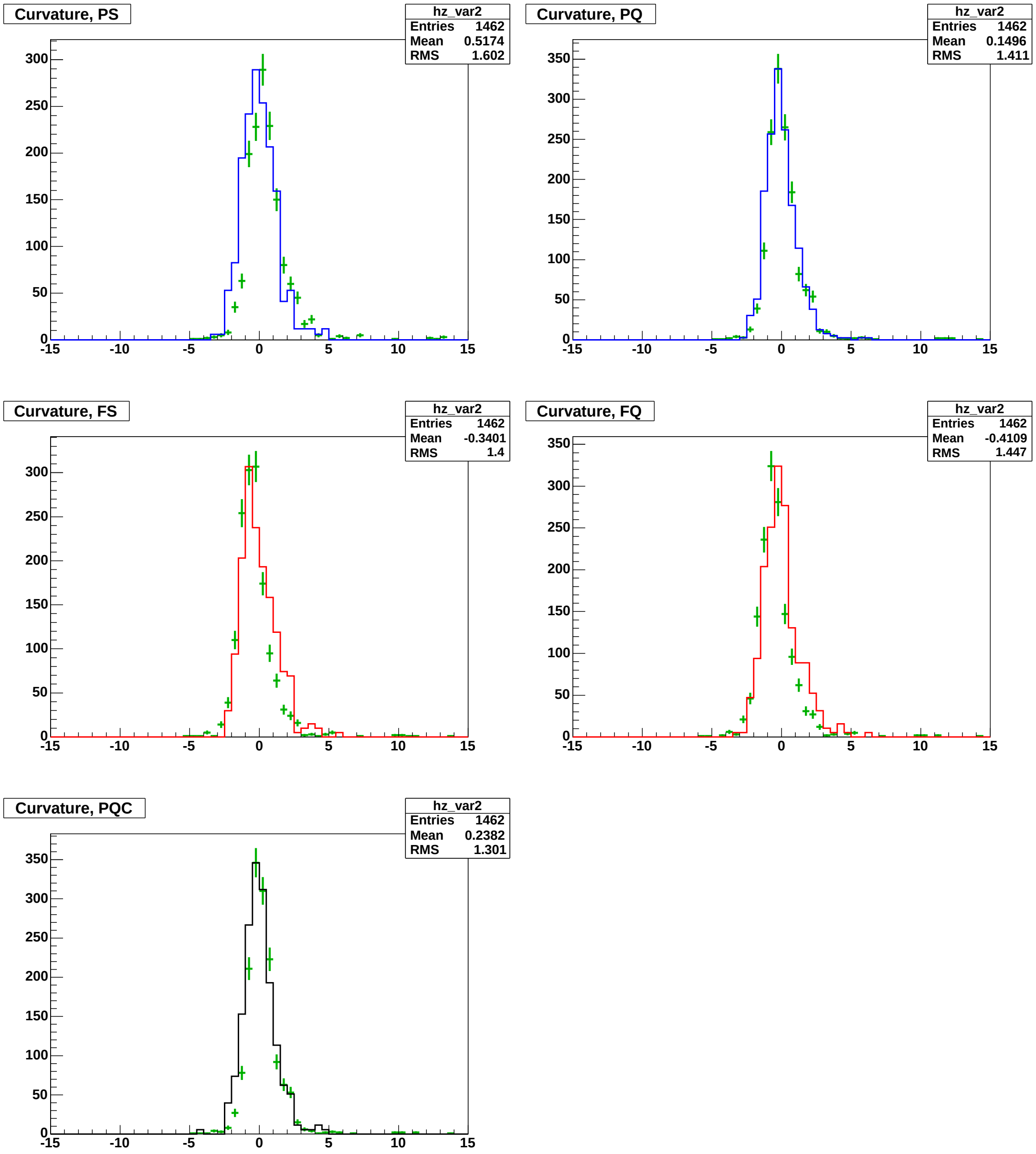


Figure 7.14: Distributions of simulated events (solid lines) and real data (green crosses) of the deviation from the predicted mean curvature in units of parameterized standard deviation. (Key: P=proton, F=iron, S=Sibyll, Q=QGSJET, C=Corsika, Aires otherwise).

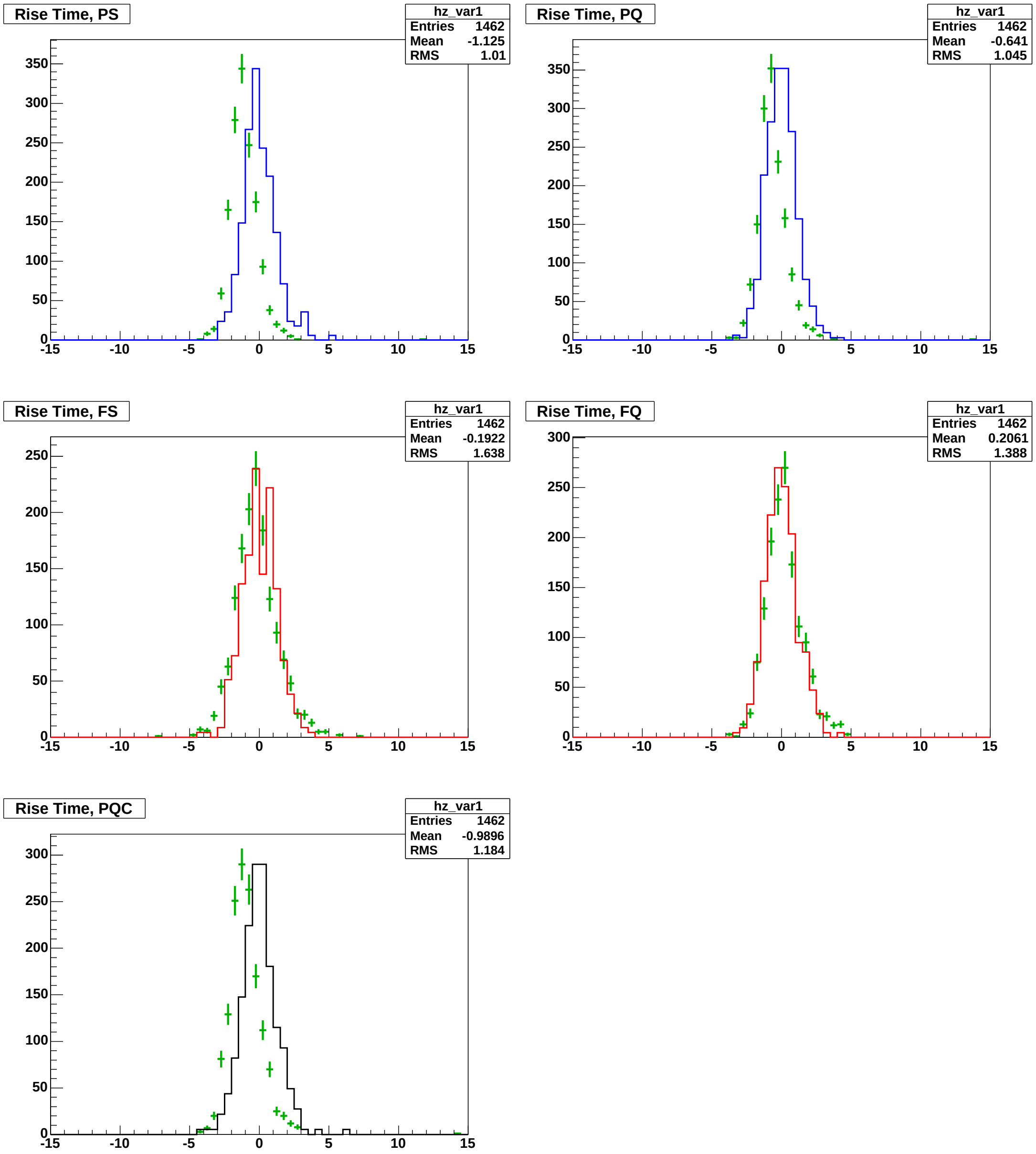


Figure 7.15: Distributions of simulated events (solid lines) and real data (green crosses) of the deviation from the predicted mean rise time in units of parameterized standard deviation. (Key: P=proton, F=iron, S=Sibyll, Q=QGSJET, C=Corsika, Aires otherwise).

Combo	Rise Time	Curvature
PS	210	36.2
PQ	156	6.53
PQC	149	9.77
FS	5.38	17.4
FQ	7.69	24.2

Table 7.1: Table of χ^2 values of different combinations in the simulations for the rise time and curvature observables. The Minimum energy is 7 EeV. P means proton, F means iron, S means Sibyll 2.1, Q means QGSJET01 and C means Corsika (otherwise it is Aires).

Table 7.1 is a summary of the results for two different observables and five different simulation combinations. As can be seen from the χ^2 values, rise time seems to be the most discriminatory variable while curvature seems less so. Rise time data alone may be sufficient to discriminate between primaries, but curvature contains information with respect to the composition as well. The next section will describe how to best combine these variables.

7.5.1 Combining Observables

On their own, the observables discussed do a fair job of discriminating between simulations, but it would be preferable to combine them to gain even more discriminating power. The χ^2 values can not simply be added together if there are any correlations between the variables. This is easily remedied by a transformation of coordinate system into a coordinate system where there is no correlation. In this new coordinate system the transformed variables are independent of each other.

The procedure for transforming two correlated variables into two independent variables can easily be expanded to many variables. First, it is necessary to construct the covariance matrix:

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{pmatrix} \tag{7.8}$$

where

$$cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{(n-1)} \tag{7.9}$$

and it is obvious that $cov(x,y) = cov(y,x)$ and $cov(x,x)$ is just the variance of x. The goal is to find a coordinate system where $cov(x,y) = 0$, which can be accomplished by a simple rotation. The angle through which to rotate is found by diagonalizing the matrix.

The procedure continues by finding the eigenvalues and eigenvectors of the covariance matrix. Using the eigenvectors (normalized such that the magnitudes are 1), the rotation matrix can be constructed as:

$$R = \begin{pmatrix} \vec{v_1} & \vec{v_2} \end{pmatrix} \tag{7.10}$$

where $\vec{v_1}$ and $\vec{v_2}$ are the two normalized eigenvectors used as the columns in the rotation matrix. The data can then be rotated using the rotation matrix, R, into a coordinate system where they are uncorrelated.

Another option, more applicable when using many variables, is reducing the dimensionality of the problem. This is called principal component analysis and is based on the preceding discussion of rotating to a new basis where all variables are uncorrelated. However, if many of the variables are correlated, not a lot is gained by including all of them in the analysis. The most significant components are those with the highest eigenvalues, and the most significant eigenvector is called the principal component. The proportional amount of “information” in each component, or the significance is:

$$S_j = \frac{\lambda_j}{\sum_{i=1}^{n} \lambda_i} \tag{7.11}$$

where λ is the eigenvalue. The dimensionality of the problem then can be reduced by choosing only the p most significant eigenvectors out of n total eigenvectors.

Performing these transformations on the monte carlo data and analyzing the real data again, rise time and curvature information can be combined to further distinguish the disagreement, or the agreement, between monte carlo predictions and real data. In Figs. 7.16 through 7.20, four plots are presented in each figure. The upper left plot is a plot of the deviation of the real data from the monte carlo prediction (green points) along with the monte carlo data (blue, red, and black markers) for curvature and rise time. The gray lines represent the eigenvectors (with the length proportional to the magnitude) where the solid line is the principal component and the dashed line is the secondary component. The plot in the upper right is the rotated comparison between monte carlo and real data. The rotation is computed only using the monte carlo data to get rid of any correlations. The bottom plots are the histograms of the two components with the percentage of the total significance in the title of the plot, where the green crosses are the real data and the solid lines are the monte carlo predictions. It is then possible to compute the distance from the prediction, according to (7.7), for the two components and then sum them to see the total deviation from the model predictions. These results are summarized in Table 7.2. According to the sum of the χ^2 values of components 1 and 2, the best fit to the data is an iron primary, on average.

The preceding tables and figures were based on requiring that the minimum energy be 7 EeV for all events according to the constant intensity cut based S(1000) to energy conversion. It is interesting to investigate the dependence on

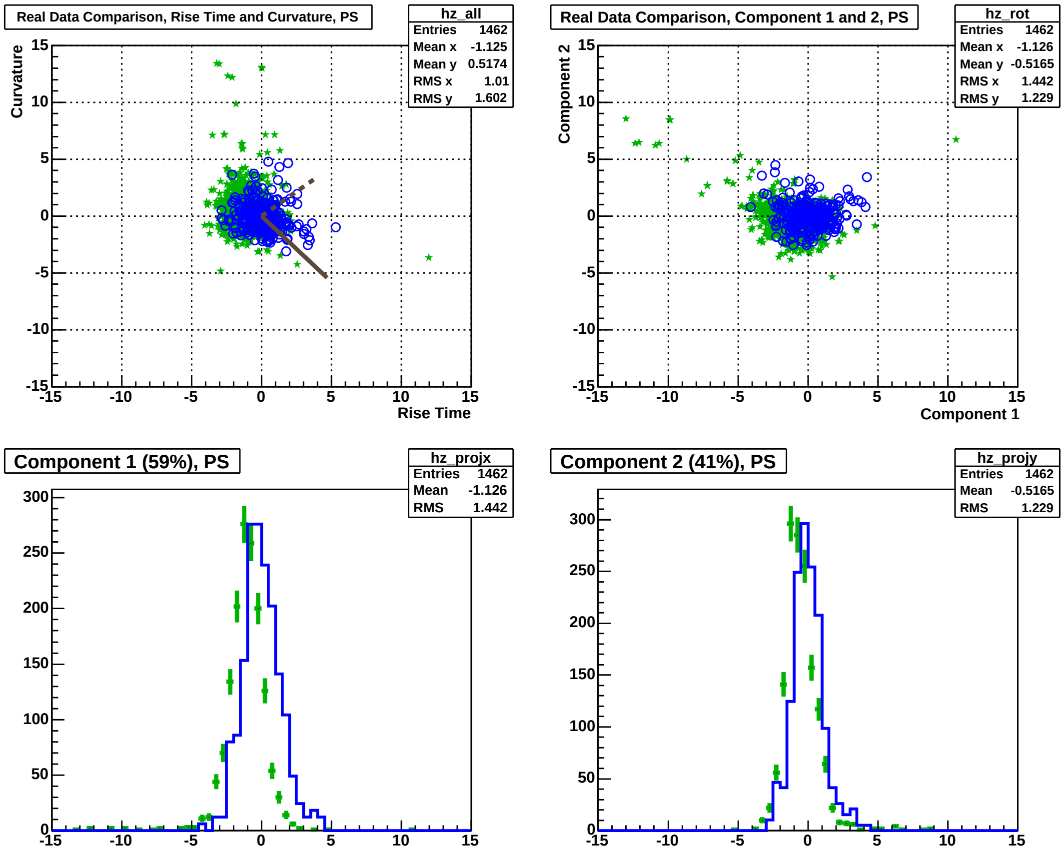


Figure 7.16: Aires+Proton+Sibyll, Top left: Rise time and curvature deviation from prediction. Top right: Rotated analysis. Bottom left: Component 1 histogram. Bottom right: Component 2 histogram.

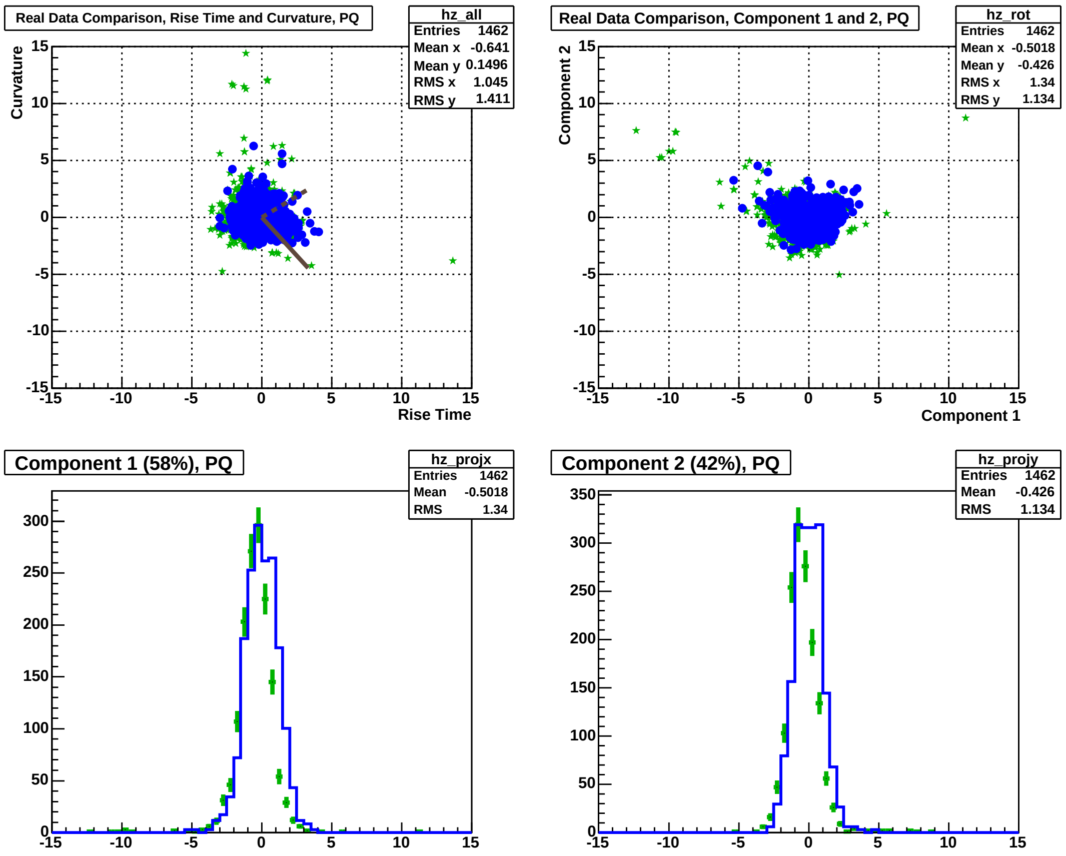


Figure 7.17: Aires+Proton+QGSJET, Top left: Rise time and curvature deviation from prediction. Top right: Rotated analysis. Bottom left: Component 1 histogram. Bottom right: Component 2 histogram.

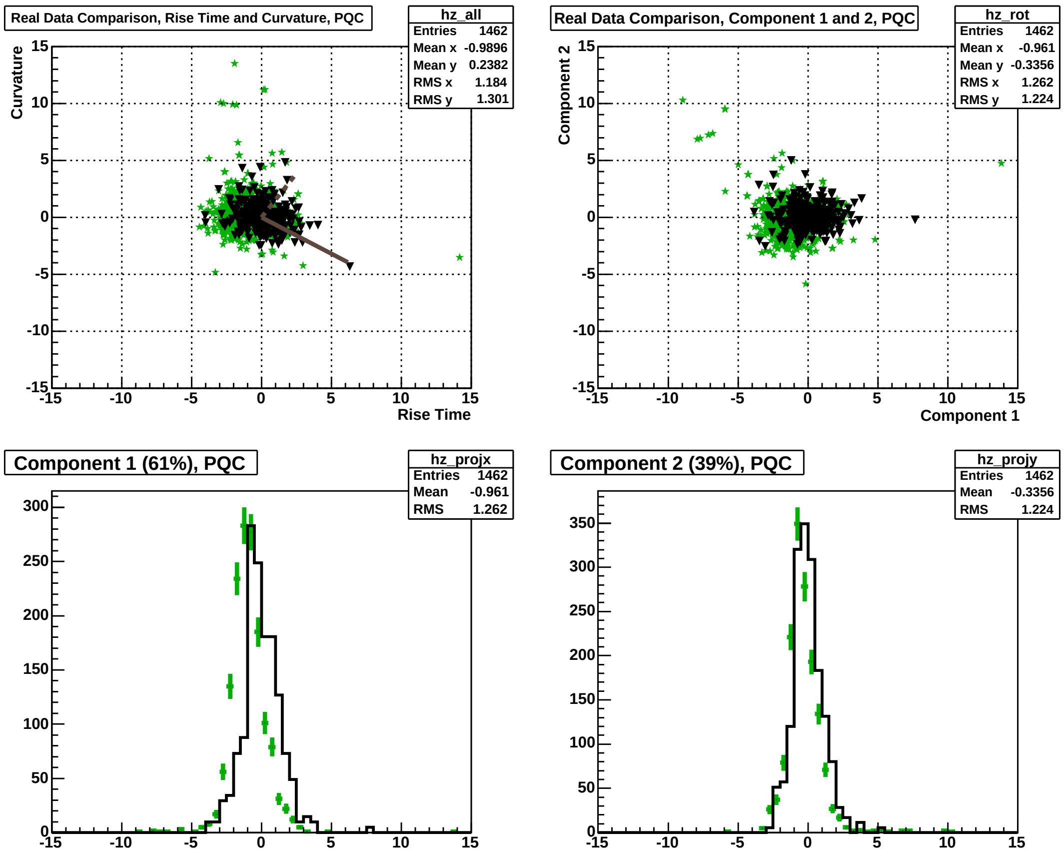


Figure 7.18: Corsika+Proton+QGSJET, Top left: Rise time and curvature deviation from prediction. Top right: Rotated analysis. Bottom left: Component 1 histogram. Bottom right: Component 2 histogram.

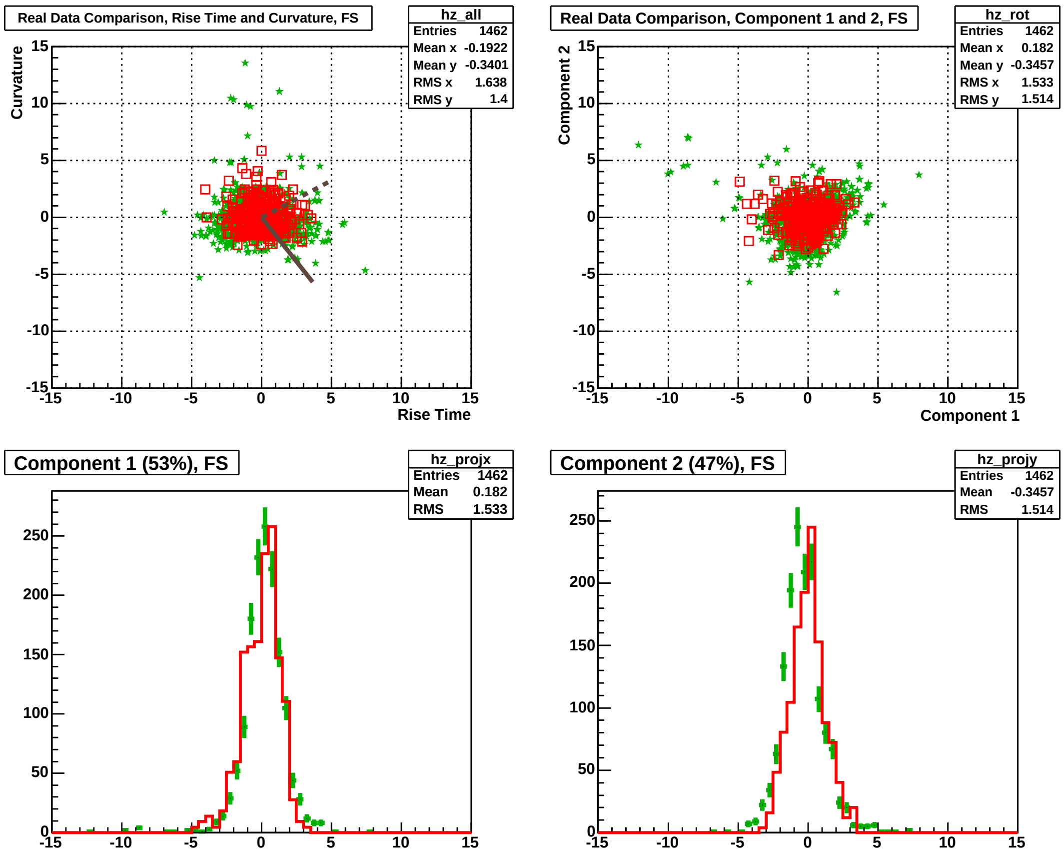


Figure 7.19: Aires+Iron+Sibyll, Top left: Rise time and curvature deviation from prediction. Top right: Rotated analysis. Bottom left: Component 1 histogram. Bottom right: Component 2 histogram.

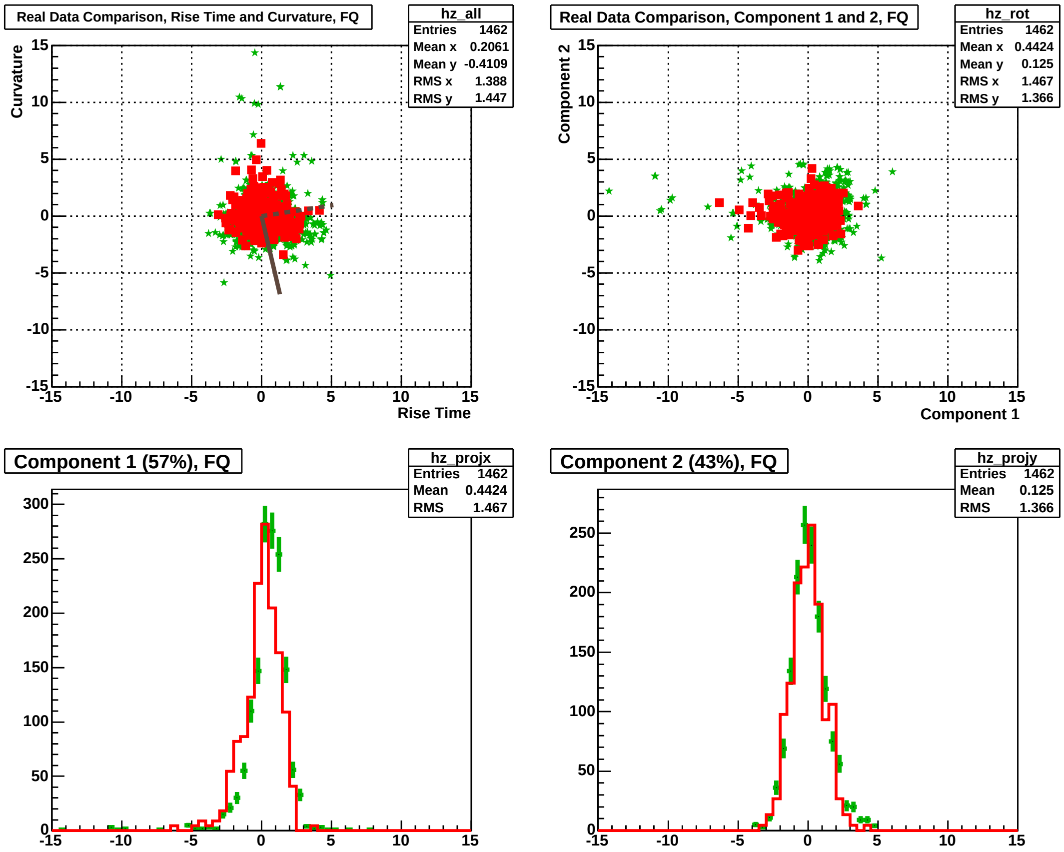


Figure 7.20: Aires+Iron+QGSJET, Top left: Rise time and curvature deviation from prediction. Top right: Rotated analysis. Bottom left: Component 1 histogram. Bottom right: Component 2 histogram.

Combo	Rise Time	Curvature	Component 1	Component 2	Sum/2
PS	210	36.2	161	48.0	105
PQ	156	6.53	71.6	71.6	71.6
PQC	149	9.77	125	22.5	73.8
FS	5.38	17.4	4.68	18.6	11.6
FQ	7.69	24.2	27.7	2.87	15.3

Table 7.2: Table of χ^2 values of different combinations in the simulations for the rise time and curvature observables, rotated component 1 and component 2, and their sum divided by 2 degrees of freedom. The minimum energy is 7 EeV. P means proton, F means iron, S means Sibyll 2.1, Q means QGSJET01 and C means Corsika (otherwise it is Aires).

energy of the χ^2 values to see any shift in the behavior of the real data. In Table 7.3, the χ^2 values are presented at 6 different minimum energies for the shower front curvature, the rise time at 1 km, and the sum of the two rotated components divided by 2. Focusing on the sum of the two components, the trend is that the real data is best approximated by iron-Sibyll simulations at lower energies which then shifts to iron-QGSJET at around 13 EeV. This could indicate that real cosmic ray showers are very muon rich because one characteristic of QGSJET showers is that they produce more muons than Sibyll showers. All the χ^2 values become smaller with higher energy on average, but that is just due to the number of events becoming smaller, thus making the uncertainty on the mean larger resulting in smaller χ^2 values for all combinations. The key is to notice the relative changes amongst the different combinations.

Since the χ^2 values are measures of distance from the mean prediction of monte carlo showers, it is not obvious where the distribution of real events falls, whether

Curvature						
Combo	5 EeV	7 EeV	9 EeV	11 EeV	13 EeV	15 EeV
PS	50.5	36.2	24.2	24.9	23.2	22.8
PQ	7.10	6.53	3.65	5.92	7.09	9.14
PQC	14.2	9.77	6.64	8.42	6.82	8.03
FS	16.5	17.4	17.2	7.75	3.97	1.26
FQ	23.3	24.2	25.1	14.3	5.49	2.51
Rise Time						
PS	189	210	183	172	191	202
PQ	147	156	123	108	115	117
PQC	143	149	123	102	118	125
FS	3.98	5.34	4.38	4.01	8.78	11.9
FQ	8.04	7.69	9.56	6.44	2.46	1.01
Sum/2 of 2 rotated components						
PS	106	104	88.0	76.7	78.0	72.0
PQ	75.8	71.6	55.6	43.2	41.7	37.0
PQC	73.4	74.0	60.9	48.8	52.9	53.1
FS	9.56	11.6	10.7	5.88	6.44	6.58
FQ	15.1	15.3	16.7	9.56	3.54	1.54

Table 7.3: Table of χ^2 values of different combinations in the simulations for the rise time and curvature observables, and the sum of 2 rotated components divided by 2 degrees of freedom. P means proton, F means iron, S means Sibyll 2.1, Q means QGSJET01 and C means Corsika (otherwise it is Aires).

it be between iron and proton predictions or whether it falls outside of both. In Fig. 7.21 a plot is shown of the monte carlo mean predictions for a fixed energy of 10 EeV (according to the constant intensity cut S(1000) to energy converter) as a function of $\sec\theta$. For each data point in the monte carlo predictions, the S(1000) that corresponds to a 10 EeV shower as predicted from the constant intensity cut S(1000) to energy converter was used as input in the parameterized function for each ground observable, as explained in the beginning of Section 7.4. Real data is superimposed with allowed energies being greater than 8 EeV and less than 12 EeV using the same energy conversion. It is apparent that the real data falls between the iron and proton predictions as expected.

7.6 Analysis Results: Photon Primaries

The previous section dealt with baryon primaries only, but the analysis methods presented have a much stronger distinguishing power when comparing monte carlo predictions of photon primaries to the real data. As an example, Fig. 7.21 includes only baryon primaries, but if photon primaries are included in the plots, Fig. 7.22, it is obvious that photon induced cosmic ray showers have very different characteristics. The necessity of combining observables is also apparent because there may be a few "photon-like" events when only considering one observable, but they do not persist in the other observable. An upper limit on the photon flux will be derived in this section based on the difference between real data and the photon monte carlo predictions.

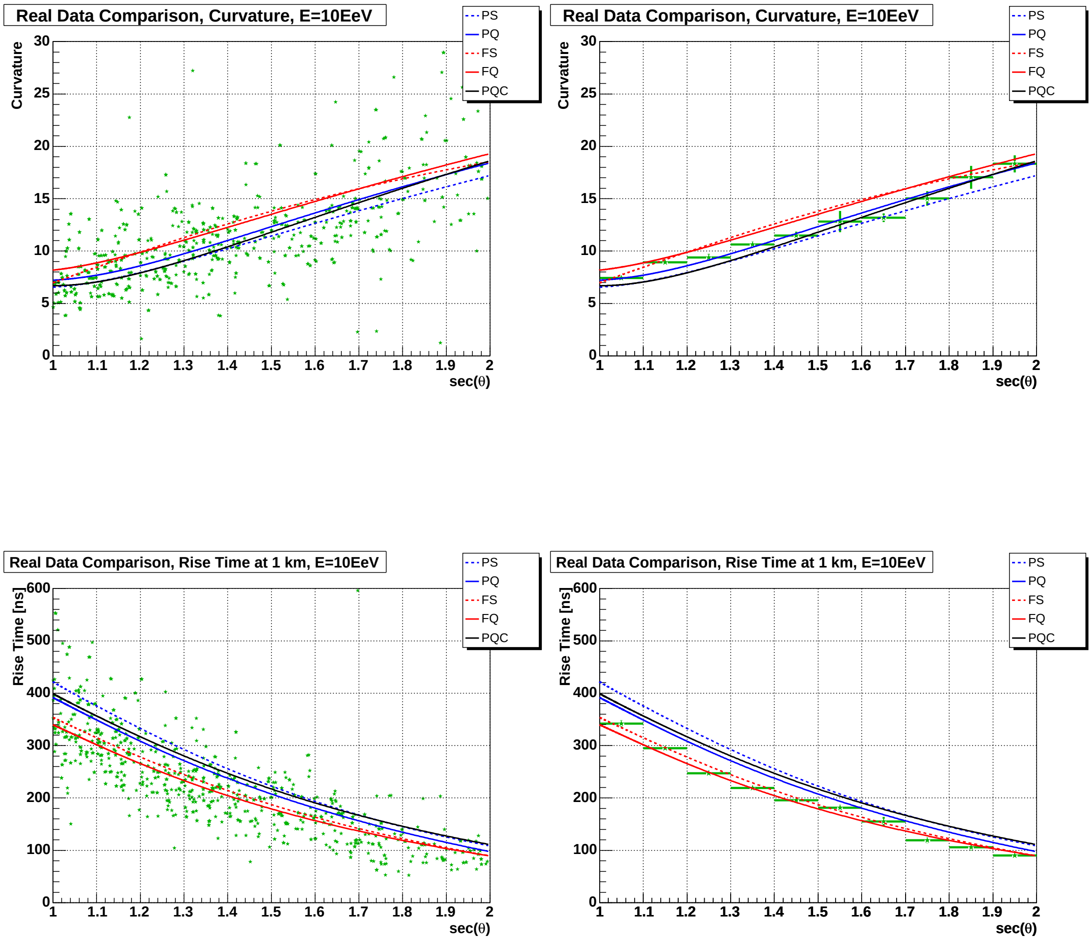


Figure 7.21: Predictions of simulations (lines) and real data (green crosses) for a fixed energy in monte carlo showers and $8 < E < 12$ EeV for real data. On the left are the scatter plots of the raw data, and on the right are binned plots of that same data. On the top are plots of curvature, and rise time on the bottom. (Key: P=proton, F=iron, S=Sibyll, Q=QGSJET, C=Corsika, Aires otherwise).

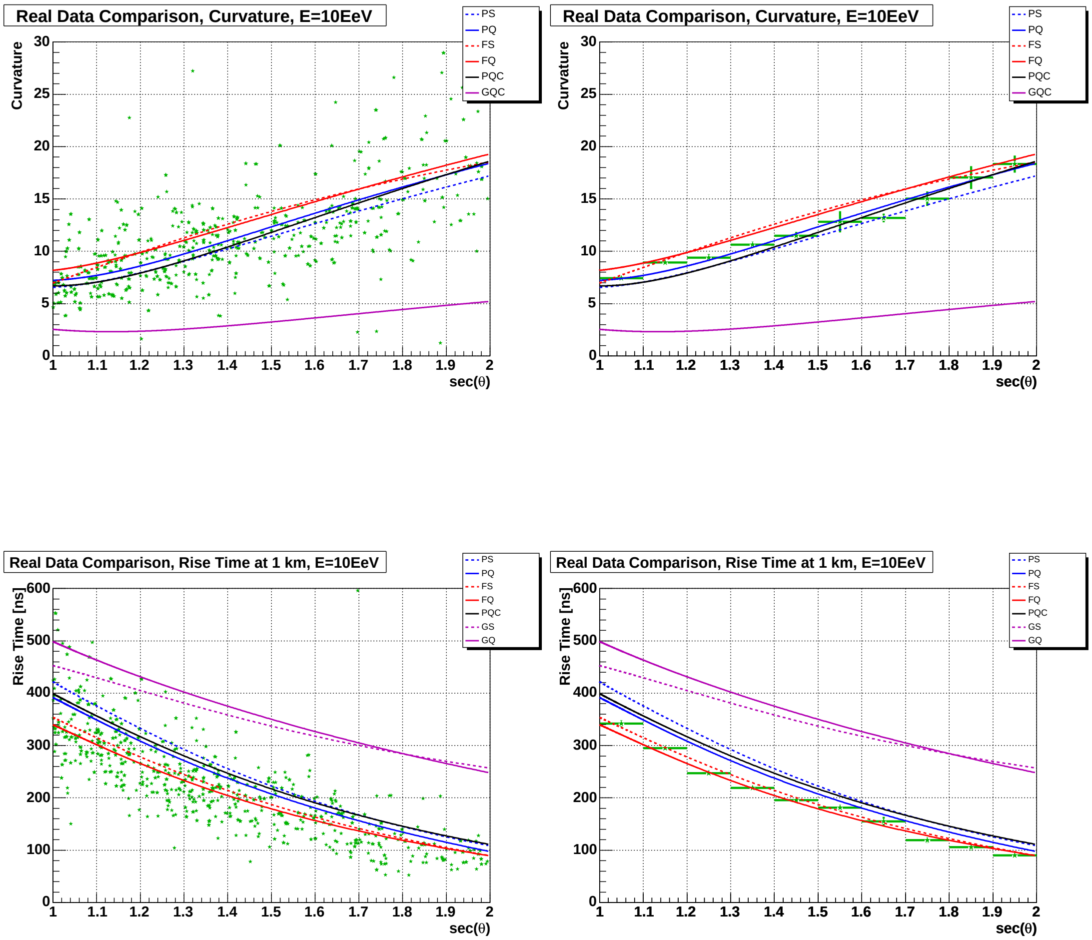


Figure 7.22: Predictions of simulations (lines) and real data (green crosses) for a fixed energy in monte carlo showers and $8 < E < 12$ EeV for real data. On the left are the scatter plots of the raw data, and on the right are binned plots of that same data. On the top are plots of curvature, and rise time on the bottom. (Key: P=proton, F=iron, G=Photon, S=Sibyll, Q=QGSJET, C=Corsika, Aires otherwise).

7.6.1 Monte Carlo Showers Induced by Photons

Special attention must be paid to the monte carlo showers induced by photons. Depending on the initial energy of the photon cosmic ray and the incoming direction, there is a probability that the photon will convert in the magnetic field of the earth [87]. At high enough energies, around 10 EeV, there is a probability that a photon will interact with the magnetic field and convert into an e^+e^- pair. These particles can subsequently lose energy by magnetic bremsstrahlung, creating more photons. If these photons have high enough energies, they can convert as well. This cascade can result in a shower of particles before entering the atmosphere, or a preshower. Thus, when entering the atmosphere, the single initial photon is now a superposition of a number of lower energy particles that initiate an extensive air shower.

A converted photon shower will differ from an extensive air shower initiated by a photon in the atmosphere due to the Landau-Pomeranchuk-Migdal (LPM) effect [88, 89]. The LPM effect is a suppression of the Bethe-Heitler cross-section which results in larger fluctuations in number of particles in the shower and deeper X_{max}. The threshold for this effect depends on the density of the medium, and for the upper atmosphere the energy is around 10^{19} eV. A converted shower will contain many low energy photons and electrons rather than one high energy photon and the result is that the converted shower will be less affected by the LPM effect than a similar unconverted photon shower, i.e. the X_{max} will be shallower and the fluctuations will be smaller.

It is necessary to include this possibility in the monte carlo simulations of cosmic ray air showers. For Corsika and Aires, preshowering programs have been written that interface with the original code to correctly account for photon preshowering [90, 91], and shower libraries have been produced using this code at a computing center in Lyon. The Corsika photon showers are simulated only at 3 zeniths (0, 45, and 60°) while the Aires photon showers were simulated at the usual 6 zenith angles (0, 25, 36, 45, 53, and 60°). For that reason, the Corsika simulations at three zenith angles and three energies are to be used solely for comparison with the Aires photon simulations. The Aires photon simulations were done at seven different energies (in $\log E$: 19, 19.3, 19.5, 19.8, 20, 20.2, and 20.5). The Aires simulations will be used in the analysis and the Corsika showers will only be used to compare with the Aires predictions to investigate any systematic difference between them. This comparison is shown in Fig. 7.23. There is no significant difference between Aires and Corsika in the observables of interest.

7.6.2 Results of Mean Behavior

The analysis done comparing real data to the predictions of baryon primaries for certain ground observables can be repeated for photon primaries. However, not much information is gained from this analysis as it is obvious from Fig. 7.22 that the majority of events are not induced by photon primaries. Therefore, the mean behavior of the real data is very different from a prediction that all events are induced by photon primaries; but, for the sake of continuity, it is presented here. Figure 7.24 is the deviation of real data from the predicted mean of photon induced showers for curvature and rise time. For comparison, the χ^2 values for rise time and curvature compared to photon-QGSJET showers in Aires are 1080 and 6240, respectively. The sum of the χ^2 values of the two uncorrelated components (divided by 2) is 3600, clearly indicating that the mean behavior of cosmic ray showers is not photon-like.

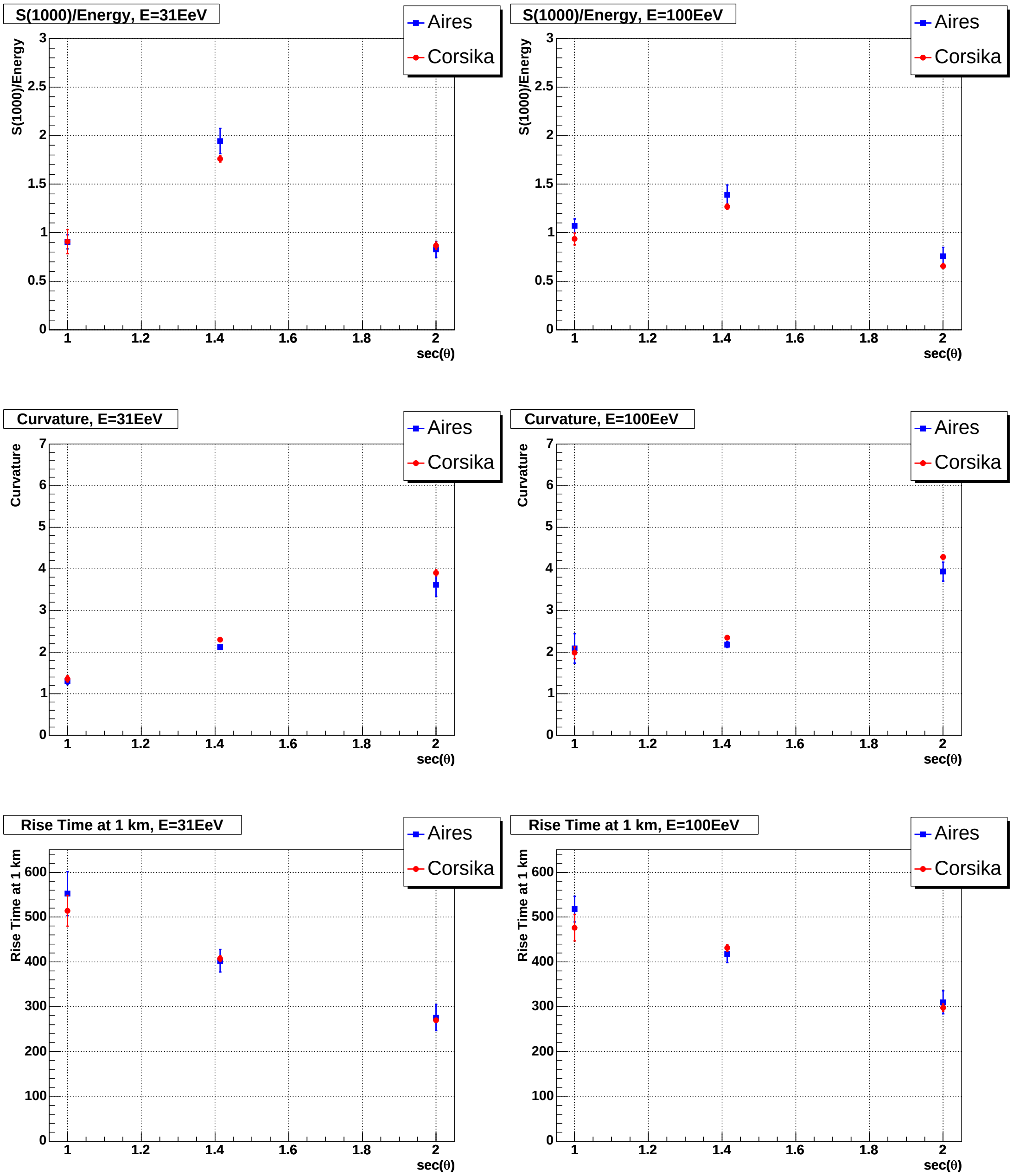


Figure 7.23: Aires and Corsika photon monte carlo showers at 31 EeV and 100 EeV. Top: S(1000). Middle: Shower front curvature. Bottom: Rise time at 1 km.

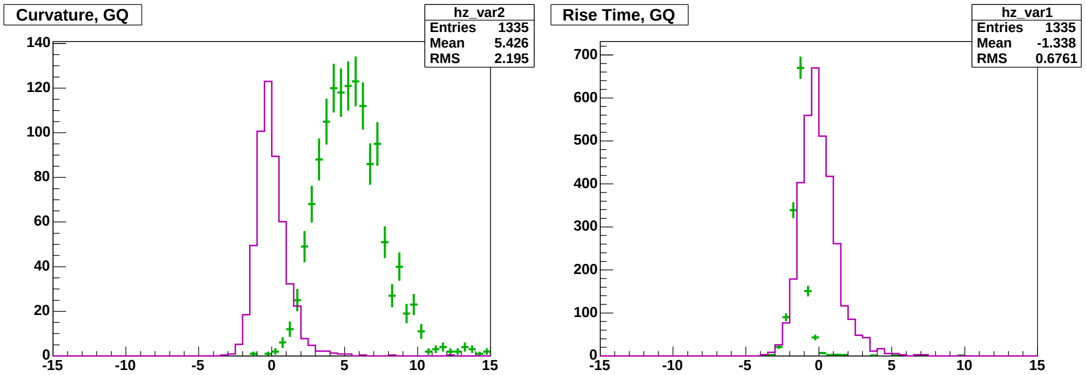


Figure 7.24: Distributions of photon simulated events (solid lines) and real data (green crosses) of the deviation from the predicted mean curvature (left) and rise time (right) in units of parameterized standard deviation Simulations are Aires+Photon+QGSJET.

7.6.3 Deriving an Upper Limit

As was expected, the mean behavior of real data is not compatible with all showers being induced by photons. There is a possibility, however, that a fraction of the events are induced by photons and that requires a different analysis.

Traditionally, what is calculated is a photon fraction, or the fraction of cosmic rays that are initiated by photons [92, 93]. This is energy dependent, and depends on the "top-down" model used to predict that fraction. The difficulty in using the array of surface detectors for the determination of the photon fraction is that a photon shower deposits less energy on the ground than a baryon shower of the same primary energy. This is shown in Fig. 7.25.

The problem that this causes can be illustrated by a simple example. Assume that the true photon fraction is 50% and that three energy bins are being considered, each bin is centered on an energy that is twice the energy of the previous

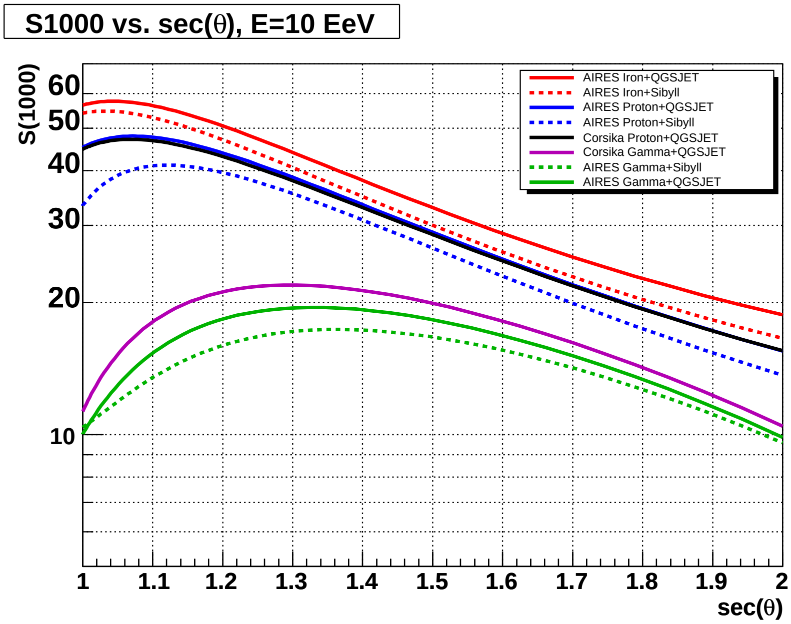


Figure 7.25: Average energy deposited in units of VEM for 10 EeV showers as a function of sec(θ).

bin. In the first bin there are 32 events, 16 protons and 16 photons, the second bin has 16 events, 8 protons and 8 photons, and the final bin has 8 events, 4 protons and 4 photons. In performing the analysis, each event is reconstructed assuming a baryonic S(1000) to energy converter because the composition is not known *a priori*. The photon S(1000) to energy converter gives energies that are half the baryonic energies (for the sake of simplicity) and that leads to the first energy bin containing the original 16 protons and the 8 photons from the higher energy bin. Thus, the reconstructed photon fraction would be 33%, lower than the true fraction of 50%.

It is fairly straight forward, however, to calculate an upper bound on the flux of photons. In each energy bin, an upper bound can be placed on the number of photons in that bin and the flux limit can be calculated directly knowing the aperture of the detector. This flux limit can then be compared to predicted photon fluxes based on the AGASA and HiRes spectrums, as presented in Chapter 2.

The method used to calculate an upper bound on the number of photon primaries in a given energy range is based on the distribution of an observable, call it X, from monte carlo predictions. A cut is determined from the distribution of X such that any event falling above (or below) that cut is deemed to be a photon candidate. This does not take into account any "background" from baryonic primaries, thus the derived limit is a conservative one. The background is difficult to calculate because the monte carlo simulations would have to reliably predict the tails of distributions of observables initiated by baryonic primaries. At this time, there is not enough confidence in the monte carlo simulations to accurately predict the tails of the distributions.

All events above the cut, then, are considered photons and from the simulated distribution the expected number of photons can be calculated for the entire data set. The fraction f of monte carlo photon events above the cut and the number of real events above the cut N are used to figure out the total number of photon candidates in the data set via the equation $N_\gamma = N/f$ [94].

The number of photon candidates N_γ will fluctuate from experiment to experiment, even under identical conditions. This fluctuation is expected to follow a Poisson distribution:

$$P(N_\gamma|\mu) = \frac{\mu^{N_\gamma} e^{-\mu}}{N_\gamma!} \tag{7.12}$$

where μ is the signal mean. The upper limit is determined with a specified confidence level (α) by requiring that a fraction $(1-\alpha)$ of identical experiments would measure a number that is less than or equal to the observed number of photon candidates:

$$1 - \alpha = \sum_{k=0}^{N_\gamma} \frac{\mu^k e^{-\mu}}{k!} \tag{7.13}$$

Solving this equation for μ gives the upper limit with confidence level (α). This upper limit and associated confidence level is consistent within the definition set by the analysis. However, due to the unknown background as explained above, the actual upper limit is conservative.

This analysis must also take into consideration the efficiency of the cuts required for photon events. Using the simulated photon showers, the efficiency of the trigger and the reconstruction is calculated by simply counting the number of simulated events that trigger the surface detector and that the reconstruction succeeds. A successful reconstruction is one where the resulting event passes the cuts required in the analysis, such as the rise time and curvature are reconstructed to physical values. A plot of the efficiency versus zenith angle for seven different primary energies is shown in Fig. 7.26. Considering the efficiency of the detector and reconstruction with respect to photon primaries, the analysis was restricted to events with a zenith angle greater than 30 degrees and a reconstructed primary

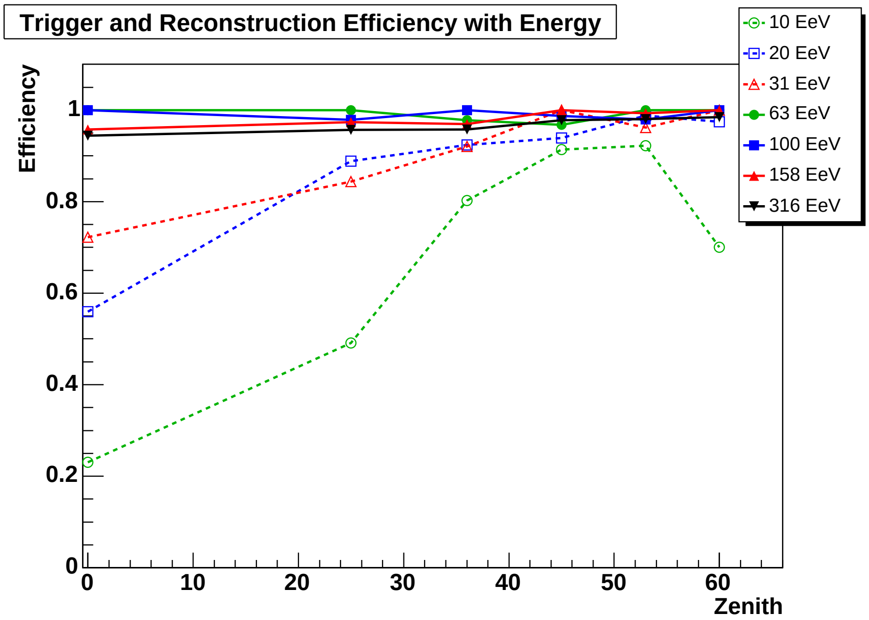


Figure 7.26: Triggering and reconstruction efficiency of photon events versus zenith angle for seven different primary energies using Aires. The triggering and reconstruction is fully efficient above 60 EeV for all zenith angles.

energy (using the photon S(1000) to energy converter) greater than 20 EeV, both in real data and monte carlo data. This restricts the analysis to a region where the efficiency is greater than 90%.

Restricting the analysis to events with a zenith greater than 30° and less than 60° confines the search for photons to a certain region in the galaxy. The resulting coverage is shown in Fig. 7.27. The region of the galaxy that is visible as defined by the zenith cuts includes both the galactic center and the southern galactic pole. That means the coverage is sufficient to detect photons coming from the galactic center or from the galactic halo.

To illustrate the analysis technique presented above, monte carlo proton showers were mixed with monte carlo photon showers. There were 304 proton events

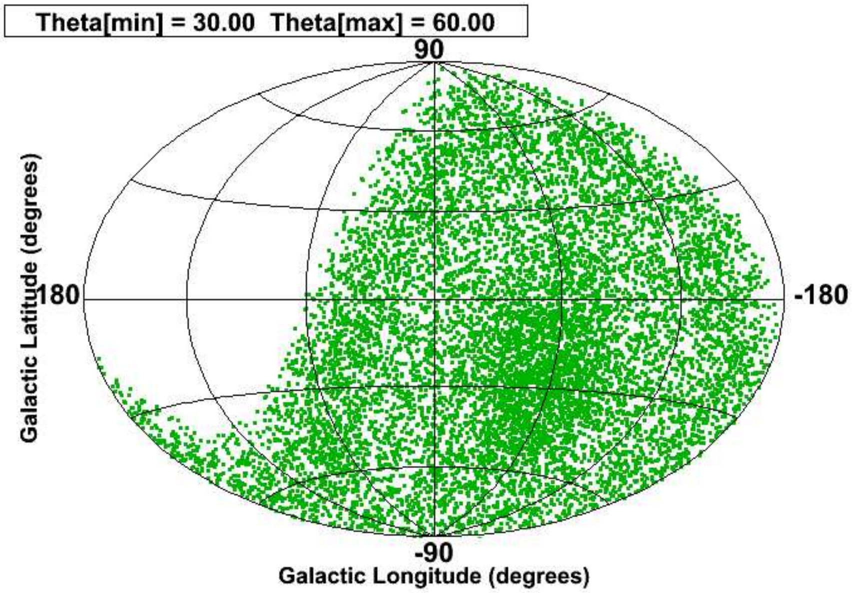


Figure 7.27: Resulting galactic coverage when restricting the events to zeniths between 30 and 60 degrees. The galactic center is visible as well as the southern galactic pole.

and 16 photon events (roughly a 5% photon fraction). In figure 7.28, the top two plots are the rise time and curvature with the cut in the discriminant variable being 0. For both rise time and curvature, there were 9 events that fell above (or below for curvature) the cut. The predicted fraction, f, for rise time was 0.45 and for curvature was 0.46, which leads to $N_\gamma = 20$. Setting the upper limit as prescribed above, the upper limit on the number of photons in this monte carlo data sample with a 90% (95%) confidence level is 26.6 (28.8). In the bottom plot in figure 7.28, the same method is applied but using both the rise time and the curvature and setting the cut at anything falling below the line $x = y$. In this case, 9 events fall below the cut and the monte carlo fraction f is 0.5, thus N_γ is 18. The same upper limits would then be 24.3 and 26.4 at 90% and 95% confidence levels respectively. The upper limits are conservative as pointed out before, but the limits improve when combining the information from two variables.

7.6.4 Upper Limit Results

The real data was analyzed as described in the preceding section, comparing the distribution of rise time and curvature to the monte carlo predictions of Aires simulations with photons as the primaries and QGSJET01 and Sibyll2.1 as the hadronic models. The analysis was restricted to zeniths greater than 30 degrees and energies (according to photon S(1000) to energy converter) greater than 20 EeV to restrict the analysis to a region with efficiency greater than 90%.

To compare the results with the photon flux predictions from Chapter 2, the analysis was divided into energy bins. Each bin has a width of 0.2 in $\log E$, starting at $\log E = 19.3$, or 20 EeV. In each energy bin, the number of photon candidates was determined from the rise time versus curvature plot, where a photon candidate was any event that fell below $x = y$, i.e. any event whose x

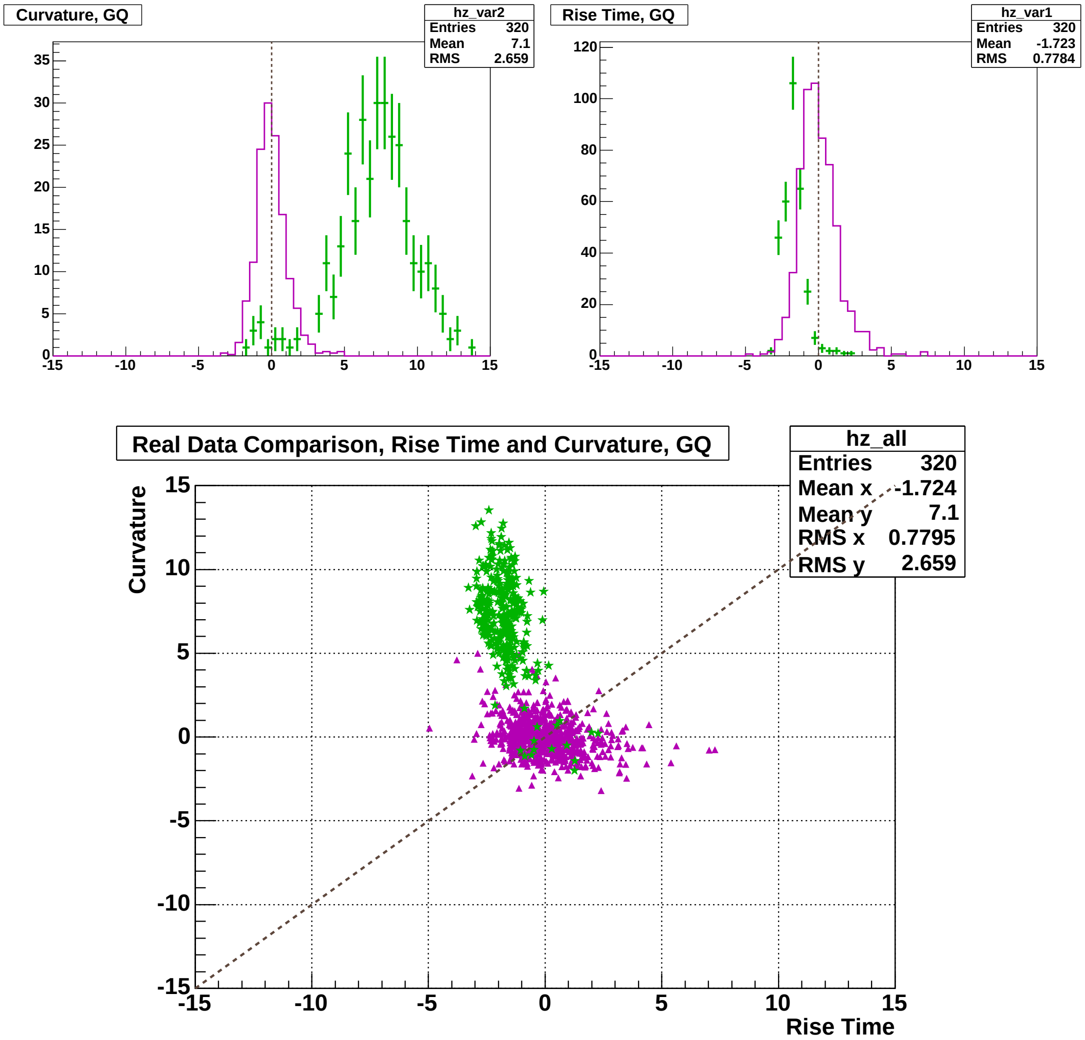


Figure 7.28: Top plots: Distributions of photon simulated events (solid lines) and proton mixed with ~5% photons (green crosses) of the deviation from the predicted mean curvature (left) and rise time (right) in units of parameterized standard deviation. Bottom plot: 2-dimensional plot of rise time and curvature with purple triangles being the photon monte carlo and the green stars the proton+photon mixed sample. Dotted lines are the cuts to determine photon candidates.

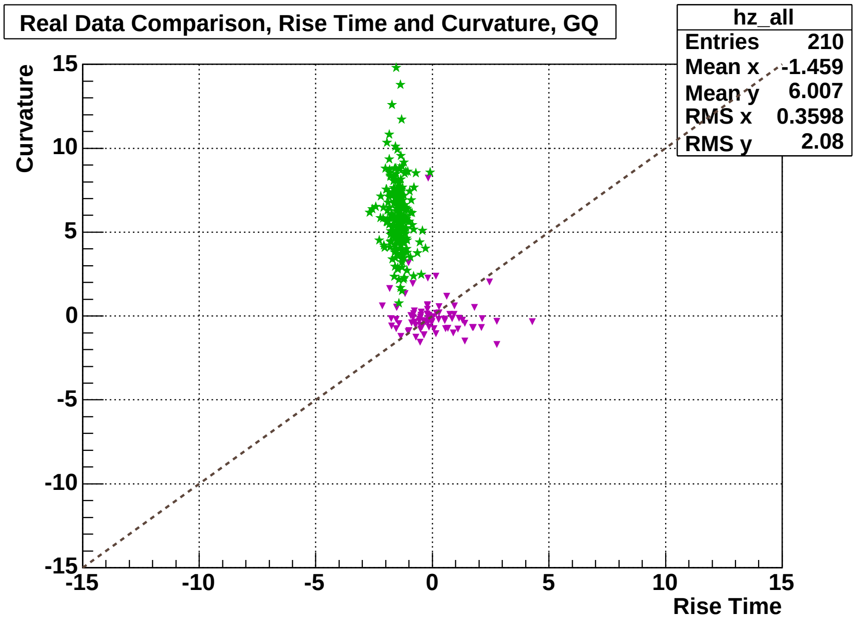


Figure 7.29: A plot of the monte carlo predictions for photon events compared to real data in the energy range 19-31 EeV and zenith range 30-60°. The cut for possible photon candidates is the dotted line at $x = y$. There are 0 events that pass this cut.

value is greater than its y value. For the lowest energy bin (20-31 EeV) a plot of the results is shown in Fig. 7.29, where there are 0 events that fall below the cut.

To calculate the upper limit on the photon flux, the aperture must be considered. From [85] the integrated aperture for the Pierre Auger Observatory from January 1 2004 to June 5 2005 was 1750 km^2 sr yr. That aperture includes from 0 to 60 degrees in zenith while the analysis here is only from 30 to 60 degrees, giving an integrated aperture for the same period 1170 km^2 sr yr.

To calculate the differential flux: $dF/dE = kE^{-\alpha}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{eV}^{-1}$, the exponent α must be assumed. From the predictions of top down models, a reasonable assumption is α=2. So, a specific bin centered on the energy 10^c eV will have N events, where N is given by:

$$N = kA \int_{10^{c-0.1}}^{10^{c+0.1}} E^{-\alpha} dE \tag{7.14}$$

Again, the size of the energy bins are 0.2 in $\log E$. Knowing the number of events in a given bin, N, and the aperture, A, it is possible then to calculate the differential flux in the bin centered at 10^c eV:

$$dF/dE = \frac{N}{A} \frac{\alpha - 1}{10^{0.1(\alpha-1)} - 10^{-0.1(\alpha-1)}} 10^{-c} \tag{7.15}$$

For α=2, the middle term is equal to 2.15 (for α=3, it is 2.1). The results are in Table 7.4, which shows the energy bin, the number of photon candidates, and the upper limit on that number at 90% and 95% confidence levels.

To compare the photon flux limits with the observed cosmic ray energy spectrum from the PAO, both are plotted in Fig. 7.30. Figure 7.30 plots the observed spectrum when using the constant intensity cut method normalized to the monte carlo prediction for protons with QGSJET01 hadronic interaction model (black squares) and the constant intensity cut normalized to fluorescence data (blue squares). These two spectra illustrate the difficulty in calculating the photon fraction as it is dependent on the energy conversion used. The photon flux limit presented here is independent of the energy conversion and is therefore very robust.

7.7 Conclusions

In this chapter, a study of the composition of ultra high energy cosmic rays was presented. Monte carlo predictions based on different primary particle and hadronic model assumptions was presented. Comparing the monte carlo pre-

$log(E_c/\mathrm{eV})$	N_γ	f	90% CL	95% CL	Flux (90%)	Flux (95%)
Aires + Sibyll						
19.4	0	0.49	2.3	3	5.33×10^{-40}	6.95×10^{-40}
19.6	0	0.48	2.3	3	3.36×10^{-40}	4.39×10^{-40}
19.8	1	0.42	5.3	6.3	4.89×10^{-40}	5.81×10^{-40}
20.0	0	0.42	2.3	3	1.34×10^{-40}	1.75×10^{-40}
20.2	0	0.51	2.3	3	8.45×10^{-41}	1.10×10^{-40}
20.4	0	0.55	2.3	3	5.33×10^{-41}	6.95×10^{-41}
Aires + QGSJET						
19.4	0	0.52	2.3	3	5.33×10^{-40}	6.95×10^{-40}
19.6	0	0.42	2.3	3	3.36×10^{-40}	4.39×10^{-40}
19.8	1	0.46	5.3	6.3	4.89×10^{-40}	5.81×10^{-40}
20.0	0	0.48	2.3	3	1.34×10^{-40}	1.75×10^{-40}
20.2	0	0.45	2.3	3	8.45×10^{-41}	1.10×10^{-40}
20.4	0	0.49	2.3	3	5.33×10^{-41}	6.95×10^{-41}

Table 7.4: Number of photon candidates in the different energy bins and the resulting upper limits on the number of photons in each bin where f is the fraction of monte carlo events that pass the cut as explained in the text and E_c is the center of the log energy bin. The upper limit at a 90% and 95% CL on the photon flux is shown in the last two columns in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{eV}^{-1}$.

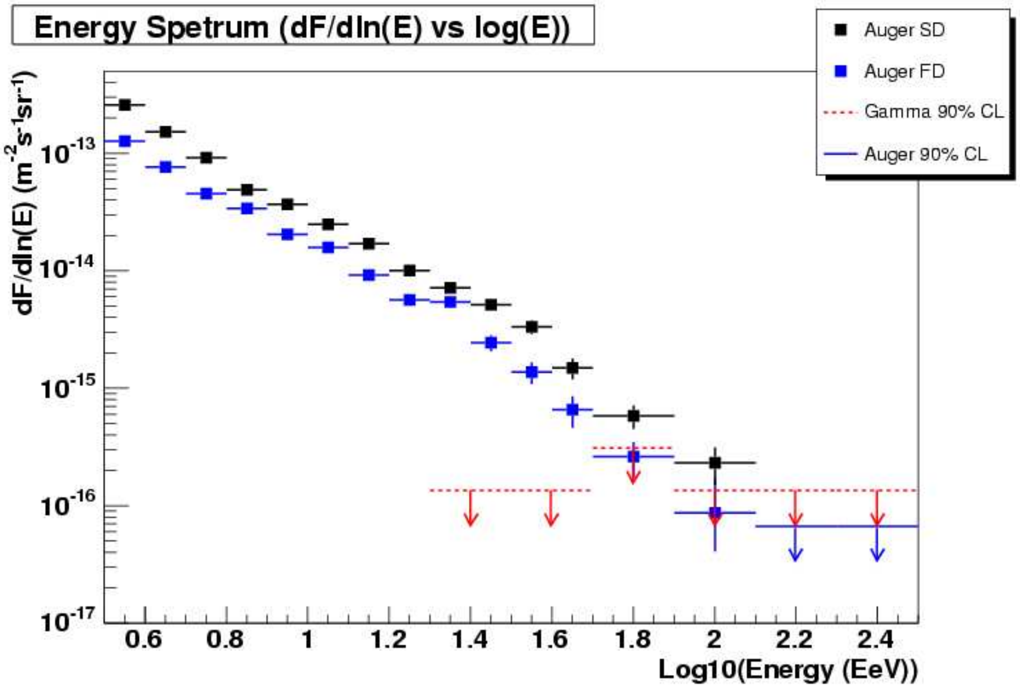


Figure 7.30: A plot of the cosmic ray spectrum using two different energy estimators. One energy estimator is based on monte carlo simulations (black squares labeled Auger SD) while the other energy is derived from hybrid data (blue squares labeled Auger FD). The energies based on monte carlo simulations are higher than the hybrid based ones.

dictions for the ground array observables rise time and curvature, the real data behaved like a heavy nucleus, similar to iron.

The difference in the characteristic values of the rise time and curvature observables between baryonic primaries and photons led to a study of photons as a possible source of the highest energy cosmic rays. Again, monte carlo predictions were compared to real data to set an upper limit on the number of photons seen in the data. Only one candidate event passed the defined cut, and that candidate event is shown in Fig. 7.31. The one "candidate" event is not truly considered a candidate as a possible photon. The prescribed cut was too relaxed and deserves a more in depth study. A more intelligent cut would not accept such events as possible photon candidates and could be more efficient in accepting the monte carlo photon events, i.e. accepting more than 50% while still rejecting possible background. This one event is due to a poor reconstruction that still passed the quality cuts imposed on the data. A more detailed study will be made in tuning this cut to maximize the signal to noise ratio and calculating the sensitivity limit to photons.

It is apparent from the photon study that the PAO is a virtually background free detector of photons. With time more events will come and the aperture will grow meaning the limit will continue to decrease or that photons will eventually be seen. If and when they are seen, the method described in this chapter will be very efficient in discovering any possible photons in the data.

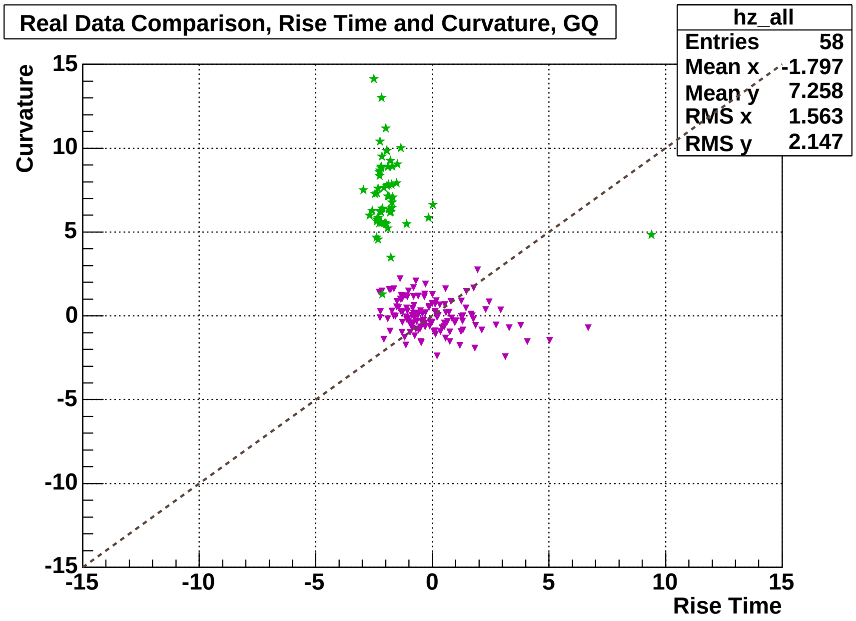


Figure 7.31: A plot of the 1 event that passed the photon candidate cut. It is obvious that this event is not photon-like, and a more intelligent cut would have rejected this event. This one event is the result of a poor event reconstruction.

CHAPTER 8

Conclusions and Discussion

The Pierre Auger Observatory is being built to investigate the mystery of the origin and composition of the highest energy cosmic rays. Many different models of the origin of ultra high energy cosmic rays have been presented over the years, with the most exotic models invoking physics beyond the standard model. Some of these exotic models predict a certain flux of high energy photons at the earth, to which the Pierre Auger Observatory is sensitive.

The array of water Cherenkov detectors on the ground at the Pierre Auger Observatory are sensitive to the muon component of extensive air showers. Because of this sensitivity, and because the muon component of extensive air showers depends on the composition and hadronic models, data from the surface detector can be used to compare monte carlo predictions of shower properties. Indeed, when comparing real data to monte carlo predictions, the behavior of the real data was most closely modelled by a heavy nucleus assumption.

Comparisons of this type were presented in this thesis, focusing on the difference between real data and a primary photon assumption, as predicted by exotic physics models. From this study, an upper limit on the flux of photons was calculated. These results can then be compared to the predicted flux of photons by various top down models fitted to the AGASA spectrum and the HiRes spectrum [51]. In Fig. 8.1, the upper limit (90% CL) on the photon flux (multiplied by E) is shown with three top down model predictions fit to the AGASA

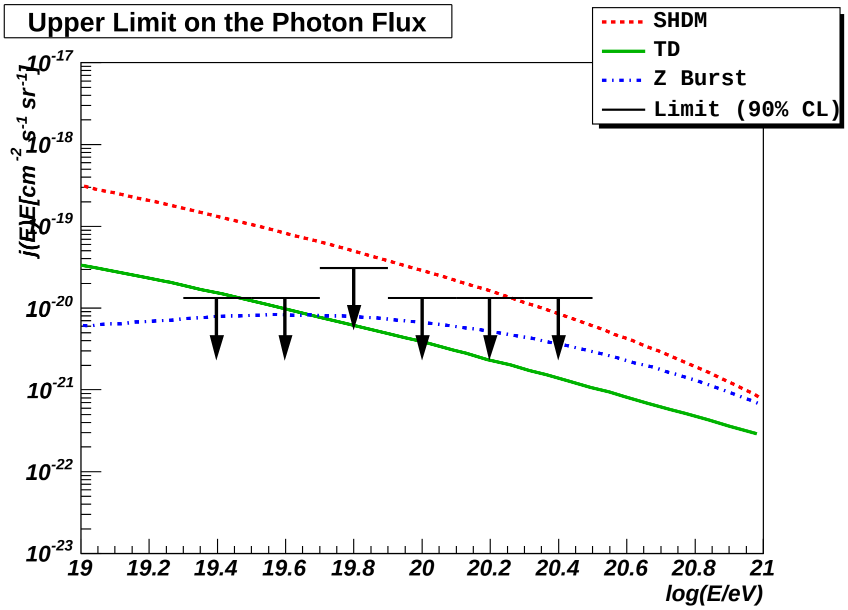


Figure 8.1: Upper limit (90% CL) on the flux of photons compared to top down model predictions based on the AGASA spectrum.

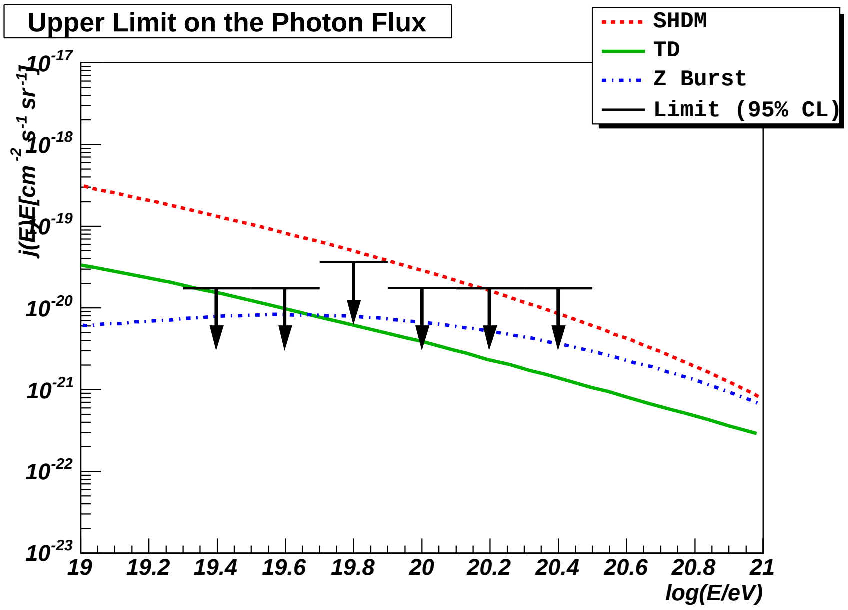


Figure 8.2: Upper limit (95% CL) on the flux of photons compared to top down model predictions based on the AGASA spectrum.

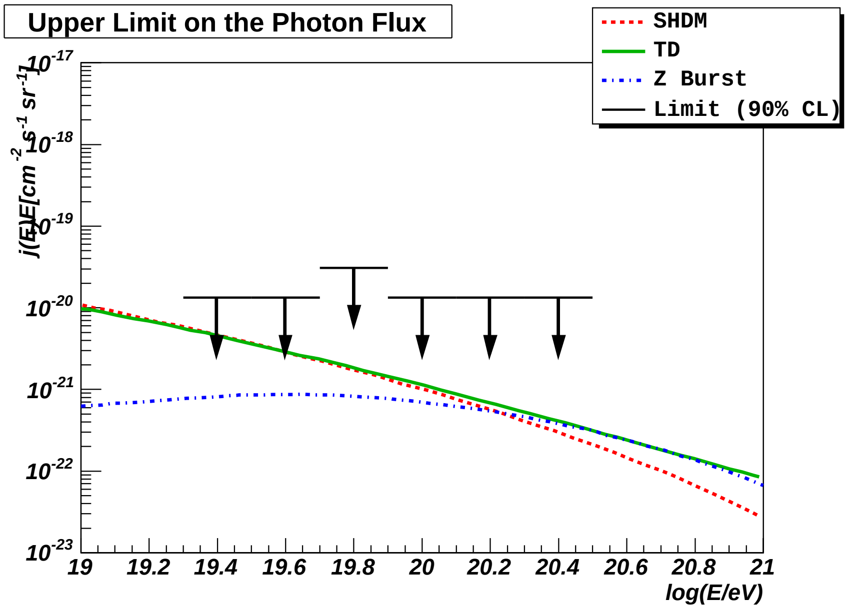


Figure 8.3: Upper limit (90% CL) on the flux of photons compared to top down model predictions based on the HiRes spectrum.

spectrum. In Fig. 8.2 the same plot is shown but with the 95% confidence level limits. Fig. 8.3 is a plot of the 90% CL limits along with the top down model predictions fit to the HiRes data. The photon flux limit set in this paper would rule out super heavy dark matter as the source of ultra high energy cosmic rays to explain the AGASA data according to the predictions made by Gelmini *et al.*, and puts limits on topological defects as well as Z-bursts.

As a final comparison, the photon flux limit is compared with the cosmic ray fluxes from HiRes, AGASA, and PAO. This is shown in Fig. 8.4 where it can be seen that the AGASA spectrum is inconsistent with the PAO spectrum as well as the HiRes spectrum. Also, the highest energy AGASA points can not be explained by assuming that they are photons to attempt to reconcile the difference in the spectrum with the reported spectrum from the PAO.

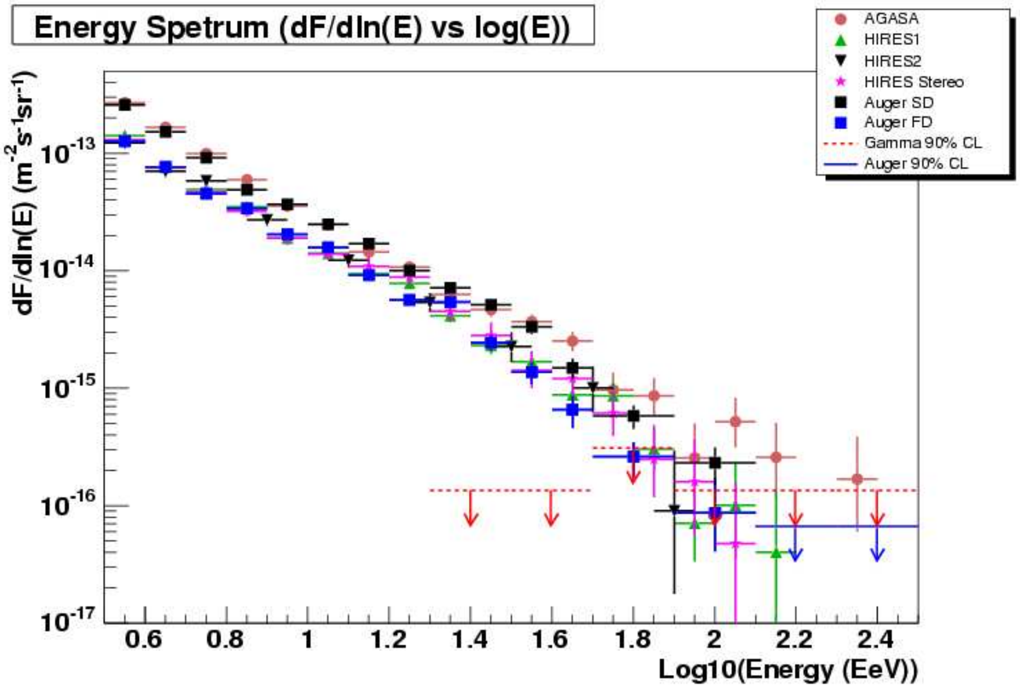


Figure 8.4: Upper limit (90% CL) on the photon flux compared with energy spectra from HiRes and AGASA along with the PAO spectrum based on energies from monte carlo estimators (Auger SD) and hybrid data estimators (Auger FD).

These results come with the caveats that these top down models have many parameters that may be adjusted to lower the predicted photon flux below the set limits. Also, the photon air cross section used in this paper for the photon simulations is the one from the Particle Data Group which is an extrapolation of data. Other extrapolations may lead to photons producing more mouns in extensive air showers, and thus behaving more like baryonic showers reducing the limit set. Model uncertainties may play a role as well, but the results here were independent of the hadronic model used, i.e. QGSJET01 or Sibyll2.1.

However, considering all these caveats, the PAO is a very good ultra high energy photon detector. It is virtually background free when restricting the analysis to events with zenith angles greater than 30° and less than 60°. If the flux of photons is high enough, the PAO will eventually detect them and the identification of photons will be clear.

References

[1] V. F. Hess, Observations in Low Level Radiation During Seven Free Balloon Flights, Phys. Z. **13**, 1084 (1912).

[2] C. D. Anderson, Early Work on the Positron and Muon, Am. J. Phys. **29**, 825 (1961).

[3] D. Griffiths, *Introduction to Elementary Particles*, John Wiley and Sons, Inc., 1987.

[4] C. M. G. Lattes, Process Involving Charged Mesons, Nature **159**, 694 (1947).

[5] G. D. Rochester and C. C. Butler, Evidence for the Existence of New Unstable Elementary Particles, Nature **160**, 855 (1947).

[6] P. Auger, Extensive Cosmic-Ray Showers, Reviews of Modern Physics **11**, 288 (1939).

[7] P. Bassi, G. Clark, and B. Rossi, Distribution of Arrival Times of Air Shower Particles, Phys. Rev. **92**, 441 (1953).

[8] J. Linsley, Evidence for a Primary Cosmic-Ray Particle with Energy 10^{20} eV, Phys. Rev. Lett. **10**, 146 (1963).

[9] M. Nagano and A. A. Watson, Observations and Implications of the Ultrahigh-Energy Cosmic Rays, Reviews of Modern Physics **72**, 689 (2000).

[10] T. Hara, S. Kawaguchi, S. Mikamo, M. Nagano, K. Suga, G. Tanahashi, K. Uchino, and H. Akiyama, Acta Phys. Acad. Sci. Hung. **29**, 361 (1970).

[11] H. E. Bergeson, G. L. Cassiday, T. W. Chiu, D. A. Cooper, J. W. Elbert, E. C. Loh, D. Steck, W. J. West, J. Boone, and J. Linsley, Measurement of Light Emission from Remote Cosmic-Ray Air Showers, Phys. Rev. Lett. **39**, 847 (1977).

[12] P. Sokolsky, *Introduction to Ultrahigh Energy Cosmic Ray Physics*, Addison–Wesley, 1984.

[13] J. Cronin, T. Gaisser, and S. Swordy, Cosmic Rays at the Energy Fontier, Sci. Amer. **276**, 44 (1997).

[14] M. S. Longair, *High Energy Astrophysics*, Cambridge Univ. Press, 1981.

[15] The High Resolution Fly's Eye Collaboration, A Study of the Composition of Ultra High Energy Cosmic Rays Using the High Resolution Fly's Eye, Astrophys. J. **622**, 910 (2005).

[16] E. Fermi, On the Origin of the Cosmic Radiation, Physical Review **75**, 1169 (1949).

[17] L. O. Drury, Acceleration of Cosmic-Rays, Contemp. Phys. **35**, 232 (1994).

[18] F. W. Stecker and M. H. Salamon, Photodisintegration of Ultra-High-Energy Cosmic Rays: A New Determination, Astrophys. J. **512**, 521 (1999).

[19] G. Bertone, C. Isola, M. Lemoine, and G. Sigl, Ultrahigh energy heavy nuclei propagation in extragalactic magnetic fields, Phys Rev D **66**, 103003 (2002).

[20] K. Greisen, End to the Cosmic-Ray Spectrum?, Phys. Rev. Lett. **16**, 748 (1966).

[21] Z. T. Zatsepin and V. A. Kuz'min, Upper Limit of Spectrum of Cosmic Rays, JETP Lett. **4**, 78 (1966).

[22] L. Bergström and A. Goober, *Cosmology and Particle Astrophysics*, John Wiley and Sons, Inc., 1999.

[23] J. W. Cronin, The Highest-Energy Cosmic Rays, Nucl. Phys. Proc. Suppl. **138**, 465 (2005).

[24] J. Wdowczyk, A. W. Wolfendale, and W. Tkaczyk, Primary Cosmic Gamma-Rays Above 10^{12} eV, J Phys A **5**, 1419 (1972).

[25] F. Halzen, R. A. Vázquez, T. Stanev, and H. P. Vankov, The Highest Energy Cosmic Ray, Astropart. Phys. **3**, 151 (1995).

[26] D. J. Bird *et al.*, Evidence for correlated changes in the spectrum and composition of cosmic rays at extremely high energies, Phys. Rev. Lett. **71** (1993).

[27] N. Hayashida *et al.*, Observation of a Very Energetic Cosmic Ray Well Beyond the Predicted 2.7 K Cutoff in the Primary Energy Spectrum, Phys. Rev. Lett. **73**, 3491 (1994).

[28] M. A. Lawrence, R. J. O. Reid, and A. A. Watson, The cosmic ray energy spectrum above 4×10^{17} eV as measured by the Haverah Park array, J. Phys. **G17**, 733 (1991).

[29] M. Takeda *et al.*, Energy determination in the Akeno Giant Air Shower Array experiment, Astropart. Phys. **19**, 447 (2003).

[30] T. Weiler, Resonant Absorption of Cosmic-Ray Neutrinos by the Relic-Neutrino Background, Phys. Rev. Lett. **49**, 234 (1982).

[31] D. Fargion, B. Mele, and A. Salis, Ultra-High-Energy Neutrino Scattering Onto Relic Light Neutrinos in the Galactic Halo as a Possible Source of the Highest Energy Cosmic Rays, Astrophys. J. **517**, 725 (1999).

[32] Particle Data Group, Review of Particle Physics, Phys. Rev. D **54**, 187–188 (1996).

[33] T. J. Weiler, Cosmic-ray neutrino annihilation on relic neutrinos revisited: a mechanism for generating air showers above the Greisen-Zatsepin-Kuzmin cutoff, Astropart. Phys. **11**, 303 (1999).

[34] D. Fargion, M. Y. Khlopov, R. V. Konoplich, and R. Mignani, Bounds on Very Heavy Relic Neutrinos by their Annihilation in the Galactic Halo, Phys. Rev. D **52**, 1828 (1998).

[35] F. W. Stecker, *Cosmic Gamma Rays*, Mono Book Corp., 1971.

[36] F. W. Stecker, Diffuse fluxes of cosmic high-energy neutrinos, Astrophys. J. **228**, 919 (1979).

[37] P. Bhattacharjee and G. Sigl, Origin and Propagation of Extremely High-Energy Cosmic Rays, Physics Reports **327**, 109 (2000).

[38] M. Hindmarsh and T. W. B. Kibble, Cosmic strings, Rep. Prog. Phys. **58**, 477 (1995).

[39] C. T. Hill, Monopolonium, Nucl. Phys. B **224**, 469 (1983).

[40] D. N. Schramm and C. T. Hill, in *Proceedings of the 18th ICRC*, volume 2, page 393, 1983.

[41] M. B. Hindmarsh and T. W. B. Kibble, Monopoles on Strings, Phys. Rev. Lett. **55**, 2398 (1985).

[42] V. Berezinsky and A. Vilenkin, Cosmic Necklaces and Ultrahigh Energy Cosmic Rays, Phys. Rev. Lett. **79**, 5202 (1997).

[43] R. J. Protheroe and T. Stanev, Limits on Models of the Ultrahigh Energy Cosmic Rays Based on Topological Defects, Phys. Rev. Lett. **77**, 3708 (1996).

[44] T. Gherghetta, Goldstino decoupling in spontaneously broken supergravity theories, Nucl. Phys. **B485**, 25 (1997).

[45] M. Birkel and S. Sarkar, Extremely high energy cosmic rays from relic particle decays, Astropart. Phys. **9**, 297 (1998).

[46] V. Berezinsky, M. Kachelrieß, and A. Vilenkin, Ultra-high energy cosmic rays without GZK cutoff, Phys. Rev. Lett. **79**, 4302 (1997).

[47] R. Kallosh, A. Linde, D. Linde, and L. Susskind, Gravity and global symmetries, Phys. Rev. D **52**, 912 (1995).

[48] V. A. Kuzmin and V. A. Rubakov, Ultra-High Energy Cosmic Rays: a Window to Post-Inflationary Reheating Epoch of the Universe?, Phys. Atom. Nucl. **61**, 1028 (1998).

[49] P. Blasi, R. Dick, and E. W. Kolb, Ultra-high energy cosmic rays from annihilation of superheavy dark matter, Astropart. Phys. **18**, 57 (2002).

[50] E. W. Kolb, D. J. H. Chung, and A. Riotto, *Dark Matter in Astrophysics and Particle Physics 1998*, IoP Publishing, 1999.

[51] G. Gelmini, O. Kalashev, and D. V. Semikoz, GKZ Photons as Ultra High Energy Cosmic Rays, arXiv:astro-ph/0506128.

[52] The Auger Collaboration, *Pierre Auger Project Design Report*, second edition, November 1996, Revised March 14, 1997.

[53] A. Cordero *et al.*, Proposal for the Optical System of the Fluorescence Detectors of the Auger Project, Internal GAP Note 1996-039, 1996.

[54] J. A. Bellido for the Auger Collaboration, Performance of the Fluorescence Detectors of the Pierre Auger Observatory, in *Proceedings of the 29th ICRC*, pages 101–106, 2005, aus-bellido-J-abs1-he14-oral.

[55] J. Brack *et al.* for the Auger Collaboration, Absolute Calibration of the Auger Fluorescence Detectors, in *Proceedings of the 29th ICRC*, 2005, arg-rovero-AC-abs1-he15-poster.

[56] C. Aramo *et al.* for the Auger Collaboration, Optical Relative Calibration and Stability Monitoring for the Auger Fluorescence Detector, in *Proceedings of the 29th ICRC*, 2005, ita-insolia-A-abs1-he15-poster.

[57] M. Roberts *et al.* for the Auger Collaboration, Atmospheric Aerosol Monitoring at the Pierre Auger Observatory, in *Proceedings of the 29th ICRC*, 2005, usa-roberts-M-abs1-he15-poster.

[58] V. Rizi, M. Iarlori, S. Petrera, R. Caruso, and A. Grillo, The Use of Raman LIDAR for Atmospheric Monitoring at Pierre Auger Observatory, Internal GAP Note 2002-004, 2002.

[59] B. Keilhauer *et al.* for the Auger Collaboration, Atmospheric Profiles at the Southern Pierre Auger Observatory and their Relevance to Air Shower Measurement, in *Proceedings of the 29th ICRC*, 2005, ger-keilhauer-B-abs2-he14-poster.

[60] M. Malek *et al.* for the Auger Collaboration, The Central Laser Facility at the Pierre Auger Observatory, in *Proceedings of the 29th ICRC*, 2005, usa-malek-M-abs1-he15-poster.

[61] The Auger Collaboration, *The Pierre Auger Project Technical Design Report*, March 2004.

[62] X. Bertou for the Auger Collaboration, Performance of the Pierre Auger Observatory Surface Array, in *Proceedings of the 29th ICRC*, 2005, arg-bertou-X-abs1-he14-oral.

[63] P. Allison *et al.* for the Auger Collaboration, Calibration of the Surface Array of the Pierre Auger Observatory, in *Proceedings of the 29th ICRC*, 2005, usa-allison-PS-abs1-he14-poster.

[64] M. Aglietta *et al.* for the Auger Collaboration, Response of the Pierre Auger Observatory Water Cherenkov Detectors to Muons, in *Proceedings of the 29th ICRC*, 2005, fra-suomivi-T-abs1-he14-poster.

[65] A. Tripathi, S. Akhanjee, K. Arisaka, D. Barnhill, C. D'Pasquale, C. Jillings, T. Ohnuki, and P. Ranin, A systematic study of large PMTs for the Pierre Auger Observatory, Nucl. Instr. and Meth. A **497**, 331 (2003).

[66] A. Tripathi, K. Arisaka, D. Barnhill, C. Howard, C. Jillings, T. Ohnuki, and P. Ranin, Tests of New Extended Dynamic Range PMTs for the Auger Project, Internal GAP Note GAP-2001-049, August 2001.

[67] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments: A How to Approach*, Springer-Verlag, second revised edition, 1994.

[68] C. Jillings, T. Ohnuki, D. Barnhill, P. Ranin, F. Suarez, K. Arisaka, C. Howard, A. Tripathi, P. Payan, S. Pollock, M. Rogers, and T. Willse, Results and Procedures of Testing Pre-Production PMTs for the Surface Detector, Internal GAP Note GAP-2002-037, August 2002.

[69] D. Barnhill, K. Arisaka, A. Tripathi, T. Ohnuki, F. Suarez, B. Garcia, M. Videla, and A. Risi, Results of Testing Pre-Production and Production PMTs for the Surface Detector in the New PMT Test Facility in Malargüe, Internal GAP Note GAP-2003-037, April 2003.

[70] P. Allison, X. Bertou, and C. Grunfeld, Pulse shape agnostic methods of measuring the dynode-anode ratio, Internal GAP Note GAP-2004-033, June 2004.

[71] P. Allison and D. Barnhill, Calculation of the number of photoelectrons produced per tank based on PMT test data and station monitoring data, Internal GAP Note GAP-2004-046, July 2004.

[72] M. Healy, D. Barnhill, J. Lee, T. Ohnuki, A. Tripathi, and K. Arisaka, Observation of the Long Term Stability of Water Tanks in the Pierre Auger Surface Detector, Internal GAP Note GAP-2005-028, 2005.

[73] K. Arisaka, D. Barnhill, A. Tripathi, T. Ohnuki, J. Lee, W. Slater, A. Chou, and G. R. Fernandez, Determination of Absolute Energy and Composition with Minimum Systematic Uncertainties, Internal GAP Note GAP-2004-037, July 2004.

[74] S. J. Sciutto, *AIRES: A system for air shower simulations*, Departamento de Fisica Universidad Nacional de la Plata, Available from http://www.fisica.unlp.edu.ar/auger/aires/.

[75] R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev, in *Proceedings of the 26th ICRC*, volume 1, page 415, 1999.

[76] N. N. Kalmykov, S. S. Ostapchenko, and A. I. Pavlov, Quark-gluon-string model and EAS simulation problems at ultra-high energies, Nucl. Phys. B (Proc. Suppl.) **52B**, 17 (1997).

[77] D. Heck, J. Knapp, J. N. Capdeville, G. Schatz, and T. Thouw, 1998, Forschungszentrum Karlsruhe; available from http://www-ik.fzk.de/~heck/corsika/physics_description/corsika_phys.html.

[78] D. Veberic and M. Roth, SD Reconstruction, Internal Gap Note 2005-035, May 2005.

[79] S. Argiro *et al.* for the Auger Collaboration, The Offline Software Framework of the Pierre Auger Observatory, in *Proceedings of the 29th ICRC*, 2005, usa-paul-T-abs1-he15-poster.

[80] P. Ghia for the Auger Collaboration, Statistical and systematic uncertainties in the event reconstruction and S(1000) determination by the Pierre Auger surface detector, in *Proceedings of the 29th ICRC*, 2005, ita-ghia-P-abs1-he14-oral.

[81] D. Newton, J. Knapp, and A. A. Watson, The Optimum Ground Parameter, $S(r_{opt})$, Internal GAP Note GAP-2005-013, February 2005.

[82] B. D. Hieu and A. Watson, Analysis of the risetime and its uncertainty for events recorded by the surface detector in the Pierre Auger Observatory, Internal Gap Note, April 2005.

[83] A. Fassó, A. Ferrari, J. Ranft, and P. R. Sala, FLUKA: Status and Prospective of Hadronic Applications, in *Proc. Monte Carlo 2000 Conf.*, edited by A. Kling, F. Barao, M. Nakagawa, and P. Vaz, page 955, Springer (Berlin), 2001, http://www.fluka.org/heart/rh.html.

[84] P. Billoir, Reconstruction of showers with the Ground Array: status of the"prototype" program, Internal GAP Note GAP-2000-025, 2000.

[85] P. Sommers for the Pierre Auger Collaboration, First Estimate of the Primary Cosmic Ray Energy Spectrum Above 3 EeV from the Pierre Auger Observatory, in *Proceedings of the 29th ICRC*, 2005, usa-sommers-P-abs1-he14-oral.

[86] D. Allard *et al.* for the Pierre Auger Collaboration, The trigger system of the Pierre Auger surface detector: operation, efficiency, and stability, in *Proceedings of the 29th ICRC*, 2005, usa-lhenry-yvon-I-abs1-he14-poster.

[87] B. McBreen and C. J. Lambert, Interactions of high-energy ($E > 5\times10^{19}$ eV) photons in the Earth's magnetic field, Phys. Rev. **D24**, 2536 (1981).

[88] L. D. Landau and I. Y. Pomeranchuk, Dokl. Akad. Nauk SSSR **92**, 535, 735 (1953).

[89] A. B. Migdal, Bremsstrahlung and Pair Production in Condensed Media at High Energies, Phys. Rev. **103**, 1811 (1956).

[90] P. Homola and D. Góra and D. Heck and H. Klages and J. Pekala and M. Risse and B. Wilczyńska and H. Wilczyński, Simulation of Ultra-High Energy Photon Propagation in the Geomagnetic Field, arXiv:astro-ph/0311442.

[91] D. Badagnani and S. J. Sciutto, AIRES photon shower production at Lyon, Internal GAP Note GAP-2005-030, March 2005.

[92] M. Ave and J. A. Hinton and R. A. Vázquez and A. A. Watson and E. Zas, Constraints on the ultrahigh-energy photon flux using inclined showers from the Haverah Park array, Phys. Rev. D **65**, 063007 (2002).

[93] K. Shinozaki *et al.*, Upper limit on gamma-ray flux above 10^{19} eV estimated by the Akeno Giant Air Shower Array experiment, Astrophys. J. **571**, L117 (2002).

[94] C. Roucelle, J. Hamilton, and P. Billoir, More on statistical methods for photon upper limit evaluations a short overview, Private communication, April 2005.