Candidacy Proposal: Observing Ultrahigh Energy Cosmic Rays with CHICOS

Elina Brobeck

May 27, 2004

Abstract

For my thesis project, I will use the data collected by the California HIgh school Cosmic ray ObServatory (CHICOS) to investigate the properties of ultrahigh energy cosmic rays (UHECRs). This will include a measurement of the cosmic ray flux across the energy range $E \sim 10^{18}$ to $E \sim 10^{21}$ eV. I will also perform a clustering analysis on the data, with the goal of identifying astrophysical sources of UHECRs. During the period of data acquisition I will share responsibility for the operation and maintenance of the detector array.

Thesis Commitee: Ed Stone, Bob McKeown, Fiona Harrison, Tony Readhead, and Marc Kamionkowski.

1 The CHICOS Project

The California HIgh School Cosmic Ray Observatory, or CHICOS, is a collaboration between U.C. Irvine, C.S.U. Northridge, and Caltech. The Project Director is Dr. Robert McKeown. The Education Director is Dr. Royichi Seki of C.S.U. Northridge. Other members of the project at Caltech include Dr. Theresa Lynn (Project Coordinator), Dr. Chris Jillings and Brant Carlson, a Caltech Physics Junior. Financially, the project is primarily supported by an NSF grant, with hardware donations from Los Alamos National Laboratory and IBM.

1.1 Introduction

When cosmic radiation was discovered in 1911 by Victor Hess (Nobel Prize, 1936), cosmic ray sensors were only able to measure the flux at energies below 10^8 eV. In 1938, Pierre Auger observed nearly coincident hits in two widely separated sensors, and realized that there must be cosmic rays at energies too high to penetrate the atmosphere. Such high-energy cosmic rays collide with nuclei in the air, setting off a cascade of secondary particles known as an air shower. The events Auger had observed were found to have energies of $\sim 10^{15}$ eV, 10 million times higher than had previously been known [1].

The measurement of high-energy cosmic rays via sampling of extended air showers was first implemented in 1954 at the Harvard College Observatory [6]. From their work, and from the many ground-array experiments that followed it, we now know that the cosmic ray spectrum spans over 10 decades in energy, ranging from 10^8 eV to 10^{20} eV. Above 10^{20} eV, the flux is not well measured, and results from different experiments are in disagreement.

The CHICOS project is designed to observe cosmic rays with energies of about 10^{18} eV and above. These are known as 'ultrahigh energy cosmic rays', or UHECRs. Such events are exceedingly rare: the total flux of UHECRs with energies above 10^{19} eV is approximately one per square kilometer per year, and falls approximately as E^{-3} . For this reason the combined data set from previous experiments remains very small; fewer than 100 events above $10^{19.6}$ eV have been observed [18].

The CHICOS array is made up of pairs of solid-scintillator cosmic ray detectors spread throughout the San Gabriel and San Fernando valleys. Each pair of detectors is situated in a high school (or in some cases a middle or elementary school), with the detectors and a GPS antenna typically placed on the roof and a workstation in a nearby science classroom.A major advantage of using secondary schools as detector sites is that the infrastructure needed for power and data transfer is already in place, allowing for a very large array to be built with minimal cost.

The teachers who are involved with the project are encouraged to integrate it into the science curriculum. All CHICOS data is made available to teachers and students via the project web site for this purpose. The project also offers a series of week-long summer programs for students from participating schools. Other cosmic ray detector arrays have used schools as detector sites (for example, ALTA in Alberta and CROP in Nebraska), but the CHICOS array differs from these projects in its greater emphasis of science goals in addition to educational contributions.

1.2 Motivations and Goals

The motivations for the CHICOS project include educational outreach as well as research. In this proposal, however, I will be focusing specifically on the science goals.

There have been a number of similar cosmic ray detector arrays since air showers were discovered. These experiments include the Haverah Park array in England [13] , the Yakutsk array in Russia [12], the Sydney University Giant Airshower Recorder (SUGAR) in Australia [19], and the Akeno Giant Air-Shower Array (AGASA) in Japan [5]. The SUGAR, AGASA, and Haverah Park projects have made flux measurements in the ultrahigh energy regime [14]. Of these, AGASA has accumulated the largest data set, and sets a standard against which CHICOS can be compared.

The AGASA array has been in operation since 1990, and covers 100 km^2 with an array of 111 scintillation detectors [5]. As of July 2002, 72 events with energies above 4×10^{19} eV had been recorded [8]. Their data shows that the slope of the low-energy spectrum extends up to the highest observed energies. (See Figure 1.) As will be discussed in Section 2.2, previous theoretical analyses had predicted a sharp cut-off in the cosmic ray spectrum at $E \sim 10^{19.6}$ eV, and results from the HiRes air fluorescence detector are in agreement with this prediction. However, if the results of the AGASA measurements are confirmed, it could mean having to revise our understanding of the physics involved.

In addition, a sky map of the UHECR data collected by AGASA between 1990 and 2000 appears to show evidence of small-scale clustering. (See Figure 2.) Using a data set of 59 events above $E \sim 4 \times 10^{19}$ eV, one triplet and 5 doublets were found, with each of the 8 pairs having an angular separation of less than 2.5°, the angular resolution of the array. The statistical significance of the clustering is approximately 5σ, with a chance probability of 10^{-4} of arising from a random distribution [10]. As will be discussed in Section 2.3, identification of small-scale clustering may be a first step to identifying the astrophysical sources of UHECRs.

These two unexpected results, if confirmed, would have important implications for both astronomy and physics, and have spurred renewed interest in ultrahigh energy cosmic rays. A new cosmic ray detector array, AUGER, is currently under construction in Argentina, and will measure the UHECR flux in the southern hemisphere. The AUGER ground array is complemented by air fluorescence detectors, and their data will allow better calibration of other ground arrays, including CHICOS. With a 3 year run time, CHICOS will be able to observe enough events at high energies to make a useful comparison with the AGASA results of clustering and total flux. The northern-hemisphere data collected by CHICOS will form a complementary set with the southern-hemisphere data from AUGER.

1.3 The Detector Array

The CHICOS array as currently envisioned will consist of 90 sites located in the San Gabriel and San Fernando valleys. There are 67 sites already installed and operational. (See Figure 3 for a map of the current configuration.) The range of energies to which the array is sensitive is determined by the spacing of the sites, and the detection rate scales as the size of the array. The CHICOS array covers a total area of approximately 300 km^2. The San Gabriel valley currently contains 40 sites spread over an area of about 100 km^2. Included in the San Gabriel array is the set of 12 sites on or near the Caltech campus which make up the Chiquita array. The sparser San Fernando array comprises 27 sites over an area of approximately 200 km^2.

The Chiquita array consists of the detectors at Pasadena City College and the Polytechnic School, as well as 5 sites on the Caltech campus. (One of the Caltech sites is a closely spaced set of 6 individual pairs.) The Chiquita array is sensitive to lower-energy events than the larger CHICOS array; the range of sensitivity lies approximately between 10^{16} and 10^{18} eV.

Each site in the array is equipped with two scintillating detectors, a GPS antenna and receiver, and a workstation with two data acquisition cards. The software which coordinates the data collection at each site is written in Labview. It is designed to have a user-friendly interface (in order to be accessible to teachers and students) as well as to operate with minimal oversight. The workstations are designed to take data continuously, without any human intervention; data transfer and recovery from hardware glitches are done automatically. The organization and operation of the Labview software, which I rewrote last fall, is described in Section 3.1.

The scintillator detectors were donated to the CHICOS project by the Los Alamos National Laboratory, where they were previously used in the CYGNUS cosmic ray project. Some of our photomultiplier tubes were recycled from the Palo Verde Neutrino Oscillation Experiment, while others had to be purchased new. A time-over-threshold discriminator circuit, built at Caltech, measures the length of the exponential PMT pulse from each detector.

The decay constant, τ, of the photocurrent is approximately 80 ns. This varies from site to site, depending on the hardware; all sites have been calibrated and the value of tau in each detector measured.

The counter cards compare the pulse with an 80 MHz oscillator, for an internal timing accuracy of approximately 12.5 ns. Clocks at widely separated locations are kept synchronized by the GPS signal. When the GPS receiver is used set in 'stationary' mode (*i.e.*, assuming fixed location), the accuracy of the timing pulse used to resynchronize the clock is accurate to better than ±50 ns.

The first CHICOS sites were deployed in the fall of 2001, and the CHICOS array has been observing UHECR showers for over a year. At the beginning of 2004, all existing sites were upgraded. My revised version of the data collection software was installed, some hardware improvements were made, and all detectors were recalibrated. Since the upgrade, the fraction of sites reporting each day has averaged approximately 90%, with the main cause of data loss being network problems beyond our immediate control.

2 Ultrahigh Energy Cosmic Rays

The flux of cosmic rays appears to fall smoothly over all 12 orders of magnitude, decreasing as the inverse cube of the energy. There is a slight break at approximately $10^{15.5}$ eV, known as the 'knee', where the slope steepens from $E^{-2.7}$ to E^{-3}. The spectrum steepens again to $E^{-3.3}$ at $10^{17.7}$ eV, then flattens slightly to $E^{-2.7}$ at the 'ankle', around 10^{19} eV [17]. Despite the relative uniformity of the spectrum, cosmic rays are believed to come from a diversity of sources, ranging from solar to galactic to extra-galactic. In the ultra-high energy range around the 'ankle' and above, it is believed that extra-galactic particles dominate the flux for reasons discussed in Section 2.3, although the specific sources are unknown.

2.1 Cosmic Ray Airshowers

When an ultrahigh energy cosmic ray enters the atmosphere and precipitates an air shower, much of the information describing the incident particle is lost. Properties of interest include the species, energy and incident angle of the primary particle. In order to extract this information from the ground data, we first need to have a firm theoretical understanding of how air showers develop.

The precise evolution of a cosmic ray air shower can be modeled with codes such as AIRES (AIRshower Extended Simulations) and CORSIKA (COsmic Ray SImulations for KAscade). The CHICOS project is currently using AIRES, which is freely available from the Universidad Nacional de La Plata, Argentina. The simulation code in turn depends on specific models of hadronic interactions; AIRES uses the SIBYLL and QGSJET models. These simulations are extremely computationally intensive; a preliminary series of simulations at the energies observable by CHICOS are being run on computers at Caltech and PCC to model the particle distributions in air showers at our altitude. This set of simulations covers both proton and iron nucleus primary particles, with zenith angle $\cos\theta$ ranging from 0.65 to 0.95.

An air shower front is characterized by its lateral distribution function (LDF), which describes the intensity of particles $\rho(r)$ as a function of distance from the shower core. The

CHICOS reconstruction software currently uses the LDF obtained empirically by the AGASA experiment as a first approximation to the LDF at our altitude. As our simulations progress, they will be used to calculate the LDF appropriate for CHICOS.

The AGASA LDF is given by:

$$\rho(r) = C\left(\frac{r}{R_M}\right)^{-\alpha}\left(1+\frac{r}{R_M}\right)^{-(\eta-\alpha)}\left[1+\left(\frac{r}{1000}\right)^2\right]^{\delta},$$

where r is the distance in meters from the core of the shower, and C is a proportionality constant related to the energy of the primary particle. The parameters α and δ are found to be 1.2 and 0.6, respectively [20]. The Molière unit, R_M, characterizing the scattering length[1], is equal to 91.6 m at the altitude of AGASA, and 85 m at the altitude of CHICOS.

The parameter η depends on the incident angle θ, measured from the vertical:

$$\eta = (3.97 \pm 0.13) - (1.79 \pm 0.62)(\sec\theta - 1),$$

for incident angles $\theta \leq 45^\circ$. No energy dependence of η has been observed, so it is assumed that this formula for the LDF can be used to describe even the highest-energy showers [16].

The measured intensity $S(r)$ is a function of the LDF and the detector response. For scintillating detectors, the signal is determined by the average energy loss in the scintillator of electrons, photons, and muons. This function can be expressed in units of the energy loss of vertically penetrating muons, C_e, a convenient measure because they determine the peak of the spectrum of single-particle events. Thus, the measured intensity of a vertical shower is given by:

$$S_0(r) = N_e C_e\left(\frac{r}{R_M}\right)^{-\alpha}\left(1+\frac{r}{R_M}\right)^{-(\eta-\alpha)}\left[1.0+\left(\frac{r}{2000}\right)^2\right]^{\delta}$$

This function has been shown to be valid between 500 m and 3 km from the core of the shower, at energies up to 10^{20} eV [14].

AGASA finds that the energy of the incident cosmic ray is related to $S_0(600)$, the measured intensity at a distance of 600 meters from the core, by the following formula:

$$E_0 = 2.03 \times 10^{17}\ \text{eV} \ \cdot S(600)^{1.02}.$$

A shower which enters the atmosphere with an inclined trajectory passes through a greater air depth, and the shower development is correspondingly affected. To determine the energy of an air shower at incident angle θ, the measured intensity $S_\theta(600)$ must first be converted to an equivalent value of $S_0(600)$:

$$S_\theta(600) = S_0(600)\exp\left[-\frac{X_0}{\Lambda_1}(\sec\theta - 1) - \frac{X_0}{\Lambda_2}(\sec\theta - 1)^2\right].$$

Here $X_0 = 920$ g/cm^2, $\Lambda_1 = 500$ g/cm^2, and $\Lambda_2 \simeq 600$ g/cm^2. This conversion formula is valid for $\theta \leq 45^\circ$ [16].

---

[1] The Molière unit is defined by $R_M = X_R E_S / E_C$, where the radiation length X_R is the scale length for energy losses from electron bremstrahlung, the critical energy E_C is the energy at which bremstrahlung and ionization losses are equal, and the scattering energy E_S relates the mean-square scattering angle to the distance x traversed by an electron in the multiple-scattering formula $\langle\theta^2\rangle = (E_S/E_C)^2 x / X_R$ [7].

The angle of incidence of a shower is determined by fitting the relative time delays to the shape of the shower front. The curvature of the shower front determines the average time delay T_d at given distance from the core:

$$T_d(\rho, r) = 2.6\left(1 + \frac{r}{30}\right)^{1.5} \rho(r)^{-0.5} \ \text{ns},$$

where r is in meters.

The width of the shower front increases with distance from the core. The resulting time spread T_s in the hits is given by:

$$T_s(\rho, r) = 2.6\left(1 + \frac{r}{30}\right)^{1.5} \rho(r)^{-0.3} \ \text{ns}.$$

The time-delay and time-spread formulae have been modified for CHICOS by removing the $\rho(r)$ term in T_d, and replacing $\rho(r)^{-0.3}$ with $\rho(r)^{-0.5}$ (*i.e.*, pure counting statistics) in T_s. This was done because the CHICOS detectors have a much shorter time constant than AGASA; hence the CHICOS detectors can generally resolve individual particles (sufficiently far from the core of the shower), while AGASA measurements integrated multiple particles in a single pulse. The equations for T_d and T_s in their original form describe the time delay and spread of the first particle to hit the detector, while it is more appropriate for CHICOS to use the average time delay and overall spread of all incident particles.

The entire set of formulae will be recalcuated in detail for CHICOS over the next year. The current implementation these formulae in the CHICOS shower reconstruction code is discussed further in Section 3.2.

2.2 The GZK Cutoff

As UHECRs travel through space, they interact with the cosmic microwave background (CMB). There are two main types of interactions involving cosmic ray protons: pair production and photo-pion production.

Photo-pion production may proceed either as $p + \gamma \to \pi^0 + p$ or as $p + \gamma \to \pi^+ + n$. The energy threshold for photo-pion production is:

$$E^{\text{th}}_{p\gamma_{\text{CMB}}} = \frac{m_\pi(m_p + m_\pi/2)}{\langle E_{\text{CMB}}\rangle} \simeq 6.9 \times 10^{19} \left(\frac{\langle E_{\text{CMB}}\rangle}{10^{-3}\text{eV}}\right)^{-1} \ \text{eV},$$

where $\langle E_{\text{CMB}}\rangle \sim 10^{-3}$ eV represents the average energy of CMB photons.

Repeated encounters will eventually cause the energy of the proton to fall below the energy threshold for the interaction. If $y \sim 0.2$ is the average energy fraction lost by the cosmic ray per interaction, the attenuation length is:

$$(\sigma_{p\gamma} n_\gamma y)^{-1} \sim 12\,\text{Mpc},$$

where $\sigma_{p\gamma} \geq 0.1$ mb is the interaction cross-section, and $n_\gamma \simeq 410$ cm^{-3} is the number density of CMB photons [17]. A more precise calculation would have to take into account the variation of the interaction cross-section with energy, and include the effect of inelasticity (which also varies with energy) [15], but the approximate result is sufficient to show that there is a discrepancy between the theory and the observations in this energy range.

Given the limits just derived, ultrahigh energy cosmic rays would only be observable if they originate from a relatively small volume around our location. A volume 10 Mpc in radius would encompass only the Local Group of galaxies. Ultrahigh energy cosmic rays that originate farther away would be observed as an accumulation of flux just below the threshold for photo-pion production, beyond which the spectrum would drop quickly. This predicted cutoff in the cosmic ray spectrum at $E \sim 10^{19.6}$ eV is known as the GZK effect after Greisen [9], and Zatsepin and Kuz'min [21], who developed the theory independently in 1966.

The energy of cosmic ray protons can also be reduced by electron/anti-electron pair production. The pair-production energy threshold is approximately 10^{18} eV and the mean free path between events is ~ 1 Mpc (compared to ~ 6 Mpc for pion production), but the mean energy loss per encounter is only 0.1%, making it a less efficient mechanism for energy loss [14].

In the spectrum obtained by AGASA (shown in Figure 1), a handful of events were unambiguously observed above the GZK cutoff. The HiRes experiment, on the other hand, reported a non-detection in the same energy range. It is difficult to draw conclusions because of the low statistical significance, hence more observations are needed. Should the AGASA result be confirmed, we will either have to reconsider the physical interactions involved, conclude that all ultrahigh energy events originate nearby, or reconsider the identity of the primary particles.

CHICOS already has the capability to detect events in that energy range, and will become more sensitive as the array is filled in over the next year. In addition, the area covered by CHICOS is approximately 3 times the area covered by AGASA. This implies that CHICOS will be able to accumulate data at a higher rate, and at energies above the range of AGASA.

2.3 Origins of UHECRs

The observation of cosmic rays above the GZK cutoff raises questions about the origins of these particles. Particles at energies at and below the 'knee' are believed to be galactic in origin, with the primary source being supernovae. The diffusive shock acceleration of a supernova produces a power law spectrum of high-energy particles [2]. A secondary source may be OB associations, in which particles are accelerated by turbulent motion and stellar winds [3]. No individual sources have yet been identified, however.

Only a few known astrophysical phenomena are plausible sources of UHECRs. These are defined by the 'Hillas criterion' [11], which states that a particle accelerated in a magnetic field can only continue gaining energy until its Larmor radius becomes comparable to the size of the acceleration region: $E_{\mathrm{MAX}} \sim ZeBR\beta c$.

At higher energies, the particle will no longer be bound by the magnetic field, and will escape from the system. The interstellar magnetic field, for example, can accelerate galactic particles to at most $E \sim 10^{18}$ eV. For this reason, it is speculated that most UHECRs are extra-galactic in origin.

Speculated extragalactic sources of UHECRs include the following astrophysical phenomena, as well as more exotic possibilities [17].

Radio Galaxies and AGN The extended lobes of radio galaxies typically contain 'hot spots' which are interpreted to be the shock front of the relativistic jets which emanate from the active galactic nucleus, or AGN. The hot spots contain a magnetic field up to a few hundred μG in an area of a few kpc^2. Under these conditions, we have $E_{\mathrm{MAX}} \sim 10^{20}$ eV. It should be noted, however, that there are only a few AGN within 50 Mpc of our location, and none are clear candidate sources for the 10^{20} eV AGASA events.

Quasar Remnants A quasar remnant is the end-stage evolution of a luminous quasar: a spinning supermassive black hole, threaded by magnetic fields generated by currents flowing in a disc around it. The EMF generated would be sufficient to accelerate a particle to ultrahigh energies. We appear to live in an epoch where luminous quasars are rare. However, extrapolating from the number of luminous quasars at high redshift, the number of quasar remnants nearby may be large.

Starburst Galaxies and LIGs Starbursts are galaxies undergoing a period of intense star formation. Due to numerous supernovae, a cavity of hot gas can be created in the center of an active region. The hot gas will expand and form a shock front as it contacts the cooler interstellar medium. Ions such as iron nuclei can be accelerated to super-GZK energies in these conditions. Luminous Infrared Galaxies (LIGs), which may form after a collision between galaxies, are similar to starburst galaxies on a larger scale.

Gamma Ray Bursts The sources of GRBs are among the most energetic in the universe and could plausibly generate UHECRs. Unfortunately, GRBs are also among the most distant objects we can observe, there would be little chance of seeing UHECRs at this distance due to the GZK cutoff.

Being able to identify any of these phenomena with UHECR generation would contribute to our understanding of both subjects. This is a strong motivaton to reproduce or rule out the AGASA clustering analysis (shown in Figure 2). If the clustering is confirmed with sufficient statistics it may be possible to identify point sources of high-energy radiation.

3 Current Status

I joined the CHICOS project last August, at a time when the number of installed sites was increasing rapidly. As the complexity of the array increased, it became clear that the on-site software needed to be more automatic and more reliable. My work during the first half of the year was focused on rewriting the data acquisition software used by the individual detector workstations to achieve this. The new version of the software was installed on all sites in the array early this year. Following the upgrade, I have been concentrating on the data analysis of the mid-energy cosmic rays observed by the Chiquita array.

3.1 Data Collection Software

Each site consists of a pair of detectors separated by 3-5 meters. Each detector contains a 1 m^2 slab of scintillator approximately 10 cm thick and records approximately 200 single-particle events per second. Hits in both detectors are considered 'coincident' if they are separated by 100 ns or less. A coincidence event is considered a 'trigger' if both hits have an intensity equal to two minimum ionizing particles[2], or 2 MIP.

Given the large quantity of data collected at every site, it would be impractical to transfer all the data files to Caltech every day. The following system of data handling is designed to minimize the amount of data which must be sent over the internet, while keeping the reliablity of data transfer as high as possible.

Each day at 12:45 am, every site sends a file of timestamped trigger events via FTP or SCP to a server at Caltech. If not immediately successful, the local site will continue trying to send the file for up to 3 hours. At 4:00 am the trigger events from all sites in the array are combined into a 'master trigger file' by an automatic shell script. At 5:00 am, the data stations download the master trigger file.

Once the local site has downloaded the master trigger file, it compares the array-wide list of triggers with the local lists of single events for each detector. A single hit occurring within ± 50 μs of a trigger is considered a 'match'. The matching process typically takes about 45 minutes, after which the site uploads the match files to Caltech. At 8:20 am the match files from all CHICOS sites are automatically assembled into candidate shower files. In order to be considered a 'candidate shower', a set of matching hits must include 3 of the 4 sites closest to the trigger site. To eliminate some accidental matches, it is further required that the separation in time between hits in a shower be less than the light travel time between the sites.

Most of the data collection software for the remote stations was originally written Juncai Gao. My work primarily involved making the system more robust, which ended up entailing a complete re-write of much of the higher-level code. The GPS startup routine was modified to check the receiver response to each input command. This error-checking ensures that the GPS receiver will be properly reset even after a loss of power. In addition, any GPS errors are now added to the daily log file of the site, for easier detection.

Each data station maintains a 7-day archive of relevant data on site. This archive includes history files of count rates per second, and coincidence rates per minute. Since this data is useful for diagnosing problems remotely, I modified the file-handler to send the daily history files with the first data transfer the following day at 12:45 am. These files are now available through a web interface to assess the status of any site on any given date.

The data transfer processes were made more robust by adding error checking and increasing the number of attempts made if file transfer fails. Three of the file-handling programs were rewritten as a single executable, which eliminates time wasted between calls to the different programs. The overall duty cycle of the array increased from approximately 80% to 90% during the upgrade period when the new software was installed. (See Figure 4.) The success rate of data transmission is close to 100% in the absence of major network problems beyond the control of Caltech.

In the process of rewriting the Labview software, documentation was added to make it easier to maintain. The current version contains over 100 separate subroutines, which can be roughly categorized into low-level data handling, high-level data handling, GPS software, error checking, file handling, and user interface.

---

[2] 1 MIP is defined as the peak of the energy spectrum of single events in a given detector, and approximately corresponds to the average energy loss in the scintillator from a vertically propagating muon.

3.2 Reconstruction Software

The reconstruction software that CHICOS currently uses was largely written by Brant Carlson and Chris Jillings. It is based on the methods used by AGASA, and is implemented in C++. In particular, the reconstruction code works by performing a χ^2 minimization of the shower data to a model LDF parametrized by incident angle and energy. The χ^2 fit routine makes use of the CERN minimization code MINUIT, and the parametrized LDF is taken from the fit to AGASA data as discussed in Section 2.1.

The procedure used to fit a set of shower data is as follows [4]:

1. The core location is given a first estimate of 200 m from the trigger site in one of the cardinal directions. (The minimization procedure is repeated 4 times, with different starting points; the fit which produces the lowest χ^2 is used.)

2. Each set of 3 hits in the shower can be fit with a plane perpendicular to the trajectory of the shower front. A first estimate of of the incident angle is found by averaging the results from each trio of events in the data.

3. No attempt is made at a first estimate of the energy. Instead, the logarithm of the proportionality constant C in $\rho(r)$ is set to 3.

4. The direction of propagation is fit by χ^2 minimization of the hit times with the equations for the time delay T_d and the time spread T_s. (See Section 2.1)

5. The energy is fit by χ^2 minimization of the hit intensities with the LDF, $S(r)$.

6. The core location and the energy are fit simultaneously, again by χ^2 minimization of the intensities given by LDF.

7. Steps 4 through 6 are repeated until the fit converges, up to a maximum of 10 iterations.

8. The minimization procedure is repeated with a different initial core location and the best result of 4 is chosen.

It is also possible to fit shower data by assigning initial values to the parameters, or forcing a reconstruction with a given set of parameters. The output of the reconstructor is given as the (x, y, z) coordinates of the core location in meters (with the origin centered on Caltech 03, the site located on the roof of Kellogg), the zenith angle θ and the directional angle ϕ (measured counterclockwise from East), the time of impact t of the core relative to the first measured event, and the energy E. The reconstructor also estimates errors in all parameters and separate χ^2 values for angle and core location.

Finally, it is important to note that all of our shower reconstruction software is made available to participating schools via the CHICOS website. It is part of the philosophy of the project that both data and analysis tools be made available for educational purposes. The reconstruction tools are accompanied by detailed documentation as well as a tutorial aimed at high-school students written by Theresa Lynn.

3.3 Spectrum of Low-Energy Cosmic Rays

Over the last few months, I have been working on combining the event data from the Chiquita array into a spectrum covering the energy range of 10^{16} to 10^{18} eV. The flux of cosmic rays in this range has been well-measured by previous experiments, and therefore can serve as a useful calibration of the larger CHICOS array.

The Chiquita array records approximately 10-12 cosmic ray airshowers per day, when all sites are operational. The overall duty cycle of the Chiquita array (all 12 sites operational) has been 60% during the time period between March 1 and May 21, 2004. (Note that this corresponds to an average 95% duty cycle per detector, better that the average over the larger CHICOS array.) A data sample of 300 showers was selected from the time intervals when all 12 Chiquita sites were running. This makes it possible to ignore time-dependence when calculating the aperture.

The effective aperture of of the array (Figure 5), given in units of m^2 sr, is a combination of the physical size of the array, the spacing between sites, the sensitivity of the detectors, and the methods used to identify shower events. The response of the array is modeled by applying Poisson statistics to the assumed LDF for a shower at a given energy and location. The simulated data is then filtered through the same cuts as the real data, and passed into the reconstruction software.

A monte-carlo routine was used to select shower parameters for a set of 20 equally spaced values of $\log_{10} E(\mathrm{eV})$ between 16.0 and 18.0. Core locations were spread over an area of (2 km)^2 centered on the Kellogg site, and the cosine of the zenith angle was chosen to be within the range of 0.5 and 1. The aperture at a given energy is defined as the ratio of the number of showers that produce simulated data that passes the filters to the total number of showers thrown at that energy, multiplied by the area and solid angle of the simulated data set.

The priliminary calculation of the energy spectrum, shown in Figure 6, falls significantly below results from other experiments in that energy range. This would indicate that the simulated results overestimate the efficiency of the the array. Several possibilities have already been ruled out: bias in the reconstrution software, for example, exists, but does not appear sufficient to account for the discrepancy. One remaining explanation is that, because of the difference in altitude between CHICOS and AGASA, the modified AGASA LDF we employed may not be a good approximation to the CHICOS LDF. Barbara Falkowski, a grad student at CSUN, has produced simulated air showers (using AIRES) in the Chiquita energy range, which will be used to evaluate any discrepancies with the AGASA LDF.

3.4 High Energy Cosmic Ray Events

The first high-energy CHICOS air shower was observed on January 29, 2003, when only 30 sites had been installed. A total of 20 showers having been observed to date, with 8 of those occuring since the beginning of 2004. An additional set of approximately 15 showers are considered marginal detections, and are not included here. The current detection rate is approximately one shower above 10^{19} eV per month, and this rate will continue to increase as the array is expanded and filled in. The final array, with a total of 90 sites, should have a detection rate 6 to 7 times higher than the current configuration of 60 sites.

Table 1 summarizes the high energy air showers observed by CHICOS from January 2003 through April 2004[3].

Date	$\log E$ (eV)	# of Hits	$\chi^2{\cdot}_{\text{angle}}$	$\chi^2{\cdot}_{\text{location}}$
01-29-03	19.90 ± 0.280	4	0.04	0.11
03-03-03	19.02 ± 0.320	3	1.62	0.02
03-24-03	18.51 ± 0.443	4	0.11	1.38
04-19-03	17.20 ± 6.214	5	1.70	0.32
05-13-03	18.31 ± 0.624	4	97.74	0.01
05-17-03	18.82 ± 0.018	4	42.10	0.24
06-14-03	16.96 ± 0.233	7	0.92	0.98
06-18-03	20.65 ± 0.554	4	7.15	0.05
06-27-03	16.93	7	10.06	24.92
08-17-03	19.67 ± 2.612	4	0.55	0.36
11-20-03	19.98 ± 0.556	5	6.07	0.54
12-26-03	18.50 ± 0.214	7	0.90	30.06
01-06-04	18.27 ± 0.904	4	0.35	0.02
01-08-04	19.38 ± 0.511	4	0.35	0.36
03-01-04	19.04	5	108.63	607.70
03-03-04	18.48 ± 0.591	4	0.27	0.02
03-07-04	18.79 ± 0.359	6	0.26	1.49
03-21-04	19.22 ± 1.165	3	0.00	0.07
03-23-04	17.70 ± 1.749	4	0.00	0.05
04-08-04	20.61 ± 1.040	4	3.22	0.12

Table 1: High energy cosmic ray airshowers observed by CHICOS since January 2003.

As can be seen from this table, CHICOS has observed 9 events with $E > 10^{19}$ eV and an additional 7 with $E > 10^{18}$. With future improvements to the reconstructor code we expect to be able to lower the error estimates and improve the angular resolution. When the current array is filled in, the average number of detectors hit by a large air shower will increase, improving the angular resolution of the array. With all 90 sites operational the angular resolution of CHICOS will be approximately 4° or better, which is comparable to the resolution claimed by AGASA.

---

[3] The reconstruction of these showers was done by Theresa Lynn. Where no error estimate is given for $\log E$, the shower was fit by hand rather than the minimization code. These results are preliminary and will be re-evaluated as the reconstruction code is improved.

4 Future Work

CHICOS will collect enough data at high energies over the next three years of observation to compare with previous results in the ultrahigh energy regime. The analysis of this data will be the basis of my thesis work. In the short term, my work will include improving the reconstructor code and ongoing maintenance of the array.

4.1 Revision of the Reconstructor Code

There are several major refinements that should be made to the current reconstruction software. The first is to use simulated air shower data to produce a parametrized LDF for our altitude. As mentioned earlier, we are in the process of running a series of series of simulated air showers over the range of energies CHICOS can observe. A significant cause of variation between air showers of the same energy is the depth of first interaction. For this reason, multiple simulations must be run for each energy and incident angle. Enough low-energy simulations have already been done to fit a parametrized LDF at energies observable by the Chiquita array.

A second important revision will be to replace the chi-squared minimization routine with a maximum likelihood fit to this LDF, as has been done by AGASA. The likelihood function takes the form:

$$L = \prod_{i=1}^{n} \left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) \cdot \exp \left[-\frac{1}{2} \sum_{i=1}^{n} \left(\frac{\rho_i - \rho(R_i)}{\sigma_i} \right)^2 \right],$$

where ρ_i is the intensity observed by the ith detector and $\rho(R_i)$ is the intensity estimated by the LDF. The fluctuation in intensity, σ, is a combination of the shower fluctuation and the detector response. Both of these improvements will be implemented this summer.

A final improvement which is not necessary for Chiquita, but critical to CHICOS will be to estimate a time-dependent aperture. The reconstruction software already tracks the 'health' of each site in the array. The same information can be used to write a program which integrates the total aperture of the array over any given time period, taking into account which sites were active. This will require a monte-carlo simulation of the array's response in all possible configurations, for a wide range of energies. This work will also be done during the summer and fall of this year.

4.2 Automation of Shower Reconstructions

In its current form, the reconstruction software is not fully automatic. Accidental hits and detector overflows bias the fit and need to be removed manually. In addition, the χ^2 minimization occasionally misses the true minimum, which can be found by hand.

One goal of improving the reconstruction software is to make it reliable enough to fit large quantities of data automatically. Before this can be done, improvements need to be made in the way that the software identifies accidental hits and afterpulses, both major sources of noise in the data.

Significant progress has already been made in this direction, and this work can reasonably be finished during the 2004/2005 school year.

4.3 Maintenance of Array Sites

In addition to rewriting the reconstruction code, I will be working together with the Project Coordinator over the next three years to maintain and operate the CHICOS array. This involves the installation of new sites, maintenance of existing sites, periodic upgrades of hardware and software, and interaction with the associated schools and teachers.

Maintenance is typically required at 2-3 sites each week. When 90 sites are deployed this will likely increase to 3-4 site visits per week. It is reasonable to plan on allowing at least one day per week to be devoted to maintaining operations. Data management at Caltech is also an ongoing responsibility. The CHICOS server and the automatic scripts which handle the data files need oversight and occasional upgrades or improvements.

It may be necessary in the future to improve the Labview software for reasons of efficiency, reliability, increased functionality, or hardware changes. Some plans for a more robust data transfer scheme, in which a failed transfer can take place on a later day, have been discussed, and may be implemented if there is a need to improve the livetime of the array. Maintenance of the Labview software will continue to be my responsibility.

4.4 Expansion of the CHICOS Array

Ideally, once all 90 CHICOS sites have been deployed, we will be able to consider expanding the array even further. With an array of 300-500 sites at schools throughout the LA basin, we could cover an area of $\sim$ 1000 km^2. An expansion beyond the current 90 sites will necessitate the design of new hardware and software, as we have used all of the detectors originally obtained from LANL. Expanding the array to this size would greatly increase the detection rate of UHECRs. Sources of funding for such an expansion will be pursued as the 90-site CHICOS array nears completion.

4.5 UHECR Analysis

With an estimated detection rate of ~ 1 event with $E > 10^{19}$ eV per month, a 3-year run observing time will accumulate a data set of ~ 36 ultrahigh energy events. This can be considered a minimum estimate, as the detection rate is expected to increase substantially by a factor of $\sim 6 - 7$ as the array expands. Events with energies above 10^{20} eV make up 0.001% of the flux above 10^{19} eV, but this is offset by the fact that the array is more likely to detect higher-energy events. Assuming that super-GZK cosmic rays exist, CHICOS should observe about $8 - 10$ during the next 3 years.

For my thesis, I will construct a spectrum of the ultrahigh energy cosmic ray flux, using the complete set of CHICOS data. In addition, the data set will be analysed using the same methods as the AGASA project to look for small scale clustering. Both of these calculations can be done using the CHICOS data set alone, or with a combined data set of all known UHECR events, to yield greater statistical significance.

References

[1] P. Auger *et al.* Extensive Cosmic-Ray Showers. *Rev. Mod. Phys.*, 11:288–291, 1939.

[2] A.R. Bell. *Mon. Not. Roy. Astron. Soc.*, 182:147, 1978.

[3] A.M. Bykov and G.D. Fleishman. *Mon. Not. Roy. Astron. Soc.*, 255:269, 1992.

[4] B. Carlson. libCTShower/Root, and CGI Documentation. SURF Project Report, August 2003.

[5] N. Chiba *et al.* *Nucl. Instr. Meth. A*, 311(1):338–349, 1992.

[6] G. Clark *et al.* Cosmic-Ray Air Showers at Sea Level. *Phys. Rev.*, 122(2):637–654, April 1961.

[7] C.J. Crannell, H.R. Allan, and M.P. Sun. Widths of electron-photon cascades. *Phys. Rev. D*, 18(1):216–219, 1978.

[8] C.B. Finley and S. Westerhoff. *Astroparticle Physics*, 2003.

[9] K. Greisen. End to the cosmic-ray spectrum? *Phys. Rev. Lett.*, 16(17):748–750, April 1966.

[10] N. Hayashida *et al.* Possible Clustering of the Most Energetic Cosmic Rays. *Phys. Rev. Lett.*, 77(6):1000–1004, August 1996.

[11] A.M. Hillas. *Ann. Rev. Astron. Astrophys.*, 22:425, 1984.

[12] A.A. Ivanov, S.P. Knurenko, and I.Y. Sleptsov. *Nucl. Phys. Proc. Suppl.*, 122:226, 2003.

[13] M.A. Lawrence, R.J. Reid, and A.A. Watson. *J. Phys. G*, 17:733, 1991.

[14] M. Nagano and A.A. Watson. Observations and implications of the UHECRs. *Reviews of Modern Physics*, 72(3):689–732, July 2000.

[15] R. Schlickeiser. *Cosmic Ray Astrophysics*. Springer-Verlag, 2002.

[16] M. Takeda *et al.* Energy determination in the AGASA experiment. *Elsevier Science*, 2003.

[17] D.F. Torres and L.A. Anchordoqui. Astrophysical Origins of UHECRs. *Reports on Progress in Physics*, 2004.

[18] Y. Uchihori *et al.* *Astropart. Phys.*, 13:151, 2000.

[19] M.M. Winn *et al.* *J. Phys. G*, 12:653, 1986.

[20] S. Yoshida *et al.* *Nucl. Part. Phys.*, 20:651, 1994.

[21] G.T. Zatsepin and V.A. Kuz'min. *Zh. Eskp. Teor. Fiz.*, 4:114, 1966. JETP Letters 4:78, 1966.

Figure 1: AGASA UHECR events and predicted flux. [16].

Figure 2: Sky map showing AGASA UHECR clusters [10].

Figure 3: Map of installed CHICOS sites (as of May 2004).

Figure 4: Graph of success rate of data transmission for 2004.

Figure 5: Aperture of the Chiquita array, in m^2 sr.

Figure 6: Energy spectrum of Chiquita events.