THE UNIVERSITY of ADELAIDE

School of Physical Sciences

Department of Physics

Quality and Stability of Cosmic Ray Energy Assignments at the Pierre Auger Observatory

Adila Binti Abdul Halim

Supervisors:
Prof. Bruce Dawson
Dr. Jose Bellido

A thesis submitted to the University of Adelaide in fulfilment of the requirements for the degree of Master of Philosophy (Mixed)

October 2021

Abstract

One of the biggest mysteries in astrophysics is the origin and nature of ultra-high energy cosmic rays (UHECRs). Measuring the energies accurately is vital in understanding sources of cosmic rays and the physical mechanisms behind their acceleration. The energy spectrum, which describes flux of cosmic rays as a function of energy, can give insight to their possible sources. Many observatories around the world have been constructed to detect UHECRs and measure their energies to the highest detection accuracy possible. Due to the low flux of cosmic rays at the highest energies, various detectors are deployed spanning extremely large areas to detect extensive air showers initiated by cosmic rays arriving at the top of Earth's atmosphere. The signal strengths, timing information and air shower profiles can then be measured to accurately obtain energies of cosmic rays.

One of the most notable detectors of UHECRs is the Pierre Auger Observatory. It is the largest cosmic ray detector in the world, located in western Argentina. It adopts a hybrid detection technique through the use of two independent detectors known as the fluorescence detector (FD) and the surface detector (SD). The hybrid design aids in eliminating model dependencies and provides important cross-checks for accurate air shower reconstruction.

This thesis investigates the stability (with time and position) of FD and SD energy assignments at the Pierre Auger Observatory. Chapters 1, 2 and 3 describe cosmic rays, detection methods and the Pierre Auger Observatory. The SD, FD and hybrid design are outlined followed by details of energy reconstruction. Analysis begins in Chapter 4 with the FD energy over SD signal ratio. The FD stability is probed through a quantitative analysis of this ratio in a 14-year long dataset subject to varying conditions. Trends in the ratio seen by each telescope in the FD are also analysed for correlations with cleaning campaigns done to the UV filters and mirrors.

Chapter 5 presents an analysis on SD energy stability, demonstrating evolution of SD event rates above a full efficiency threshold energy with time and station position across the array. Results are compared with previous work and possible causes for the observed drift are suggested. Chapter 6 details an analysis on SD event rate evolution with time and distance from each FD site. This is done through the selection of stations involved with hybrid events. Hybrid event rates are then assigned using yearly SD rates. Observed trends are compared with structures in the SD rates and energy ratio. It was found that the time dependence in the energy ratio and rates are statistically significant. On average, energies measured by the SD are increasing with time and this contributes to the drift in

the energy ratio. Position dependence also exists across the array and this was attributed to a decay in sensitivity of some stations in the SD. Other causes for the dependencies are suggested, but further work is needed to investigate the extent of their contributions. Chapter 7 summarises the main results and outlines some further investigations.

Declaration of Originality

I, Adila Binti Abdul Halim, certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

17 February 2022

Signature ..... Date.....

Acknowledgements

First and foremost, I would like to thank my supervisors Bruce Dawson and Jose Bellido. I am so grateful to have been under your guidance and I could not have asked for anyone better. Thank you for putting up with my questions, slow thoughts, mistakes and frequent meetings. I would also like to give many thanks to Roger Clay, for all the feedback given in group meetings, for teaching me concepts I needed help with, and for proofreading all my chapters.

To Brad, Fraser, Tristan, Violet, Justin, Jarryd and Steven; thank you so much for all the help you have given me. Whether it be getting my various codes to work, giving feedback during group meetings or being supportive, it means so much. I would also like to give thanks to Phong Nguyen as I used his thesis format as a visual guide to write mine.

To my parents, thank you for all of your support through financial aid, words of encouragement and for checking in on how I am doing everyday. To my siblings, thank you so much for supporting me throughout this journey and for listening to my worries and daily mishaps. To Leonardo, thank you for helping me understand physics, proofreading all of my chapters and for believing in me. Thank you for holding my hand through it all and for the lovely walks with Ruki. I could not have done this without all of you.

Contents

Chapter 1

Cosmic Rays

Cosmic rays are charged particles originating from powerful sources in deep space. Historically, it was thought that cosmic rays were a form of electromagnetic radiation which is why it has been termed cosmic "ray" making it a misnomer. The energies can be extremely high; up to the energy of a fast tennis ball serve, which is about 10^{19} electron Volts (eV). All of this energy is in a single particle [1]. Cosmic rays exceeding 10^{18} eV are called ultra-high energy cosmic rays (UHECRs) and they can be lone protons or heavier atomic nuclei. Since cosmic rays have electric charge, their trajectories are strongly influenced by magnetic fields present in the path from their originating source to the Earth where they are observed.

Once cosmic rays arrive at Earth, they impact particles in the upper atmosphere and produce cascades of lower energy particles and detectable light. These cascades, called extensive air showers, can extend to cover an effective area of 10 km^2 or more when produced by a cosmic ray with an energy of 10^{20} eV [2]. Ever since their discovery, there has been a tremendous amount of effort to study them further. Many experiments have deployed detectors of varying sizes in orbit, on Earth, and in the atmosphere. The results from these experiments are studied extensively as high energy cosmic rays can provide insight to the physical mechanisms in their sources and the broader physics of the highest energy particles in Nature.

1.1 History of Cosmic Rays

The history of cosmic ray studies began in the late 18th century when a physicist named Charles-Augustin de Coulomb conducted experiments involving an electroscope [3]. He discovered that his instrument would discharge over time in a random manner. It was later discovered that this was because of ionisation of the air inside the instrument, but the source of ionising particles was a mystery at the time. Considerable progress was made in 1912 when Victor Hess made a breakthrough discovery whilst flying a balloon to an altitude of around 5 km [4]. The measurements he obtained during his trip indicated that ionization of the air first dropped, then strongly increased with altitude. This signified

that a source of radioactivity was not terrestrial. Since then, many experiments have been conducted to identify the sources of radiation coming from above the atmosphere. However, the mechanisms responsible for accelerating cosmic rays to such high energies are still incompletely understood.

1.2 Production of Cosmic Rays

1.2.1 Solar Cosmic Rays

Firstly, it is important to note that there are many cosmic rays present in our solar system which originate from physical processes occurring at the surface of the Sun. One of the most common processes is found in solar flares. They are described as explosions which occur on the surface of the Sun due to detangling and rapid variation of magnetic fields. When there is a solar flare, it is usually accompanied by gamma-rays, soft x-rays and even radiowave emissions. For a relatively large solar flare, there can be a burst of high energy particles containing electrons, protons or even heavier nuclei. Solar cosmic ray protons can have energies between 1 MeV to 10 GeV [6]. Figure 1.1 shows a schematic of the many kinds of phenomena which occur on the surface of the Sun, including the production of solar cosmic rays.

1.2.2 Acceleration of Cosmic Rays Beyond the Solar System

Due to the detection of cosmic rays possessing energies of far higher magnitudes than the Sun can produce, it was theorised that there must be cosmic rays being accelerated by other sources beyond the solar system. For many years, the cosmic ray acceleration mechanism had been a real mystery. In 1949, Fermi formulated a theory now known as Fermi's Original Theory of Cosmic Ray Acceleration [7]. It describes the bombardment of cosmic rays onto many different magnetised gas clouds with random velocities causing energy gain.

Figure 1.2 shows a particle entering a cloud at an angle θ_1 and exiting at an angle θ_2. In the frame of the cloud, the scattering is collisionless which indicates elastic scattering with no change in energy. Furthermore, the cosmic ray which enters a cloud scatters off in a random direction since gas clouds tend to be moving at random velocities, approximated to be around 15 km/s [8]. Firstly, consider a cloud moving at velocity v_{cloud} and a charged particle with energy, E_1. In the frame of the cloud, the energy of this particle is calculated to be

$$E_1' = \gamma_{cloud} E_1 (1 - \beta_{cloud} \cos\theta_1) \tag{1.1}$$

where $\beta_{cloud} = \frac{v_{cloud}}{c}$, c is the speed of light and $\gamma_{cloud} = \frac{1}{\sqrt{1-\beta_{cloud}{}^2}}$. Due to elastic scattering, the energy of the particle after it emerges from the cloud, E_2' is unchanged($E_2' = E_1'$). To obtain E_2 which is the energy of the particle in the lab

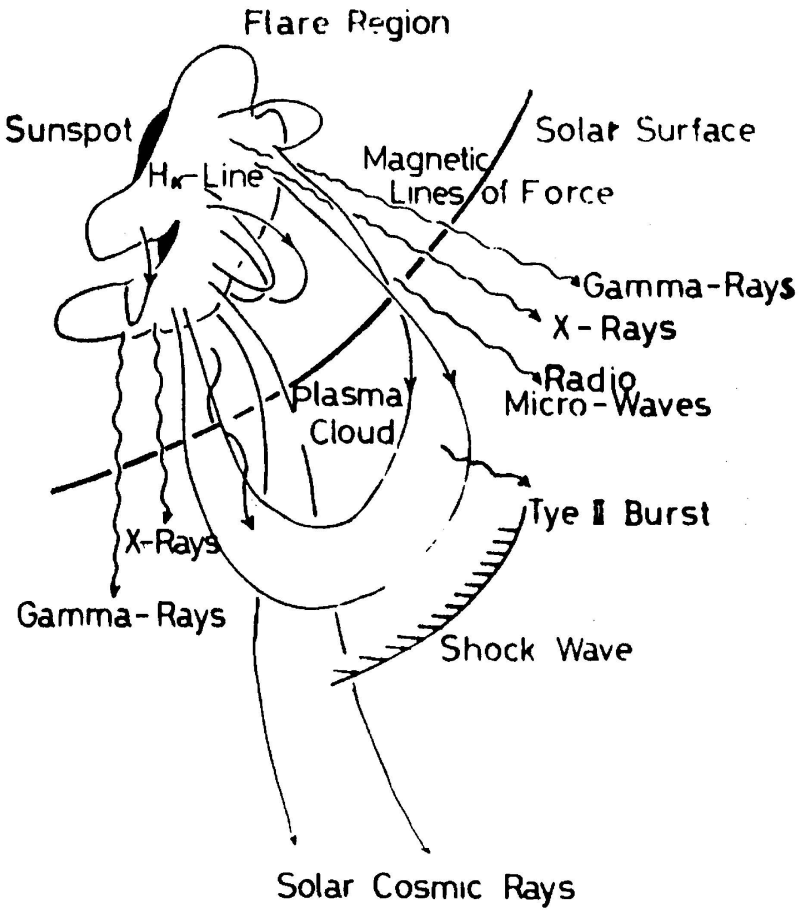


Figure 1.1: A diagram of the physical processes on the surface of the Sun including the production of solar cosmic rays [5].

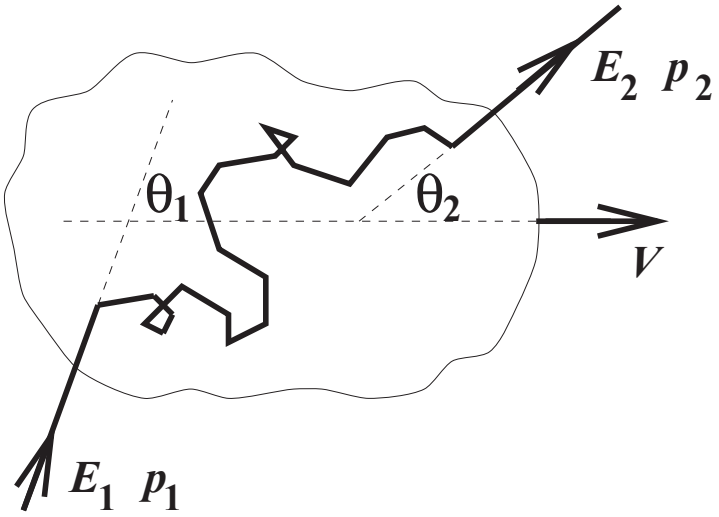


Figure 1.2: A schematic showing a particle scattering off of a cloud moving to the right at speed V which contains a magnetic field [8]. The particle enters with energy E_1 at angle θ_1 and momentum p_1 then exits with energy E_2 at angle θ_2 and momentum p_2 in the lab frame.

frame after it leaves the cloud, Lorentz transformations can be used to obtain

$$E_2 = \gamma_{cloud} E_2'(1 + \beta_{cloud} \cos\theta_2') \tag{1.2}$$

where θ_2' is the angle at which the particle exits the cloud in the frame of the cloud [8]. Since E_2' = E_1', the energy of the particle emerging from the cloud in the lab frame is calculated to be

$$E_2 = \gamma_{cloud}^2 E_1(1 - \beta_{cloud} \cos\theta_1)(1 + \beta_{cloud} \cos\theta_2') \tag{1.3}$$

Therefore, the fractional energy gain in the lab frame can be calculated and the result is shown as

$$\frac{\Delta E}{E} = \frac{E_2 - E_1}{E_1} = \frac{1 - \beta_{cloud}\cos\theta_1 + \beta_{cloud}\cos\theta_2' - \beta_{cloud}^2 \cos\theta_1 \cos\theta_2'}{1 - \beta_{cloud}^2} - 1 \tag{1.4}$$

Since the magnetised cloud causes the particle to scatter in a random direction, $<\cos\ \theta_2'>$ which is the expected value of $\cos\theta_2'$ is equal to zero. The value of $<\cos\theta_1>$ is calculated by identifying the collision rate of cosmic rays entering the cloud as a function of $\cos\theta_1$. It is found that the rate of collision is proportional to the relative velocity between the cloud and the particle [8]. So, the probability of having a collision at angle θ_1 is proportional to the difference in their relative velocities. At ultrarelativistic speeds, the speed of the cosmic rays approach *c* which means that the probability is proportional to 1 - $\beta_{cloud}\cos\theta_1$. From this result, the average value, $<\cos\theta_1>$ can be obtained in the usual way (by integrating $\cos\theta_1$ multiplied with the probability with respect to $d\cos\theta_1$) to obtain

$$< \cos\theta_1 >= \frac{-\beta_{cloud}}{3} \tag{1.5}$$

This result can be substituted into Equation 1.4 which yields a fractional energy gain per cloud interaction shown as

$$\frac{< \Delta E >}{E} \approx \frac{4}{3}\beta_{cloud}^2 \tag{1.6}$$

This result shows that the fractional energy gain is positive and proportional to β_{cloud}^2 (2nd order in β_{cloud}). However, $\beta_{cloud} \ll 1$ because the probability of a head-on collision between a cloud and a cosmic ray ($\theta_1 \sim 180°$) is only very slightly higher than the probability of a tail-on collision ($\theta_1 \sim 0°$). This results in a small average energy gain. This mechanism alone is not sufficient to explain the considerable observational flux of high energy cosmic rays in the sky. Furthermore, it had been shown that this mechanism is more feasible for protons, but is insufficient for accelerating heavier nuclei [7].

In the late 1970s, this theory was modified and a more likely mechanism was put forward called diffusive shock acceleration [9]. A 'shock' is defined as

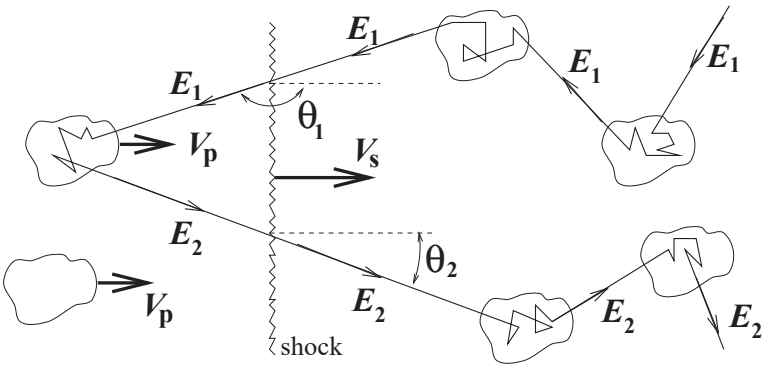


Figure 1.3: A schematic of a combination of scattering processes which occur for many particles in an accelerating shock [8]. It is shown that V_s is the speed of the shock while V_p is the speed of the particles moving behind the shock.

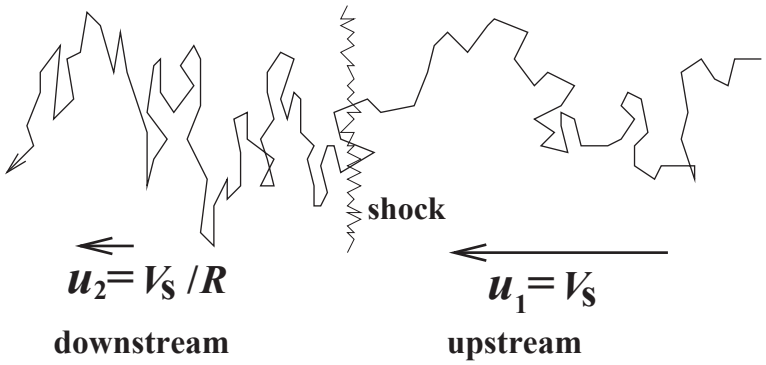


Figure 1.4: A schematic showing the speeds of media in the shock frame [8]. In this frame, it is observed that particles are moving towards the shock at speed u_1 while the particles behind the shock are moving away at speed $u_2 = V_S/R$ where R is the compression ratio.

the plane where there exists a discontinuity in speed between two interstellar media. The shocked medium moves at supersonic speed, pushing onto an unshocked medium which has zero speed relative to the shocked medium. The term 'diffusive' describes how the cosmic rays diffuse across the shock and back. As the cosmic rays undergo this process, they are increasing in speed and energy.

Figure 1.3 shows a schematic of multiple accelerating particles within a shock where V_s is the speed of the shock and V_p is the speed of particles behind the shock. An example of this process is associated with supernova explosions, where the material of a star is violently ejected outwards causing a shock to move through the interstellar medium. As the shock propagates, charged particles collide with gas clouds and they will have some probability of crossing the shock front. This process can happen repeatedly thus causing an energy gain as the particle crosses the shock front, back and forth. Figure 1.4 illustrates how this acceleration mechanism works in the shock frame. In this frame, the particles are moving towards the shock at speed u_1 while the particles behind the shock

are moving away at speed $u_2 = V_S/R$ where R is the specific heat compression ratio of shocked and unshocked media. In the lab frame, the shock would be travelling upstream at speed $u_1 = V_s$ while the particles behind the shock would have a speed $u_2 = V_s - V_s/R$ [8]. In an ideal case of strong shocks with an ionised medium, the value of R is 4 according to the theory of shock hydrodynamics and this yields the result

$$\frac{V_s}{V_p} = \frac{4}{3} \tag{1.7}$$

Consequently, the fractional energy gain is calculated to be

$$\frac{<\Delta E>}{E} \approx \frac{V_s}{c} \tag{1.8}$$

It is shown that this modified theory yields a result of a fractional energy gain proportional to β_s in the first order where $\beta_s = V_s/c$ [8]. This modification provides a more efficient acceleration mechanism since cosmic rays collide with gas clouds moving towards them at high speeds on each side of the shock front, increasing the likelihood of approximately head-on collisions. A calculation of the probability for cosmic rays to escape from the shock leads to the generation of the power law energy spectrum discussed in Section 1.7. Many experiments have detected a large number of cosmic ray events over the years whose spectra agree with the power law spectrum to a very good approximation.

1.3 Propagation

It is important to note that since cosmic rays are under the influence of magnetic fields, they experience a magnetic force. The particles move in a helical motion around the magnetic field lines. The distance between a cosmic ray at any point of its motion and the magnetic field line is known as the gyroradius. This is obtained by first noting that the force applied onto a cosmic ray by a magnetic field can be equated with the force from Newton's second law shown as

$$ZevB = \frac{mv^2}{r} \tag{1.9}$$

where m is the mass of the particle, v is its speed, B is the magnitude of the magnetic field, Ze is the amount of charge and r is the gyroradius. This can be simplified to

$$r = \frac{mv}{ZeB} \tag{1.10}$$

Since mv represents the momentum of the particle, it can be replaced by p shown as

$$r = \frac{p}{ZeB} \tag{1.11}$$

For cosmic rays that are highly relativistic, their energy can be written as $E \sim pc$ where c is the speed of light. We can multiply both sides by c to get:

$$rc \sim \frac{E}{ZeB} \tag{1.12}$$

This shows that the gyroradius can be determined by the energy of the particle as shown by

$$r \sim \frac{E}{ZeBc} \tag{1.13}$$

and the rigidity, ρ is defined as

$$\rho = \frac{E}{Ze} \tag{1.14}$$

which has units in Volts. For example, a 10 GeV (where 1 GeV $\equiv 10^9$ eV) proton has a rigidity of 10 GV. To locate sources of cosmic rays, it is useful to recognise that the gyroradius of the cosmic ray cannot be larger than the size of the accelerating structure [10]. So, if the gyroradius of a high-energy cosmic ray is comparable to the thickness of the Milky Way (with a Galactic magnetic field of 0.3 nT), the origin of said cosmic ray is extragalactic [10]. Also, Equation 1.13 shows that there is a limit to the energy of a cosmic ray given a radius and magnetic field strength it experiences in a particular source. This maximum acceleration energy limit is known as the Hillas criterion [11] shown as

$$E_{max} \leq \gamma r Z e B \tag{1.15}$$

where γ is the Lorentz factor of the shock moving relativistically. Figure 1.5 shows the Hillas plot which shows possible accelerating sources based on the strength of their magnetic fields as well as their size. They include white dwarfs, neutron stars and supernova remnants as discussed in Section 1.4.1 and 1.4.2. Other sources include radio galaxy jets and the Galactic disk halo. Overall, although there are plausible sources of cosmic rays which hold great potential, there are limits to how much energy is produced considering the scale of magnetic fields and the energies they possess.

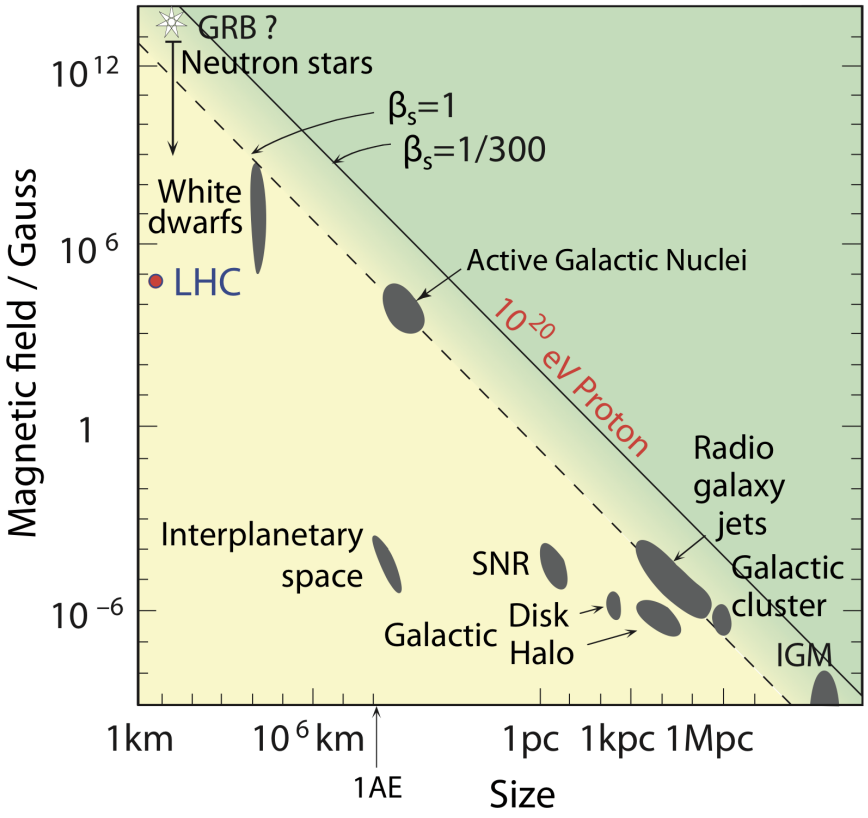


Figure 1.5: A Hillas plot which shows possible accelerating sources based on magnetic field strength and size [11, 12]. GRB stands for gamma-ray bursts, IGM stands for intergalactic medium, SNR is supernova remnant and LHC is Large Hadron Collider. β_s represents the fraction of the speed of a shock to the speed of light, i.e. $\beta_s = \frac{v_s}{c}$ and 1 AU (shown as AE on the figure) is an approximate distance between the Earth and the Sun.

1.4 Potential Sources of Cosmic Rays

The origin of cosmic rays are dominated by three key points; the particle origin, energy and site of acceleration [10]. There are multiple sources which are plausible for accelerating cosmic rays within and outside of the Milky Way galaxy.

1.4.1 Galactic sources

- **Supernova remnants (SNRs)**
  One of the potential sources of acceleration include supernovae which are events where stars undergo massive explosions. The particle origin would be the stellar ejecta, their energies would come from the explosive power causing ejecta to travel at high speeds and the site of acceleration would be the shock front travelling outwards pushing against the unshocked interstellar medium. Supernovae ejecta can possess total kinetic energies up to 10^{44} J [10]. Despite this great source of energy as well as the addition of amplification from magnetic fields [13], the maximum modelled energy that can be imparted to any one cosmic ray by this mechanism only reaches 10^{17} eV [14].

- **Neutron stars**
  Neutron stars are plausible candidates that could be responsible for the acceleration mechanisms in high energy cosmic rays. Neutron stars are very dense, compact objects made largely of degenerate nucleon material with densities of about 500 trillion times that of water [15]. A neutron star is formed when a very massive star explodes at the end of its lifespan leaving behind its core. The angular momentum of the contracting core is conserved in the resulting neutron star spanning a radius of a few tens of kilometres. This causes the neutron star to spin rapidly with periods of about 60 revolutions per second. In some cases, these neutron stars are observed as 'pulsars' when their rotating magnetic fields of massive strength result in a pulsing radio source observable from Earth, provided their magnetic poles are oriented correctly towards Earth. Although there have been studies on the possibilities that neutrons stars could be plausible candidates for cosmic ray acceleration due to their large magnetic fields [16], there are insufficient known neutron star sources found in the sky which explains the arrival of cosmic rays with energies above 10^{19} eV in an isotropic manner [14].

- **Galactic centre**
  There has been an extensive amount of effort to observe the Galactic centre since it can be studied from the radio to gamma-ray band. In particular, observations of the Galactic centre confirm the presence of a

supermassive black hole called Sagittarius A*. From this, there have been recent discoveries which indicate that Sagittarius A* can be a source of PeV (1 PeV $\equiv 10^{15}$ eV) galactic cosmic rays which have been accelerated in the Fermi bubbles formed by tidal disruption of stars by the black hole [17, 18]. There have also been observations of a high flux of cosmic ray events with energies below $\sim$ 1 EeV (1 EeV $\equiv 10^{18}$ eV) coming from the direction of the Galactic centre [19–21].

- **Massive star clusters**
  Massive stars with powerful winds have been proposed as a viable supplier of accelerated cosmic rays. Recently, a detection of gamma-ray photons at energies ranging from 100 TeV (1 TeV $\equiv 10^{12}$ eV) up to 1.4 PeV from many ultra high-energy Galactic gamma-ray sources were found [22, 23]. This indicates the presence of cosmic ray PeVatrons in the Milky Way galaxy. These sources were detected using the Large High Altitude Air Shower Observatory (LHAASO) and Tibet Air Shower-γ, located in Sichuan Province and Tibet in China. The positional coincidence of one of the sources with the 'Cygnus Cocoon'; a superbubble that surrounds a region of massive star formation, can be treated as evidence of the operation of massive stars as hadronic PeVatrons [22]. The High-Altitude Water Cherenkov observatory (HAWC) located in Sierra Negra in Mexico have also reported observations of 1-100 TeV gamma-rays coming from the Cygnus Cocoon [24].

As these sources still fail to explain the flux of even higher energy particles, an extragalactic source becomes likely for ultra-high energy cosmic rays.

1.4.2 Extragalactic sources

- **Active Galactic Nuclei (AGN)**
  An AGN is a galaxy with a supermassive black hole in its centre actively accreting matter and producing extremely powerful jets of particles [4]. These jets are perpendicular to the plane of the accretion disk of the black hole and are believed to accelerate particles to relativistic energies. Although not much is known about AGNs and there is a lack of evidence on a strong correlation between an AGN and a high flux of cosmic rays coming from its direction, AGNs are still considered to be a likely source of high energy cosmic rays [25].

- **Starburst Galaxies**
  Starburst galaxies are galaxies which have rates of star formation that are much higher than that of a typical galaxy. An example of an infrared selected starburst galaxy is M82 [26]. The high rate of activity due to supernova explosions and massive stellar winds lead to the collision of many particles, forming hot gas clouds. Furthermore, the cooling times

of these gas clouds are long in comparison with the age of the starbust. The winds from supernovae can be so strong that they have sufficient power to expel the interstellar medium resulting in the formation of a shock front travelling at ~ 1000 km/s. This is an ideal environment for the acceleration of particles.

- **Gamma-ray Bursts (GRBs)**
  GRBs are short flashes of gamma-rays that are extremely energetic. They are also very bright, outshining other gamma-ray sources and can last from less than a second to sometimes hundreds of seconds. There have been several models proposed which are based on the acceleration of particles through diffusive shock acceleration from GRBs [27,28].

It is difficult to determine the origin of cosmic rays due to their randomised trajectories. Only at the highest energies are where their deviations from their paths are not as large. Anisotropy studies can provide powerful tools to search for sources of the highest energy cosmic rays. The Pierre Auger Observatory (introduced in Chapter 3) found that there is a dipole modulation with an amplitude of $6.6^{+1.2}_{-0.8}$% and pointing $\sim 125^{\circ}$ away from the Galactic centre for cosmic ray energies above 8 EeV [29–31]. These results indicate an extragalactic origin. Figure 1.6 shows the cosmic ray flux on a sky map for events above 8 EeV. There has also been evidence of two intermediate scale anisotropies from two observatories. The Telescope Array (TA) (Chapter 2) located in Utah, USA measured an anisotropy of high-energy cosmic ray events in the Northern Hemisphere originating from the direction of the Ursa Major cluster while the Pierre Auger Observatory observed a hotspot in the Southern Hemisphere in the direction of Centaurus A. The TA hotspot is shown in Figure 1.7, where there seems to be a high flux of cosmic rays with energies above 57 EeV. A local significance of 5.1σ was found at this hotspot with its R.A.= 146.7° and Dec.=43.2°. However, the chance probability that this hotspot appears in this location of the entire sky was calculated (using Monte Carlo methods) and found to be 3.4σ (post-trial) which indicates that these results do not have a statistical significance on the discovery level [32–34].

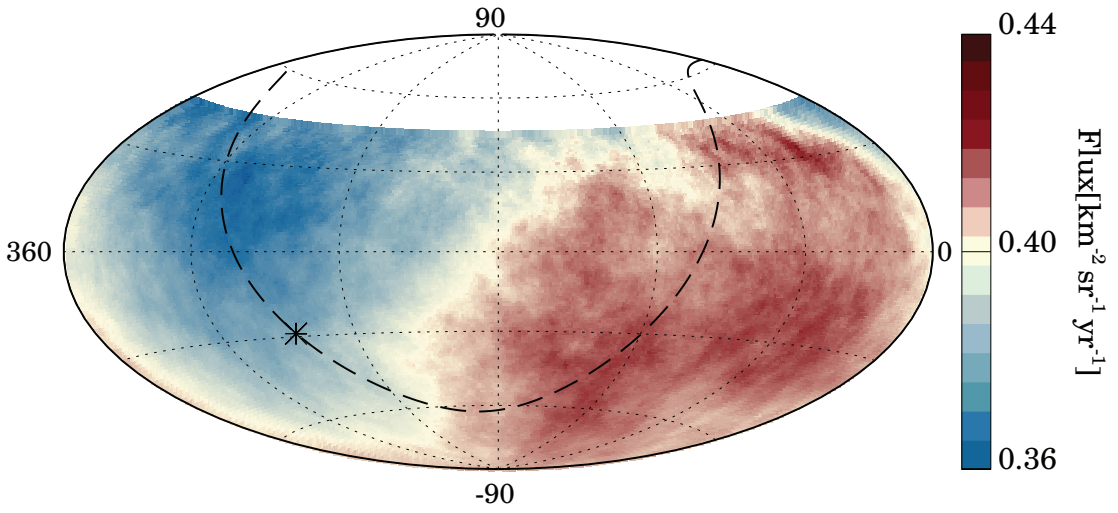


Figure 1.6: A sky map in equatorial coordinates showing a dipole anisotropy for cosmic rays above 8 EeV [29]. The dashed line indicates a Galactic plane and the star is the Galactic centre. The z-axis represents the flux of cosmic ray events.

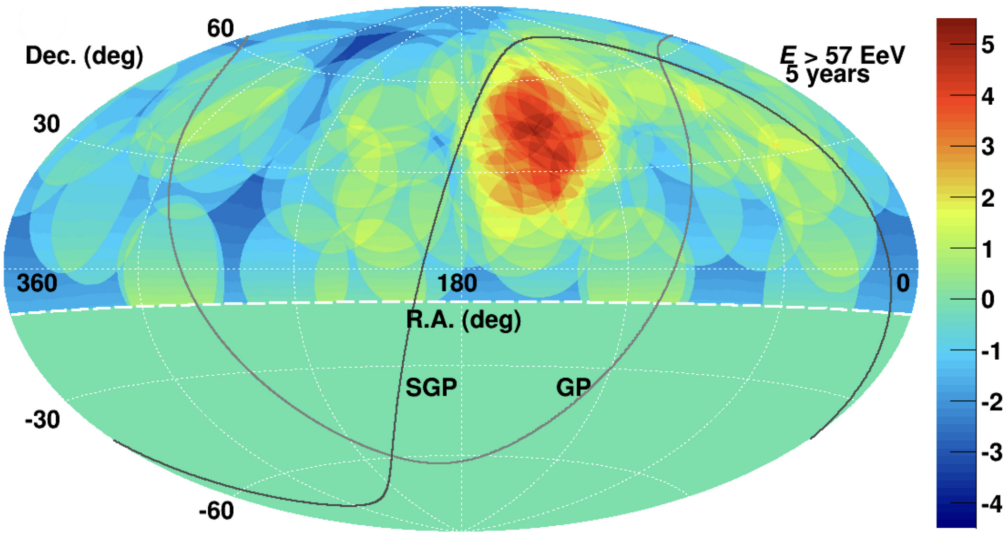


Figure 1.7: A sky map in equatorial coordinates showing an apparent hotspot in the Northern hemisphere for cosmic ray events above 57 EeV measured by Telescope Array (TA) in the Northern Hemisphere [35]. The z-axis represents the statistical significance. The hotspot seen in the Ursa Major cluster was calculated to have local statistical significance of 5.1σ. The Galactic plane (GP) and Supergalactic plane (SGP) are also shown.

1.5 Extensive Air Shower Development

When a cosmic ray reaches Earth, it hits the top of Earth's atmosphere and undergoes a series of interactions. The cosmic ray which initiates these interactions is known as the *primary* particle. Since the Earth's atmosphere contains a variety of different nuclei due to the presence of gas compounds and aerosols, the primary particle initiates a chain of interactions with them resulting in many cascades of secondary particles. These cascades form what is called an extensive air shower (EAS). The secondary particles produced in an EAS are divided into three categories; the electromagnetic component, hadronic component and muonic component. Figure 1.8 shows the kinds of interactions which form the three components in an EAS. Secondary particles produced in an EAS can include neutral or charged pions, kaons, electrons, muons, neutrinos and anti-neutrinos. Figure 1.9 shows a schematic of an air shower with a comparison to the evolution of the number of particles as the shower develops and reaches ground level.

1.5.1 Electromagnetic Interaction

There is a simple model for the complex series of electromagnetic interactions in an air shower - called Heitler's model [37]. It explains how an electromagnetic shower evolves as more and more particles interact in the atmosphere. The particles involved in electromagnetic interactions are electrons, positrons and gamma-ray photons. This component is the most abundant in an air shower and spreads significantly from the shower axis [38]. Moreover, it is also attenuated rather rapidly preventing the bulk of the interactions from reaching the ground [38].

The model begins by having two particles interacting with one another. One of them is the primary particle with energy, E_0 while the other is an atmospheric nucleus. They interact within one interaction length, λ. Then, two particles are produced where each carries half of the energy of the primary particle, $E_0/2$. These two particles than interact with other atmospheric nuclei and forming the start of a large cascade. Mathematically, if n is the number of interaction lengths, then the total number of particles produced at n is 2^n. The atmospheric depth traversed is denoted as d where $d = \lambda \ln 2$ and λ is the interaction length. In this simple model, a gamma-ray with initial energy E_0 interacts with an atmospheric nucleus. This initiates the production of an electron-positron pair. This is shown as

$$\gamma + N \rightarrow N + e^- + e^+ \tag{1.16}$$

where γ is the gamma-ray photon with initial energy E_0, N is an atmospheric nucleus and e^- as well as e^+ are the electron-positron pair. The electron and positron each carry half of the initial energy. They travel a further distance d and

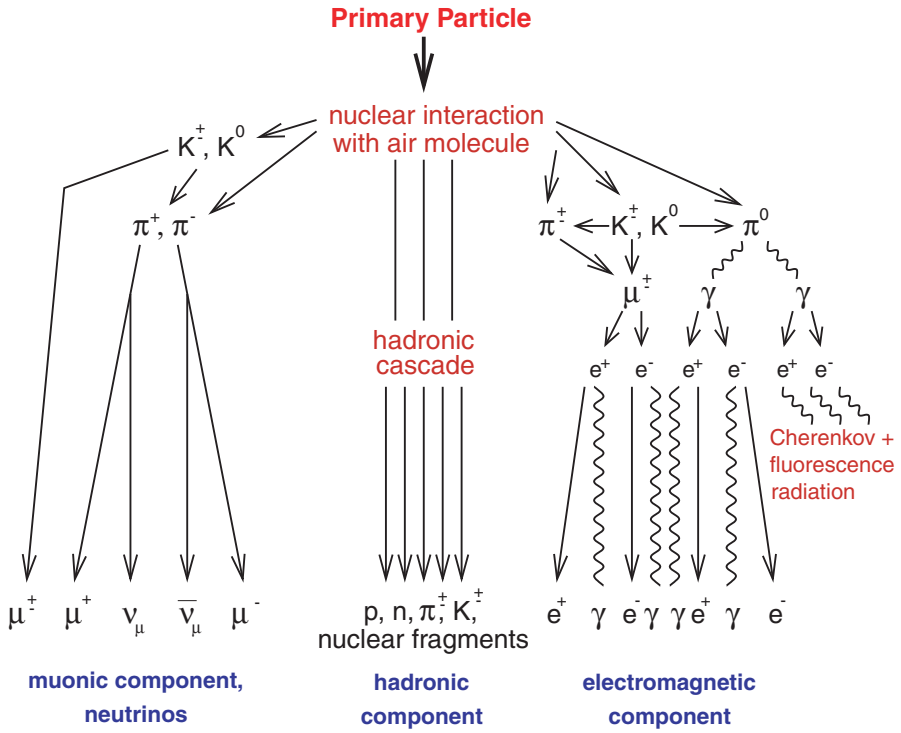


Figure 1.8: A schematic showing three components of an extensive air shower with examples of particle interactions [36]. It is noted that muons can also produce Cherenkov radiation in the atmosphere.

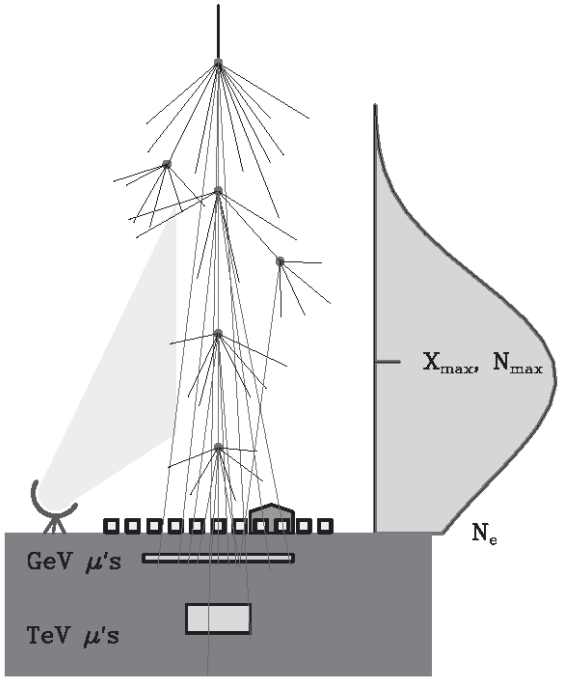


Figure 1.9: Air shower illustration and its comparison with the change in the number of secondary particles produced throughout the shower [4]. The term N_{max} denotes the maximum number of secondary particles produced in the shower. This indicates the depth in the atmosphere at which the shower reaches maximum size, called X_{max}.

lose energy through Bremsstrahlung radiation. This is shown as the following

$$e^{\pm} + N \rightarrow N + e^{\pm} + \gamma \tag{1.17}$$

The new particles are modelled to possess half the energy from the parent particle. The gamma-ray then undergoes the same interaction as before to produce an electron-positron pair. This process repeats itself sequentially to produce many particles, where, after every interaction length, the number of particles is doubled while their energy is halved. This is shown in Figure 1.10. However, there is a limit to these interactions as the number of particles cannot increase indefinitely. At a certain depth in the atmosphere, the number of particles will then reach a maximum. The number of particles produced in the interactions start to decrease after this maximum, since the ionization losses for the electrons dominate over Bremsstrahlung radiation. The energy at which this occurs is called the critical energy, E_c. The value of E_c for electrons in air is $\sim$ 85 MeV [39]. Figure 1.9 shows a curve on the right representing the number of particles and it reaches a maximum value of N_{max} indicative of when the shower reaches maximum size. The depth in the atmosphere at which this happens is widely known as X_{max}. The value of X_{max} can then be measured and used to calculate what the energy of the primary particle would be. An equation showing

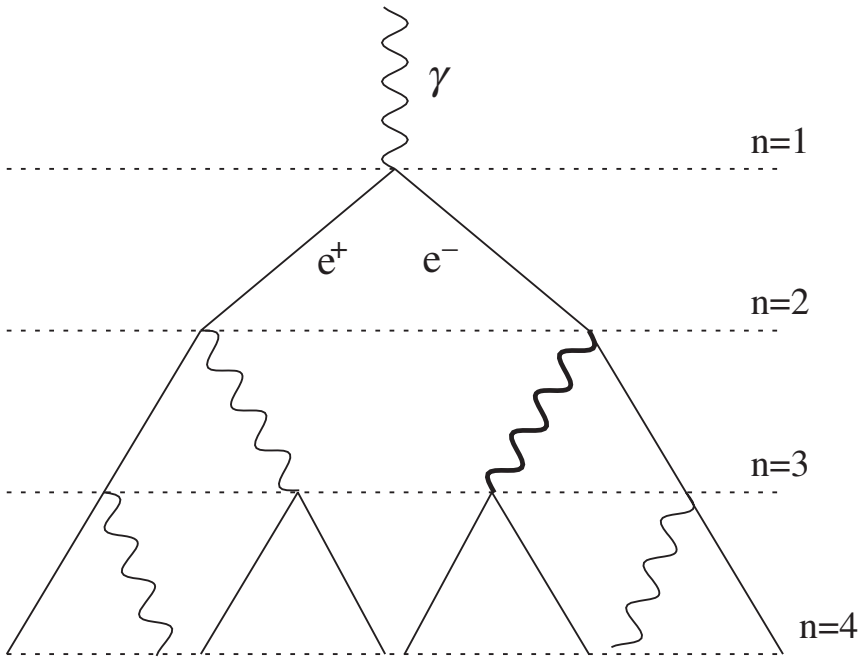


Figure 1.10: A schematic of the start of an electromagnetic cascade as described by the Heitler model [39]. The number of interaction lengths is represented by n. The first interaction that occurs is when a gamma-ray undergoes electron-positron pair production after interacting with an atmospheric nucleus.

how X_{max} is calculated is shown in the following:

$$X^{\gamma}_{max} = \lambda \ln\left(\frac{E_0}{E_c}\right) \tag{1.18}$$

where γ indicates that this only applies to electromagnetic showers. This simple model had been confirmed by detailed simulations and experimental data. The most notable feature is that X^{γ}_{max} is proportional to the natural logarithm of the initial energy E_0.

1.5.2 Hadronic Interaction

Hadronic interactions are initiated by hadrons which are subatomic particles composed of either two or more quarks held together by the strong force. There are many examples of hadrons such as protons, neutrons and pions. In an EAS, hadrons are formed within a relatively close distance to the shower axis [38]. Hadronic interactions can be modelled by Heitler's theory but with a few modifications. From this modified model [39], the cosmic ray will deposit approximately half of its energy to a secondary meson. Then the other half is deposited to another hadron which is subsequently created in the atmosphere. The mesons are hadrons with only two quarks and they include charged and

neutral pions. The modified model starts with an interaction between a primary hadron (p) and an atmospheric nucleus N in the upper atmosphere. The products of these interactions include neutral and charged pions. This is shown in the following as

$$p + N \rightarrow p' + N' + \pi^0, \pi^+, \pi^- + ... \tag{1.19}$$

The neutral pions possess a very low mean lifetime (around 0.897×10^{-16} s [40]) so they rapidly decay to two gamma-rays of equal energy shown as

$$\pi^0 \rightarrow \gamma + \gamma \tag{1.20}$$

and the gamma-rays will then interact with other particles in the atmosphere contributing to the electromagnetic component of the shower. On the other hand, charged pions have a much higher mean lifetime (around 26 ns [41]) than that of neutral pions. Consequently, the probability of interaction in the atmosphere is high causing the production of other particles. The charged pions can interact via the weak interaction by decaying to produce muons and neutrinos. An example of a weak interaction is shown below for a positive pion.

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{1.21}$$

Other types of secondary particles produced can include π^-, μ^- as well as neutrinos, ν_e and ν_τ. This process is repeated until the charged pions drop below their own critical energy, E_c^π similar to the electromagnetic interactions. This is when the pions start to only undergo weak interactions to produce muons. The muons which make up the muonic component lose less energy than electrons allowing them to reach the ground [38]. Hadronic interactions form a complex evolution of the electromagnetic component of the shower due to continuous π^0 decays. Considering only the first generation of electromagnetic showers, an estimation of X_{max} can be done in a simple manner. Figure 1.11 shows a schematic of hadronic interactions up to 3 interaction lengths with some cascades represented by the neutral and charged pions. The depth of shower maximum, X_{max} can be determined from this model. The X_{max} for a cosmic ray proton of primary energy E_0 and interaction length λ is calculated as shown

$$X_{max}^p = X_0 + \lambda \ln\left(\frac{E_0}{3N_{ch}E_c^e}\right) \tag{1.22}$$

where X_0 is the atmospheric depth where the first interaction occurs, N_{ch} is the multiplicity of charged particles produced and E_c^e is the critical energy of the electrons [39]. This result can be compared with the depth of shower maximum for electromagnetic interactions. This provides a simplification to

$$X_{max}^p = X_{max}^\gamma + X_0 - \lambda \ln(3N_{ch}) \tag{1.23}$$

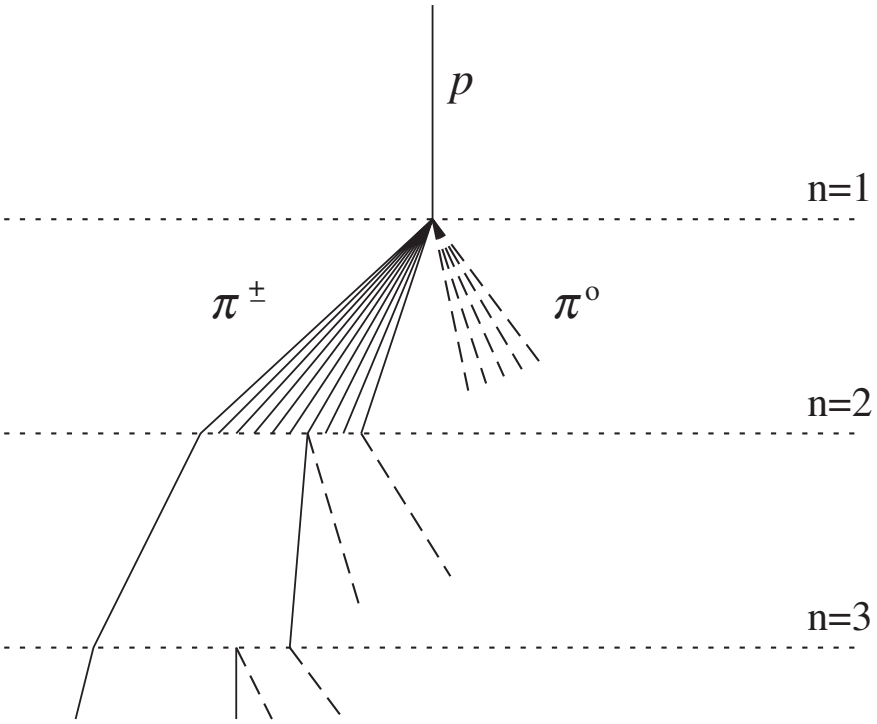


Figure 1.11: A schematic of a Heitler model modification for hadronic interactions [39]. The dashed lines indicate the neutral pions which quickly decay. The decay products then contribute to the electromagnetic component of the shower. Note that not all pion lines are shown after n=2.

Understanding the shower development of cosmic rays from each component can aid in their detection as well as in the determination of their primary energies. Also, the X_{max} value is a very significant observable which aids in the determination of mass composition of cosmic rays.

1.6 Mass Composition

Mass composition measurements of cosmic rays aid in the understanding of the features in the energy spectrum discussed in Section 1.7. The distribution of observed values of X_{max} can provide estimates of the average mass composition with the aid of hadronic models. The study of mass composition for cosmic rays above 10^{18} eV is not possible due to fluctuations in the development of the EAS. Although this is the case, an average mass composition can be determined from shower observables. Most notably, it is shown that X_{max} depends on the atomic mass of the cosmic ray [39]. Even though mass composition studies are limited by insufficient data at the highest energies, the Pierre Auger Observatory (Chapter 3), found evidence that cosmic rays heavier than pure protons are more abundant at the highest energies [42]. In contrast, the Telescope Array located in Utah, USA (Chapter 2) could not exclude the possibility of a pure proton composition at the highest energies [43].

Not only that, there are two conflicting scenarios regarding mass composition for cosmic rays above 10^{19} eV. It was found that if cosmic ray nuclei undergo photodisintegration (a process where a high energy gamma-ray is absorbed and a subatomic particle is emitted), a large fraction of protons is expected. In contrast, if a cosmic ray reaches maximum rigidity as shown from Equation 1.14, a heavier mass composition is expected due to a large gyroradius [44–48]. This emphasises the need for composition-sensitive data to allow for a better interpretation of X_{max} to aid in the understanding of the composition of cosmic rays [49, 50].

1.7 The Energy Spectrum of Cosmic Rays

The energy spectrum of cosmic rays describes the flux of cosmic rays as a function of energy. It contains information about the types of sources which accelerate cosmic rays and how they propagate from the source to Earth [38]. The energy spectrum for cosmic rays follow a series of power laws to a good approximation. It is shown as

$$\frac{dN}{dE} \propto E^{\gamma} \qquad \mathrm{m}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{eV}^{-1} \quad (1.24)$$

where N is the number of particles which possess energies in the interval between E and dE, E is energy and γ is the index of the power law. The behaviour of this power law is illustrated in Figure 1.12. In general, most literature acknowledges the prominence of three main structures called the *knee*, *second knee* and *ankle*.

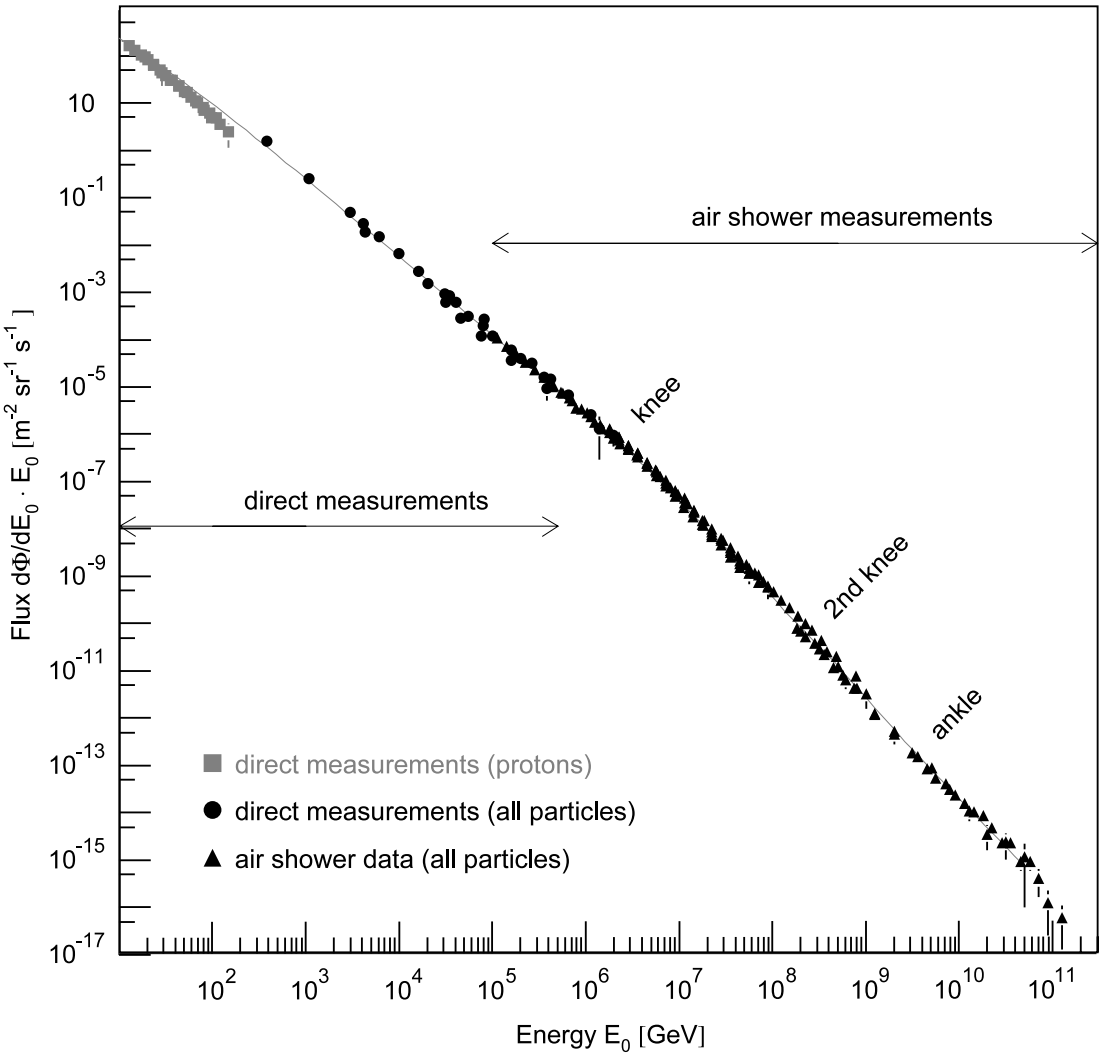


Figure 1.12: The cosmic ray energy spectrum [14]. The x-axis is in units of GeV where 1GeV $\equiv 10^9$ eV. The *knee*, *second knee* and *ankle* are as labelled as well as measurement types discussed in Chapter 2.

There has been observational evidence of a rapid reduction in flux above $10^{19.5}$ eV. This could be interpreted as the maximum acceleration energy due to photo-pion and pair production interactions between cosmic rays with these high energies and photons in the cosmic microwave background. This is called the Greisen-Zatsepin-Kuz'min (GZK) limit [51,52]. This implies that cosmic rays with energies above this suppression could never reach Earth, and if there were such high-energy cosmic rays, they are said to originate from the nearby universe at distances less than 100 Mpc with their sources accelerating them at much higher magnitudes than other known messengers [53]. Another interpretation from this observed reduction in flux would be that sources are incapable of accelerating cosmic rays to such high energies.

The structures in the spectrum can provide hints regarding the types of sources which power cosmic rays and this is governed by changes in the value of γ. It had been found that $\gamma \approx$ -2.7 for energies up to a few PeV (where 1 PeV $\equiv 10^{15}$ eV) and above this, $\gamma \approx$ -3.1 [14]. The steep decrease in flux gives rise to many challenges in cosmic ray detection experiments. Further, there is a lack of agreement between different cosmic ray observatories regarding the handful of high energy cosmic ray events located beyond the *ankle* in the spectra. Therefore, it is essential that the precision and accuracy in the detection of cosmic rays is very high as it affects the derived energy spectrum and the information it conveys.

Chapter 2

Detection of Cosmic Rays

2.1 Detection Methods

Cosmic rays can be detected using either direct or indirect methods, depending on their energy. Direct detection is the detection of primary cosmic rays going through a controlled medium where it interacts with other particles. Direct detection of cosmic rays can only be done feasibly for energies up to $\sim 10^{15}$ eV on high-altitude balloons or space-based observatories [55]. This is because the flux at $\sim 10^{15}$ eV is about one particle per square metre per year and it decreases even more at higher energies making direct detection impractical [14]. Examples of experiments which implemented direct detection include PAMELA which is a space-based cosmic ray observatory launched in 2006 and the AMS-02 launched in 2011 which is currently mounted on the International Space Station [56].

For cosmic rays with energies above 10^{15} eV, indirect detection is more feasible. Indirect detection involves using the atmosphere as a medium for secondary particle interactions which cover a much greater area. This makes detection feasible over reasonable time scales. The interactions between the primary cosmic ray and atmospheric particles initiate large showers of secondary particles and emissions which can be detected by several instruments. The use of ground arrays in combination with fluorescence methods and radio detectors can form a hybrid detector to not only observe the air shower development but determine the energy of primary cosmic rays. Figure 2.1 shows how a combination of different detectors can detect different components and emissions from a cosmic ray air shower.

2.1.1 Ground Arrays

Ground arrays are networks of detectors on the ground, typically with fixed spacing. The aim of ground arrays is to create a large detection area for extensive air showers caused by high energy particles. Since very high energy cosmic rays reach Earth infrequently, a small detection area is unlikely to observe a timely event. Many ground arrays all over the globe have been built with large detection areas to increase the frequency of indirect detections of high energy cosmic rays.

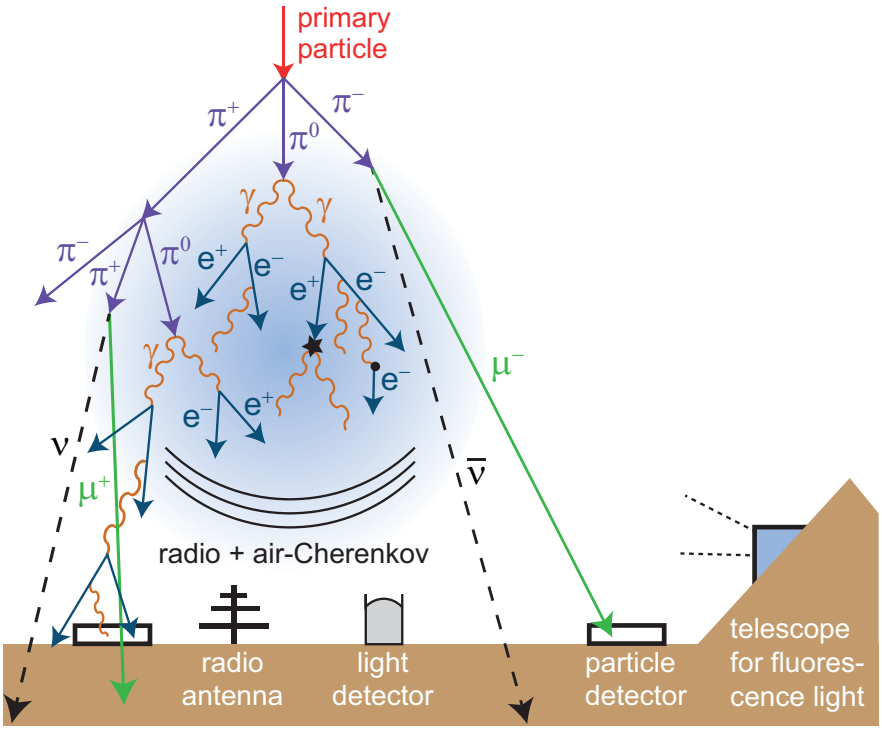


Figure 2.1: A diagram of an air shower observed by a variety of detectors [54].

2.1.1.1 Scintillation Detectors

Scintillation is a process where a material emits light pulses when exposed to ionizing radiation. This process causes the phenomenon of fluorescence in which light is immediately emitted. A scintillation detector or counter is made of either plastic, liquid hydrocarbon or inorganic material and is then coupled to a photomultiplier tube (PMT). The medium in the scintillator is excited by secondary air shower particles, causing the emission of radiation. The PMT serves to amplify faint levels of radiation and perform conversions into electrical signals. It is important to note that the light emitted is proportional to the energy deposited by secondary particles. Additionally, the intensity curve of light as a function of time can depend on the type of incident particle or radiation allowing its identification. Usually, there is a very fast rise in the pulse followed by an exponential decay with a typical decay time [56]. Scintillation detectors are used by many cosmic ray observatories such as Yakutsk (Section 2.2.5) and Telescope Array (Section 2.2.8).

2.1.1.2 Water-Cherenkov Detectors

Due to the extreme energies of primary cosmic rays, the secondary particles produced in a shower possess large amounts of energy, causing them to travel at high speeds. Cherenkov light is emitted when particles travel faster than the speed of light in a dielectric medium. This type of light emission, which is

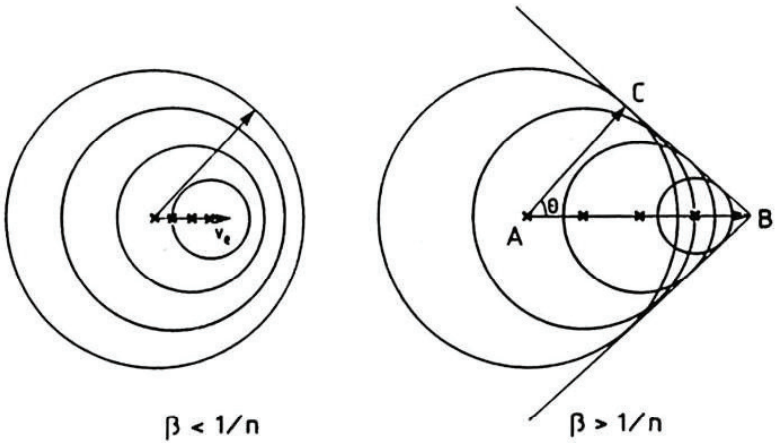


Figure 2.2: Two scenarios illustrating the effects of when a particle travels through a medium at two largely different speeds [57]. The left panel illustrates a scenario of when the particle is travelling much slower than the speed of light in the medium. The wavefronts interfere destructively causing no residual field to exist. The right panel shows a case of when the particle travels faster than the speed of light in the medium resulting the coherence of wavelets forming a plane wave front and the emission of Cherenkov light. The angle θ is shown as the angle of Cherenkov light emission. In this diagram, $\beta = v/c$ where v is the speed of the particle and c is the speed of light in a vacuum. n is the refractive index of the medium.

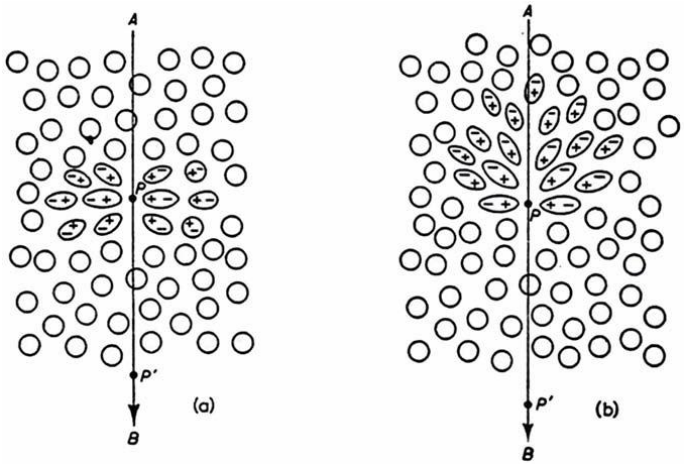


Figure 2.3: Two scenarios, (a) and (b), illustrating the effects of when a particle travels from point A to point B through a medium at two largely different speeds [57]. Panel (a) illustrates a scenario of when the particle is travelling much slower than the speed of light in the medium. Panel (b) shows a case of when the particle travels faster than the speed of light in the medium resulting in Cherenkov light emission. The behaviour of atoms in the medium surrounding the particle in its line of travel in both panels differ at point P.

typically blue in colour, was first discovered by P. Cherenkov [4]. In particular, the light would travel at speed *c/n* where *c* is the speed of light in a vacuum which is $\sim 2.998 \times 10^8$ ms^{-1} [4], while *n* is the refractive index of the medium. As the particle travels through the medium, Cherenkov light is emitted in the shape of a cone at an angle θ to the line of particle trajectory. For a particular angle θ and speed of a particle $v > c/n$, the angle of emission can be obtained as

$$\cos\theta = \frac{1}{\beta n} \tag{2.1}$$

where $\beta = v/c$. As $\beta \to 1$ which represents the ultra-relativistic limit, the maximum angle of Cherenkov light emission is calculated as

$$\theta_{max} = \cos^{-1}(1/n) \tag{2.2}$$

Figure 2.2 shows how the emission angle is obtained through Huygen's construction. For example, in air at sea level, the refractive index, n = 1.00029 which gives a value of $\theta_{max} = 1.3^\circ$. This angle varies between 0.3° and 1.3° in going from 15 km altitude to sea level due to the change in refractive index with air density. In water, n = 1.33 which makes $\theta_{max} = 42^\circ$ [58]. This emission can be explained in detail by understanding the polarisation effect among atoms.

Figure 2.3 shows two scenarios. Panel (a) illustrates the first scenario where the undisturbed atoms in the medium are represented by perfect circles. If a particle travels much slower than light does in the medium, the travelling particle would cause the atoms which are completely surrounding it to elongate since the electric field of the particle distorts the charges. The net electric field would be zero since the charges distorted are symmetric about the particle at every point of its trajectory. However, if the particle travels faster than the speed of light in the medium, there is an asymmetry in the polarised atoms as the distortion only occurs in the trailing edge of the particle trajectory as shown in panel (b). This generates a non-zero electric field causing the atoms in the medium to oscillate. This oscillation causes the emission of Cherenkov light. This light can be used as an indicator of air showers where their intensity can be measured experimentally. An observatory can exploit this technique by deploying water-Cherenkov detectors consisting of tanks filled with purified water (n = 1.33) fitted with photomultiplier tubes to detect secondary air shower particles via the emission of Cherenkov light [59].

2.1.1.3 Atmospheric Cherenkov Detectors

Cherenkov detectors can also be used to view Cherenkov radiation emitted in the atmosphere from extensive air shower particles. The first such atmospheric Cherenkov detectors were used to view the lateral distribution of Cherenkov radiation. An example of this is the Yakutsk experiment located in Yakutsk, Russia where vertically mounted photomultiplier tubes with a diameter of 15

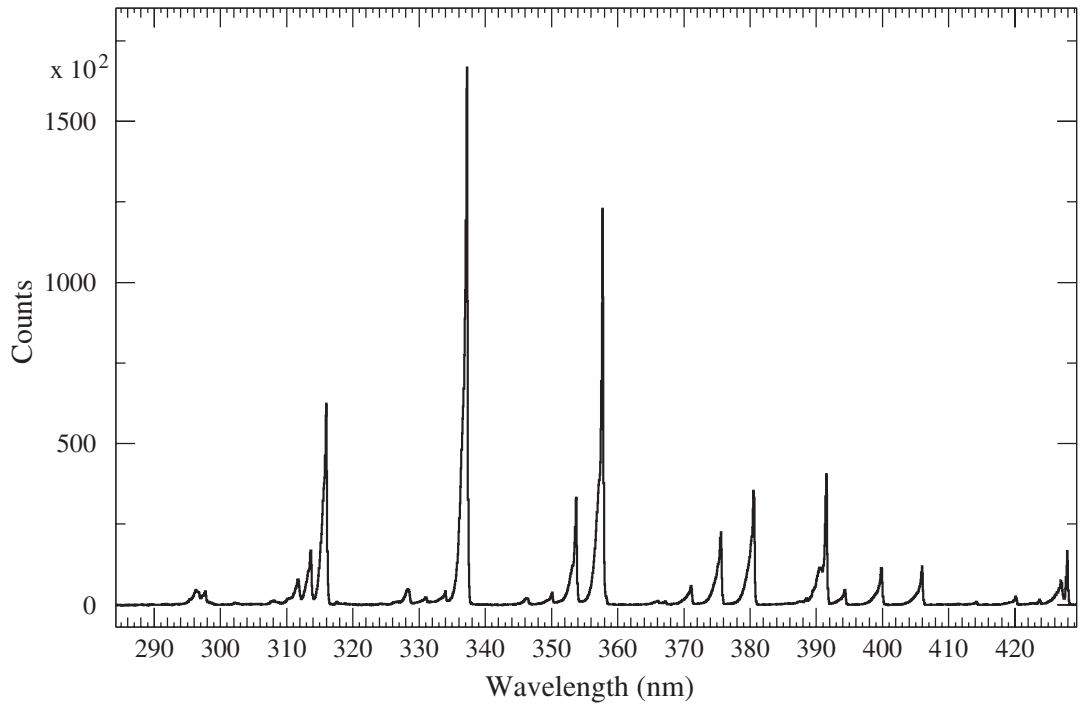


Figure 2.4: The fluorescence spectrum measured in dry air composed of 78 % N_2, 21 % O_2 and 1 % Ar [62]. The pressure of air was measured to be 800 hPa and had a temperature of 293 K.

cm were used. This experiment is detailed in Section 2.2.5. Then, Cherenkov detectors progressed to viewing the angular or longitudinal distribution of Cherenkov radiation by using imaging telescopes. This is known as the Stereoscopic Atmospheric Imaging Technique [60]. One of the reasons why this technique is particularly suited to gamma-ray astronomy is that the Cherenkov angle of emission in the atmosphere is very small, so much so that the Cherenkov light retains the original direction of the primary photon; making the angular resolution high [61]. Examples of experiments which employ this technique include H.E.S.S. (High Energy Stereoscopic System), MAGIC-II (Major Atmospheric Gamma Imaging Cherenkov II) and VERITAS (Very Energetic Radiation Imaging Telescope Array System) [60].

2.1.2 Fluorescence Detectors

Fluorescence detectors serve as another instrumentation method to detect cosmic ray air showers. This technique uses telescopes consisting of mirrors, cameras and photomultiplier tubes to collect light emitted from the shower initiated by a primary cosmic ray. The light is emitted when the energetic secondary particles in the air shower interact with nitrogen molecules in the atmosphere. These molecules enter an excited state and then de-excite, emitting light in the ultraviolet range. Figure 2.4 shows a fluorescence spectrum that was measured by

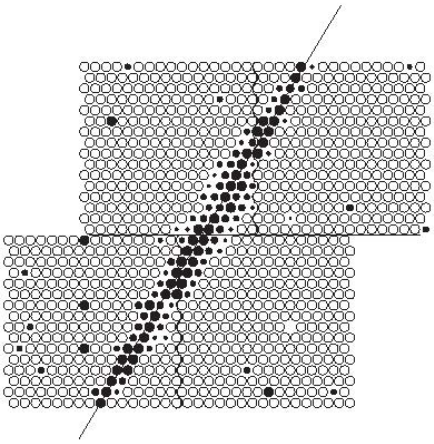


Figure 2.5: Outline of a shower trajectory imaged by photomultiplier tubes [4].

the AIRFLY collaboration in 2007 [62]. There are peaks in the spectrum indicating emission of fluorescence light in the range between 290 nm to 420 nm.

It is possible for particles to travel faster than the speed of light in air and as mentioned in Section 2.1.1.2, the maximum angle of emission for Cherenkov light in air is $1.3°$. This indicates that the Cherenkov light is intensely forward-directed. Additionally, Cherenkov light emits in a cone-like shape at a small angle from the line of trajectory unlike nitrogen fluorescence which emits isotropically. The atmospheric conditions play a vital role in the intensity of fluorescence light detection. Fluorescence can only be detected well on a clear and moonless night. Therefore, the use of fluorescence detectors is limited which results in a low duty-cycle. When there is an air shower, some of the photomultiplier tubes of the fluorescence detector will trigger successively and this aids in the determination of shower trajectory. The timings and signal strength of light can then be observed and recorded.

Figure 2.5 shows how a shower trajectory can be imaged by several photomultiplier tubes in a fluorescence detector. Most importantly, the fluorescence yield is proportional to the energy deposited in the atmosphere as a result of ionisation losses by charged particles. This fluorescence yield is determined by the number of photons emitted when a charged particle goes through one metre of air. The total photon yield between 300 nm and 406 nm in air excited by a 0.85 MeV electron is 3.81 $\pm$ 0.13 photons per metre at a pressure of 1013 hPa and a temperature of 20 °C and this is typically used at most observatories [63]. The fluorescence yield is the most fundamental information for estimating the primary energy of cosmic rays.

2.1.3 Radio Detectors

Radio emission from particle cascades is generated by relativistic electrons and positrons in the electromagnetic component of the shower. The radio detection

method for cosmic rays exploits the process of geosynchrotron emission caused by deflection of electrons and positrons. This is due to the Lorentz force exerted by the Earth's magnetic field, which causes the electrons and positrons to be accelerated in opposite directions while travelling at speeds close to that of c [64]. There are several other physical processes which can cause the emission of radiowaves in cosmic ray air showers. One of them is the Askaryan effect [65]. This effect takes place due to the continuously ionising atmosphere during shower evolution. As ionisation takes place, free electrons propagate with the air shower while the much heavier positive ions are left behind causing a charge imbalance [66]. So, as the shower evolves, the net charge grows with the total number of particles, reaches a maximum, then declines. The propagating negative charges cause a longitudinal current that varies with time, resulting in the emission of radiowaves [66].

The variation in the atmospheric electric field can also cause the emission of radiowaves [54]. During normal weather conditions, the electric field in the atmosphere is a few 100 V/m and its force is neglible to that of the geomagnetic field. However, during thunderstorms, the field can exceed several tens of kV/m. Consequently, the force from the electric field becomes dominant and the acceleration of particles is increased, causing any radio emission to be significantly enhanced [54]. One example of an experiment which exploits these radiowave emissions is the LOFAR Prototype Station (LOPES) experiment. This experiment implements a low-frequency array (LOFAR) technique using dipole antennas with digital electronics [54]. Another example of a radio detector is the Taiwan Astroparticle Radiowave Observatory For Geo-synchrotron Emissions (TAROGE) [67]. Figure 2.6 shows an example of a radio detector at TAROGE and a schematic showing how the radio emissions can be detected.

The use of radio detection methods alone makes it difficult to determine the exposure and cosmic ray flux [68]. To fully exploit the radio detection techniques, understanding the emission process is essential. This can be done by creating simulations of radio emission in an air shower and comparing the results with experimental data. This can also be accompanied by hybrid measurements to obtain an even higher accuracy of cosmic ray air shower reconstruction and determination of energy.

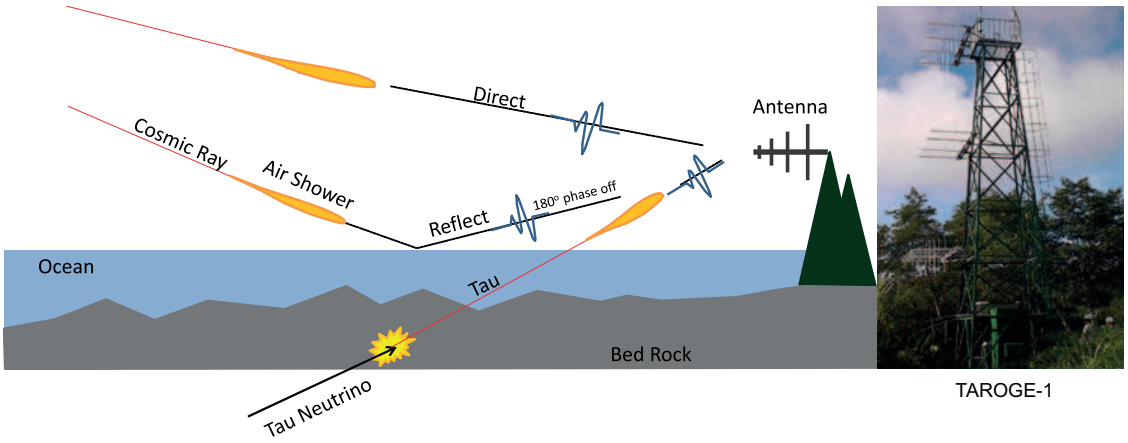


Figure 2.6: A schematic and photo illustrating an antenna of the TAROGE experiment [67]. It is located on a tall mountain close to the coast overlooking the ocean. Radio emissions can either be detected directly or from reflections off of the surface of the ocean. The tau particle originates from charged-current interactions of tau neutrinos.

2.2 Past and Current Cosmic Ray Experiments

Many experiments have been conducted over the years aiming to understand cosmic rays. This section presents a description of past and current experiments which are well-known.

2.2.1 Volcano Ranch

The first giant air shower array called Volcano Ranch was constructed in Albuquerque, New Mexico. A group led by Bruno Rossi from the Massachusetts Institute of Technology operated the experiment between 1959 and 1963 [69, 70]. The array was made up of 19 plastic scintillation counters with an area of 3.3 m^2 each. The spacing between each detector was 442 m, making a total area of 2.2 km^2. Then in 1960 the spacing increased to 884 m covering an area of 8.1 km^2. Every scintillation counter was accompanied with a 12.7 cm diameter photomultiplier tube. Oscilloscopes were used to display the signals produced aiding in measurements of pulse amplitude and relative arrival time. Figure 2.7 shows a layout of the Volcano Ranch array.

Data from this array yielded the first measurements of cosmic ray energies above 10^{18} eV. This array gave the earliest indication of a flattening of the spectrum above this energy [73]. It was also found that the high energy events detected did not exhibit a noticeable anisotropy [73]. The event with the largest cosmic ray energy ever recorded by Volcano Ranch was found to be 1.4×10^{20} eV [72]. It remains as one of the highest energy events ever detected to this day. Figure 2.8 shows a map of this event in detail, with the position of the shower core, defined as the intersection between the axis of the shower and the ground.

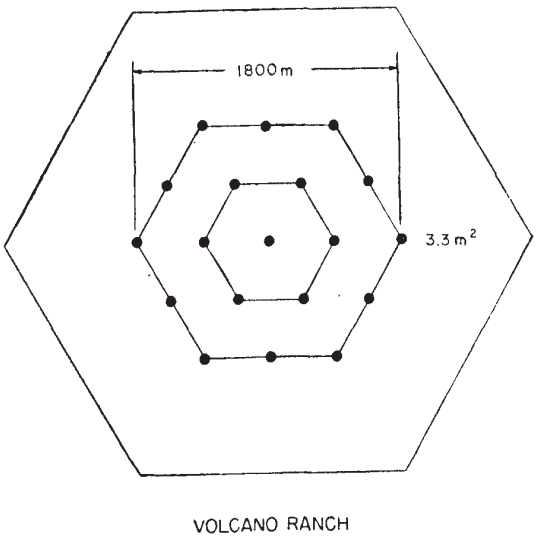


Figure 2.7: The Volcano Ranch array layout [71]. The black circles represent the 19 scintillation counters. The expansion is shown as the outer most hexagon.

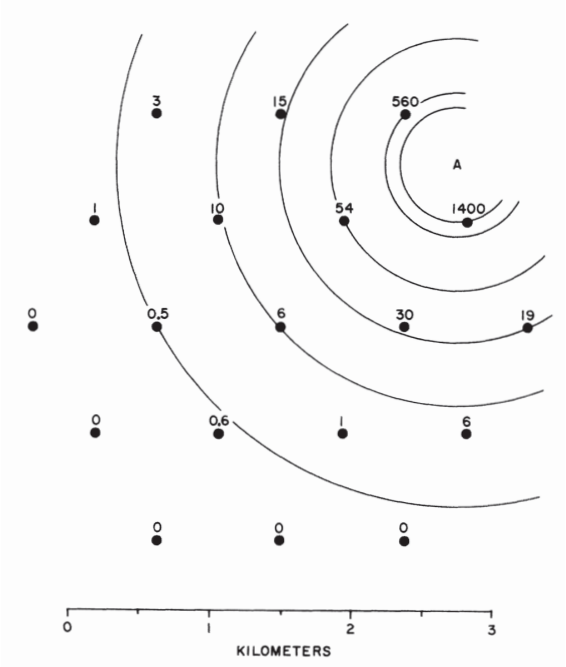


Figure 2.8: Particle densities and shower core location measured by the Volcano Ranch array from a very high energy event (1.4×10^{20} eV) [72]. The black dots represent the 3.3 m^2 scintillation counters. The numbers labelling each one represent particle densities in units of particles per m^2. The concentric circles represent contours verifying the core location for this event located at A.

Then, detection of fluorescence light and Cherenkov emission was achieved using an optical system that was built separately and in coincidence with the Volcano Ranch array [74]. The aim was to detect the emission from extensive air showers using optical systems and to see if the signals agree with those observed by the scintillation counters. The success of these synchronised detection methods led to the construction of the Fly's Eye observatory.

2.2.2 Fly's Eye

The Fly's Eye observatory consisted of fluorescence detectors positioned at two stations; Fly's Eye I and Fly's Eye II. This observatory was constructed for the purpose of detecting fluorescence light emitted from cosmic ray air showers. This observatory was located at the Dugway Proving Ground in Western Utah and operated from 1981 to 1992. It was the first cosmic ray detector to exclusively use the air fluorescence technique to successfully measure fluorescence from air showers. Fly's Eye I and Fly's Eye II were separated by 3.3 km [75].

Fly's Eye I consisted of 67 spherical section mirrors with a diameter of ~ 1.58 m, Winston light collectors, photomultiplier tubes (PMTs) and data acquisition electronics. The PMTs and Winston light collectors were hexagonally packed in groups of either 12 or 14 in the focal plane of each mirror. Fly's Eye II is a smaller array of identical units, 8 in all, with 120 total light collectors and PMTs. The projection of each hexagonal light sensing light collector and PMT onto the celestial sphere resemble the compound eye of an insect which gives it the name, Fly's Eye [75].

The Fly's Eye had taken measurements of cosmic ray energies near 3×10^{18} eV. Not only that, it contributed to the knowledge on mass composition of cosmic rays through measurements of the depth of shower maximum, X_{max}. Also, the highest energy event detected had an energy of 320 ± 38 (statistical) $\pm$ 85 (systematic) EeV [76]. There were several sources identified that might have been associated with this event. However, it was found that the sources could not have a magnetic field strong enough to accelerate the cosmic ray to this energy, leaving the origin to remain a mystery [76].

2.2.3 High Resolution Fly's Eye (HiRes)

The High-Resolution Fly's Eye cosmic ray detector (HiRes) consist of two FD stations called HiRes I and HiRes II which are 12.6 km apart located at the U.S. Army Dugway Proving Ground in Utah. This experiment was operational from 1997 to 2006 [77]. It served as an extension to the Fly's Eye experiment aimed to increase the detection rate of cosmic ray events above 10^{19} eV, improve the depth of shower maximum resolution and enhance the angular resolution and acceptance in the EeV range to increase sensitivity to point sources [77]. The FD stations consisted of 22 telescope modules at HiRes I and 42 modules at HiRes II pointing at different parts of the sky.

Figure 2.9: An aerial view of the Fly's Eye I detector located on the top of Granite Mountain, Dugway, Utah [75].

HiRes was the first cosmic ray detector to observe the GZK limit which is the proposed upper limit to the cosmic ray energy spectrum, known to be 6×10^{19} eV. Throughout its nine years of operation during clear, moonless nights, HiRes had detected many events between $10^{17.2}$ to $10^{20.5}$ eV. Most notably, there are two breaks in the energy spectrum obtained by HiRes consistent with the GZK limit and the ankle. The statistical significance of the break identified with the GZK limit is 5σ. The energy representing the GZK limit was measured to be 5.6 ± 0.5 (statistical) $\pm$ 0.9 (systematic) $\times 10^{19}$ eV [78]. HiRes also determined that above 10^{18} eV, the results are consistent with a light, protonic composition [78].

2.2.4 Haverah Park

The Haverah Park array was constructed in northern England. It was built as a collaboration between Universities of Durham, Leeds and Nottingham as well as the Imperial College. This experiment was in operation between 1968 and 1987 [79]. The array consisted of water-Cherenkov detectors located 220 m above sea level made up galvanised iron tanks. The detectors were filled with deionised water to inhibit the growth of fungus. Each detector had a 5-inch (12.7cm) diameter photomultiplier tube fitted within them. The Haverah Park array layout is shown in Figure 2.10. The data collected from the Haverah Park array showed that the measurements in lateral distribution function (LDF)

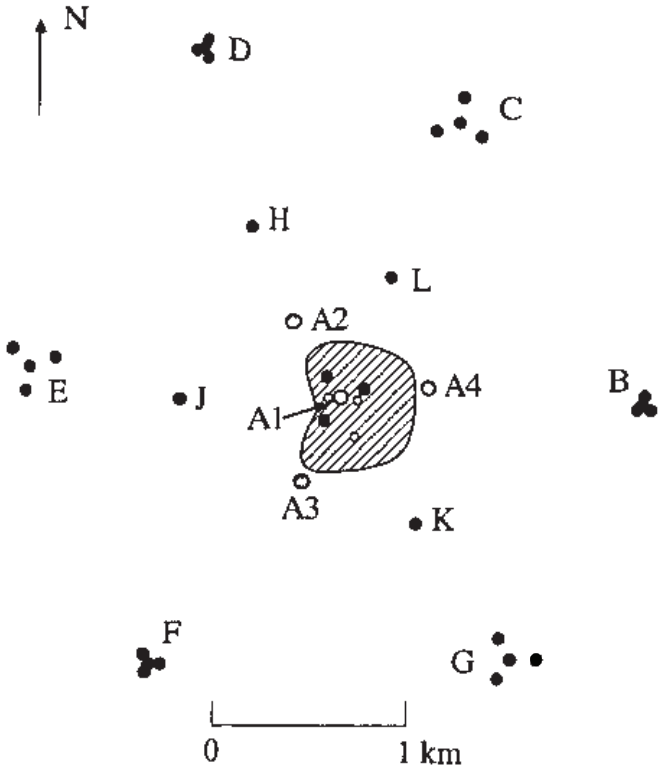


Figure 2.10: A schematic showing the Haverah Park array layout [79]. Circles marked A1 up to A4 are water-Cherenkov detectors covering a total area of 34 m^2. The three black circles represent detectors placed 150 m from A1 and they have an area of 9 m^2. The shaded area contains a lattice of 30 $\times$ 1 m^2 detectors which makes up an 'infilled' array. The other three open circles surrounding A1 are muon detectors. The sub-arrays made up of clusters of detectors marked B up to G make up 4 $\times$ 13.5 m^2 detectors. At H is a detector covering an area of 13.5 m^2, and at J,K and L are detectors which cover 2.25 m^2.

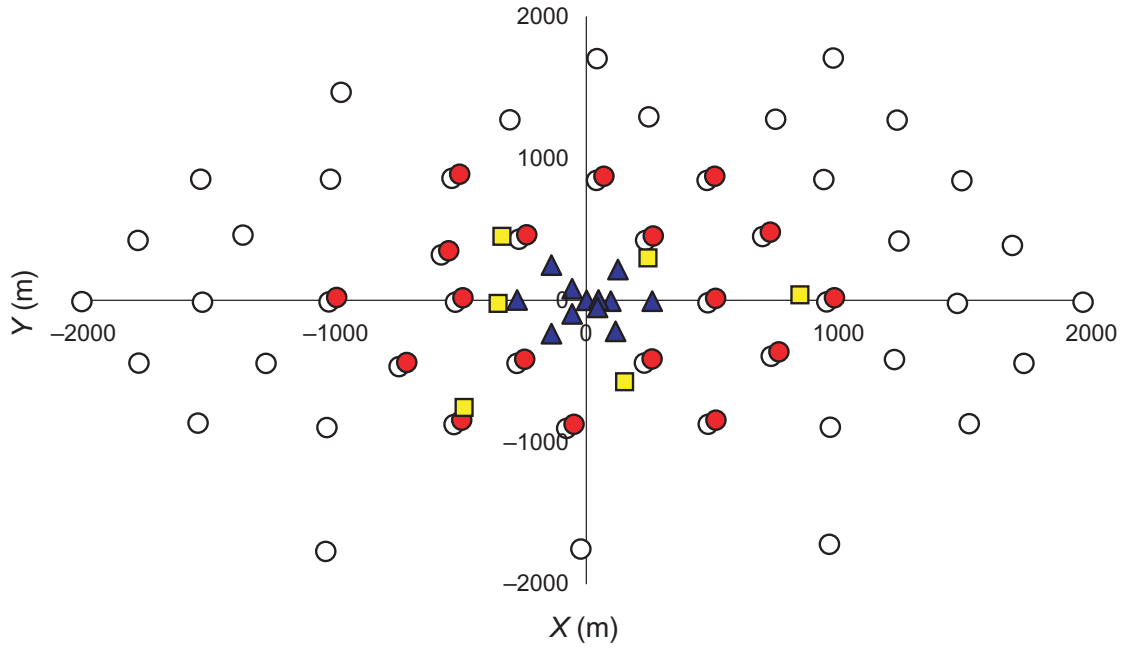


Figure 2.11: Map of the Yakutsk array [81]. Open circles are charged particle detectors, the red circles and blue triangles represent the atmospheric Cherenkov light detectors with a spacing of $\sim$ 500 m and 50-200 m respectively while the yellow squares are muon detectors.

and energy calibration were in good agreement with Volcano Ranch and another experiment at Yakutsk [80].

2.2.5 Yakutsk

The Yakutsk array is an array located at Oktyomtsy near Yakutsk, Russia. It had covered an area of 17 km^2 from 1974 to 1991. Later, the array contracted and it currently only covers 8 km^2. This was done to obtain detailed measurements of lateral distributions from cosmic ray air showers. Currently, Yakutsk contains 58 ground-based and 6 underground scintillation detector stations, with 48 detectors consisting of PMTs in shuttered housing for observing atmospheric Cherenkov light as well as 6 radio detectors with a frequency of 32 MHz. [81, 82]. Figure 2.11 shows a map of the Yakutsk array. There have been many contributions made by Yakutsk in the study of the cosmic ray energy spectrum in the range of 10^{15}-10^{20} eV. The array is currently undergoing a modernization process to enhance the accuracy of measurements which include the determination of mass composition of cosmic rays [83].

2.2.6 Sydney University Giant Air-shower Recorder (SUGAR)

The Sydney University Giant Air-shower Recorder (SUGAR) was a giant array built by the University of Sydney located near Narribri which is in northern New South Wales, Australia. The array consisted of 54 autonomous detectors deployed over an area of more than 60 km^2 with its own local power source. Each detector which is made up by a pair of 6 m^2 scintillators were buried 1.7 m below ground level to reduce the background rate and increase the sensitivity of muon detection. This allows for a more sensitive observation to the muon component of an air shower. SUGAR operated between 1968 and 1979 and the large area coverage aided in the detection of air showers initiated by very high energy cosmic rays at the expense of resolution [84]. Figure 2.12 shows a map of the array. The spacing between each detector in SUGAR was $\sim$ 1.61 km and this was proved to be too great as even in the largest events, the number of stations that recorded the particle densities from air showers was too small [58]. This caused a difficulty in the determination of signals from the stations which dictate the energy of the primary cosmic ray.

In addition, since the precision achieved by SUGAR was not as high compared to other arrays, the data which recorded events with energy above 5×10^{19} eV was not very reliable. In spite of this, it is important to note that since SUGAR was the first giant air shower array built in the Southern Hemisphere, this initiated a study on whether there is a preference for arrival directions of cosmic rays to be close to the supergalactic plane, as this was found by the Northern Hemisphere arrays [85]. It was found that this is not the case for events detected by SUGAR [86]. In spite of this, the novel method of autonomous data recording paved the way for the construction of large arrays such as the Pierre Auger Observatory (Chapter 3) which takes advantage of 30 years of technological development [58].

2.2.7 Akeno Giant Air Shower Array (AGASA)

The Akeno Giant Air Shower Array (AGASA) is another example of an experiment dedicated to detecting air showers initiated by high energy cosmic rays using scintillation detectors. AGASA consists of an array with 111 scintillation detectors covering an area of about 100 km^2 located in Akeno about 130 km west of Tokyo, Japan [88–90] It was the largest cosmic ray observatory during its period of operation between 1990 and 2004. Each surface detector consists of a plastic scintillator which detects charged particle densities on the surface and contains muon counters underneath to measure the muon component of an air shower. Figure 2.13 shows a map of the AGASA array. Each scintillation detector is separated by $\sim$ 1 km and sequentially connected with a pair of optical fibre cables [88]. AGASA had been divided into branches to aid data acquisition and trigger mechanisms. They include the Akeno Branch (AB), the Sudana Branch (SB), the Takane Branch (TB) and the Nagasaka Branch (NB).

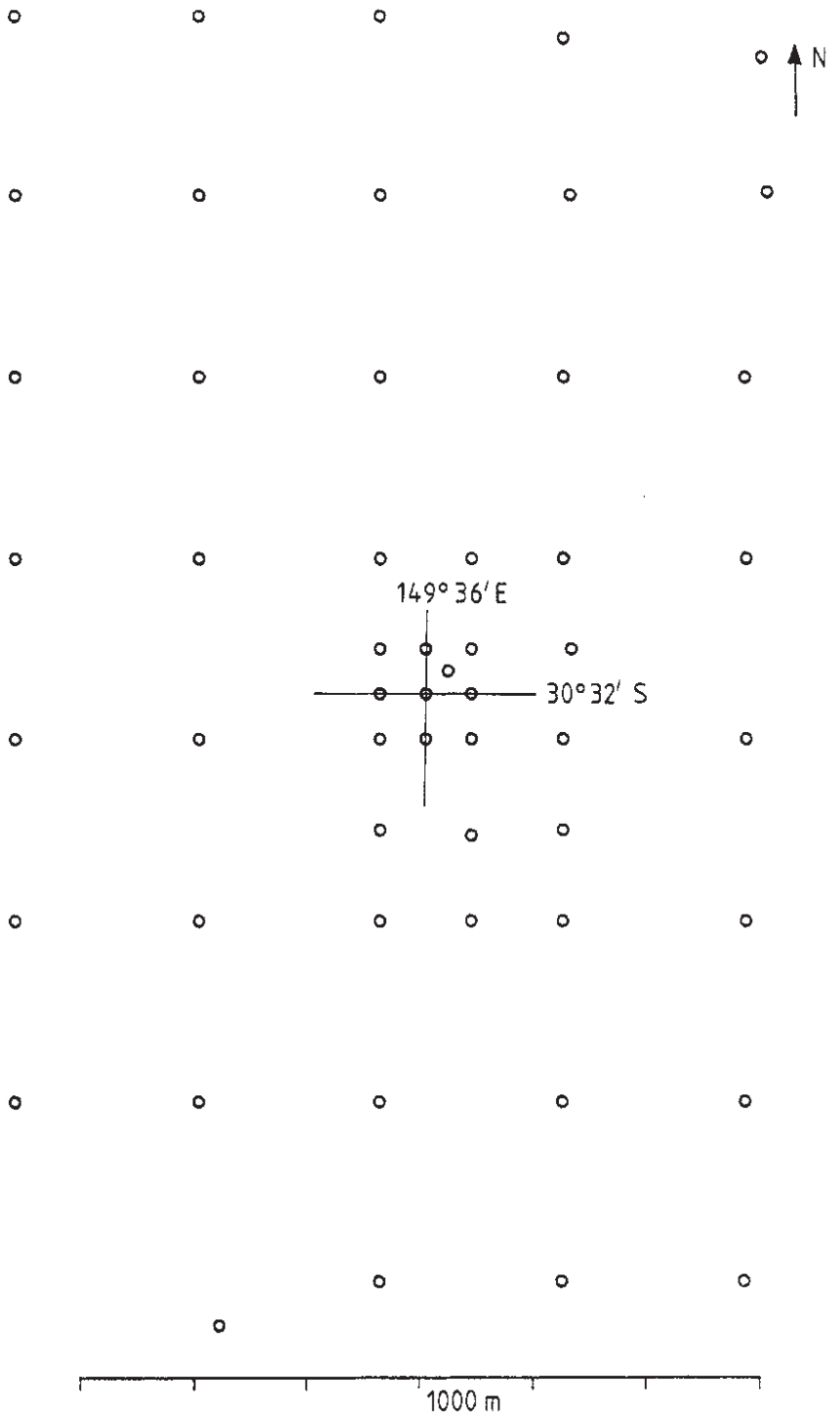


Figure 2.12: Map showing the array of SUGAR [87]. Each circle represents the 54 autonomous detectors. Most of the detectors have a spacing of $\sim$ 1.61 km. In addition, each detector had two liquid scintillator tanks buried underground, 50m apart in the north-south direction. The cross represents the central position of SUGAR with coordinates 30° 32′ S and 149° 36′ E.

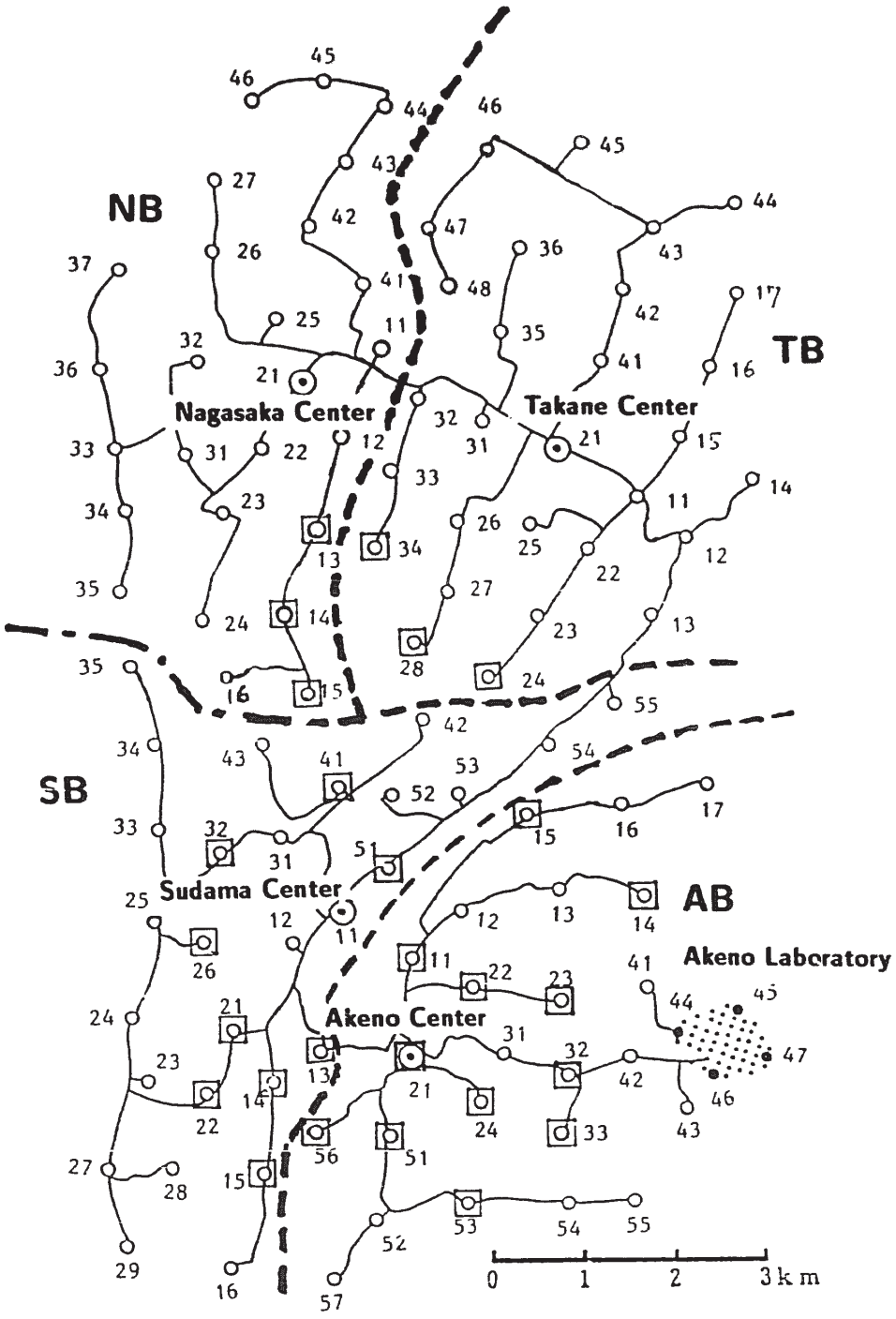


Figure 2.13: A map of AGASA [88]. Open circles represent the scintillation detectors each with an area of 2.2 m^2 and the open squares represent muon detectors. Solid lines show the route of the fibre optic cables used for data transfer. Dotted lines show the borders between the four branches. Big open circles with dots in the middle show the centres of the respective branches. There are four closed circles which represent scintillation counters each with an area of 1 m^2 forming the 1 km^2 array.

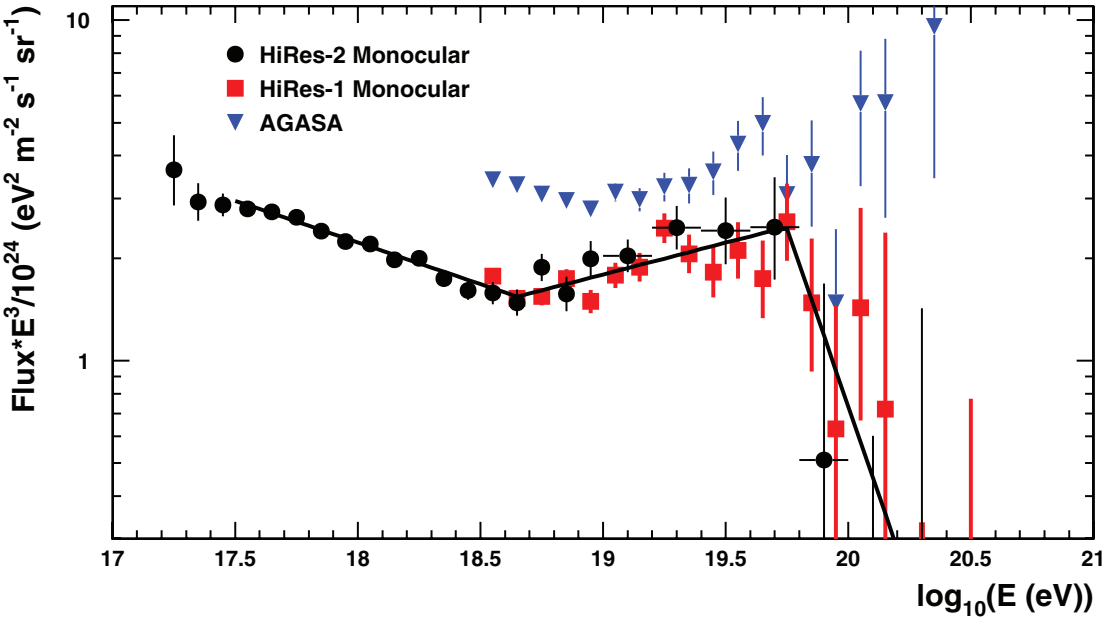


Figure 2.14: A comparison of both energy spectra for AGASA and HiRes I and II [78]. The highest two energy bins in the spectra of HiRes I and II are empty. The lines in those bins are shown as the 68% confidence levels. The black trend line is a binned maximum likelihood fit applied to the data from both HiRes I and II. There is no suppression observed by AGASA at the expected GZK limit.

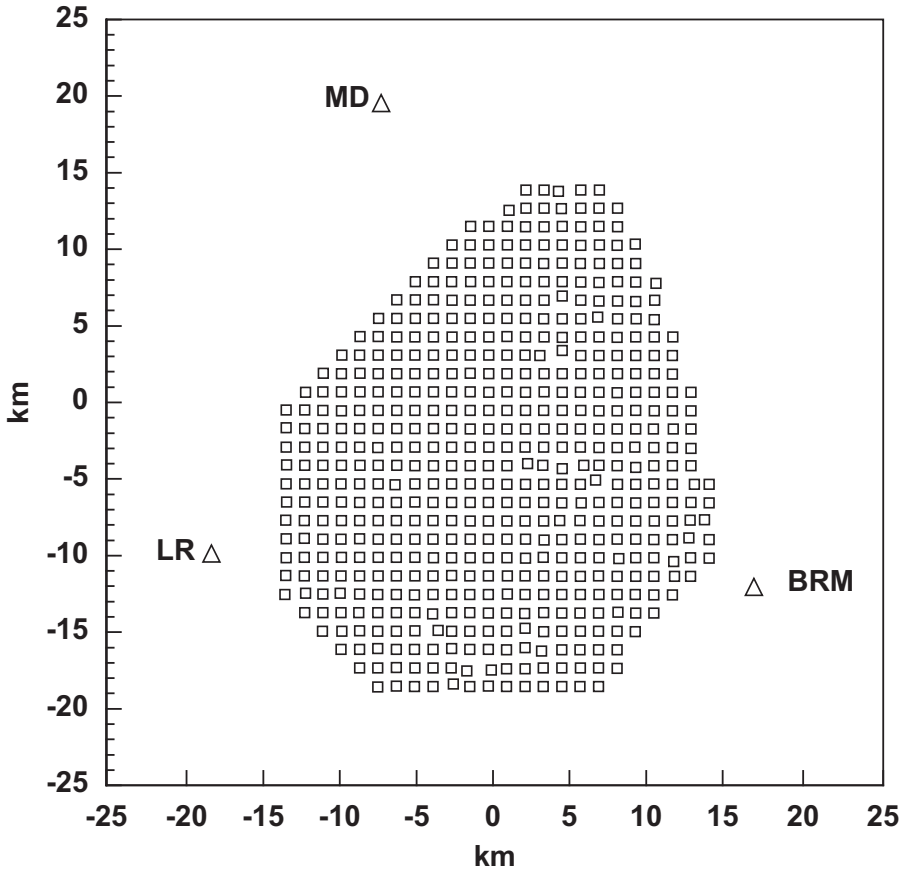


Figure 2.15: Map of the Telescope Array in Utah, USA [91]. The squares represent the 507 scintillators while the triangles are the FDs. They are the Middle Drum (MD), Black Rock Mesa (BRM) and Long Ridge(LR) FD sites.

Then in 1995 these four branches merged providing a significant increase in the effective area.

In 1998, it was published that AGASA had observed a handful of cosmic ray events with energies above 10^{20} eV [89]. This result was unexpected since it generated an energy spectrum that was inconsistent with the GZK limit. In 2004, this result was re-visited and the events were plotted on an energy spectrum as shown in Figure 2.14. This spectrum was compared with the one of HiRes and it is observed that there is a clear lack of suppression in the AGASA spectrum past the GZK limit located at $\log(\text{E (eV)}) \sim 19.8$. This disagreement inspired the construction of a new generation of detectors such as the Telescope Array and the Pierre Auger Observatory (Chapter 3).

2.2.8 The Telescope Array (TA)

The Telescope Array (TA), based in Utah, USA is currently the second largest cosmic ray observatory in the world and the largest in the Northern Hemisphere.

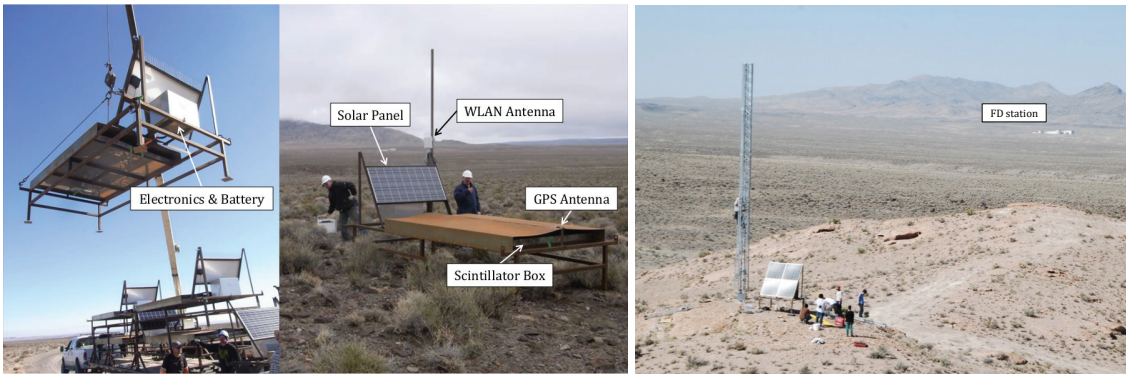


Figure 2.16: The two left photos show a deployed TA scintillation detector in the field [92]. An electronics unit is stored under the solar panel and the scintillator is placed on a platform under a roof. The photo on the right shows a view of the FD station from the Middle Drum communication tower for the SD TALE array.

TA is designed to observe ultra high energy cosmic rays with energies in excess of 10^{18} eV [43]. It has been operating since 2007 and employs a hybrid method of detection as it includes surface detectors and fluorescence detectors. A map of the array is shown in Figure 2.15. There are 507 scintillation detectors with a spacing of 1.2 km in triangular positions covering an area of approximately 700 km^2 [91]. This is approximately 7 times larger than the coverage of AGASA [93]. The SD is made up of two-layered plastic scintillators and each of them have an area of 3 m^2. There are 36 fluorescence telescopes distributed among three FD stations placed at the perimeter of the SD array on three sites called Black Rock Mesa (BRM), Long Ridge (LM) and Middle Drum (MD). Figure 2.16 shows a photo of one of the scintillation detectors as part of the array as well as the FD station overlooking the array at its perimeter.

As mentioned in Chapter 1, TA had detected a cosmic ray hotspot with an energies greater than 57 EeV located in direction of the Ursa Major cluster positioned approximately 19° from the supergalactic plane. From using the data throughout an 11-year acquisition, it was found that the hotspot had a chance probability in an isotropic sky of 2.9σ (post-trial) [32] which is far from discovery level. TA had also observed a suppression with a GZK limit of 5.4×10^{19} eV [94].

An extension to the instrumentation at TA is the Telescope Array Low-energy Extension (TALE). It serves to detect cosmic ray air showers with lower energies, typically below $10^{16.5}$ eV with reasonable resolution. In 2012, the TALE extension was implemented by the installation of 10 FDs pointing at a higher elevation, to observe lower energy cosmic rays than the TA FDs with an expected event rate of 5000 hybrid events per year [95]. Additionally, at lower energies, TALE uses air-Cherenkov light instead of fluorescence light for air-shower observations.

TA is currently undergoing an upgrade called TAx4 which began in 2015 [96]. It entails the extension of its SD array to reach an area coverage of 3000 km^2 as well as construction of new FDs, similar to that of the Pierre Auger Observatory

(Chapter 3). The aims of TAx4 are to clarify the source of the anisotropy hotspot detected, gain a better understanding of the energy spectrum through the increased rate of event detection and determine the mass composition of cosmic rays [96]. Figure 2.17 shows the TAx4 extension. There is an ongoing deployment of an additional 500 detectors to extend the SD while increasing the spacing to 2.08 km. Currently, there are 257 assembled SDs which have been deployed in February and March of 2019 [97]. Also, four FD telescopes have been built in the north site and first light was observed on 16 February 2018 [97]. Eight telescopes will be built in the Black Rock site and its construction is ongoing. The field of view of new FDs overlaps the detection area of new SDs to observe hybrid events. Test runs of data acquisition at each communication tower were started from the end of April 2019. Consequently, the TAx4 upgrade will enable the hotspot and energy spectrum anisotropy to be studied with 4 times larger statistics and three times the number of hybrid events at the highest energies. Mass composition can also be studied extensively from these hybrid events [97].

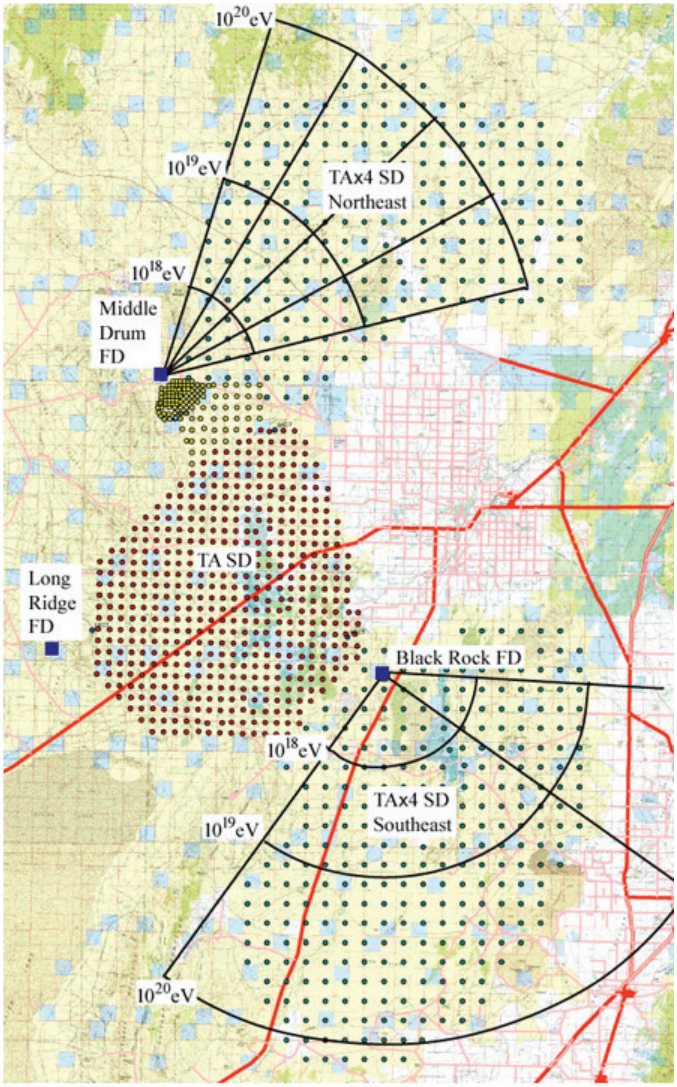


Figure 2.17: A map showing the TAx4 extension in the northeast and southeast part of the TA array [97]. The TALE extension is shown as the yellow circles in between the TA SD and TAx4 SD Northeast. The two fan shapes drawn with black lines show the expected field of view of the TAx4 FDs.

Chapter 3

The Pierre Auger Observatory

The Pierre Auger Observatory is the largest cosmic ray observatory in the world and it is composed of various detectors aimed to detect and measure the energies of high energy cosmic rays. It is located near Malargue which is in the province of Mendoza in Argentina. The observatory was named after a renowned physicist Pierre Auger, who discovered the phenomenon of an extensive air shower being generated by a single cosmic ray bombarding the top of Earth's atmosphere. He also made observations indicating that cosmic rays were reaching energies in the order of 10^{15} eV [99]. The high energy nature of cosmic rays raised many questions about the origins and physical processes which result in particles of such high energies in Nature. The idea and initiative to build the Pierre Auger Observatory was established by James Cronin from the University of Chicago and Alan Watson from the University of Leeds [98]. The Pierre Auger Observatory covers an area of 3000 km^2 and a map of the observatory is shown in Figure 3.1.

Two different types of detectors are used to detect cosmic rays; the use of over 1660 instrumented water tanks called the surface detector (SD) as well as 27 telescopes which make up the fluorescence detector (FD). Both instruments provide a hybrid design which is a key feature used to detect ultra-high energy cosmic rays simultaneously [98]. There are also a number of laser facilities and other instruments used to monitor atmospheric conditions. In particular, they are used to improve the accuracy of air shower detection by taking atmospheric changes into account as they affect shower development and light transmission. The Pierre Auger Observatory has been operational since 2004 with its detectors observing many millions of events all over the array. The purpose of a hybrid design is to provide important cross-checks, to enable measurement redundancies and to eliminate model-dependence during data acquisition [98, 100].

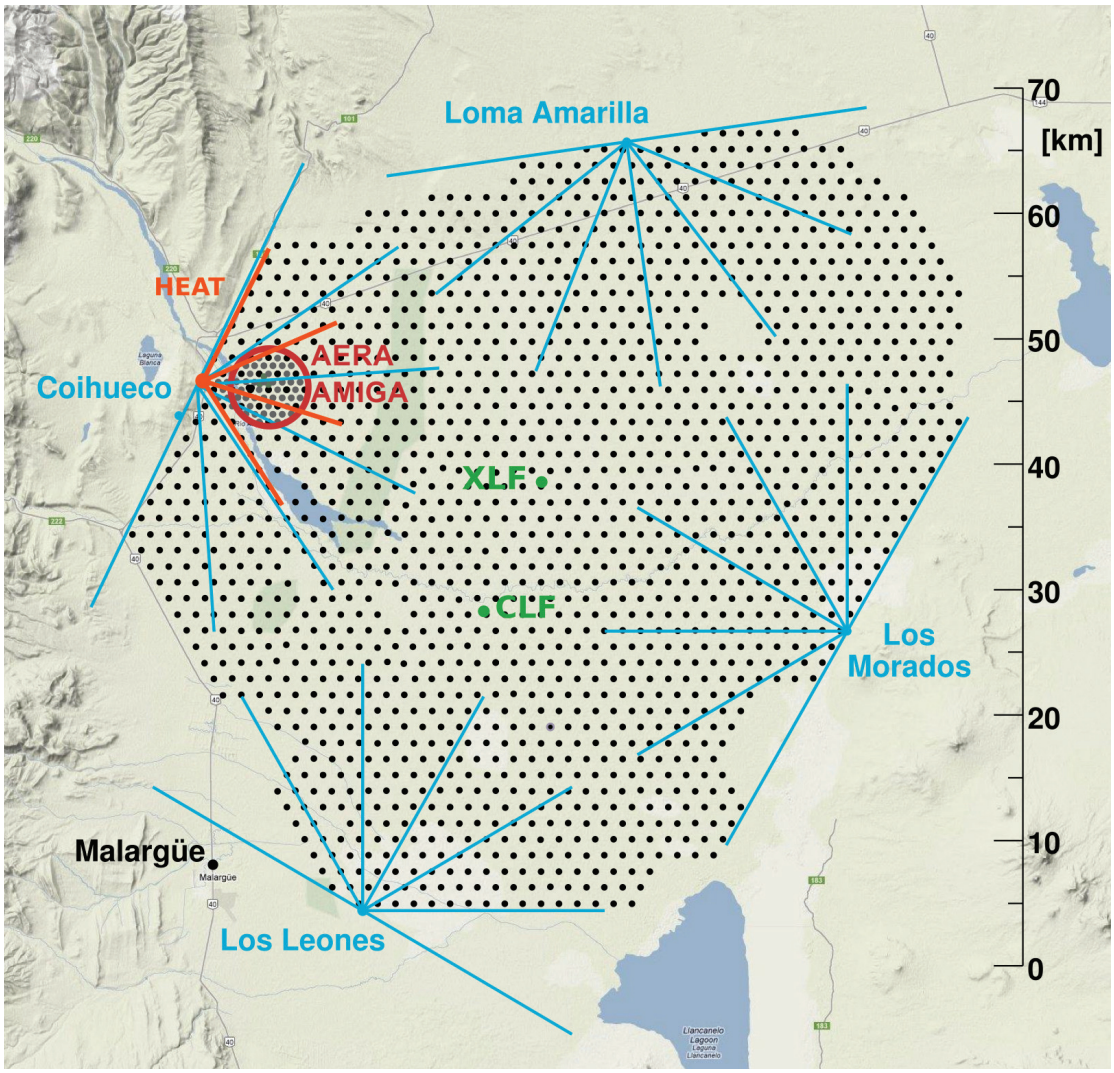


Figure 3.1: Map of the Pierre Auger Observatory [12, 98]. Each black dot represents a water-Cherenkov detector making up the surface detector while the four sites representing a fluorescence detector are shown as Los Leones, Los Morados, Loma Amarilla and Coihueco. Each site has 6 telescopes each, governing the area as depicted by the space between the blue lines. The XLF and CLF are laser facilities for calibration. HEAT are high elevation fluorescence telescopes located at Coihueco, AERA is a radio array and AMIGA is a denser array of muon and water-Cherenkov detectors with a spacing of 750 m amidst the larger ground array with a spacing of 1500m.

3.1 Surface Detector (SD)

The ground array at the Pierre Auger observatory has over 1660 water-Cherenkov stations. Each station is in a triangular grid spacing of 1.5 km. In addition, there is a denser array called the Auger Muons and Infill for the Ground Array (AMIGA) shown in Figure 3.1 which consists of 61 stations. These detectors have a spacing of 750 m. Each water-Cherenkov station has a diameter of 3.6 m, with a tank inside up to a height of 1.2 m. They are each filled with 12,000 litres of purified water [98]. The outer surface is made of polyethylene while the inner surface has a sealed liner which is diffusively reflective. This reflective surface allows the Cherenkov light emitted to be easily detected by three 9-inch (229 mm) diameter, photomultiplier tubes (PMTs). Each tank is self-sustaining and powered by a solar panel placed on top of the tank which provides an input to two 12V batteries [98]. In addition, each tank also connects wirelessly to the central data acquisition system [98]. Figure 3.2 shows an image of a water-Cherenkov station as part of the array while Figure 3.3 shows a schematic of what is inside.

3.1.1 Station Calibration

Since the stations are laid out at such an immense scale, the calibration process is done locally and automatically by each station. Firstly, the photomultiplier tubes (PMTs) of the SD record signals generated from Cherenkov light emitted by relativistic particles going through the detector. This light generates a signal and the unit chosen to express this signal is called the *vertical equivalent muon* (VEM). It is defined as the signal produced by a muon entering the centre of a tank at a vertical trajectory [98]. This unit allows for an easy comparison with other stations and Monte Carlo simulation studies [101]. The aim of the calibration process is to determine the amount of charge generated by 1 VEM. The conversion factor to convert ADC to VEM is denoted as Q^{peak}_{VEM}. This is evaluated by every station in real-time through the measurement of background muon signals.

In particular, each station measures the integrated charge and amplitude of muon signals, thereby generating charge and amplitude distributions. Figure 3.4 shows two panels depicting histograms of charge and amplitude distributions measured by 3 PMTs. Integrated ADC channels and ADC channels are the hardware units for charge and amplitude, which is equivalent to ADC counts, but for 3 PMTs instead of just one. The first peak depicted by the black 3-fold histogram is generated by low energy particles while the second peak is due to muons travelling in a vertical trajectory going through the centre of the tank. The red histogram represents the VEM signal generated by muons travelling through the centre of the tank.

The amplitude histogram is used in relation to surface detector triggers in the next Section (Section 3.1.2). There is a threshold and time-over-threshold trigger which are set in electronic units called channels which is a measure of

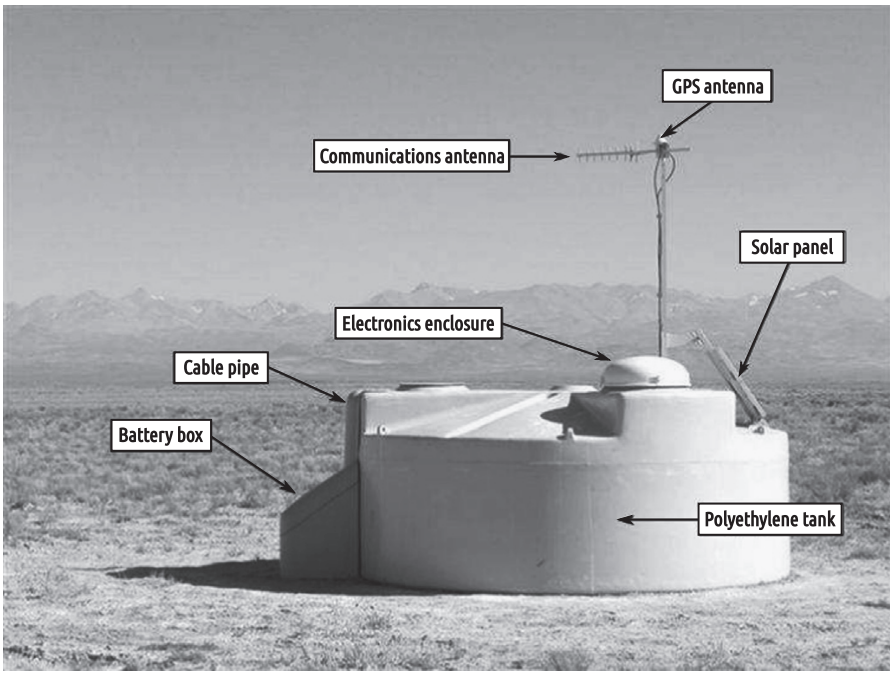


Figure 3.2: A water-Cherenkov station located on site [98].

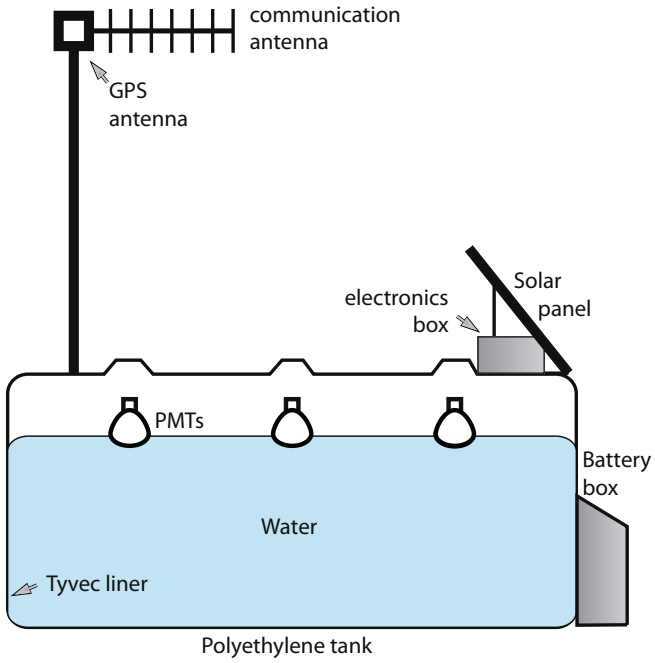


Figure 3.3: A schematic showing the parts inside a water-Cherenkov station [2]. There are three 9-inch (229mm) diameter photomultiplier tubes (PMTs) on the surface of the water.

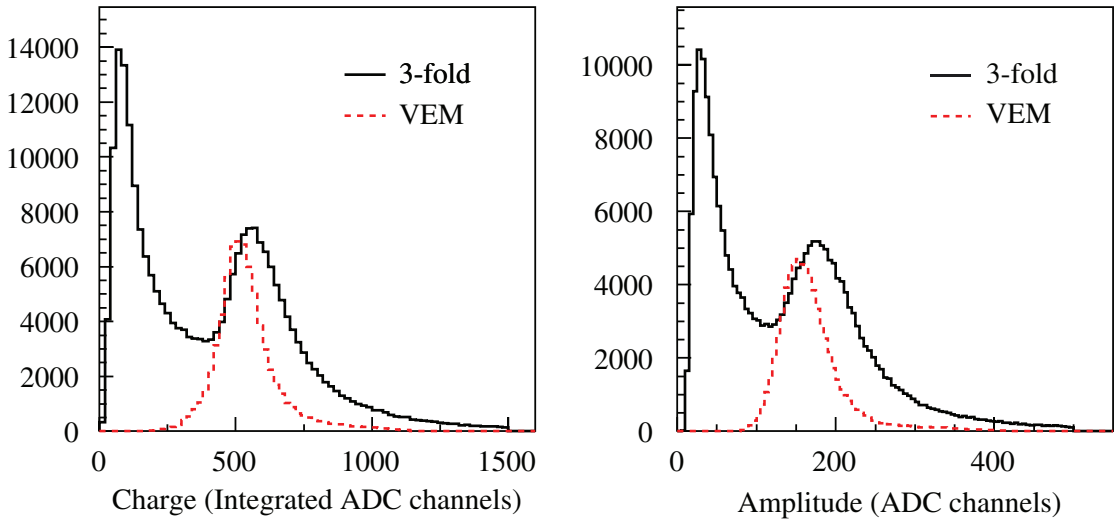


Figure 3.4: Example of charge and amplitude distributions measured by 3 PMTs [101]. The 3-fold histogram represents the sum of signals between 3 coincident PMTs. The first peak is generated by low energy particles while the second peak is generated by muons coming from all directions. The red dashed histogram represents the signal generated by muons travelling vertically through the centre of the tank as measured by a muon counter.

the current from the PMTs [101]. Similar to how muons in the atmosphere produce a peak in the charge histogram, a peak is also generated in the amplitude histogram and this is denoted I_{VEM}^{peak} [101]. This is then used as a common reference unit for threshold levels. The conversion from electronic units to I_{VEM}^{peak} must be continuously updating to maintain the proper trigger level [101].

3.1.2 Surface Detector Triggers

The surface detector has triggers depicted as a hierarchy when it comes to collecting data. These triggers are schematically shown in Figure 3.5. There are two levels of triggers called **T1** and **T2** formed at each detector. Firstly, T1 triggers data acquisition in each water Cherenkov detector. The T2 triggers are combined with those from other detectors which leads to whether there are coincidences in signal within a spatial and temporal scale. This leads to a formation of the **T3** trigger sending a signal to the main Central Data Acquisition System (CDAS) [102].

3.1.2.1 Station Level Triggers

At the station level, there exists four types of triggers and they are the following:

1. **Threshold (TH)**
   This trigger requires three PMTs in a single detector to have a summed

signal above 1.75 I_{VEM}^{peak} where I_{VEM}^{peak} is an electronic unit representing the peak amplitude generated by a vertically travelling muon. In particular, Th-T1 trigger is used to select large signals that are not necessarily spread in time. They are effective for detecting very inclined showers that are dominantly muonic. Ultimately, the Th-T1 trigger reduces the rate due to atmospheric muons from ≈ 3 kHz to ≈ 100 Hz [103]. The **T2** trigger acts as a controller for every station to reduce the number of events detected per unit time. This is to cope with limited bandwidth of the communication system between the detectors and the main campus [102].

2. **Time-over-threshold (ToT)**
   This trigger requires at least 13 time bins in a window of 3μs to be above a threshold of 0.2 I_{VEM}^{peak} in coincidence for two out of three PMTs. In contrast to Th-T1, the role of ToT-T1 is to select sequences of small signals spread in time, particularly due to more vertical showers where the signals are very close to the shower axis [104, 105].

3. **Time-over-threshold de-convolved (ToTd)**
   This trigger is one of the new triggers implemented in 2013. It applies an algorithm where the exponential tail generated by the Cherenkov light from electromagnetic interactions is removed through a de-convolution process. Pulses from muons or other particles in the shower will generate a sudden increase in signal and a prolonged exponential decrease due to reflections from the inner lining of the tank [12]. The aim in removing the exponential tail caused by Cherenkov light is to surpress changes in the signal that are far from the decay time which may have short spikes that could trigger the ToTd algorithm [106]. The de-convolved signal at time bin *i* is given by

$$D_i = \frac{S_i - S_{i-1}e^{-\Delta t/\tau}}{1 - e^{-\Delta t/\tau}} \tag{3.1}$$

   where S_i is the signal obtained at bin *i*, S_{i-1} is the signal at bin *i-1*, Δt is the width of one time bin which is 25 ns and τ is the average decay time of Cherenkov light which is 67 ns. The ToT algorithm is applied to this deconvolved signal. Then, the trigger is promoted to the T2 level if the electronic threshold of 0.2 I_{VEM} is satisfied. This results in a trigger rate of 0.3 Hz [12, 106].

4. **Multiplicity of Positive Steps (MoPS)**
   This trigger was also implemented in mid-2013 following the ToTd trigger [106]. This algorithm is completely independent from other triggers. The number of consecutive positive steps in the signal is counted over time and the amplitude *j* is monitored. For each group of consecutively increasing signal bins, *j* must lie between 3 and 31. This is so that the signal is above the average noise and below the average bin increase caused by a vertical

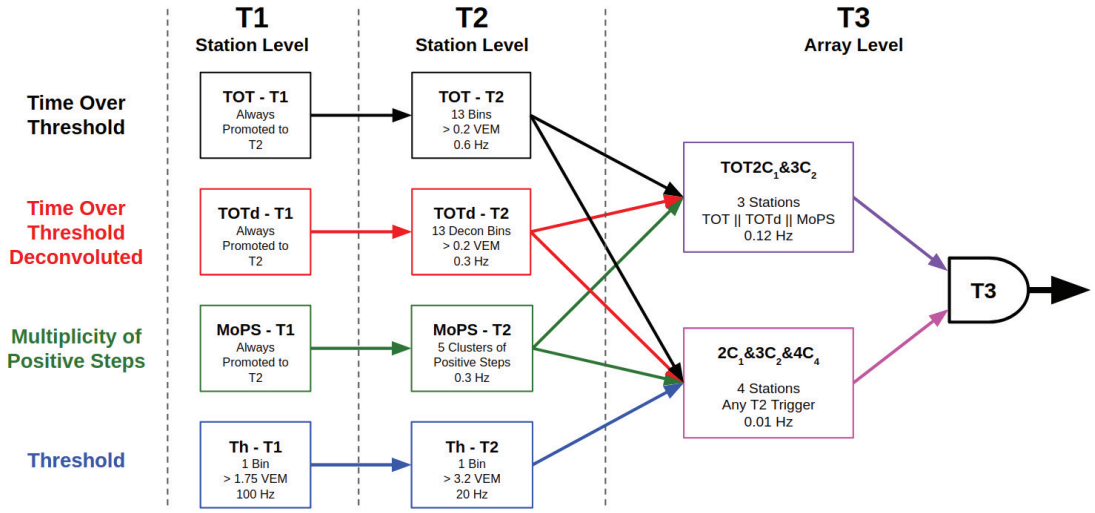


Figure 3.5: A schematic of trigger types in each level [106].

muon. Every time a group satisfies this condition it contributes to the multiplicity m. The trigger is satisfied when $m > 4$ within a 120 bin window where each bin is 25 ns. MoPS triggers are automatically promoted to the T2 level. Like the ToTd, the trigger rate which results from these conditions is 0.3 Hz [106].

If there are signals observed which satisfy the *threshold* trigger conditions, the traces are composed of spikes in a short duration of time. This can occur particularly in inclined air showers where particles go through a higher amount of atmospheric depth. In contrast, the ToT triggers are satisfied by signal traces that are low in amplitude but last for long durations of time and this can occur mostly in vertical air showers. The ToTd and MoPS triggers were introduced to significantly increase the sensitivity of the array to low energy air showers.

3.1.2.2 Array-level Triggers

The third trigger, **T3** is based on coincidences of signals between stations within a temporal and spatial scale which pass the T2 trigger. There are two modes in the T3 trigger both of which depends on the number of crowns over which the signal is spread over. The first crown denoted C_1 represents the six closest neighbouring stations. The second crown, C_2 is the next outer concentric, hexagonal crown of stations surrounding C_1 and so on. The crowns are labelled in Figure 3.6. The two modes are:

1. **ToT2C$_1$&3C$_2$**
   This is the first mode and it requires at least three Tot triggered stations where two stations are within one crown (ToT2C$_1$) and all three stations within two crowns (3C$_2$). This is shown in the left panel of Figure 3.6.

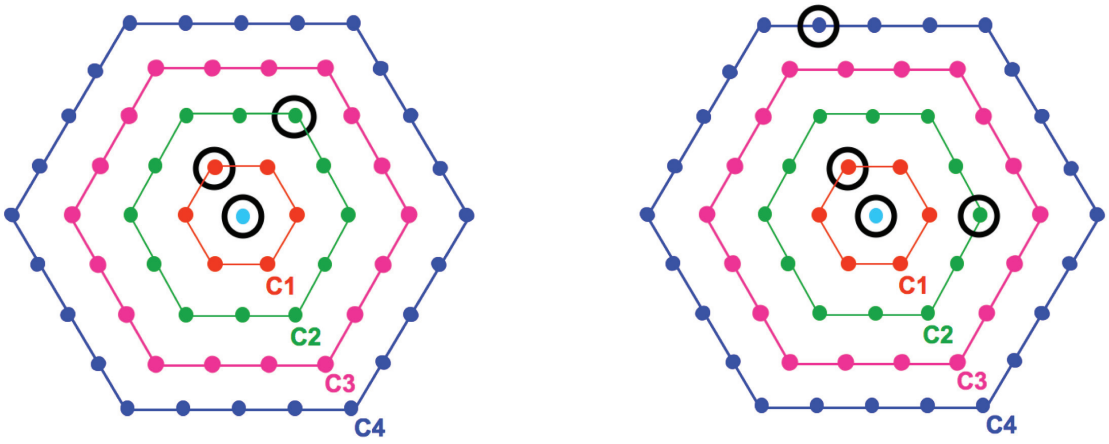


Figure 3.6: A schematic showing the two modes of the T3 trigger [102]. The left panel shows the condition which satisfies ToT2C_1&3C_2 and the right panel satisfies 2C_1&3C_2&4C_4.

2. **2C_1&3C_2&4C_4**
   This mode allows any type of T2 trigger. It is similar to the first mode except that it requires the fourth station to be contained within the fourth crown. This is shown on the right panel in Figure 3.6.

When a T3 trigger is satisfied, the signals are sent to CDAS as well as any T1 triggers which are temporally coincident within 30μs of the T3 trigger [12].

3.1.2.3 Event Selection

There are two triggers above T3 denoted as T4 and T5. These are performed offline as they are involved in the reconstruction process of cosmic ray air showers. The T4 trigger is also called a physics trigger and it is needed to select real showers from stored T3 data [102]. There are two criteria for these triggers and they are 3ToT and 4C1.

1. **3ToT**
   This criterion requires three nearby stations satisfying the T2-ToT trigger in a triangular pattern. In addition, it needs the times of the signals in the three stations to fit a plane shower front traversing at the speed of light. This trigger has a rate of 0.02 Hz. Since the stations detecting the signals are in a particularly compact orientation, events with zenith angles less than 60° are selected with high efficiency [102].

2. **4C1**
   The 4C1 criterion requires four nearby stations, without necessarily the need to satisfy the T2 trigger condition. Similar to 3ToT, the signal times must fit to a plane shower front traversing at the speed of light. This allows

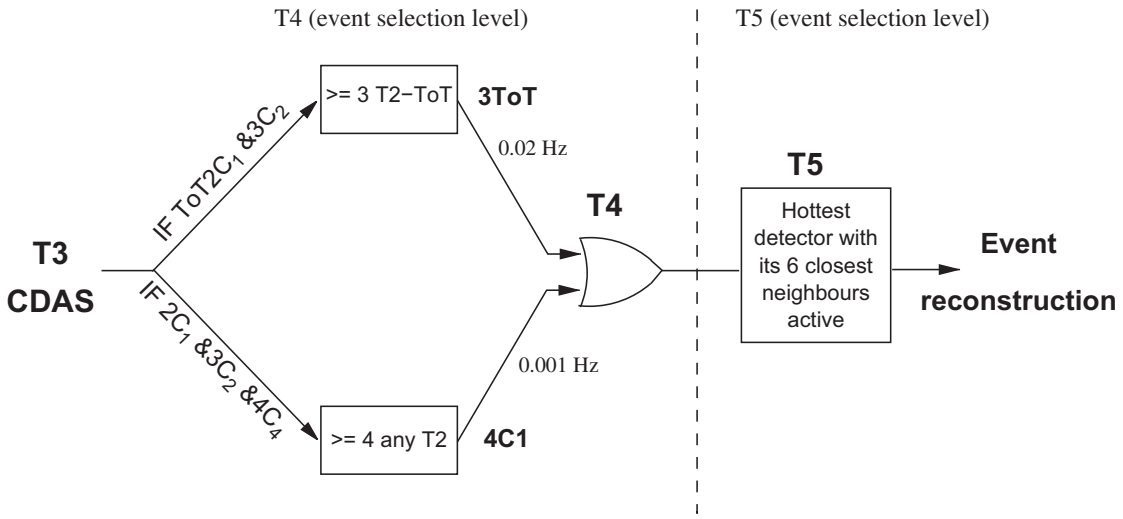


Figure 3.7: A schematic showing hierarchy of event selection trigger levels, T4 and T5 after satisfying T3 [102].

a detection efficiency of $\sim 100\%$ for showers with zenith angles greater than 60°.

Figure 3.8 shows the zenith angle distribution of events from the T4 selection criteria. The two criteria are shown to be complementary since the 4C1 tends to select events with greater zenith angles as shown in the left panel. On the right panel, the median energy of events selected by 3ToT is $\sim 6 \times 10^{17}$ eV while the median energy of 4C1 selected events is $\sim 3 \times 10^{18}$ eV [102]. Then, the identification and rejection of accidental events and detectors are removed. In particular, accidental detectors are detectors which have signals that are, by chance, in time with others but not part of the event. Finally, around 99.9% events successfully undergo the full event reconstruction algorithm, where arrival direction, core position and S(1000) are calculated [102].

3.1.2.4 Fiducial Trigger

The **T5** trigger is also known as the fiducial trigger. This trigger eliminates events falling close to the border of the array where a section of an air shower is missing [102]. Figure 3.9 shows an example of a hybrid event where information on part of the shower is missing due to its position near the border of the array. From this, many properties of the cosmic ray event such as core position and primary cosmic ray energy could be incorrectly reconstructed. In the example shown in Figure 3.9, if the core position were to be outside the array as determined by only the SD, the energy calculated would be four times larger than the energy determined by the FD. Therefore, the aim of the T5 trigger is to only select events which are well-contained in the array to ensure shower core positions and energies are accurately reconstructed. Not only that, the T5 trigger also ensures that there are six working

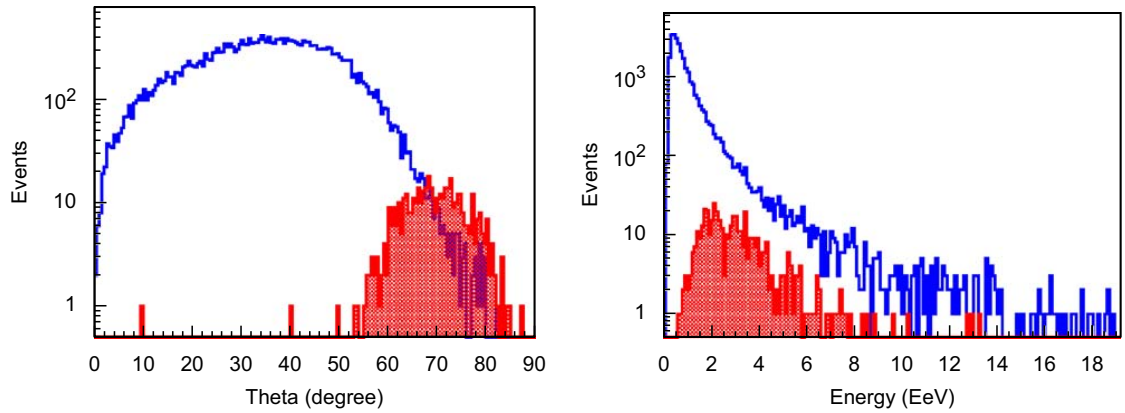


Figure 3.8: The left plots shows the zenith angle distribution of events, where zenith angle is represented by theta in degrees. The right plot is the energy distribution of events selected by the T4 triggers [102]. The blue unfilled histogram is represented by 3ToT and the red filled histogram is represented by 4C1 with no ToT.

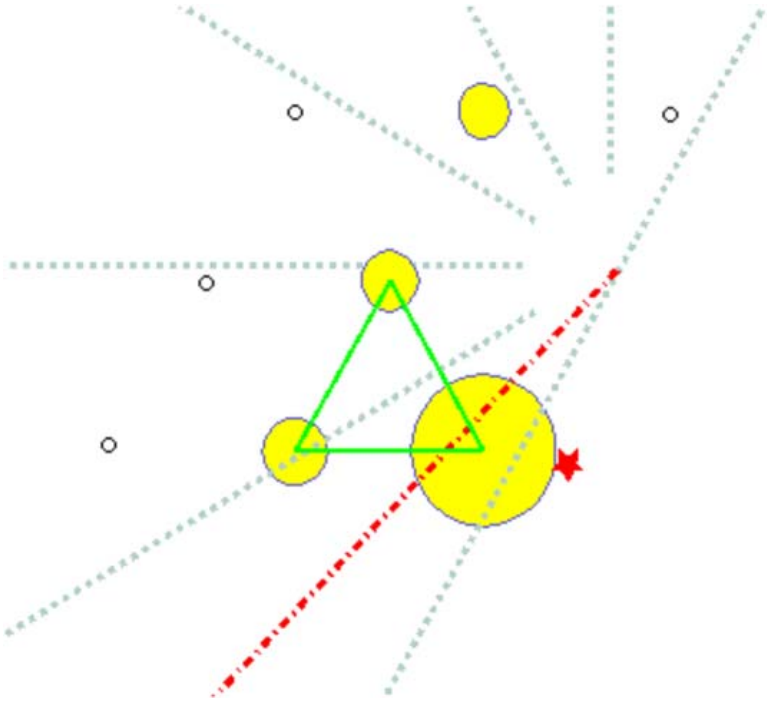


Figure 3.9: A hybrid event that falls close to the border of the array. The yellow circles indicate the four detectors which detected the event. The size of the circle if proportional to the logarithm of the signal strength. The shower detector plane reconstructed by the FD is indicated by the red dashed line. It indicates that the position of the shower core is within the green triangle. However, the SD reconstruction indicates that the core position lies outside of the array as indicated by the red star. In this example, the energy of the event is misreconstructed to be four times larger [102].

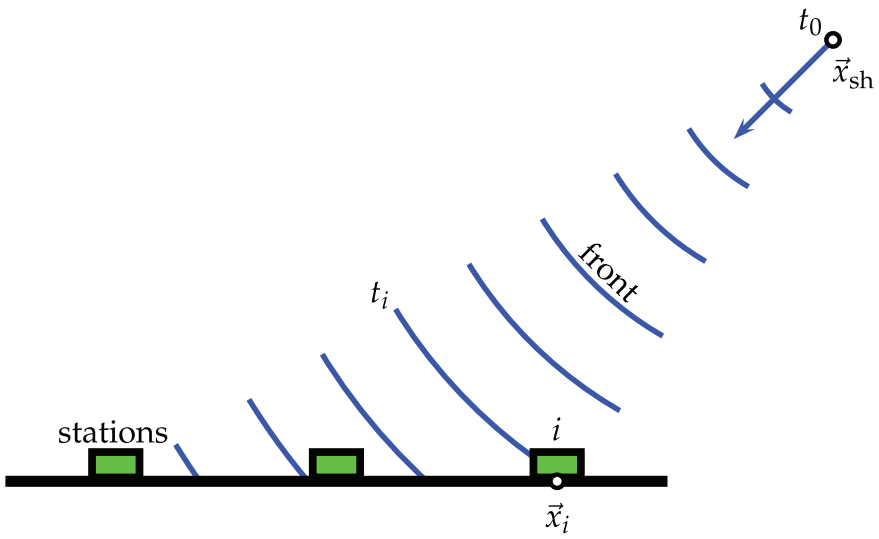


Figure 3.10: A schematic showing a slightly concentric-spherical shower front [98]. t_i is the start time of the signal at station i. The vector $\vec{x_{sh}}$ is the *virtual* origin of the shower with start time t_0.

stations surrounding a station which possesses the highest signal whilst detecting an event. SD event reconstruction depends on the sizes and timing of signals detected by individual stations. Various characteristics such as arrival direction, shower geometry and size are obtained from sampling information from stations triggered by an incoming event. Even in a situation where an event falls close to the border of the array causing part of the data to be missing, the seven detectors closest to the shower can ensure proper reconstruction of the event.

To ensure the events are contained within the array, the trigger 6T5 ensures that the station with the highest signal is surrounded by 6 active stations. Furthermore, even if an event lies well-contained in the array, this fiducial trigger filters out events detected by a station surrounded by non-working stations. Due to the large coverage of the array, about 1% of the stations are not working at any time even with regular maintenance. Assuming that all stations are working, the application of the T5 trigger reduces the effective area by 10% [102].

3.1.3 SD Event Reconstruction

After satisfying the trigger conditions, the timing information and signal strengths of extensive air showers are collected and ready for reconstruction. During reconstruction, the shower geometry is calculated. The shower geometry conveys valuable information about the shower such as the core position and the arrival direction of the cosmic ray [98].

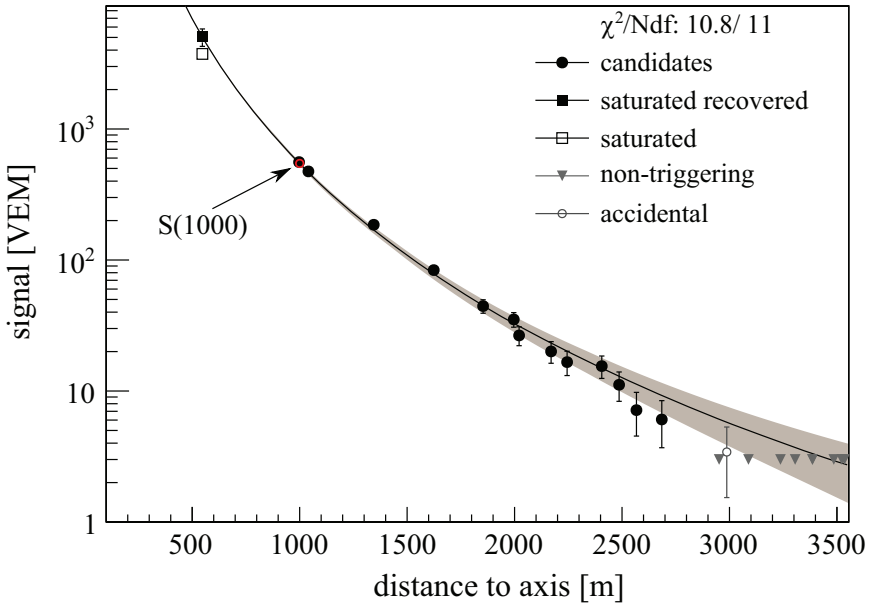


Figure 3.11: Dependence of signal on the distance to the shower axis [98].

3.1.3.1 Shower Geometry

To obtain a rough approximation for the arrival direction of a shower, the start times of the signals, t_i for an individual station are fitted to a plane front. If there are events with a sufficient number of triggered stations, the start times are depicted by a more detailed concentric-spherical model [98]. Figure 3.10 depicts this behaviour. The evolving shower fronts are travelling at the speed of light. A calculation of this moving shower front as a vector is as shown:

$$c(t_i - t_o) = |\vec{x_{sh}} - \vec{x_i}| \tag{3.2}$$

where x_{sh} is the *virtual* origin, x_i is the position of a station on the ground and t_o is the start time of the shower. This 4-parameter fit results in the determination of the radius of curvature of a spherical wavefront due to the inferred time at which the shower core hit the ground [98]. The arrival direction of the shower can be obtained as the result of the difference between the virtual shower origin and the shower core position. The core position, $\vec{x_c}$ is obtained as a result of fitting a lateral distribution function discussed in Section 3.1.3.2. This arrival direction is represented by the shower axis vector, $\hat{a}$. It is shown as the following:

$$\hat{a} = \frac{\vec{x_{sh}} - \vec{x_c}}{|\vec{x_{sh}} - \vec{x_c}|} \tag{3.3}$$

3.1.3.2 Lateral Distribution Function

To estimate the energy of a cosmic ray from an extensive air shower, a lateral distribution function (LDF) is fitted to the signals recorded by each individual station. The LDF describes shower particle density as a function of perpendicular distance to the shower axis [98]. This relationship is described in Figure 3.11. The LDF used at Auger is known as the modified Nishimura-Kamata-Greisen function [107]. It is described as

$$S(r) = S(r_{opt})(\frac{r}{r_{opt}})^{\beta}(\frac{r + r_1}{r_{opt} + r_1})^{\beta+\gamma} \tag{3.4}$$

where r_{opt} represents the *optimum distance*. For the Pierre Auger Observatory, the optimum distance for a 1500m spaced grid of detectors is 1000 m and r_1 = 700m [98]. As for the detectors spaced at 750 m, the optimum distance is 450m. $S(r_{opt})$ is the signal obtained at r_{opt} denoted $S(1000)$ for stations in the 1.5km-spaced ground array. The parameters β and γ depend on the *zenith angle* (explained in Section 3.1.3.3) as well as shower size [98]. From these fits, the core position of the shower can be determined.

3.1.3.3 Energy Reconstruction

Firstly, to measure the shower energy, $S(1000)$ must be obtained. This signal depends on two main factors and they include the energy of the primary cosmic ray and the *zenith angle*. The zenith is defined as the line pointing into the sky directly above the observer. The zenith angle is the angle subtended between the zenith and the shower axis. In particular, as the zenith angle θ increases, the signal S(1000) decreases for a given cosmic ray energy. This is due to geometry and effects of attenuation from the atmosphere. To remove this dependence, the shape of the attenuation curve was extracted from the data using the Constant Intensity Cut (CIC) technique [108]. This curve defined as $f_{CIC}(\theta)$ is fitted with a third degree polynomial [98] shown to be

$$f_{CIC}(\theta) = 1 + ax + bx^2 + cx^3 \tag{3.5}$$

where a = 0.980 $\pm$ 0.004, b = -1.68 $\pm$ 0.01, c = -1.30 $\pm$ 0.45 and

$$x = \cos^2\theta - \cos^2\overline{\theta} \tag{3.6}$$

given that $\overline{\theta}$ is a reference or median angle. Figure 3.12 shows how the signal S(1000) changes as a function of *sec* θ for a fixed rate of events in all equal-exposure angular bins. The fitted solid line is represented by Equation 3.5. In particular, the angle $\overline{\theta}$ = 38° was chosen as it represents the median zenith angle which can be used for all showers specifically with an array spacing of 1500 m at the Pierre Auger Observatory [109]. To remove the zenith angle dependence, $S(1000)$ is converted into S_{38}. This conversion represents the signal $S(1000)$ which

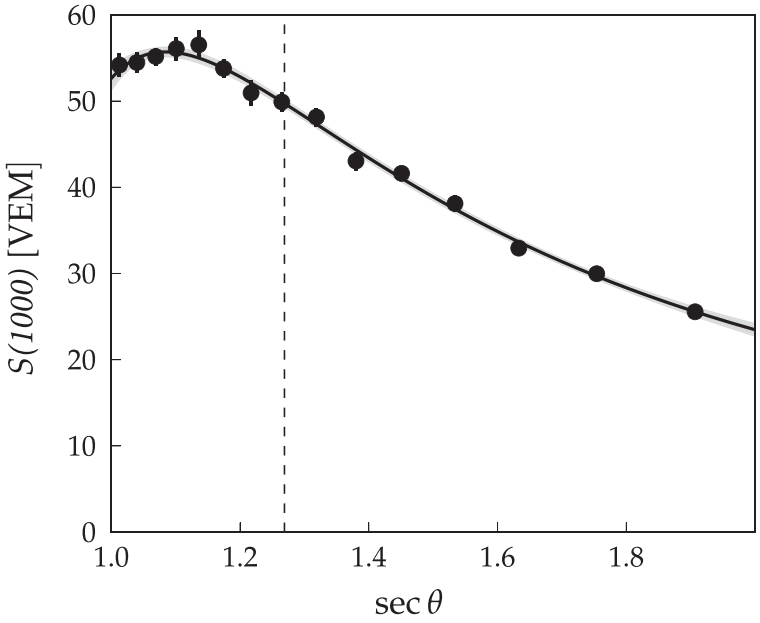


Figure 3.12: Attenuation curve depicted by the third degree polynomial $f(_{CIC}\theta)$ with the dashed line being the value of *sec* θ when $\theta = 38°$ [98].

would have been produced if the shower arrived at a zenith angle of $38°$. The conversion is done as follows:

$$S_{38} = \frac{S(1000)}{f_{CIC}(\theta)} \tag{3.7}$$

Finally, a cross-calibration of this signal with the FD allows the determination of the primary cosmic ray energy and this is further explained in Section 3.3.

3.1.4 Effects of Weather and Geomagnetic Field and the Need for SD Energy Stability

The detection of showers created from cosmic rays can be affected by the atmosphere and the geomagnetic field. Firstly, variations in atmospheric density and pressure can impact air shower evolution and, consequently, the signal measured by the SD [110]. On the other hand, the geomagnetic field affects shower propagation [111]. The paths travelled by charged particles will be curved according to the magnetic field lines of the Earth and this should be accounted for when reconstructing air showers to obtain the energies. From both weather and geomagnetic field variations, corrections have been made by the Pierre Auger Collaboration for air shower reconstruction. Examining the SD energy stability subject to the aforementioned conditions is essential especially in the long term as it can determine whether energy assignments are done precisely.

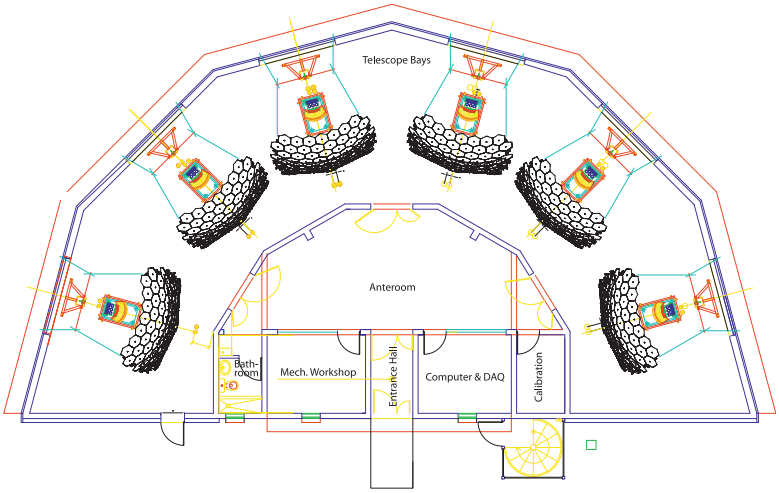


Figure 3.13: A top-view schematic of six fluorescence telescopes at an FD site [103].

3.2 Fluorescence Detector (FD)

The observatory has 27 fluorescence telescopes which collectively make up the fluorescence detector on the array. There are four sites which house these fluorescence telescopes called Los Leones, Los Morados, Loma Amarilla and Coihueco. Each site has 6 telescopes. It is important to keep in mind that data acquisition using the fluorescence technique is not done continuously. Due to the need for good weather and dark nights, the duty cycle of these fluorescence telescopes is only $\sim 15\%$ [98]. However, they can image the longitudinal shower profile. The fluorescence technique provides a near-calorimetric method in determining the energy of a primary cosmic ray since only a small percent of its energy is deposited into the ground [98].

Each telescope has a $30^\circ \times 30^\circ$ field of view, all facing inwards towards the array. Figure 3.13 shows a schematic of an FD site while Figure 3.14 shows an example of what the FD site looks like at Los Leones in a photograph. Figure 3.15 shows the side-view and scale of the telescope. At each site, the telescopes are housed in climate-controlled buildings. The shutters close during the day and when there is high-speed wind or rain detected. Each telescope detects nitrogen fluorescence through a glass window of radius 1.1 m and an ultra-violet (UV) filter to improve the signal-to-noise ratio. To correct for optical aberration and provide larger effective aperture, an annular corrector ring is used. This is mounted on the inside of the UV filter. Finally, the light is focussed by a set of segmented mirrors with a total area of ~ 10 m^2 onto a 22 x 20 pixel camera [98, 103].

Maintenance and cleaning campaigns have been done throughout the years of data acquisition at each site. Although the telescopes are sheltered, there are various segmented mirrors present which can unfortunately accumulate dust. The accumulation of dust layers may cause the absorption of light by the dust

Figure 3.14: A photo taken of the FD site at Los Leones [103].

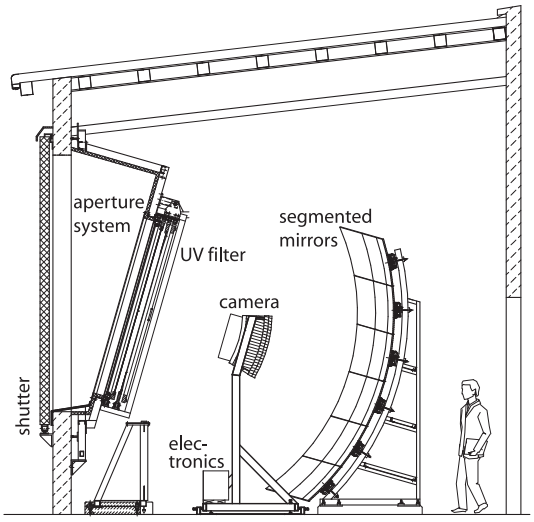


Figure 3.15: A side-view illustration of one fluorescence telescope on site with a human as scale [103].

particles. In addition, the dust seems to decrease specular reflection and increase diffusive reflection. This causes the image spot to increase in size, so not all the light is captured by the cameras of the FD. To counter this, cleaning campaigns have been done to clean the mirrors using a dry and wet method [98]. It was found that after a cleaning campaign, the reflectivity increased by $\sim 1\%$ in the case of mirror segments in the upper rows and $\sim 5\%$ in the lower rows after six years of installation [98]. Similarly, the UV filter can also accumulate dust affecting light transmission. It was found by Nguyen [110] that the discontinuity in energy scale ratio, defined as the ratio of FD energy to SD signal is coincident with the time at which the filters and mirrors were cleaned. A detailed study on the behaviour of these telescopes is crucial in investigating the extent of a possible correlation between cleaning campaigns and the FD energy assignments.

3.2.1 Fluorescence Detector Calibration

During event reconstruction, signals in terms of analogue to digital converter (ADC) counts are converted to a flux of light for each pixel on the camera detecting a shower. But first, the pixels must be calibrated to determine the response of each PMT to any given photon flux. There are three calibrations performed for the fluorescence detectors and they include a semi-regular absolute calibration, multi-wavelength calibration and nightly relative calibrations [98].

3.2.1.1 Absolute Calibration

Absolute calibration, a process involving a drum with an embedded light source is done to provide an absolute, end-to-end, calibration for every pixel on the camera [98]. This is also referred to as *drum calibration*. The drum-shaped source emits a pulsed flux of light with known intensity and uniformity all throughout the aperture including all 440 pixels on the FD camera. This triggers the pixels on the camera then calibrates them. Figure 3.16 illustrates how the drum calibration works.

The drum is made of a lightweight aluminium frame. It has a diameter of 2.5m and it is 1.4m deep. The interior lining is made up of Tyvek which is a diffusively reflective material in the ultraviolet range. The drum provides diffusive illumination to the telescope optical system by having a front face made up of a thick Teflon sheet [112].

3.2.1.2 Relative Calibration

Relative calibration involves illumination of the optical system for each telescope from various positions. The aim of this calibration is to monitor the stability of the optical system between absolute calibration campaigns [110]. They occur regularly, typically before and after every night of data acquisition for each

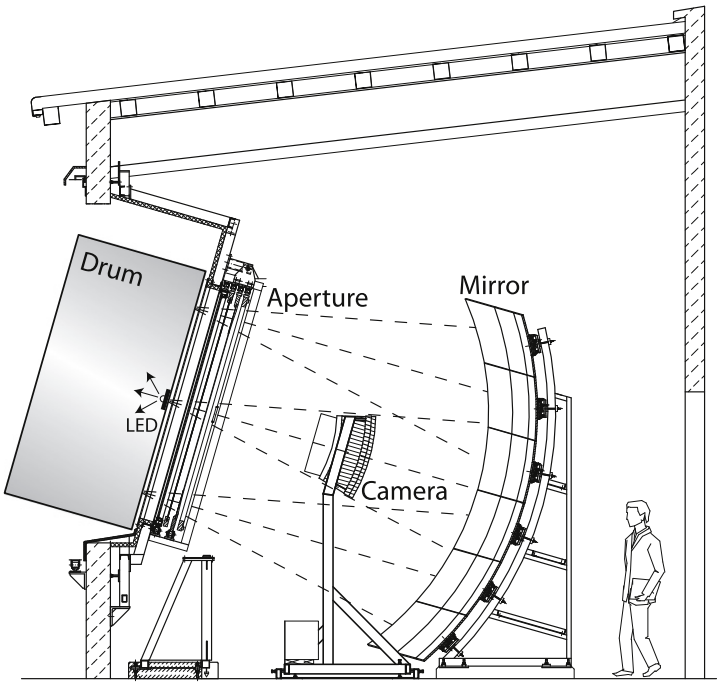


Figure 3.16: A side-view illustration of the drum calibration [103]. There is an LED embedded in the drum and it provides uniform illumination throughout the aperture.

fluorescence telescope. Figure 3.17 shows three positions A,B and C at which the light source has been placed for relative calibration.

Changes in the calibration process for the telescopes can affect the event reconstruction for the primary cosmic ray energies. For instance, at the start of 2015, the calibration process being done to illuminate the cameras on the telescopes with the calibration A light source was modified. The cameras were illuminated every 30 minutes during data acquisition rather than just before and after every night [110]. The purpose of this new calibration technique is to track the gain evolution of each individual pixel on the FD camera. This can ensure the accuracy of the camera's light detection. After this calibration was done, the subsequent data collected for every cosmic ray event showed that the energy scale ratio, defined as the ratio of FD energy to SD signal had an amplitude change of about 0.25 % throughout the year 2015 [110]. *Cal B* is another relative calibration process where two Xenon flash lamps on the sides of the camera body illuminate the mirror. This allows a measurement of the response from the mirror and camera simultaneously. Lastly, *Cal C* consists of a light outside the telescope aperture which is directed towards a reflective surface. This diffuses the light through the filter and provides end-to-end calibration of the entire optical system.

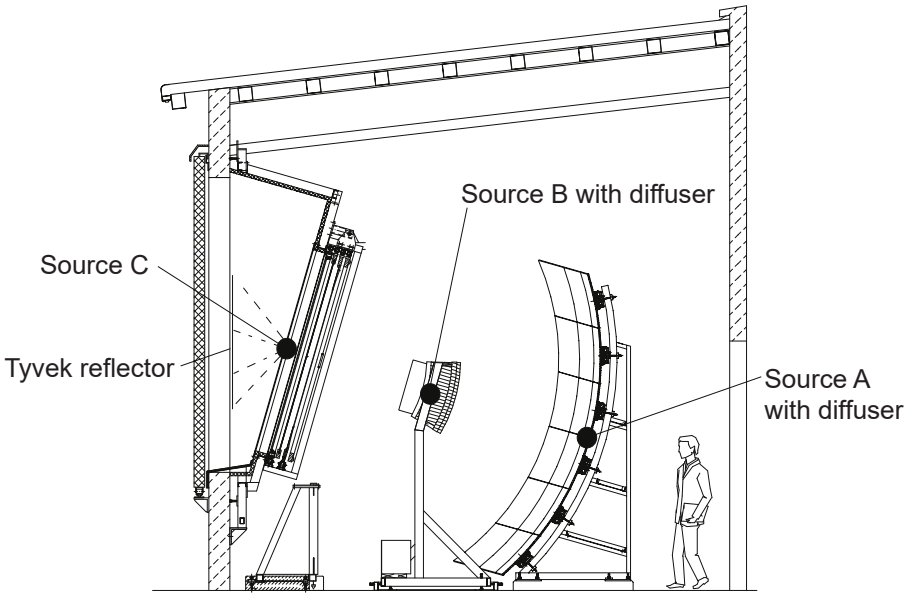


Figure 3.17: A side-view illustration of three different positions, A,B and C where the light sources have been placed for relative calibration [103].

3.2.2 Effects of Aerosols on Energy Assignments

Apart from effects of weather and the geomagnetic field, small particles such as dust or droplets also known as atmospheric aerosols may be in suspension in the outside air. Aerosols can hinder the propagation of ultra-violet light emitted from air showers [113]. Ultimately this impacts the amount of light detected by the FD. Not only that, the quantity of aerosols can vary depending on the place, time, wind and weather conditions. Since aerosol scattering can predominantly affect air shower detection, the Pierre Auger Collaboration conducted an extensive study and maintains a database on aerosol measurements. These measurements can be used to implement corrections on the data acquired from air showers [114].

3.2.3 HEAT

High Elevation Auger Telescopes (HEAT) is comprised of three telescopes located near the Coihueco FD. The aim of HEAT is to provide an enhancement in the detection of low energy cosmic rays [115]. HEAT can be configured to extend the elevation field of view of Coihueco from 30° to 60° degrees in *upward* mode or provide calibration with the Coihueco telescopes by switching to *downward* mode. Figure 3.18 shows a photograph taken of one of the HEAT telescopes in the *upward* mode. One of the advantages of HEAT is that its field of view can be combined with Coihueco's to observe high quality events for energy spectrum and mass composition studies [98, 115].

Figure 3.18: A photograph of a HEAT telescope in upward mode [12].

3.2.4 FD Event Reconstruction

The reconstruction procedure of the FD begins with fluorescence light being emitted by air showers and detected by a series of triggered pixels across the FD camera. Figure 3.19 shows an example of this phenomenon. The reconstruction is done in several steps. Firstly, the pixels on the FD camera are calibrated to ensure that there is no strong influence from the night sky background and electronic noise. This is then converted into a signal and photon flux. Next, the FD pixel trace is searched to find a signal corresponding to an air shower. A full description of pixel calibration with the requirements of high signal to noise ratio is shown in this reference [98]. Then, the geometry of the shower is determined. The sequence of pulse times from triggered camera pixels in the FD aids in the calculation of the shower detector plane (SDP). This is the plane that contains the location of the telescope observing the event as well as the shower axis. Figure 3.20 shows a schematic of the shower detector plane.

Then, individual pixels' timing information can be used to determine the position of the shower axis. The shower axis can be characterised from two parameters: χ_0 which is the angle between the shower axis and ground level within the SDP and R_p which is the closest distance between the shower and the telescope. This is also shown in Figure 3.20.The light arrival time at the *i-th* pixel in the FD can be shown as

$$t_i = t_0 + \frac{R_p}{c} \tan(\frac{\chi_0 - \chi_i}{2}) \tag{3.8}$$

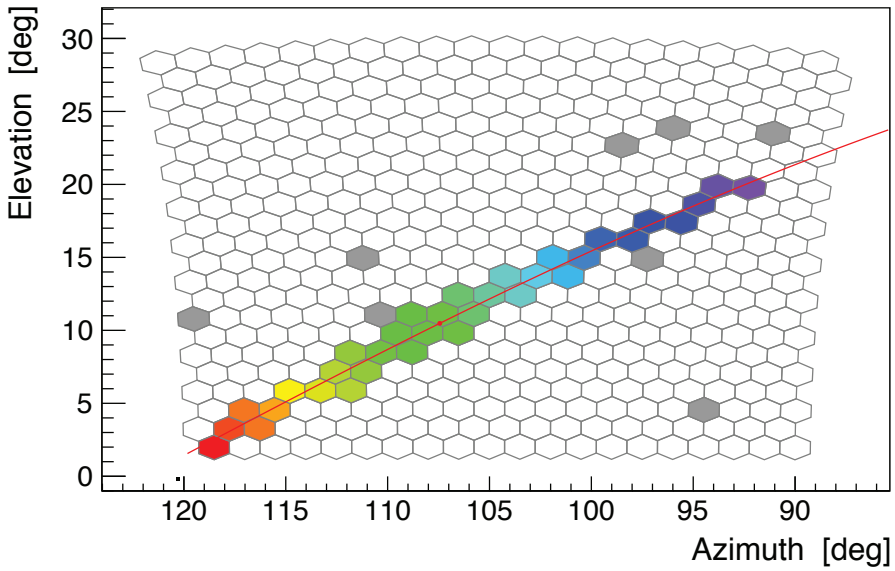


Figure 3.19: An example of the illumination of pixels from an air shower [12].The red line shows the intersection of the shower-detector plane with the field of view of the telescope. The colour scale from purple to red indicates the timing. Early signals are purple and they evolve to red indicating late signals. The grey pixels represent those which have triggered but did not pass the reconstruction algorithm due to its angle and timing being too far to fit to the shower-detector plane [12].

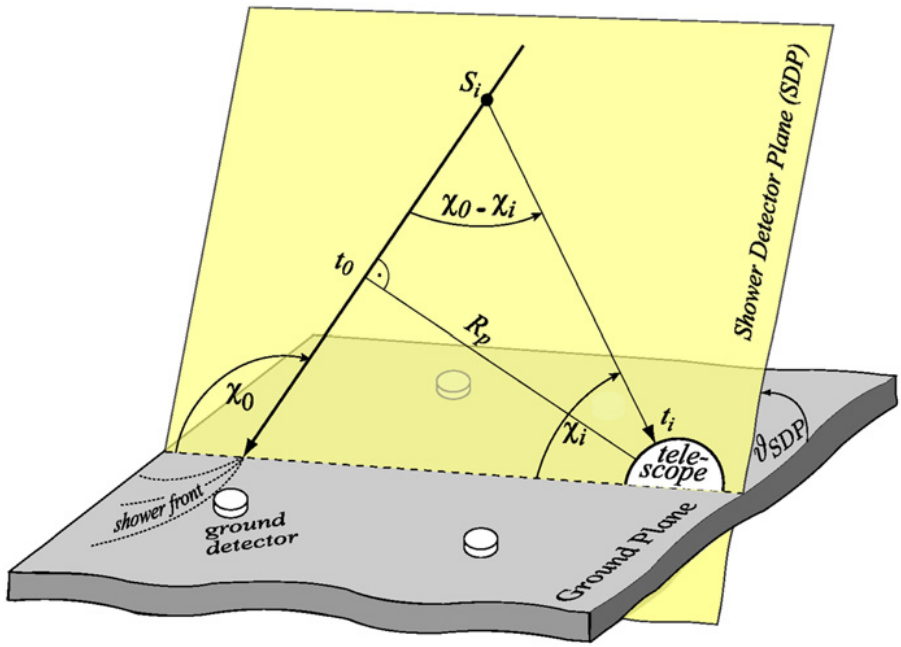


Figure 3.20: A diagram of the shower geometry during reconstruction and the shower detector plane (SDP) [103].

where *c* is the speed of light, χ_i is the elevation angle of the *i-th* pixel within the SDP, R_p is the distance from the point of closest approach to the camera and t_0 is the time when the shower front passes R_p. If the air shower is detected by a single fluorescence detector, the shower parameters would be dependent on the measured angular speed denoted as $\frac{d\chi}{dt}$ over the length of the track. In some cases, $\frac{d\chi}{dt}$ is small and this can lead to fit degeneracies, causing increased uncertainties. This degeneracy can be broken only if the shower reconstruction is performed in a hybrid method, where the SD can provide timing information which contributes to the determination of shower geometry [103].

After the shower geometry is established, the energy of the shower can be determined. It is most important to note that the fluorescence light produced in a shower is proportional to the amount of energy deposited in the atmosphere. Therefore, the light collected at the aperture of each telescope as a function of time is converted to an energy deposit of the shower as a function of atmospheric depth, known as $\frac{dE}{dX}$. However, light is attenuated from the shower to the detector due to many factors. The atmosphere has a big impact on shower detection, so a good knowledge of atmospheric conditions is crucial in event reconstruction. Some factors include atmospheric density, pressure, temperature and humidity and they can contribute to the fluorescence yield. Therefore, careful monitoring of the atmospheric conditions is essential for event reconstruction. Each air shower has a particular amplitude profile in the longitudinal direction. This profile can be fitted with a function called a Gaisser-Hillas function [116]. From this, the full longitudinal shower profile as well as X_{max}, which is the depth in the atmosphere where the shower reaches maximum size, is determined.

$$f(X) = \frac{dE}{dX}(X) = (dE/dX)_{max}(\frac{X - X_0}{X_{max} - X_0})^{\frac{X_{max}-X_0}{\lambda}} e^{\frac{X_{max}-X}{\lambda}} \tag{3.9}$$

The Gaisser-Hillas function is shown in Equation 3.9 where X is atmospheric depth while X_0 and λ are two shape parameters. For the case at which part of the shower is outside the field of view of the FD, the shape parameters are constrained to a range of typical values. From this, the calorimetric energy of the primary cosmic ray is obtained by performing an integration of Equation 3.9. The other $\sim$10% is carried away by other particles such as neutrinos and high energy muons which deposit this energy into the ground [98]. Thus, a correction for this "invisible" energy is made and taken into account during event reconstruction.

3.3 Hybrid Design and Energy Reconstruction

Since the Pierre Auger Observatory employs both fluorescence and ground array techniques to detect air showers, it exploits the hybrid design to perform event reconstruction and obtain energies of primary cosmic rays. The purpose of a hybrid design is to allow for cosmic ray detection, and measurement of their properties, using the same instruments. This hybrid design enables precise

determination of the shower axis position with a higher accuracy than what would be achieved using only the surface detector or fluorescence telescope [98].

The hybrid design gives an advantage when deriving the energy of the primary cosmic ray. Hybrid events are events such that the reconstruction of the energy estimator can be derived independently from the SD and FD [98]. Only a subsample of the events is used, such as those which pass some strict conditions called *quality* cuts. They involve having a substantially good fit to the Gaisser-Hillas function as well as ensuring the position of maximum energy deposition is present in the telescope field of view. Also, the fluorescence technique is intrinsically calorimetric, thus providing an energy measurement without the need to use shower simulations. This energy measurement can then be used in the calibration of the signal S_{38} obtained by the SD. After passing the strict quality cuts, the energy can be obtained using a known relation between the energy from the FD and S_{38} as follows:

$$E_{FD} = A \times S_{38}^{B} = E_{SD} \tag{3.10}$$

In Equation 3.10, A = (1.90 $\pm$ 0.05) $\times$ 10^{17} eV and B = 1.025 $\pm$ 0.007 [117]. These values are extracted from the fit depicted in Figure 3.21. The energy of the primary cosmic ray is obtained in this manner and later accompanied with systematic and statistical uncertainties. This fitting is one of the most vital jobs as it contributes to the current understanding of the cosmic ray energy spectrum.

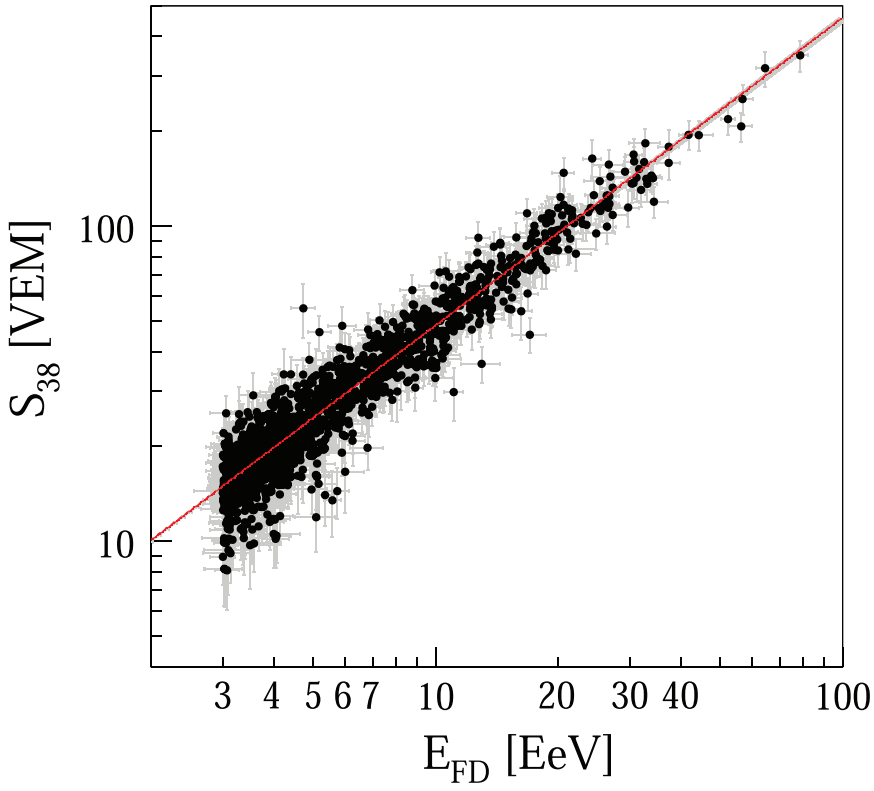


Figure 3.21: Correlation between S_{38} and E_{FD} with a red line representing the fit. The fit yields coefficients A and B in Equation 3.10 [103]. This provides a calibration of the SD energy estimator, S_{38}.

3.4 Atmospheric Monitoring

The atmospheric conditions can strongly influence the detection of extensive air showers at the Pierre Auger Observatory. Therefore, extensive atmospheric monitoring campaigns had been conducted to quantify the fraction of light being attenuated and scattered by atmospheric particles on site. Several atmospheric instruments had been deployed at the observatory to measure the atmospheric effects on recorded air showers. Figure 3.22 shows these instruments and they consist of weather stations, lidars, infrared cloud cameras, laser facilities and other instruments vital in contributing to the accuracy of air shower reconstruction.

3.4.1 Ground-based Weather Stations

The weather stations at each FD site are equipped with temperature, pressure, humidity and wind speed sensors. These sensors record data every 5 minutes. The data is then transferred, processed and stored in databases for atmospheric monitoring information. The stations at Los Leones, Coihueco and the laser facilities are each additionally equipped with a wind direction sensor [98].The Balloon Launching Station (BLS) had a weather station which serves as a base unit for an electric field meter. This has now been moved to the AERA site (Section 3.5.2). The electric field is recorded every second to detect lightning and thunderstorms. Since air shower reconstruction is affected by atmospheric conditions, these measurements are essential for improving the accuracy of cosmic ray energies.

3.4.2 Extreme Laser Facility (XLF) and Central Laser Facility (CLF)

There are many aerosols in the atmosphere which can cause attenuation of light (through scattering) emitted by extensive air showers. To quantify their attenuation, vertical aerosol optical depth (VAOD) profiles are measured at each of the four FD sites using data provided by the XLF and CLF [118]. The CLF had been in operation since 2004 while the XLF was built in 2008 aiming to service the Loma Amarilla FD site. These facilities house a pulsed laser with a wavelength of 355 nm that can be operated remotely. The XLF and CLF can not only monitor atmospheric conditions and improve geometric reconstructions, they also satisfy many calibration purposes [118].

3.4.3 Lidars

Every FD site has lidars (light detection and ranging instruments) aimed to monitor cloud cover on site. They consist of a laser with wavelength 351 nm

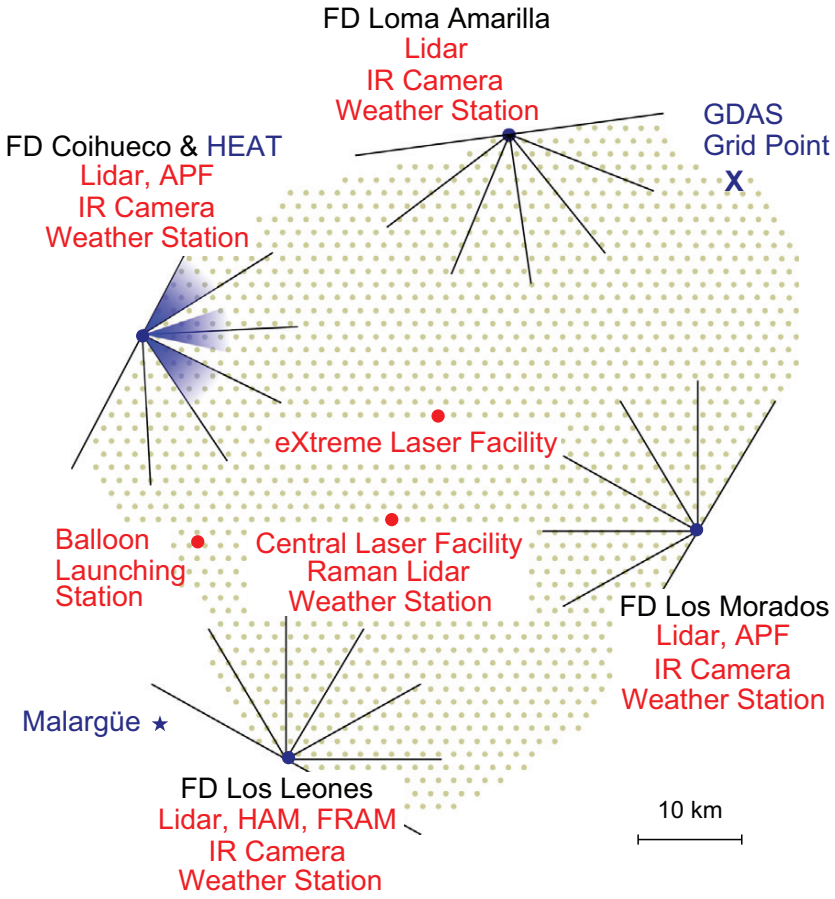


Figure 3.22: Schematic showing the atmospheric monitoring instruments at the Pierre Auger Observatory [98]. At each FD site, there is a lidar station, an infrared cloud camera and a weather station. There are also several other devices such as the Aerosol Phase Function (APF) at Coihueco and Los Morados, as well as the Horizontal Attenuation Monitor (HAM) and the ph(F)otometric Robotic Atmospheric Monitor (FRAM) at Los Leones. At the centre of the array are two laser facilities; the Extreme Laser Facility (XLF) and the Central Laser Facility (CLF), accompanied by the Raman Lidar. A Balloon Launching Station is located in the western part of the array. This was built together with a weather station which is a base unit for an electric field meter.

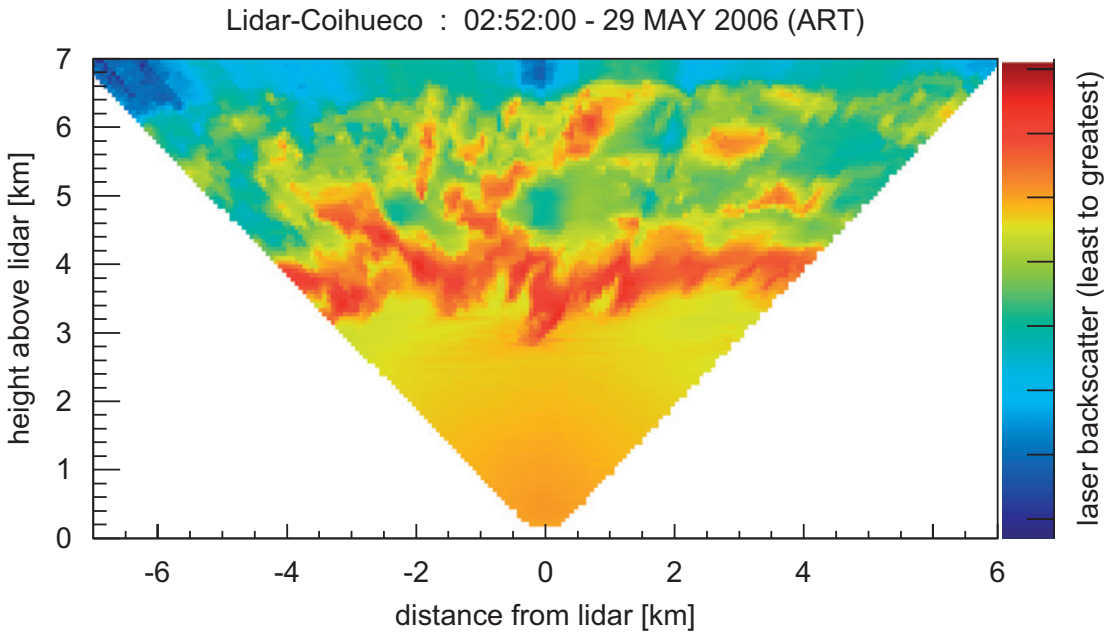


Figure 3.23: A diagram showing an example of a lidar scan at Coihueco [103]. The z-axis shows the intensity of backscattered light, the y axis is the height above the station and x axis is the horizontal distance to the lidar station [119].

which automatically scans up to 45° from zenith [98, 119]. The laser reflects off clouds in the sky and returns a signal that can be used to determine the distance to clouds as well as their density above the observatory. There is a Raman lidar located at the Central Laser Facility which measures the vertical aerosol optical depth (VAOD), aerosol backscatter coefficient and water vapour mixing ratio profiles routinely [120].

3.4.4 Cloud Cameras

In addition, there are infrared cloud cameras installed on the roof of every FD building aimed to capture images of the cloud conditions in the field of view of the telescopes during data acquisition. These images are used to analyse the fraction of pixels dominated by cloud and how much it compares to the night sky background [121].

3.5 Auger Upgrades

The Pierre Auger Observatory has had several enhancements in terms of instrumentation, since the successful deployment of the SD and FD. For example, the observatory has extended its lower energy threshold to improve sensitivity to showers in the knee-to-ankle region of the spectrum. This section describes several upgrades performed on the observatory.

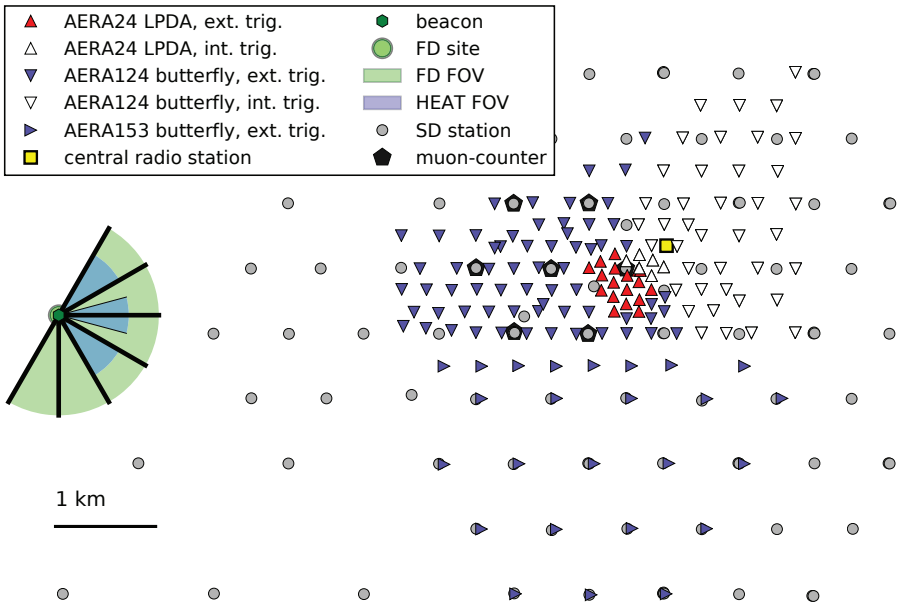


Figure 3.24: A map showing the positions of deployed radio stations as marked with triangles [124].

3.5.1 AMIGA

The Auger Muon and Infill for the Ground Array (AMIGA) aims to directly measure the muon content of cosmic ray air showers [122]. AMIGA comprises of an area filled with 61 detector pairs. Each one consists of a water-Cherenkov detector and a muon counter buried underneath it. The aim of AMIGA is to provide a direct measure of the muon content of extensive air showers. This can contribute to the study of the mass composition of cosmic rays as well as hadronic interactions in a range of energies which correspond to either Galactic or extragalactic sources [123].

3.5.2 AERA

The Auger Engineering Radio Array (AERA) is a radio detector system at the Pierre Auger Observatory aimed to detect radiowave emissions from cosmic ray air showers [124]. In addition, AERA can be used to perform several cross-checks with the FD and SD as well as measure the electromagnetic component for highly inclined showers. Also, the radio detectors and the SD can operate together to measure properties of cosmic rays with energies above $10^{17.5}$eV [123]. There are 153 radio stations spaced at several distances covering an area of 17 km^2. A map of the deployed radio detectors are shown in Figure 3.24.

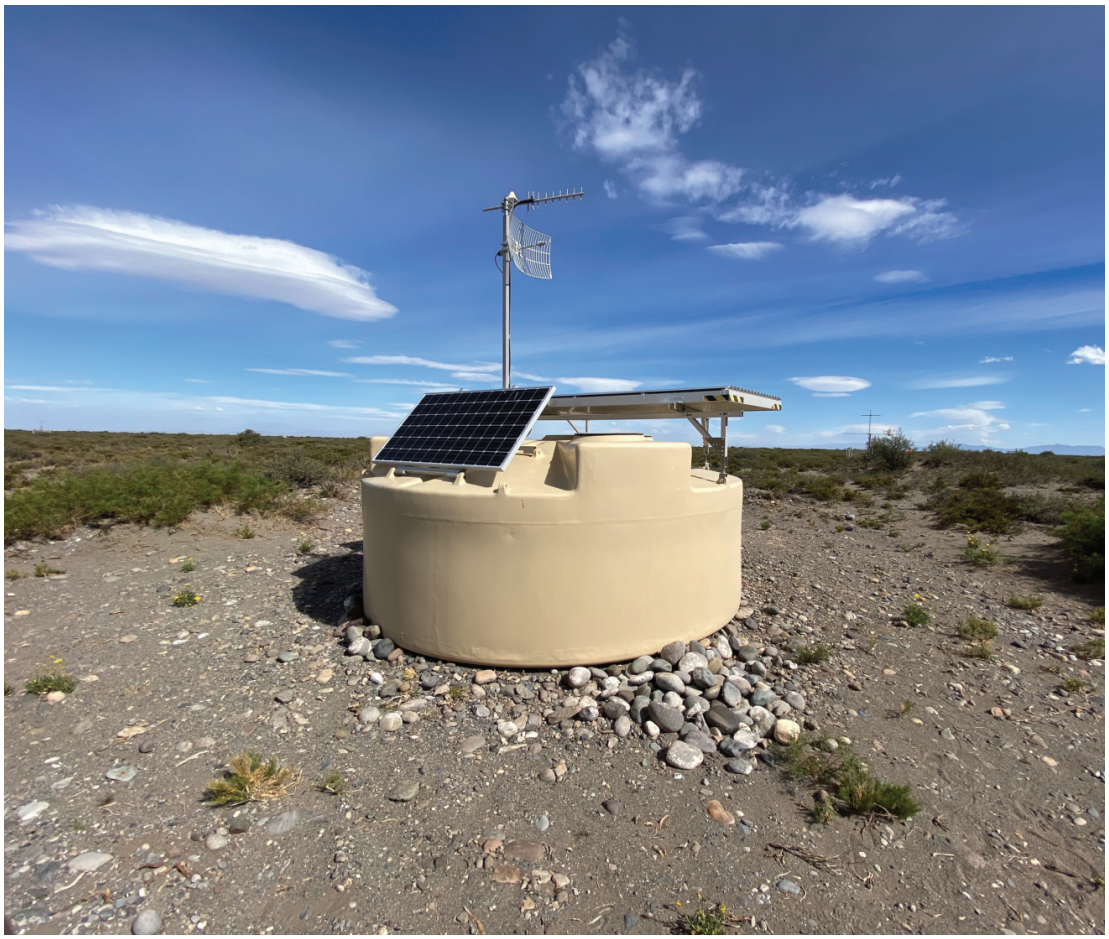


Figure 3.25: A photo taken of an upgraded water-Cherenkov station with a plastic scintillator installed as part of the SSD of AugerPrime [12].

3.5.3 AugerPrime

AugerPrime is an upgrade which is currently being constructed at the Pierre Auger Observatory. The aim of this upgrade is to improve the measurement of the flux suppression as well as mass composition at high energies in the spectrum [123]. The features of the upgrade are the following:

3.5.3.1 Surface Scintillation Detectors (SSD)

One very significant component in the upgrade is the installation of scintillators on the top of all the existing water-Cherenkov stations. These plastic scintillators make up the Scintillation Surface Detectors (SSD) and these in combination with the water-Cherenkov stations will enhance the ability of the array to distinguish between the muon and electromagnetic component of cosmic ray air showers. Figure 3.25 shows a photo taken of an upgraded station where a plastic scintillator has been installed on the roof of a water-Cherenkov station.

3.5.3.2 Electronics of the SD

The upgraded unified board (UUB) is a single electronics board that will help process signals from all three PMTs of the water-Cherenkov detector as well as the scintillator at an increased digitisation rate [123].

3.5.3.3 Extended FD Operation

The duty-cycle of the FD only goes up to $\sim$ 15% due to the requirement of clear, moonless nights. With AugerPrime, the duty-cycle of the FD can increase by about 50% if the FD operation is extended during nights with larger night sky background. It had been shown through preliminary tests that a reduction in supplied high voltage of the PMTs satisfied the required FD performance for extended operation [123].

3.5.3.4 The Underground Detector and Radio Upgrade

AugerPrime also comprises of a full upgrade of AMIGA, which can allow cross-checks to be done on the methods used to derive muon information from the SD. Furthermore, with the operation of AERA currently having 153 radio detectors, an upgrade had been proposed where *every* water-Cherenkov detector would be equipped with a radio antennas mounted on the top surface. It was demonstrated that horizontal showers (zenith angle θ = 60°-84°) can be measured feasibly using radio detectors due to their relatively large footprint on ground level. Once this upgrade has reached completion, the Pierre Auger Observatory would become the largest cosmic ray radio detector in the world [125].

Chapter 4

Monitoring the Energy Scale of the Fluorescence Detector at the Pierre Auger Observatory

The Pierre Auger Observatory has an enormous advantage in measuring the cosmic ray energy scale since it deploys a hybrid combination of detectors. The FD provides a nearly calorimetric energy measurement as the light produced is proportional to the energy deposited, while the SD measures the distribution of particles on the ground. By analysing showers detected by the FD in coincidence with the SD, the signal, S(1000) can be calibrated against the calorimetric energy measured by the FD. Thus, hybrid detection allows energy assignment to be largely independent of air shower simulations [117]. A large number of air showers detected over 14 years allows for the analysis of long-term changes in the energy scale. This chapter begins with studies of the stability of the FD energy and SD signal ratio as a function of time subject to a new calibration technique applied, as a continuation from previous analysis done by Nguyen [110]. This will be followed by a study which extends this analysis to individual FD sites and telescopes as well as a quantitative approach to measuring its evolution through fitting algorithms.

4.1 Energy Scale Ratio

The energy scale ratio (ESR) is defined as the ratio of FD energy to signal S_{38} (obtained from S(1000) as shown in Section 3.1.3.3). This can be used to monitor the stability for both detectors at the observatory. Since the ESR is a ratio of two independent energy estimates, it is expected that it yields a constant value. It is noted that since the ESR consists of a value in units of energy divided by signal, the value expected is not necessarily equal to 1. Previous work had been done on the analysis of the long term behaviour of E_{FD}/S_{38} [110]. Figure 4.1 shows a plot of how the ESR evolves with time. The data used in this plot consists of 8895 high quality hybrid showers spanning between January 2005 and December 2015 with a zenith angle $\theta < 60^{\circ}$. Only hybrid showers with FD energies above $3 \times$

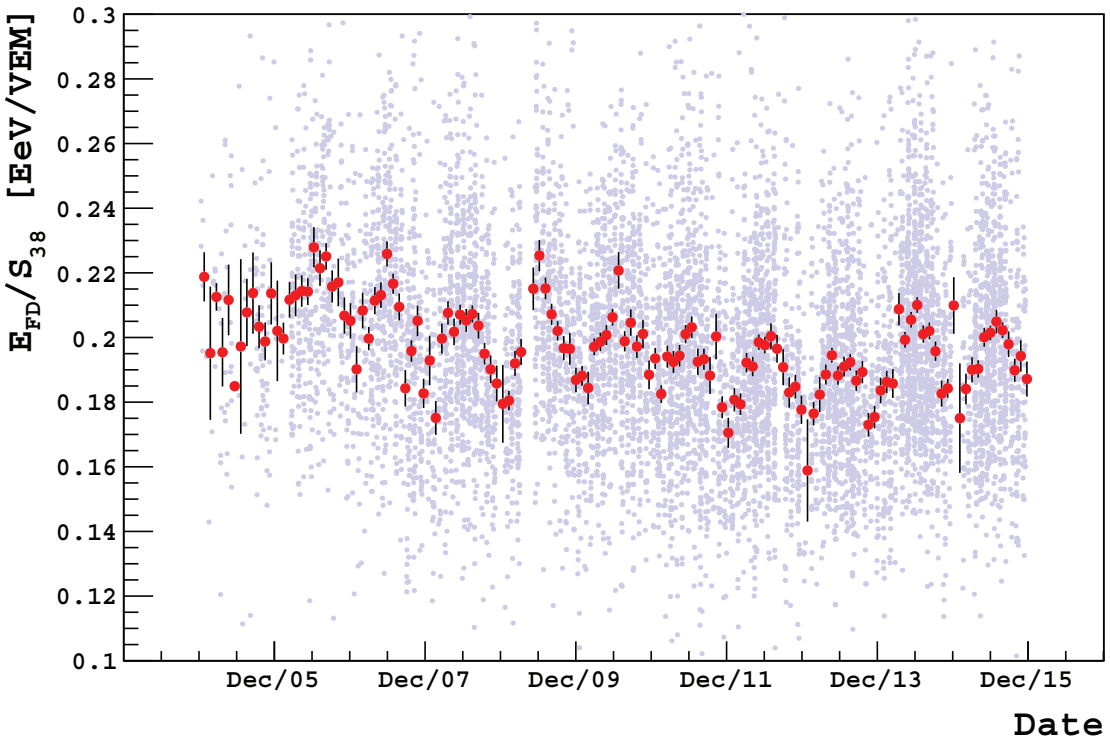


Figure 4.1: The energy scale ratio E_{FD}/S_{38} as found by Nguyen [110]. The scatter plot showing the light purple dots represent the ESR for every event between January 2004 and December 2015. A profile plot is shown as red points averaging the events in monthly time bins where one bin is 29.53 days representing a synodic month. The error bars represent the standard error on the mean for each month.

10^{18} eV were selected as this is the threshold energy at which the SD operates at 100% efficiency. It can be seen that the ESR does not exhibit a constant value as a function of time. The behaviour shows a seasonal modulation which peaks during the winter months in the Southern hemisphere. In addition, there is a slight downward drift. This may be caused by either the SD, FD or a combination of factors from both [110]. A fit was performed onto the ESR which is described by the following functional form with respect to time, t as a first guess.

$$\frac{E_{FD}}{S_{38}}(t) = B + m\left(\frac{t-t_s}{T}\right) + A\sin\left(\phi + 2\pi \times \left(\frac{t-t_s}{T}\right)\right) \tag{4.1}$$

where T is the number of seconds in a year, ϕ is the phase and t is the time in seconds since the start time, t_s which was chosen to be the 1st of January 2006. The terms B, m and A represent parameters applied to the fit. This fit was done using *MINUIT*, a software toolkit used to apply a minimisation of χ square algorithm. However, it was shown in the work of Nguyen [110] that a poor goodness of fit value (reduced χ^2 = 3.54) was obtained. This indicates that the behaviour of the ESR cannot be represented fully by the proposed functional form [110]. It was

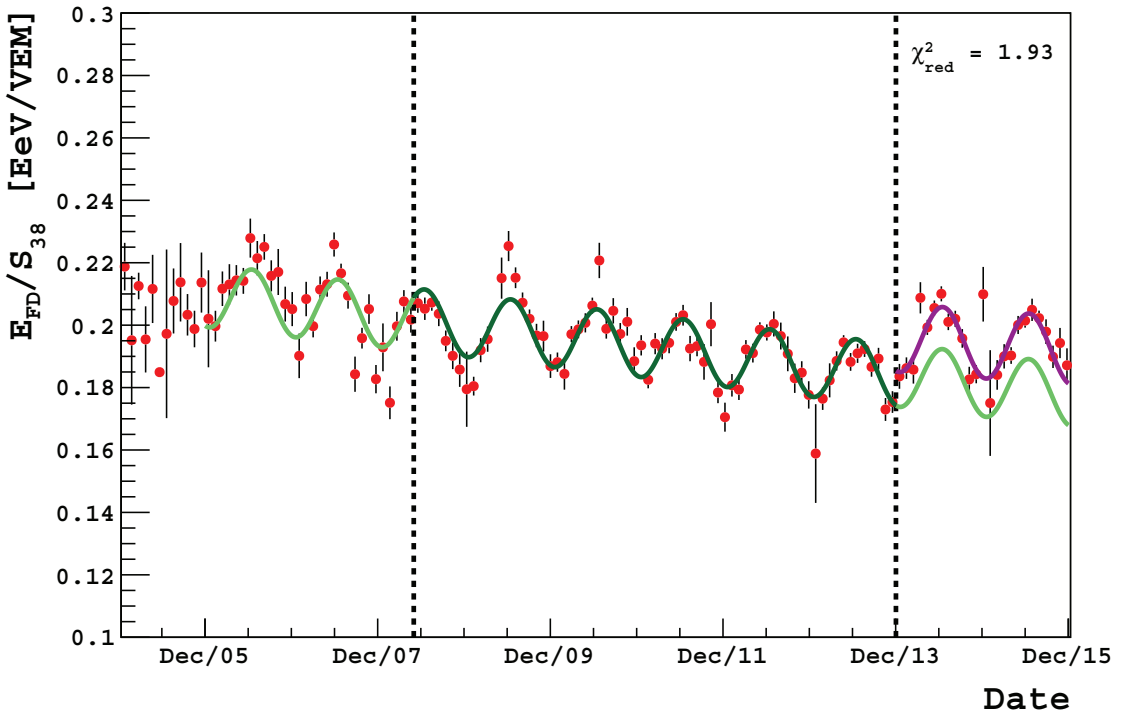


Figure 4.2: The energy scale ratio E_{FD}/S_{38} with a fit as done by Nguyen [110]. There are two distinct fits displayed. The first fit is in dark green; applied to the datapoints from mid-2008 to 2014 and extrapolated to either side of the vertical dashed lines representing the breakpoints. A second fit is shown in magenta after the second empirical breakpoint at the start of 2014. Both the dark green and magenta fit are combined to generate a reduced χ^2 value shown on the top right as 1.93.

inferred that this may be because of the presence of discontinuities in the ESR which had not been applied in the functional form of the fit.

From this, the ESR was analysed with respect to significant epochs in the operation of the observatory as well as from empirical observations of the ESR. Although there are several significant changes such as the introduction of more stringent operations due to large night sky background, the completion of the FD as well as the completion of the SD, the χ^2 value is optimised when there are only two breakpoints where one of them is empirical. The first breakpoint is placed in mid-2008 to mark the completion of the SD array. The second breakpoint is empirical due to a noticeable discontinuity present at the beginning of 2014. This result showing a best fit is illustrated in Figure 4.2, where the dark green fit applies to data in the range between mid-2008 to 2014 and the magenta fit represents the best fit post-2014. The overall reduced χ^2 value is found to be 1.93. The fit-function from mid-2008 to 2014 was extrapolated on either side of the vertical dashed lines and this extrapolation is shown in light green. There were no significant changes done in detection technique for the SD indicating that the discontinuity at 2014 is not related to the SD [110].

One of the many factors influencing the ESR is the condition of aerosols in the atmosphere. Aerosols can scatter fluorescence light beyond the FD field of view and this weakens the detected signal and reconstructed affects the shower energy if not properly accounted for. Due to this, the Pierre Auger Observatory has acquired a large database of aerosol measurements using the CLF and XLF described in Section 3.4.2. An improvement to the aerosol database was made in recent years to improve the reconstructed shower energies. These corrections have revised the FD energy upward which in turn increases the average value of the ESR [110].

Other than that, weather plays a big role in affecting the detection of extensive air showers, with impact on the evolution of the electromagnetic and muonic components. The observatory has developed an SD weather correction algorithm to correct the signals for any weather-dependent variations. A reference for a more detailed description of the weather correction algorithm is listed here [114, 126].

The changes in weather are due to variations in pressure and air density and these can affect shower development. The pressure is a measure of the mass of a vertical column of air. The variations in pressure can determine the 'age' of the shower when it reaches ground level [114]. As pressure changes in the atmosphere, the longitudinal profile of an air shower is attenuated accordingly. This results in an air shower arriving at ground level at a more advanced stage in its evolution in an atmosphere with higher pressure than one with lower pressure, which in turn affects the signal measured by the SD. On the other hand, as air density varies, the Molière radius, which is a characteristic measure of the lateral spread of shower particles is affected. In particular, the lateral spread of the shower is caused by the multiple Coulomb scattering due to charged particles in the shower. As the air density increases, the Molière radius decreases. As a result, the measured SD signal decreases [114].

The geomagnetic field can also affect the EAS development. The trajectories of the charged particles in any air shower are curved due to the magnetic field of the Earth. This results in a broadening of the spatial distribution of particles in the direction of the Lorentz force [111]. It has been found that the geomagnetic field induces an azimuthal modulation of the estimated energy of cosmic rays up to $\sim$ 2% at large zenith angles [111]. A correction to the energies of showers is made by the Pierre Auger Observatory to take the geomagnetic field into account and this can be found in detail in this reference [111].

Corrections from all three factors were applied to the ESR to see if there are any significant changes to its behaviour. The ESR with the applied corrections is as shown in Figure 4.3. It can be seen that there is a decrease in the peak-to-peak amplitude but still a non-zero downward drift in the ESR. In particular, looking only at the data pre-2014, there is a drift of -1.6 $\pm$ 0.2 % per year [110]. The amplitude obtained in Figure 4.2 which is 5.1 $\pm$ 0.4 %, reduced to 2.0 $\pm$ 0.4 % depicted in Figure 4.3 after the combined corrections [110]. The improvements

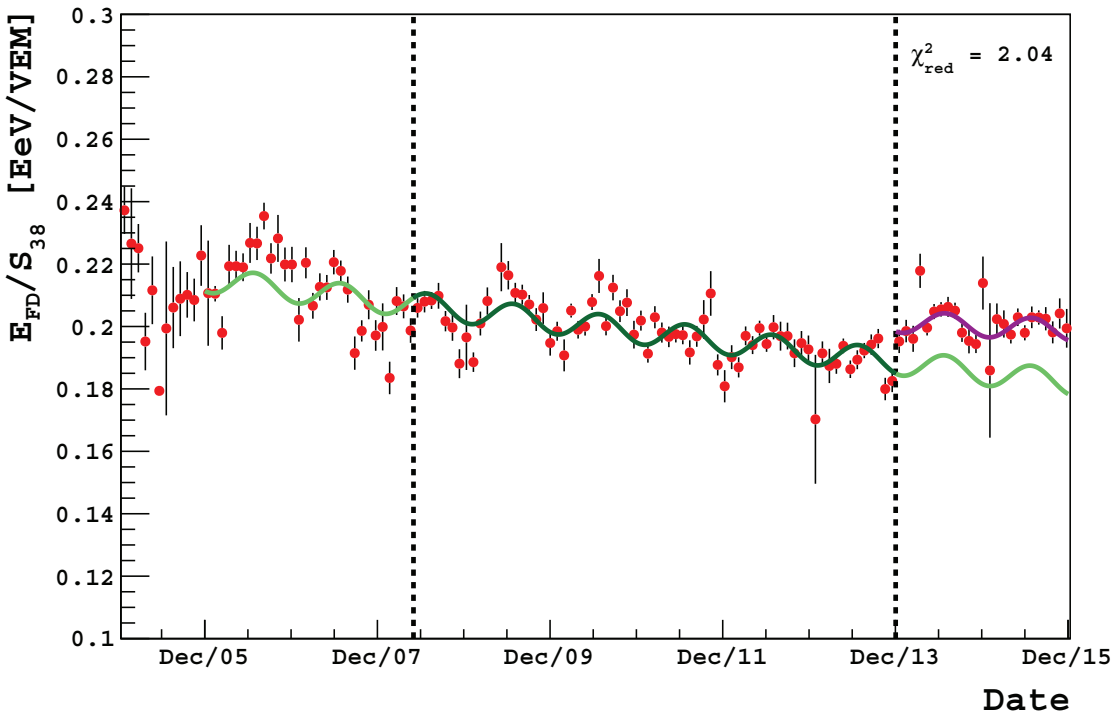


Figure 4.3: The energy scale ratio E_{FD}/S_{38} with applied weather corrections, geomagnetic field corrections and an improved aerosol database as done by Nguyen [110]. The fit resulted in a reduced χ^2 value is shown on the top right with a value of 2.04.

to the SD and FD reconstructions have reduced the seasonal modulations by a maximum of 3.9 %.

4.1.1 Results for Comparison and Discussion

Since then, new data have been acquired, as well as the production of a new enhanced aerosol database for updated corrections in energy reconstruction. An analysis to extend the current ESR is performed and compared with the work done by Nguyen [110]. The data set used in this analysis is *Golden Hybrid* which means both the FD and SD have detected the event with at least three active stations surrounding another station with the highest signal and the reconstruction of an energy estimator can be derived independently from both the SD and FD data [127]. The reconstruction used is the latest *ICRC19* (International Cosmic Ray Conference 2019) reconstruction. In addition, several *quality cuts* were applied to the data to ensure no edge effects and large uncertainties in shower parameters arise during reconstruction to obtain the primary cosmic ray energies. One of them ensures that the FD energies of all selected events are greater than 3×10^{18} eV as it represents the threshold at which the detectors operate at 100% efficiency. It is also ensured that the hybrid data used in this analysis consist of air shower zenith angles smaller 60° since

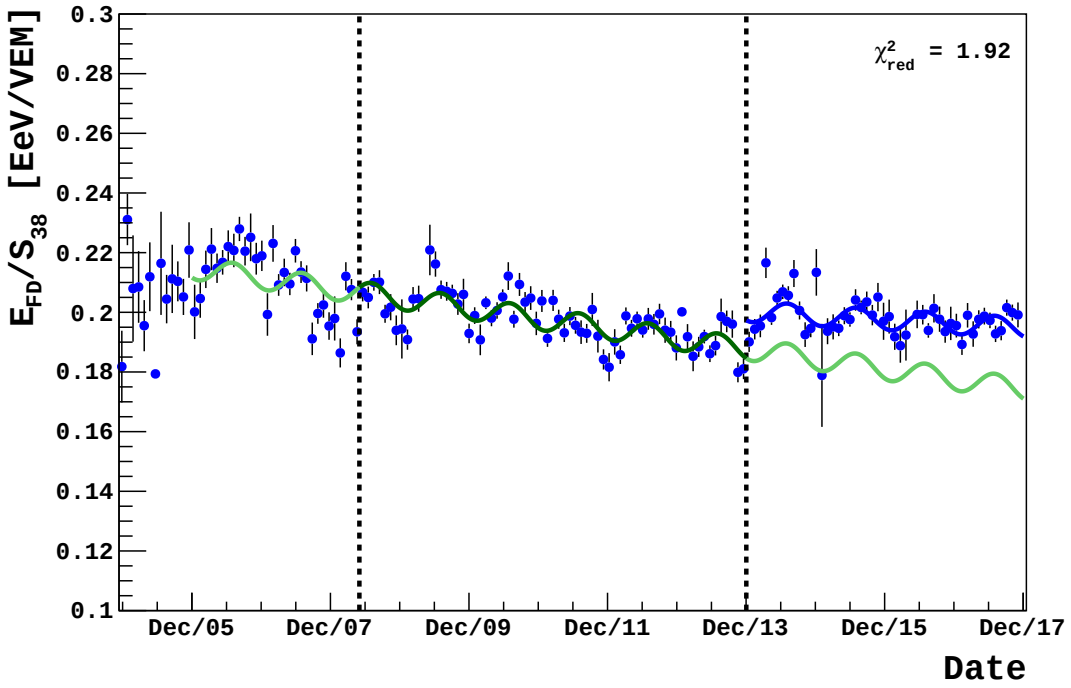


Figure 4.4: The energy scale ratio E_{FD}/S_{38} now extended to December 2017 with applied weather corrections, geomagnetic field corrections and an improved aerosol database. The fit applied to data from 2008 to 2017, results in a reduced χ^2 value shown on the top right as 1.92.

Study	Data type	χ^2_{red}	Drift (pre-2014) [% per year]	Amp (pre-2014) [% per year]	Drift (post-2014) [% per year]	Amp (post-2014) [% per year]
Nguyen [110]	Aero DB	2.04	-1.6 $\pm$ 0.2	2.0 $\pm$ 0.4	-0.7 $\pm$ 0.9	1.7 $\pm$ 0.7
This work	New Aero DB	1.92	-1.7 $\pm$ 0.2	1.9 $\pm$ 0.4	-0.7 $\pm$ 0.3	1.7 $\pm$ 0.5

Table 4.1: Results showing the drifts and amplitudes pre and post 2014 comparing results done by Nguyen in Figure 4.3 and this work shown in Figure 4.4.

the SD can detect more vertical showers with higher accuracy. The complete list of cuts are shown in Appendix A. The data used extends the ESR done in previous work by two more years. The results are shown in Figure 4.4 where it shows that the reduced χ^2 value is 1.92. This shows a very slight improvement to the analysis done in Figure 4.3 which could be caused by the improved aerosol database. The drifts and amplitudes which represent seasonal modulation were calculated and presented in Table 4.1. It can be seen from these results that the drifts and amplitudes are consistent with each other even after the extension of two more years of data.

4.2 Running Calibration

The FD is responsible for detecting the development of extensive air showers. It fully exploits the fact that fluorescence yield is proportional to the energy deposited in the atmosphere from the EAS. To ensure accuracy of reconstruction, it is important that the FD is calibrated to generate a reliable conversion between ADC counts and photon flux. From this, a more accurate estimate of E_{FD} is obtained. In 2014, there was a temporary modification in the relative calibration technique where a running "calibration A" measurement was introduced [128]. In this technique, the camera is illuminated with an LED every 30 minutes throughout the night during FD data acquisition while the telescope shutters are open. This allows for the study of pixel variation associated with the light from night sky background. This technique is in contrast to the previous relative calibration which requires manual operation and was performed only before and after data acquisition, while the shutters are closed. Details of how the running calibrations are obtained is shown in this reference [129].

A comparison between the regular cal A and running cal A measurements was done and it is shown that there is a small discrepancy of the order of 2 % [110]. Since this measurement affects pixel gain, this can in turn have implications on the cosmic ray energy, E_{FD}. Due to this, the energy scale ratio can vary due to performing running calibrations. Also, it was suspected that the residual seasonal modulations observed in the energy scale ratio are related to the FD pixel gain variations due to night sky background [110]. So, the introduction of the running calibration constants could potentially minimise this variation since the pixel gain is measured throughout the night, rather than at the beginning and end of each night as per the regular procedure. This can be used to calibrate the FD measurements hence, obtain a more accurate value for the FD energy.

4.2.1 Results for Comparison and Discussion

Previous work done by Nguyen compares the hybrid showers for three different reconstructions, comparing hybrid data with no corrections applied, with an updated aerosol database, with weather and geomagnetic field corrections as well as data with applied running calibrations. However, it was noted that the calibration constants were not recorded for Los Leones. So, only data from the three remaining FD sites were obtained for analysis. The results are shown in Figure 4.5 [110]. The amplitudes of each fit are shown in the bottom right hand corner of the plot. It can be seen that there is a reduction in amplitude; from 5.78% to 0.54% due to weather corrections. For comparison, a similar plot was generated in this work and shown in Figure 4.6 but with a new enhanced aerosol database developed by V. Harvey [121]. The key difference is the use of uncorrected data, which already uses an updated aerosol database as shown by the black points, unlike in Figure 4.5. After applying weather and geomagnetic field corrections, the amplitude decreases from $\sim 4 \pm 1$ % to 1 ± 1 %, as shown

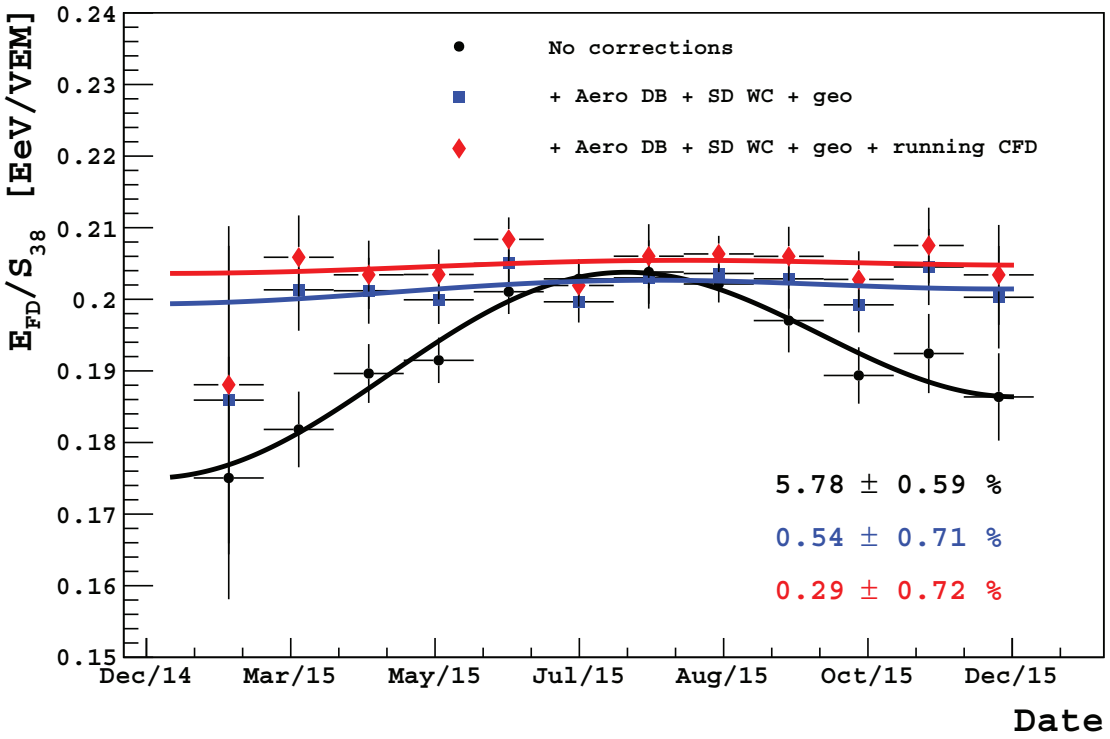


Figure 4.5: The ESR for the year 2015 to compare the three different improvements done to the hybrid showers done by Nguyen [110]. The black points represent data reconstructed with no corrections applied, the blue squares represent data using an updated aerosol database as well as weather corrections and geomagnetic field corrections (+ Aero DB + SD WC + geo). The red diamonds represent data with previous corrections and running calibrations applied (+ Aero DB + SD WC + geo +running CFD). The amplitude of the seasonal modulation is shown in the bottom right hand corner for each fit.

in Figure 4.6. This is consistent with the decrease shown in Figure 4.5. However, it is concluded that the running calibration provides no effect on the seasonal modulations, since the amplitudes of the fits are consistent with one another, similarly shown by the work of Nguyen.

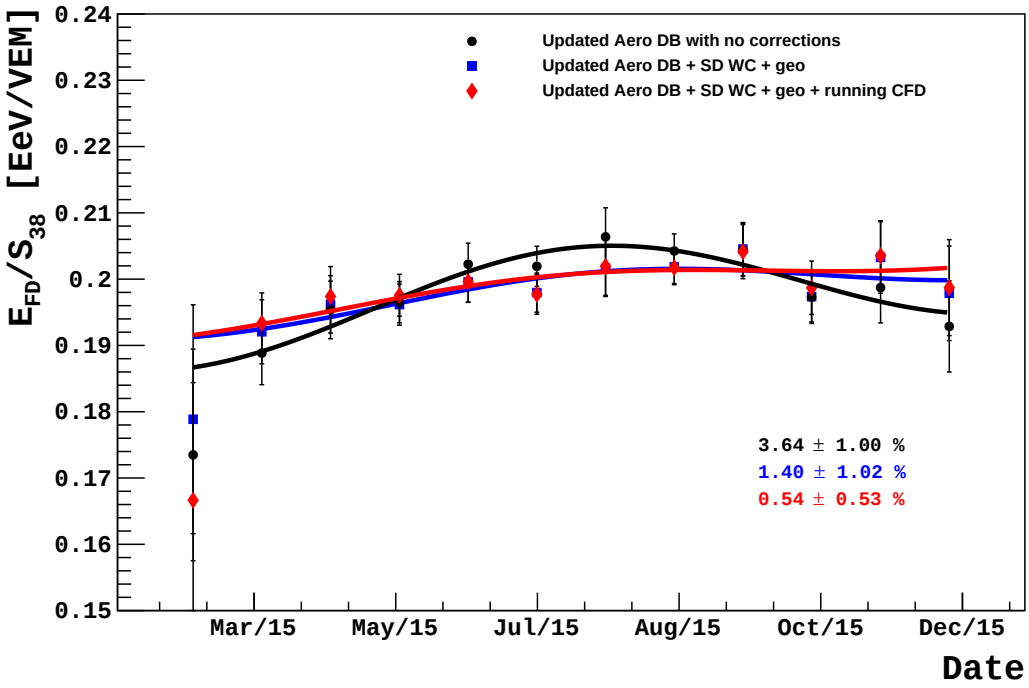


Figure 4.6: The ESR for the year 2015 to compare the three different improvements done to the hybrid showers but with an improved aerosol database done by V. Harvey. The black circles represent data reconstructed using the updated aerosol database only (Updated Aero DB with no corrections) while the blue squares represent this same data but with weather and geomagnetic field corrections (Updated Aero DB + SD WC + geo). The red diamonds represent data with previous corrections as well as running calibrations applied (Updated Aero DB + SD WC + geo + running CFD). The amplitude of the seasonal modulation is shown in the bottom right hand corner for each fit.

4.3 Monitoring the Energy Scale Ratio in the Long Term

The energy scales for the FD at the Pierre Auger Observatory are subject to many external changes during data acquisition. One of the factors that could influence the ESR is the accumulation of dust and dirt on the filters and mirrors of the FD. This can have an impact on the reflection and transmission of fluorescence light which will influence the determination of the FD energy, E_{FD}. Due to this, cleaning campaigns were introduced to maintain the optical transmission properties of the FD for accurate photon flux measurements and energy determination. This section begins with some plots from previous studies [110], showing the behaviour of a normalised ESR as a function of time, with labelled filter and mirror cleaning dates. This section then presents an extension of this analysis to look at the specific behaviour of each FD site and telescope from 2004 to the end of 2017 using a complete list of mirror and filter cleanings.

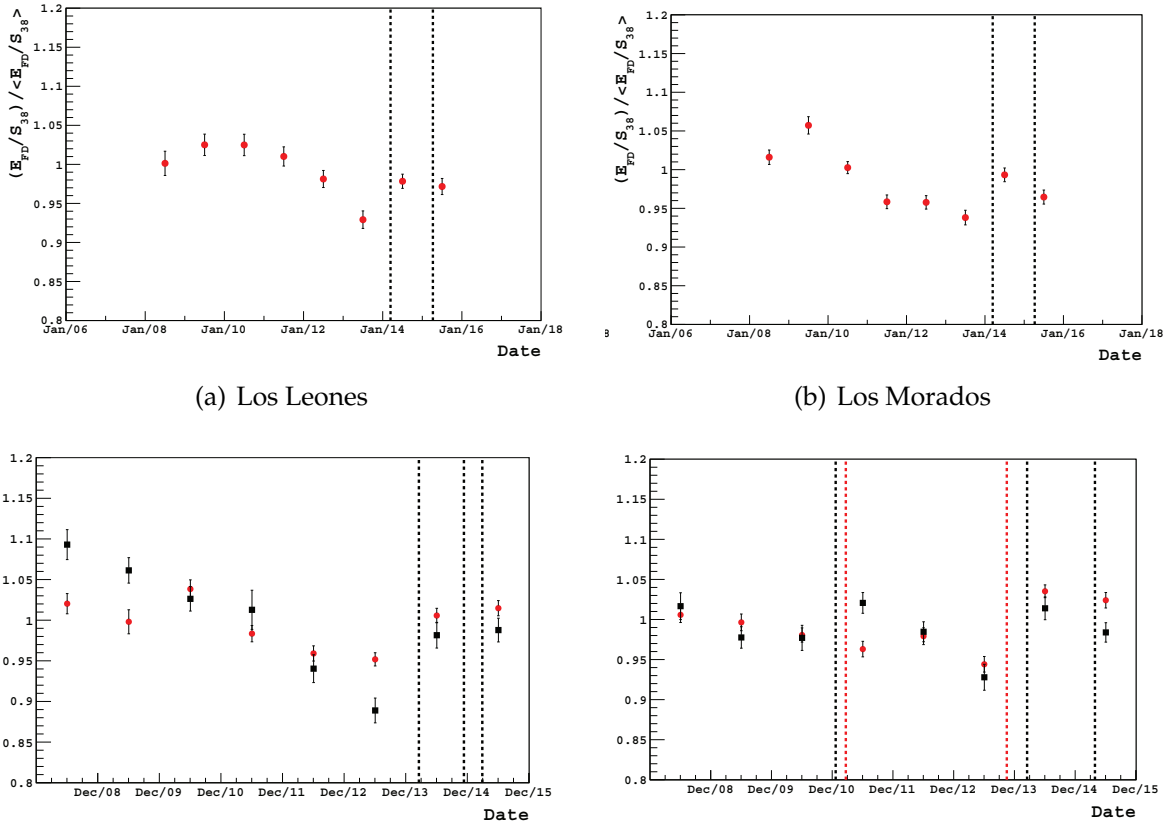
4.3.1 Energy Ratio as a Function of Time for Each FD Site

The energy ratio for four FD sites, which are Los Leones, Los Morados, Loma Amarilla and Coihueco were observed. Figure 4.7 shows a previous study of the behaviour of the ESR as a function of time for each FD site as done by Nguyen. It was found that there seems to be a correlation between the occurrence of a step in the ESR and when the filters were cleaned.

As an extension to the analysis done by Nguyen, this work uses 3344 *Golden Hybrid ICRC19* events with the same quality cuts as mentioned in Appendix A. Figure 4.8 shows the energy ratio, E_{SD}/E_{FD} for every FD site between 2004 and 2017. The ratio as a function of time for each FD site shows a different modulation around 1 for each FD site. The behaviour shown in Figure 4.8 is approximately illustrating an inverse of the behaviour in Figure 4.7 done by Nguyen since this study illustrates the ratio of SD energy to FD energy rather than FD energy to the signal, S_{38}. In Figure 4.7 only the filter and mirror cleanings near the 2010 and 2014 discontinuity were plotted to observe any possible correlations with the behaviour of the ESR [110].

4.3.2 Energy Ratio as a Function of Time for Each Telescope

To extend this study further, the energy ratio can be calculated for each individual telescope noting that there may be insufficient data for some telescopes in particular. To observe the variation in the energy ratio for each telescope, a plot showing a representative ratio as a function of telescope was generated. It is important to note that when averaged over all time and events, this ratio should give 1 as the SD energy is calibrated using the FD energy. Figure 4.9 shows the average energy ratio for each telescope. Each data point represents



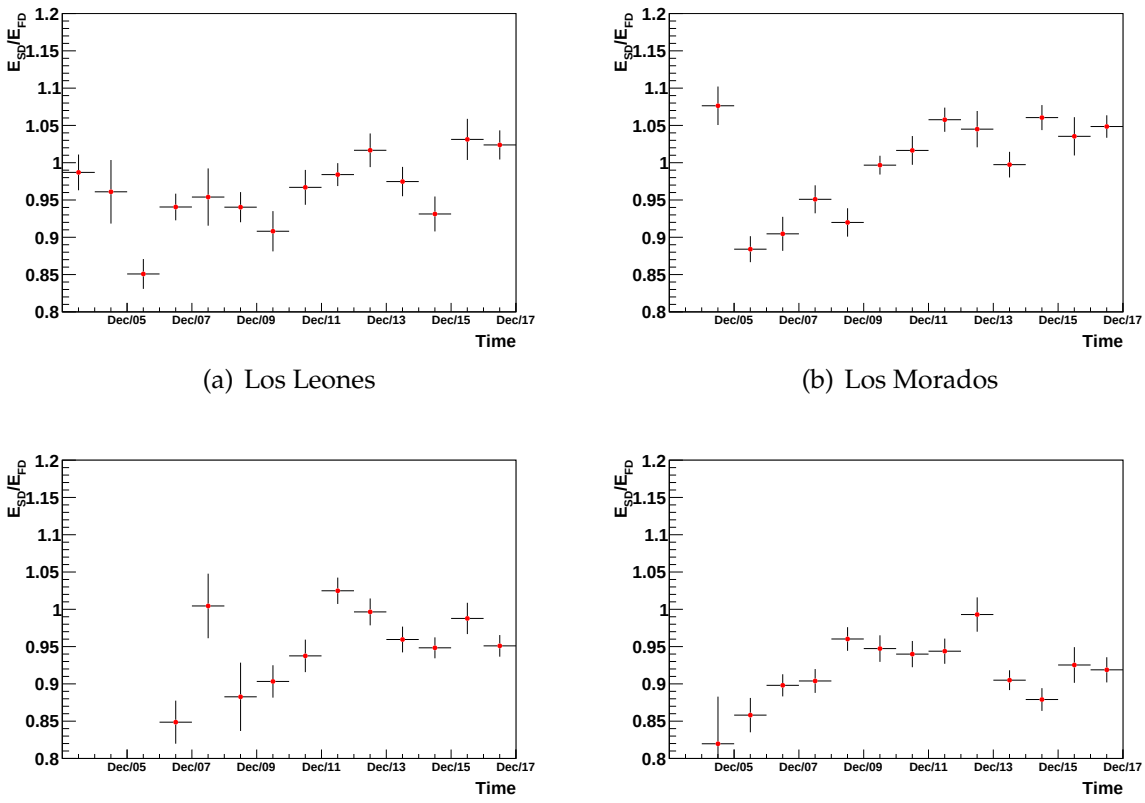
(a) Los Leones

(b) Los Morados

(c) Loma Amarilla

(d) Coihueco

Figure 4.7: Plots of normalised energy ratio vs time done by Nguyen [110]. Red circles represent the normalised ESR, defined as the ESR value in each bin, divided by the average ESR calculated as the mean of all the datapoints in all time bins. The black squares present in the plots for Loma Amarilla and Coihueco make up a star calibration profile obtained in Nguyen's work, unrelated to this work [110]. All datapoints are in yearly bins. The black dashed vertical lines represent the dates of filter cleanings around the 2010 and 2014 discontinuity. Red dashed lines represent the date of mirror cleanings for Coihueco (no mirror cleanings were performed for the other three sites near the 2010 and 2014 discontinuities in this time range).



(a) Los Leones

(b) Los Morados

(c) Loma Amarilla

(d) Coihueco

Figure 4.8: Ratio of SD energy to FD energy as a function of time for each FD site. The ratios are weighted according to the energy resolution functions for both SD and FD which will be explained in Section 4.3.2. It is also noted that since this illustrates the ratio of two energies for the same events, in theory it should yield 1 as a constant value throughout time. The filter and mirror cleanings were not plotted here in this work, as it was uncertain as to which specific dates were chosen to be plotted by Nguyen in Figure 4.7 (for comparison).

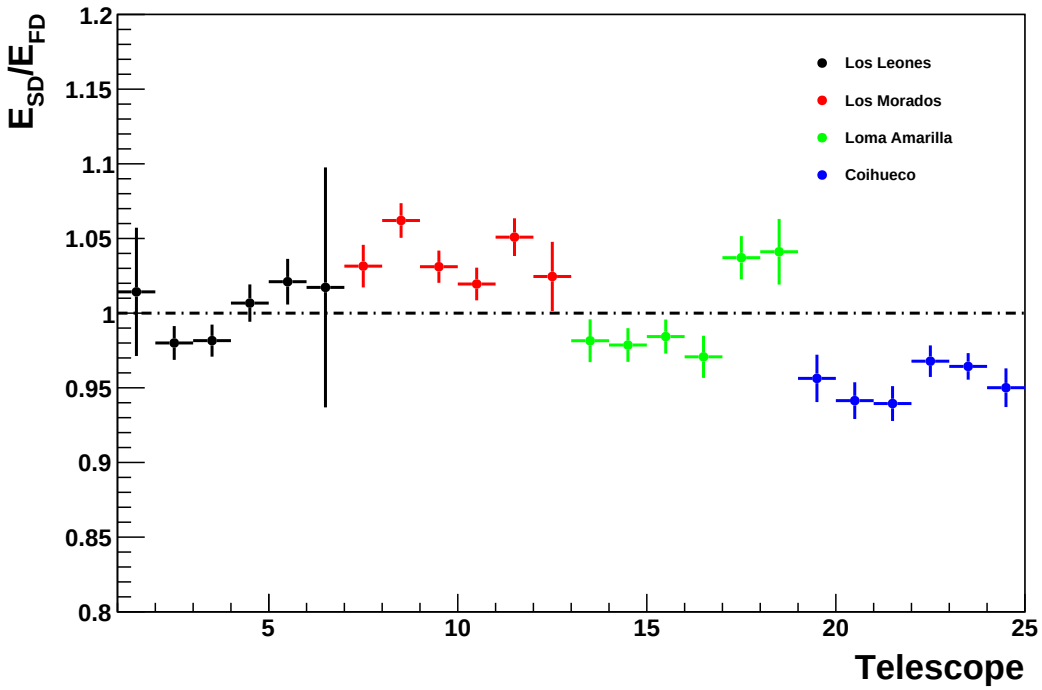


Figure 4.9: The energy ratio E_{SD}/E_{FD} (for events subject to cuts in Appendix A) of each telescope is plotted where a set of 6 telescopes from each FD site have a distinctive colour. The points showing the energy ratio for telescopes at Los Leones are black, for Los Morados; red, for Loma Amarilla; green and Coihueco; blue. The x-axis represents the telescope numbers 1 to 6 for each FD site, where telescope 1 of Los Leones is at x = 1.5 and telescope 6 at Coihueco is at x = 24.5. A black horizontal dashed line at E_{SD}/E_{FD} = 1 is shown just for display.

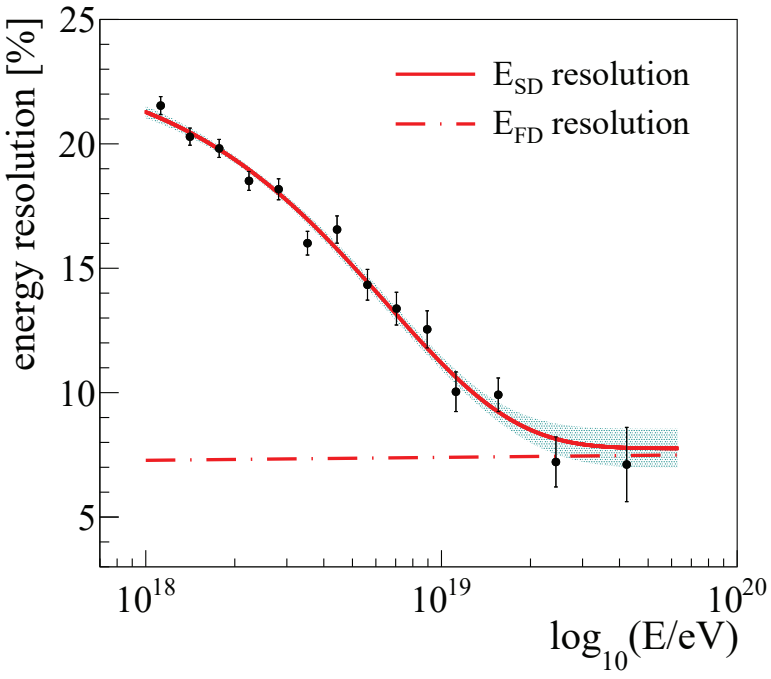


Figure 4.10: E_{SD} and E_{FD} energy resolutions as a function of energy and their fitted functional forms [131]. The green band represents the systematic uncertainty of the SD energy resolution.

an averaged value over a time period of 14 years from 2004-2017 for every event observed by that telescope while the error bar is the standard error on the mean. The behaviour shown in Figure 4.9 indicates that every telescope has a different calibration or energy threshold sensitivity. This emphasises the need to investigate any specific behaviours of individual fluorescence telescopes in terms of energy assignments in the long term. A quantitative analysis of this correlation is essential in order to reduce the systematic uncertainties present in the energy spectrum. On top of being subject to filter and mirror cleanings, the drum calibration explained in 3.2.1.1 can affect the telescope sensitivity during data acquisition.

Then, the behaviour of the energy ratio for individual hybrid events is observed as a function of time, with dates of filter and mirror cleanings obtained from an updated list in a log file available in the following reference [130]. Due to relative ease of cleaning, filters are cleaned more frequently than mirrors. Also, due to the energy resolution function for the SD and FD, the uncertainties in the energy ratio depend on SD and FD energy. There are two functions which describe the energy resolution for both the SD and FD and these are illustrated in Figure 4.10 [131]. Figure 4.10 shows the energy resolution function for both SD and FD. The functional form of the FD energy resolution, $\sigma_{E_{FD}}$ is found to be constant and yields a value of 7.4% [128]. The functional form of the SD energy

resolution is shown as

$$\sigma_{\mathrm{E_{SD}}} = 0.078 + 0.160e^{(-0.15E)} \tag{4.2}$$

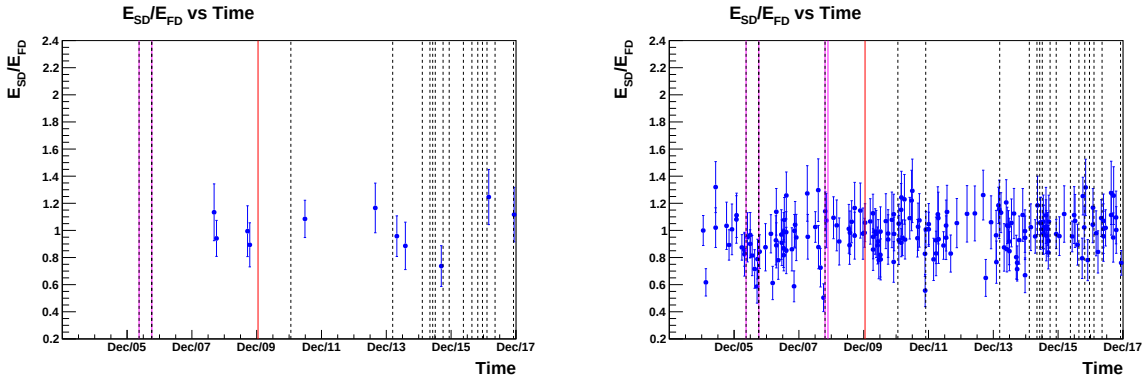
where $\sigma_{\mathrm{E_{SD}}}$ is the SD energy resolution (%) and E is energy in EeV. From this, the propagated uncertainty for the energy ratio of each event is calculated as shown in the following:

$$\sigma_i = \frac{\mathrm{E_{SD}}}{\mathrm{E_{FD}}} \times \sqrt{\frac{\sigma^2_{\mathrm{E_{SD}}}}{\mathrm{E^2_{SD}}} + \frac{\sigma^2_{\mathrm{E_{FD}}}}{\mathrm{E^2_{FD}}}} \tag{4.3}$$

where σ_i is the propagated statistical uncertainty of $\frac{\mathrm{E_{SD}}}{\mathrm{E_{FD}}}$ for each event. Figures 4.11-4.14 show the energy ratio for every individual event as a function of time for all 24 telescopes with filter and mirror cleanings plotted as black dashed and fluorescent vertical lines. The red line marks the date of when the drum calibration was conducted (January of 2010) [110]. It was noted that there are less events detected by telescopes 1 and 6 of Los Leones compared to others. This is due to the cameras which are filled with only half the number of pixels compared to the other telescopes. It was also noted that telescope 6 of Coihueco had some technical issues causing the lack of data post-2015. It is hypothesised that the energy ratio would increase as a function of time after long periods without any cleaning campaigns. This is due to the diffuse scattering of fluorescence light caused by dust particles on the filter and mirror, resulting in the loss of light detected by the cameras of the FD and a lower determination of $\mathrm{E_{FD}}$.

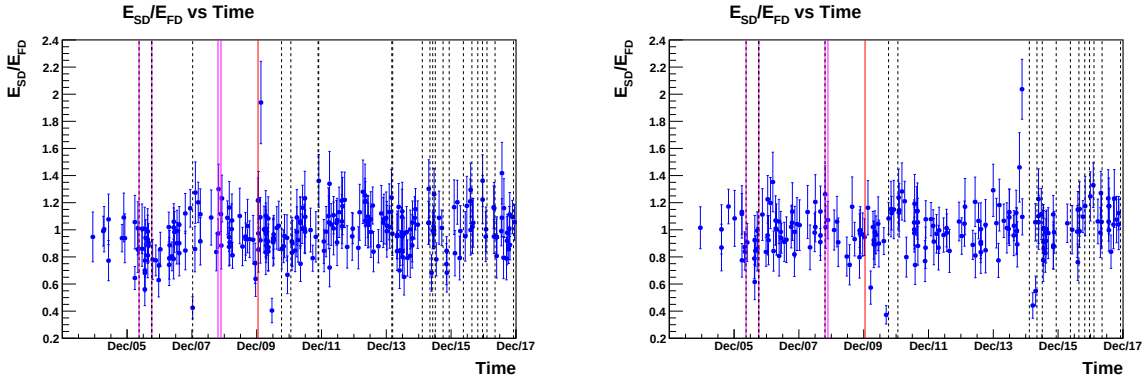
Profile plots were generated to obtain the average behaviour of the energy scale ratio for each telescope. These are shown in Figures 4.15 - 4.30 showing the weighted ratios seen by each telescope, where the calculations are shown in this section. Firstly, it is noted that in the years 2015, 2016 and 2017, the frequency of cleanings started to increase as there was a team available on-site to perform the cleanings more regularly. Due to the frequency of cleanings for most of the telescopes and to ensure there is a clear distinction between events detected before and after cleanings, the binning process for the datapoints are altered according to each telescope. Firstly, periods between cleaning campaigns are what determine the sizes of bins and hence the bin limits. Then, it was calculated that if the duration between two cleanings is less than one year, the bin size is increased until the next cleaning happens, which determines the next bin limit. On the other hand, if the duration between two cleanings is more than one year, the bin will be assigned a size of one year starting from the earlier cleaning. The mean of the energy ratio is calculated using the weighted mean method due to the energy resolution function causing variation in SD and FD energy uncertainties [132]. The weighted mean of energy ratios for every bin, μ' is calculated as

$$\mu' = \frac{\Sigma x_i/\sigma_i^2}{\Sigma 1/\sigma_i^2} \tag{4.4}$$



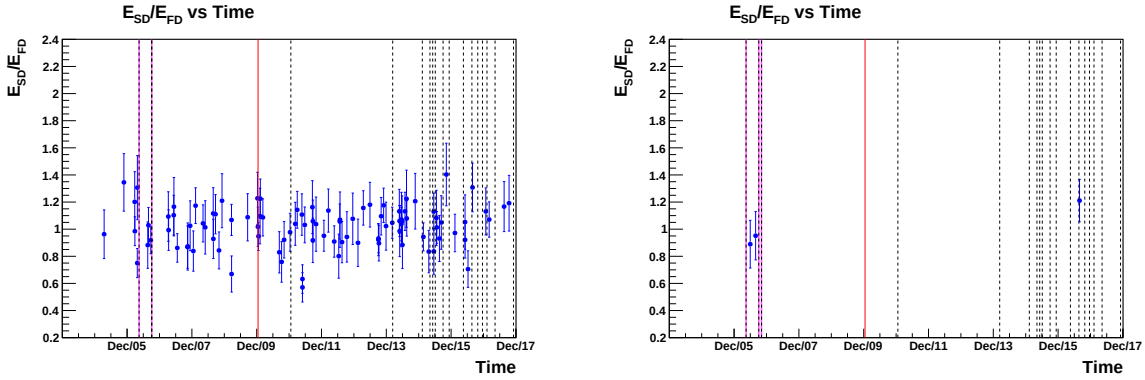
(a) Los Leones telescope 1

(b) Los Leones telescope 2



(c) Los Leones telescope 3

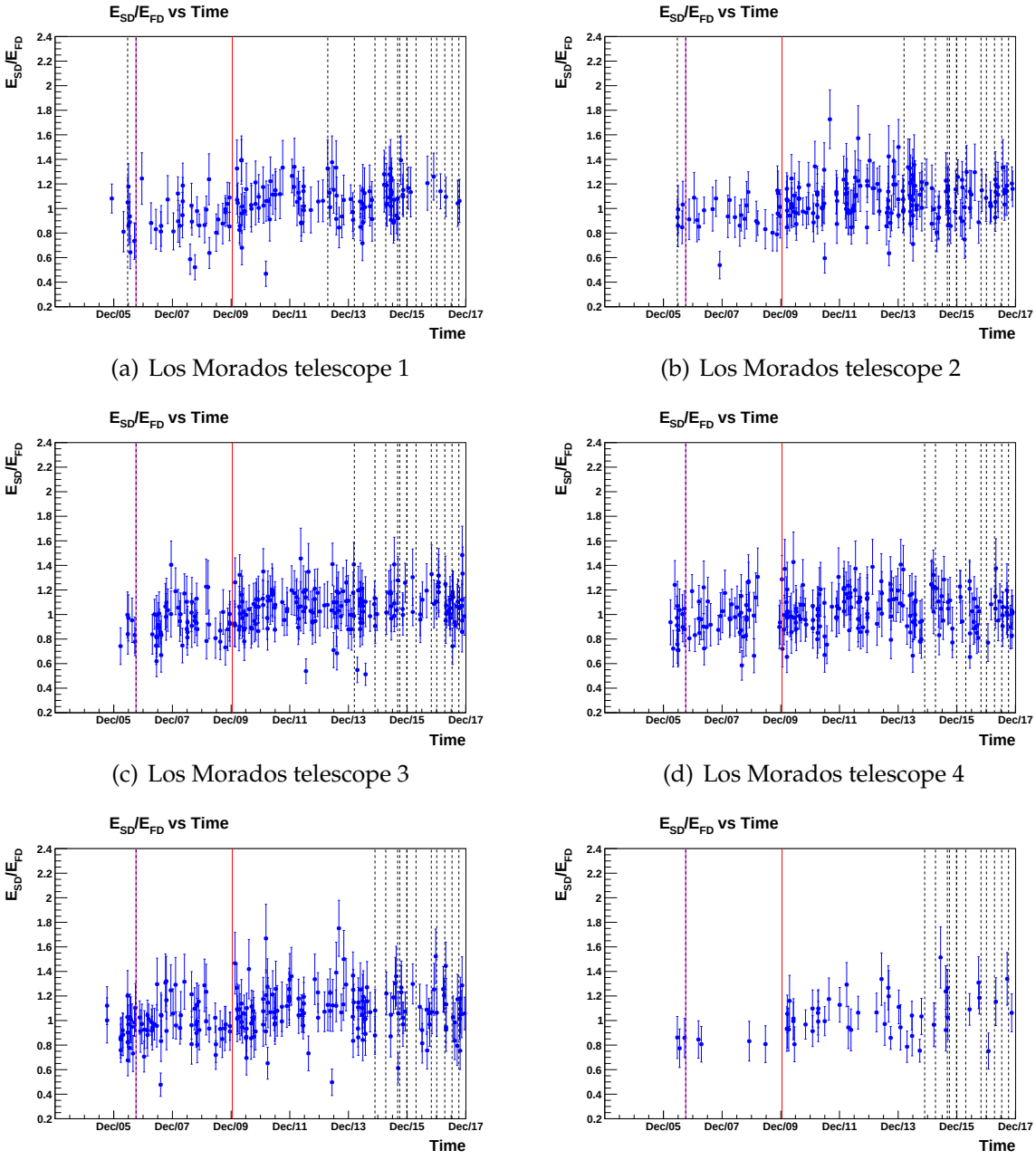
(d) Los Leones telescope 4



(e) Los Leones telescope 5

(f) Los Leones telescope 6

Figure 4.11: Energy ratio of every individual event as a function of time detected by each telescope at Los Leones. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. It is also noted that in telescopes 1 and 6, there are very few events. This is due to the field of view of the telescopes mostly pointing outside the area where water-Cherenkov detectors are present, as well as the cameras being filled with only half the number of pixels.



(a) Los Morados telescope 1

(b) Los Morados telescope 2

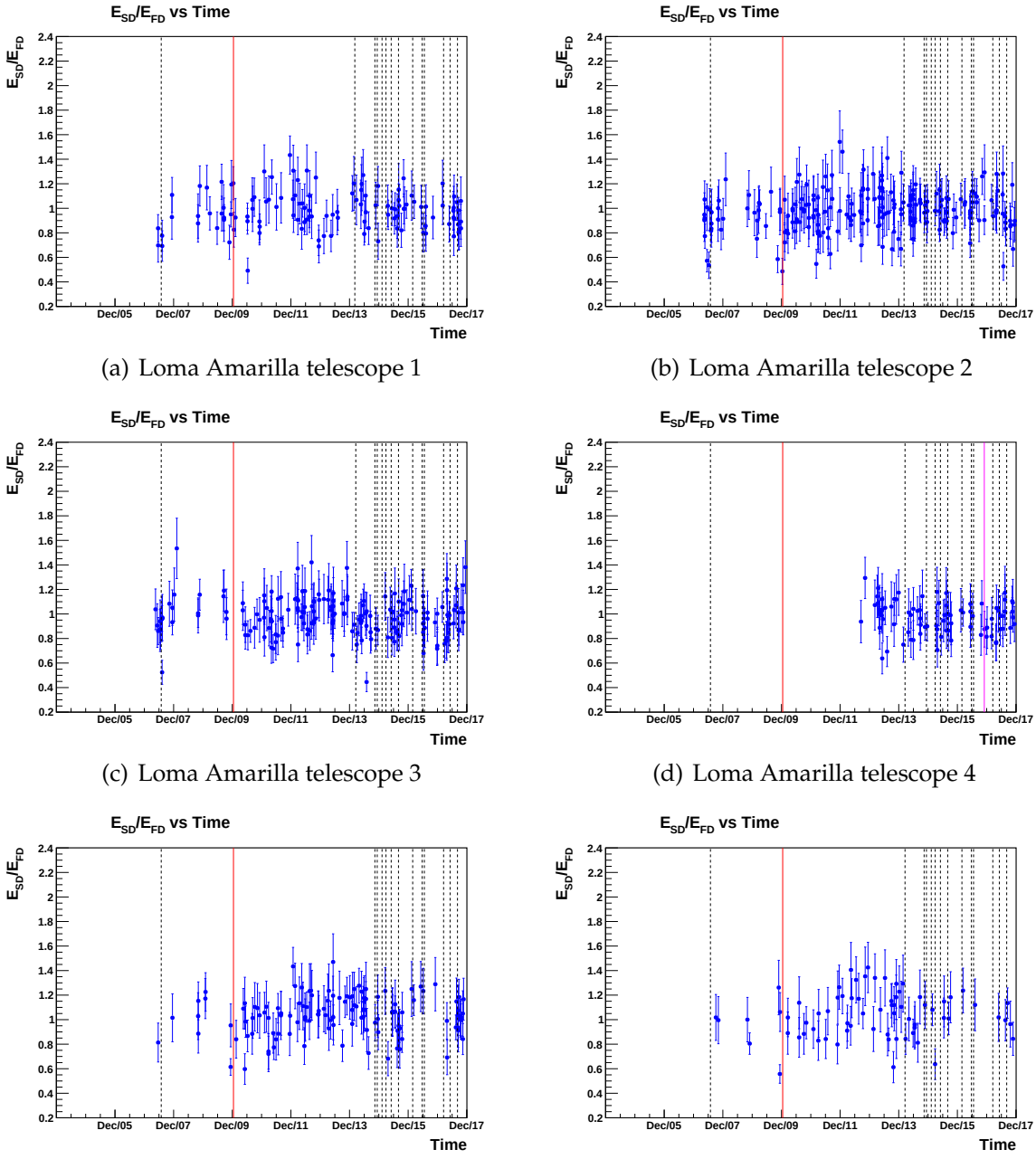
(c) Los Morados telescope 3

(d) Los Morados telescope 4

(e) Los Morados telescope 5

(f) Los Morados telescope 6

Figure 4.12: Energy ratio of every individual event as a function of time for each telescope at Los Morados. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings.



(a) Loma Amarilla telescope 1

(b) Loma Amarilla telescope 2

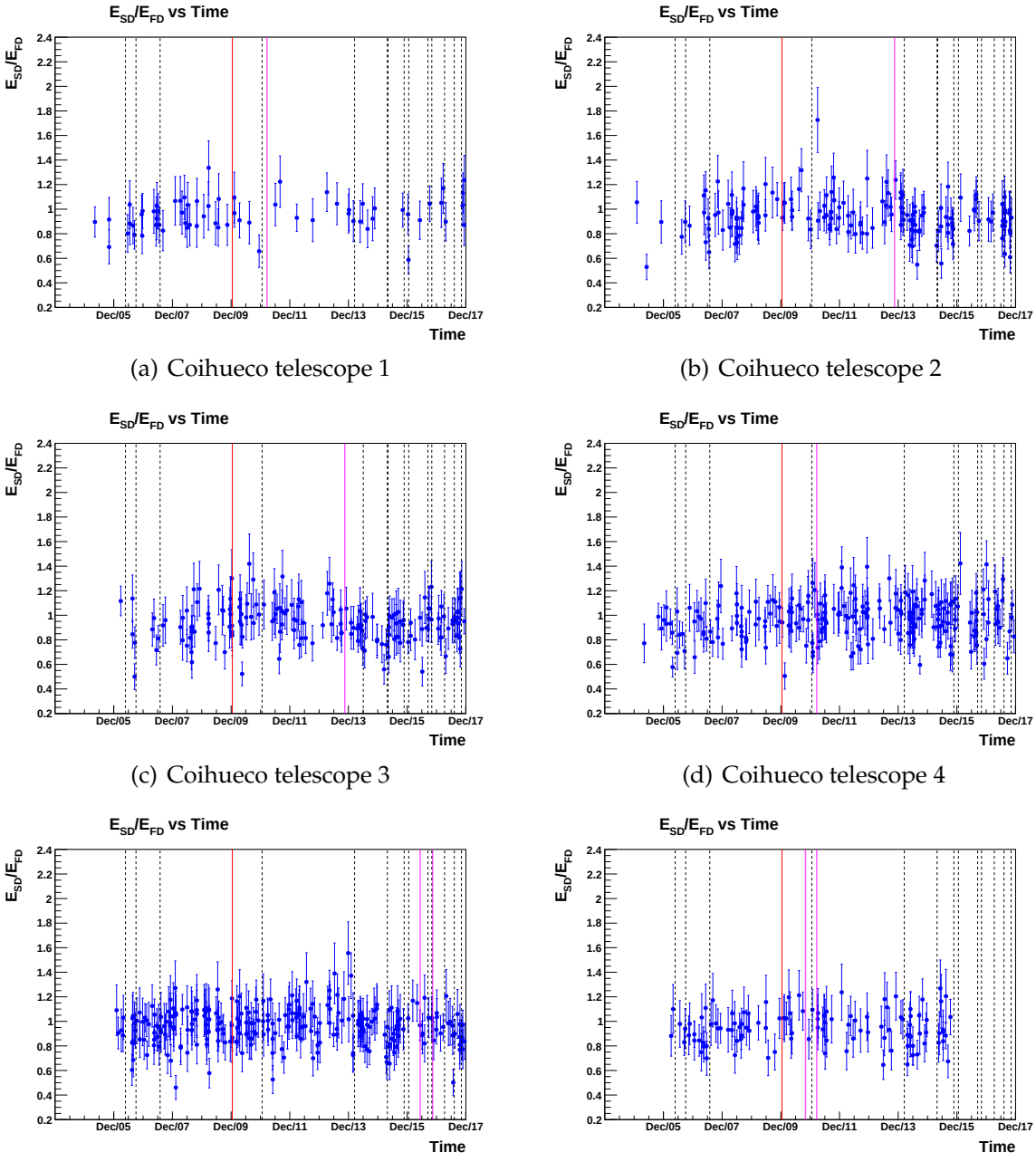
(c) Loma Amarilla telescope 3

(d) Loma Amarilla telescope 4

(e) Loma Amarilla telescope 5

(f) Loma Amarilla telescope 6

Figure 4.13: Energy ratio of every individual event as a function of time for each telescope at Loma Amarilla. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings.



(a) Coihueco telescope 1

(b) Coihueco telescope 2

(c) Coihueco telescope 3

(d) Coihueco telescope 4

(e) Coihueco telescope 5

(f) Coihueco telescope 6

Figure 4.14: Energy ratio of every individual event as a function of time for each telescope at Coihueco. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. Due to a technical issue, no data was recorded for telescope 6 post-2015.

where

$$x_i = \frac{\mathrm{E_{SD}}}{\mathrm{E_{FD}}} \tag{4.5}$$

for every event and σ_i is as shown in Equation 4.3. The weighted error of the energy ratio for every bin, σ_μ is calculated as the following:

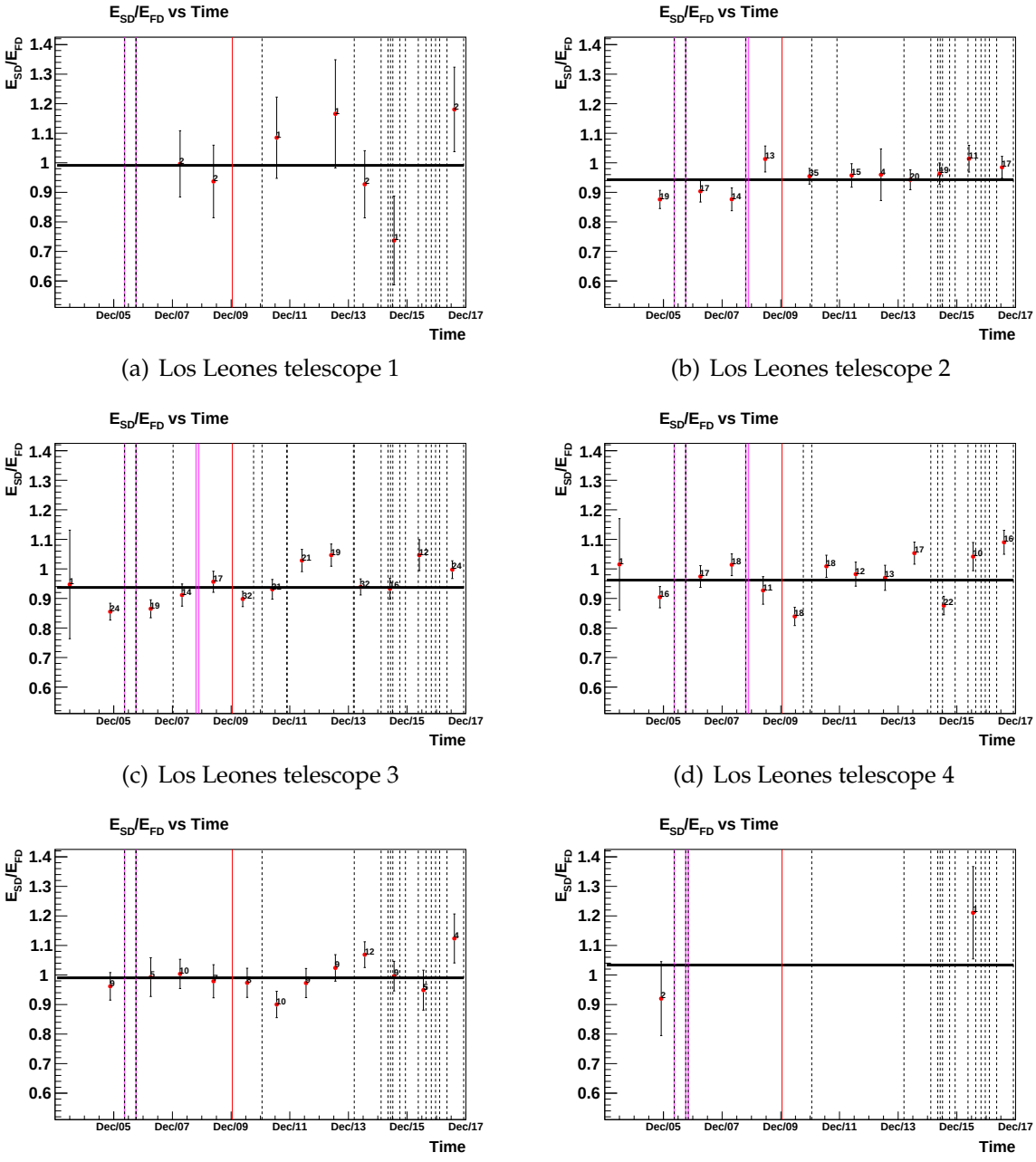
$$\sigma_\mu = \frac{1}{\sqrt{\Sigma 1/\sigma_i^2}} \tag{4.6}$$

It is expected that there would be a varying trend in energy measurement as time progresses due to the cleaning campaigns and absolute calibration. An analysis quantifying this trend was done by looking at the goodness of different types of fits applied to plots of energy ratio for each telescope.

4.3.2.1 Horizontal Fit Results

Horizontal line fits for Los Leones, Los Morados, Loma Amarilla and Coihueco are shown in Figures 4.15, 4.16, 4.17 and 4.18. It was expected that after filter and mirror cleanings are done, the energy ratio would drop, due to an increase in $\mathrm{E_{FD}}$ since the dust and dirt causing diffuse scattering of light is reduced. However, from these plots, there is a clear variation in time of the energy ratio irrespective of the cleanings and this could be due to the high frequency of cleanings in some telescopes. But there are some examples showing a clear drop such as at the start of 2014 for Los Leones telescopes 1, 3 and 4, telescopes 1, 3 and 5 at Los Morados, telescopes 3 and 5 at Loma Amarilla and all telescopes at Coihueco. Also, it seems to be inconclusive whether there is a correlation between the drum calibration and the energy ratio since for some telescopes it seems to either increase or decrease in an unpredictable manner.

The fit parameter, χ^2 and reduced χ^2 values are obtained and the probabilities are calculated for every fit. These values are available in Tables 4.2 and 4.4. It can be seen that all FD sites have at least one fit probability which has a deviation from the hypothesis that is beyond 3σ indicating an unlikely horizontal trend. The most promising horizontal fit results are shown by Coihueco where telescope 2 is the only telescope that has a deviation beyond 3σ. The normalisations (y-intercepts) of the horizontal line fits seem to vary between 0.937 ± 0.009 to 1.0 ± 0.1 which is expected since intrinsically, the FD energy and SD energy obtained from the same cosmic ray air shower event should be equal. For completeness, a plot of the energy ratio as a function of telescope as shown in Figure 4.9 was modified to include the weights induced by the energy resolutions of both the FD and the SD. This is shown in Figure B.1 in Appendix B and it can be seen that the energy ratio values match the horizontal intercepts shown in Table 4.4.



(a) Los Leones telescope 1

(b) Los Leones telescope 2

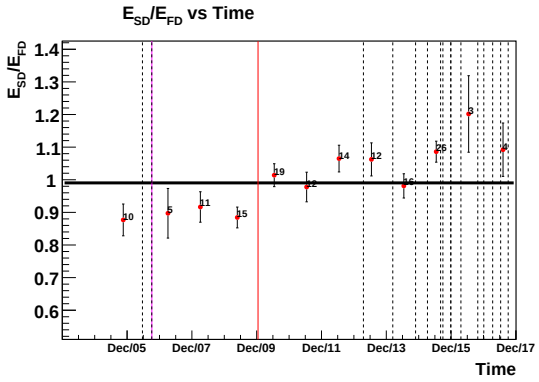
(c) Los Leones telescope 3

(d) Los Leones telescope 4

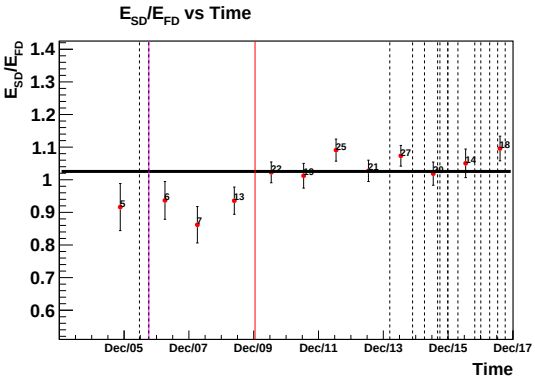
(e) Los Leones telescope 5

(f) Los Leones telescope 6

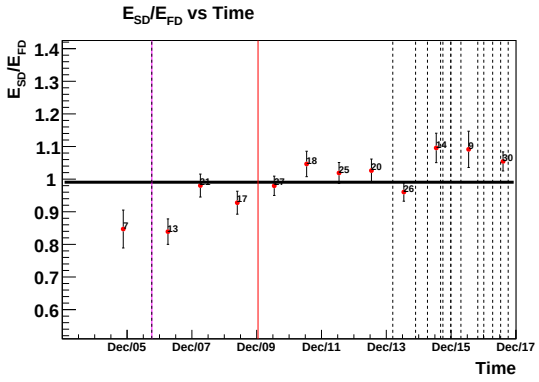
Figure 4.15: Horizontal line fits applied on the energy ratio as a function of time for each telescope at Los Leones. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



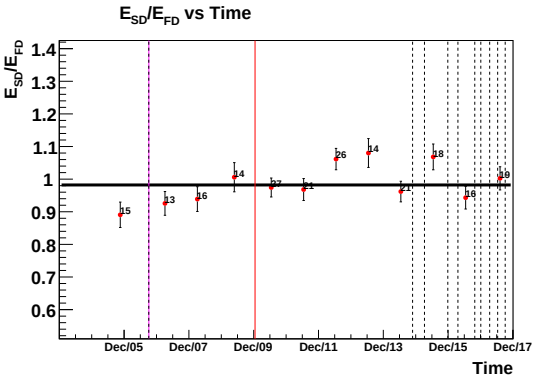
(a) Los Morados telescope 1



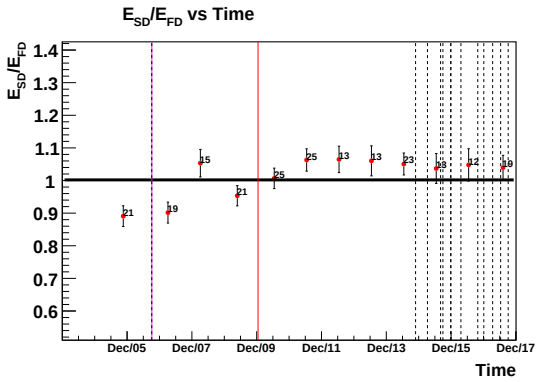
(b) Los Morados telescope 2



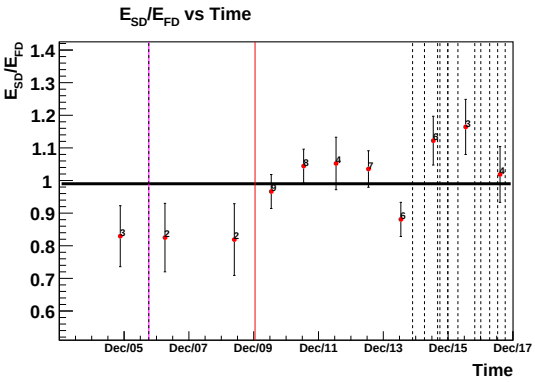
(c) Los Morados telescope 3



(d) Los Morados telescope 4

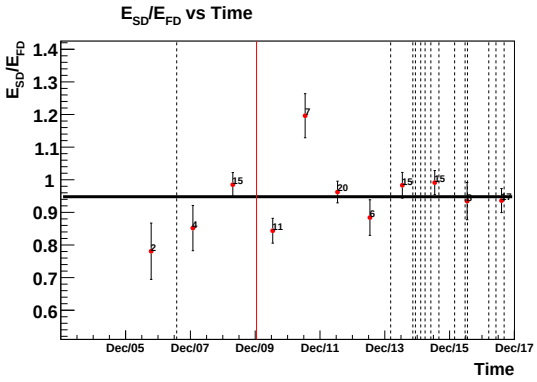


(e) Los Morados telescope 5

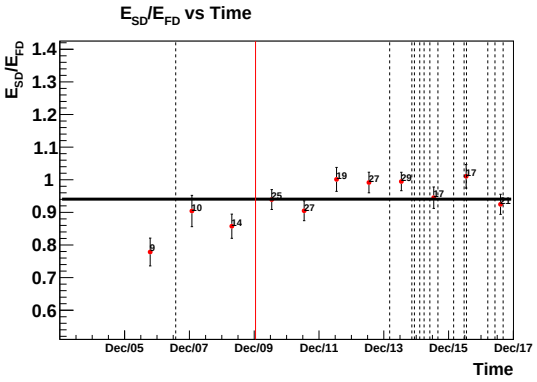


(f) Los Morados telescope 6

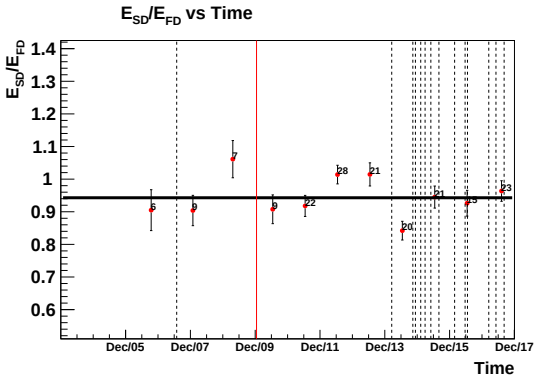
Figure 4.16: Horizontal line fits on the energy ratio as a function of time for each telescope at Los Morados. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



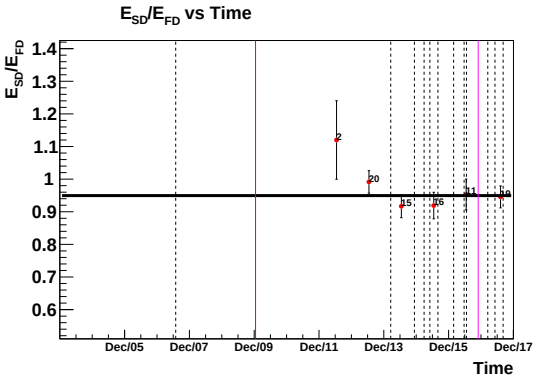
(a) Loma Amarilla telescope 1



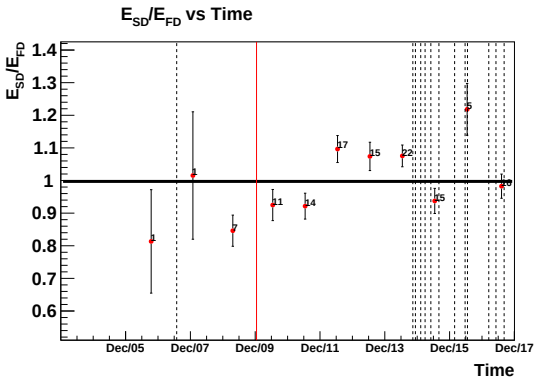
(b) Loma Amarilla telescope 2



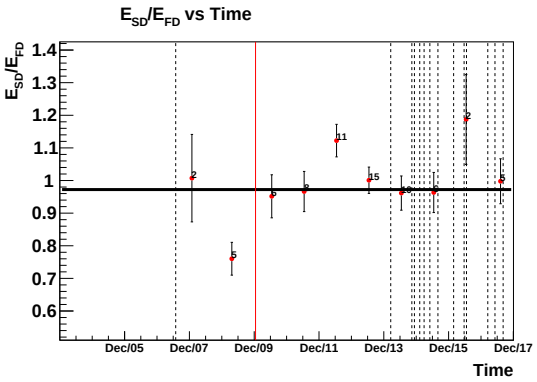
(c) Loma Amarilla telescope 3



(d) Loma Amarilla telescope 4

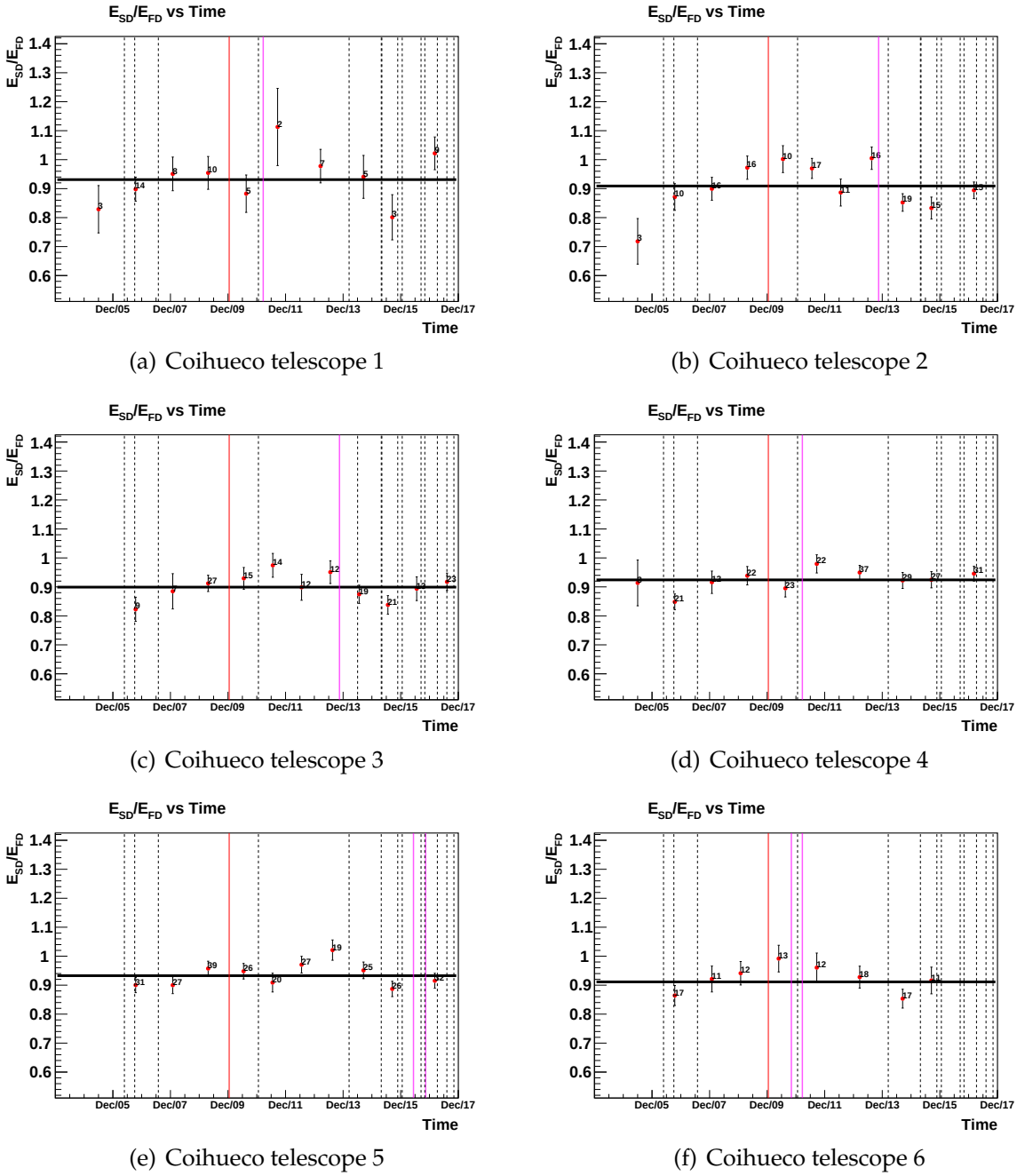


(e) Loma Amarilla telescope 5



(f) Loma Amarilla telescope 6

Figure 4.17: Horizontal line fits on the energy ratio as a function of time for each telescope at Loma Amarilla. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



(a) Coihueco telescope 1

(b) Coihueco telescope 2

(c) Coihueco telescope 3

(d) Coihueco telescope 4

(e) Coihueco telescope 5

(f) Coihueco telescope 6

Figure 4.18: Horizontal line fits on the energy ratio as a function of time for each telescope at Coihueco. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.

4.3.2.2 Linear Fit Results

A linear line fit was then applied to all telescope energy ratio plots to test for improvement of χ^2 values and probabilities. The fit equation used is as follows:

$$y(t) = m\left(\frac{t - t_0}{T}\right) + c \tag{4.7}$$

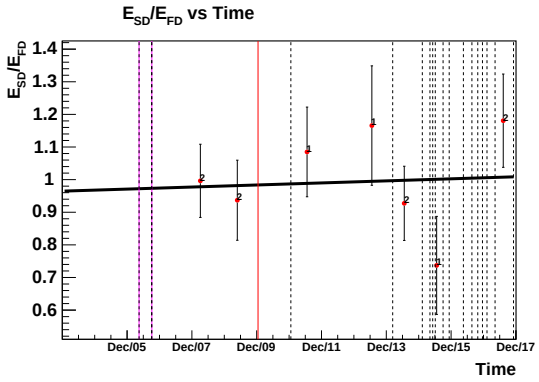
where *y(t)* is the energy ratio of the fit at time *t*, *m* is the slope parameter, *t* is time in seconds since the start time, t_0 at 2004, *T* is the number of seconds in a year and *c* is the y-intercept parameter. The results in Table 4.2 show that the linear fits for only five out of the 24 telescopes have a deviation beyond 3σ. This excludes Los Leones telescope 6 since the fit is trivial, with 2 parameters and 2 datapoints having a perfect fit. A substantial statistical improvement is prominent in Los Morados where five of the telescopes have fits with a deviation between 0-2σ and one with a deviation between 2σ-3σ, indicating that a linear fit in energy ratio is more likely than a horizontal trend. It is also shown from Table 4.4 that there is a positive slope in the energy ratio for Los Leones and Los Morados, which go up to 2.2 ± 0.4 % per year. The linear fits which include Los Leones telescope 4, Loma Amarilla telescopes 1, 3 and 5 as well as Coihueco telescope 2 which have deviations beyond 3σ indicate the presence of a complex structure in the trend which cannot be visualised with a simple linear fit. This emphasises the need for a combination horizontal-linear, sinusoid or polynomial fit to see if the reduced χ^2 (χ^2_{red}) values improve.

4.3.2.3 Linear-Horizontal Results

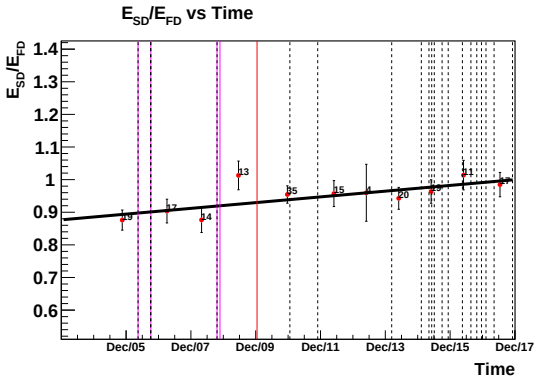
Another fit that was applied to the energy ratios consist of two lines, separated by a breakpoint where one is linear and the other is horizontal. The fit function is shown as follows

$$y(t) = \begin{cases} m\left(\frac{t-t_0}{T}\right) + c & \frac{t-t_0}{T} < b \\ mb + c & \frac{t-t_0}{T} \geq b \end{cases} \tag{4.8}$$

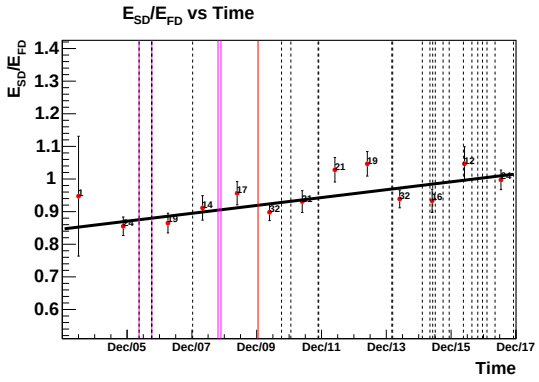
where *y(t)* is the energy ratio as a function of time, *m* is the slope in units of year^{-1}, *t* time in seconds since t_0 at 2004, *T* is the number of seconds in one year, *c* is the *y*-intercept and *b* is the breakpoint as the number of years after 2004. It is noted that different ranges of fit parameter values were allowed for all fits, to allow for a converged fit for all telescope energy ratio behaviours. Figures 4.23-4.26 show the fit results while the goodness-of-fit values and parameters were calculated and listed in Tables 4.3 and 4.5. Los Leones telescope 4 seems to have a more complex structure than a simple horizontal, linear or linear-horizontal fit, since it has a deviation beyond 3σ for all three fits. It can also be seen that the χ^2_{red} values regress from the linear fit to the linear-horizontal fit for Los Morados Telescope 1. There is also a sign of over-fitting for telescope 2. The goodness-of-fit value for Coihueco telescope 2 changes from having a deviation beyond 3σ in



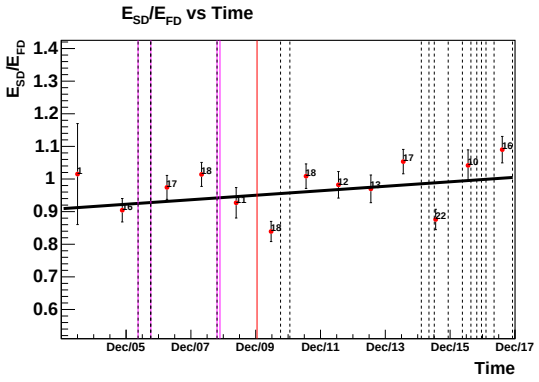
(a) Los Leones telescope 1



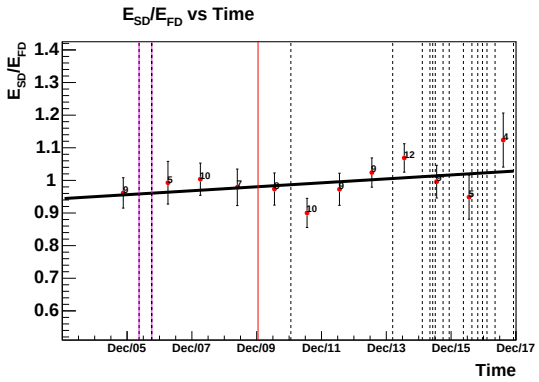
(b) Los Leones telescope 2



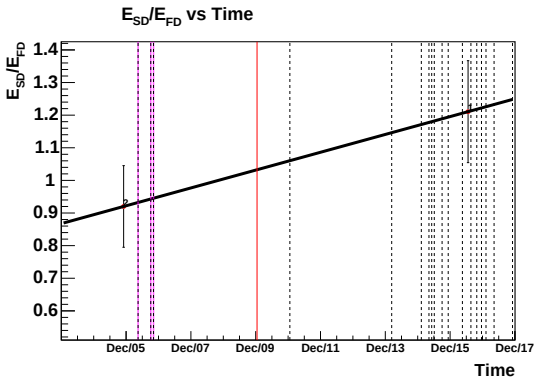
(c) Los Leones telescope 3



(d) Los Leones telescope 4

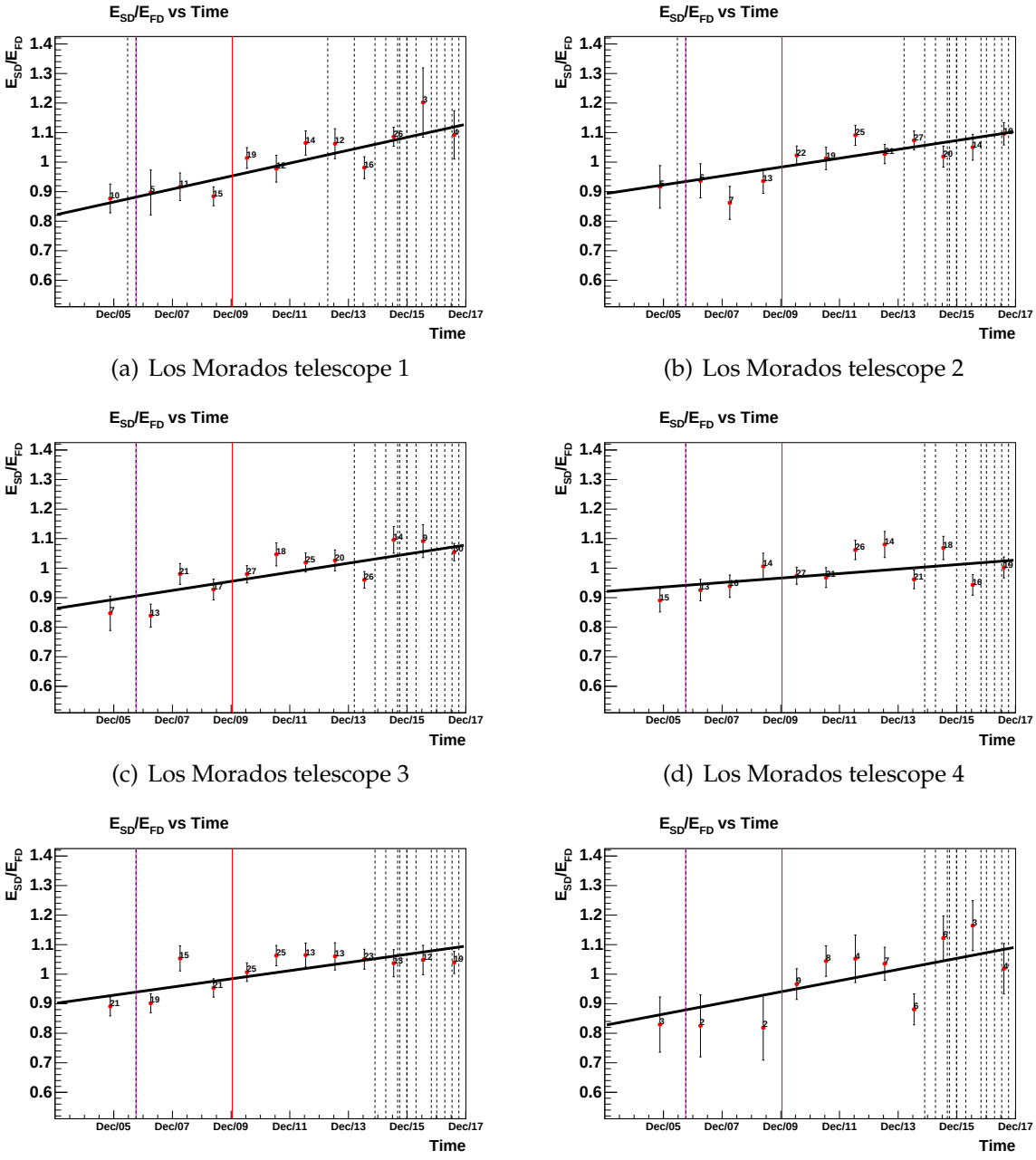


(e) Los Leones telescope 5



(f) Los Leones telescope 6

Figure 4.19: Linear fits on the energy ratio as a function of time for each telescope at Los Leones. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint. The result of the fit in telescope 6 is trivial since there are only two datapoints and 2 parameters making the χ^2_{red} is infinite.



(a) Los Morados telescope 1

(b) Los Morados telescope 2

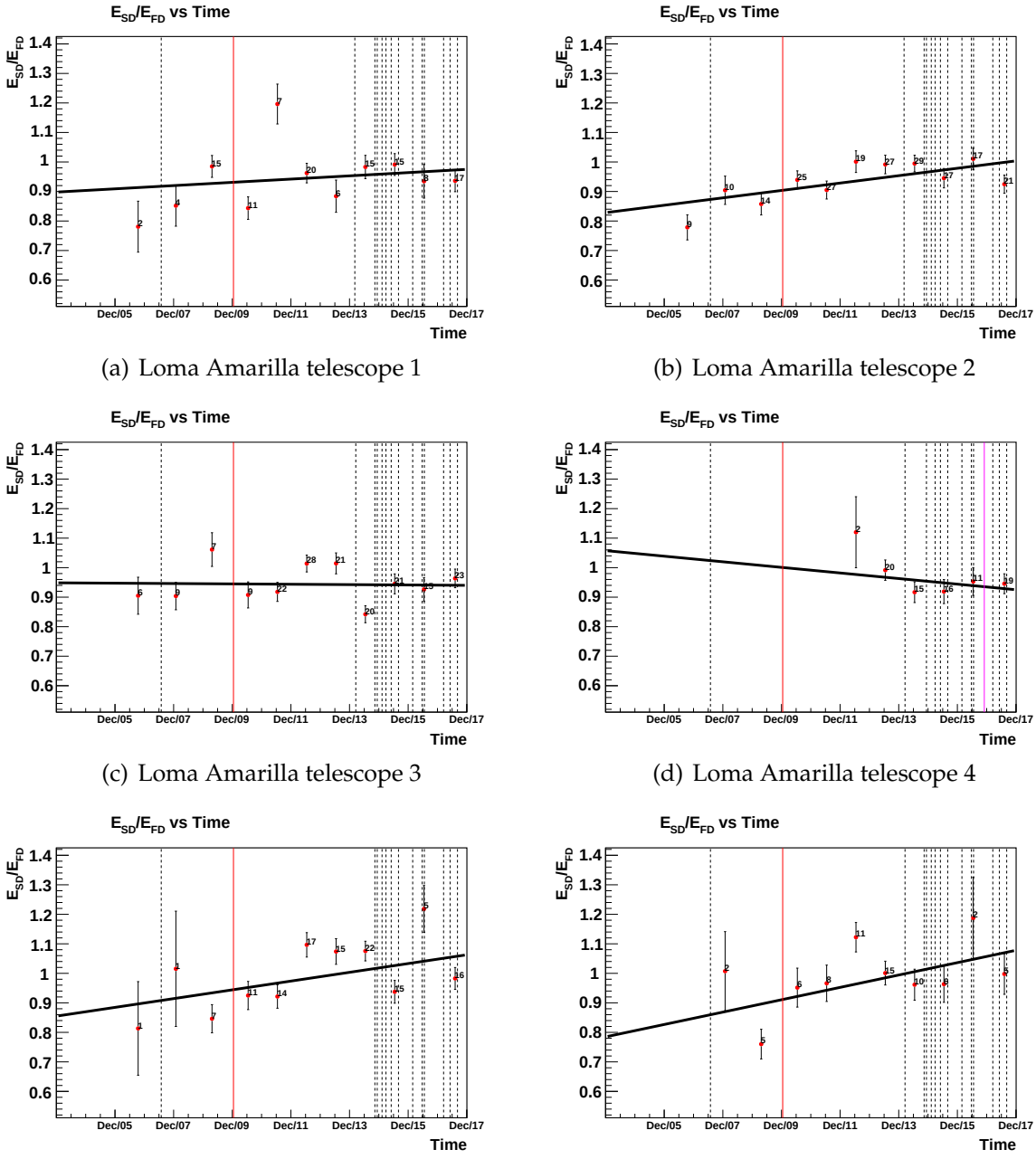
(c) Los Morados telescope 3

(d) Los Morados telescope 4

(e) Los Morados telescope 5

(f) Los Morados telescope 6

Figure 4.20: Linear fits on the energy ratio as a function of time for each telescope at Los Morados. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



(a) Loma Amarilla telescope 1

(b) Loma Amarilla telescope 2

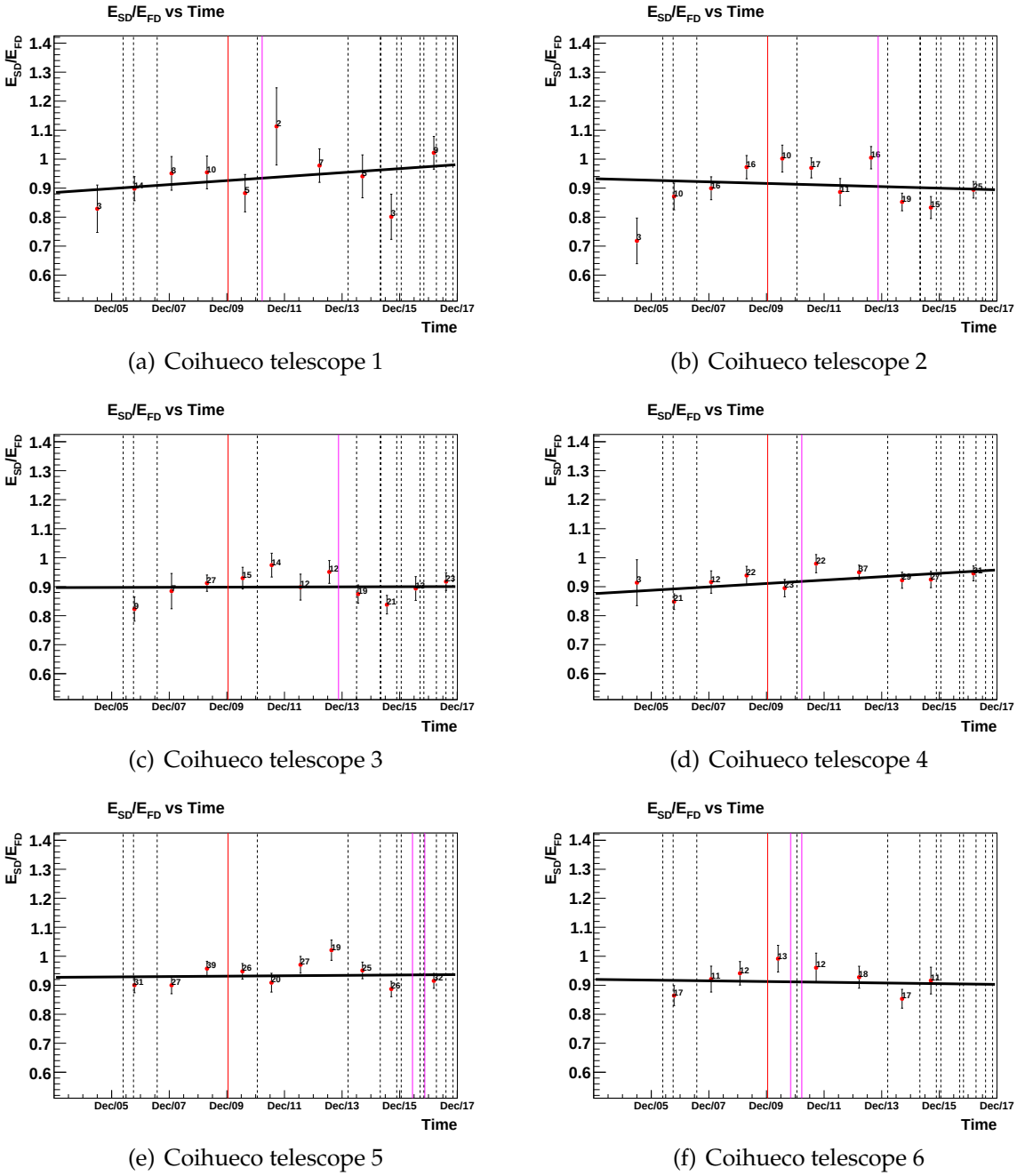
(c) Loma Amarilla telescope 3

(d) Loma Amarilla telescope 4

(e) Loma Amarilla telescope 5

(f) Loma Amarilla telescope 6

Figure 4.21: Linear fits on the energy ratio as a function of time for each telescope at Loma Amarilla. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



(a) Coihueco telescope 1

(b) Coihueco telescope 2

(c) Coihueco telescope 3

(d) Coihueco telescope 4

(e) Coihueco telescope 5

(f) Coihueco telescope 6

Figure 4.22: Linear fits on the energy ratio as a function of time for each telescope at Coihueco. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.

the linear fit, to a deviation between 2-3σ in the linear-horizontal fit. All other telescopes at Coihueco have a deviation which is unchanged between the linear and linear-horizontal fits.

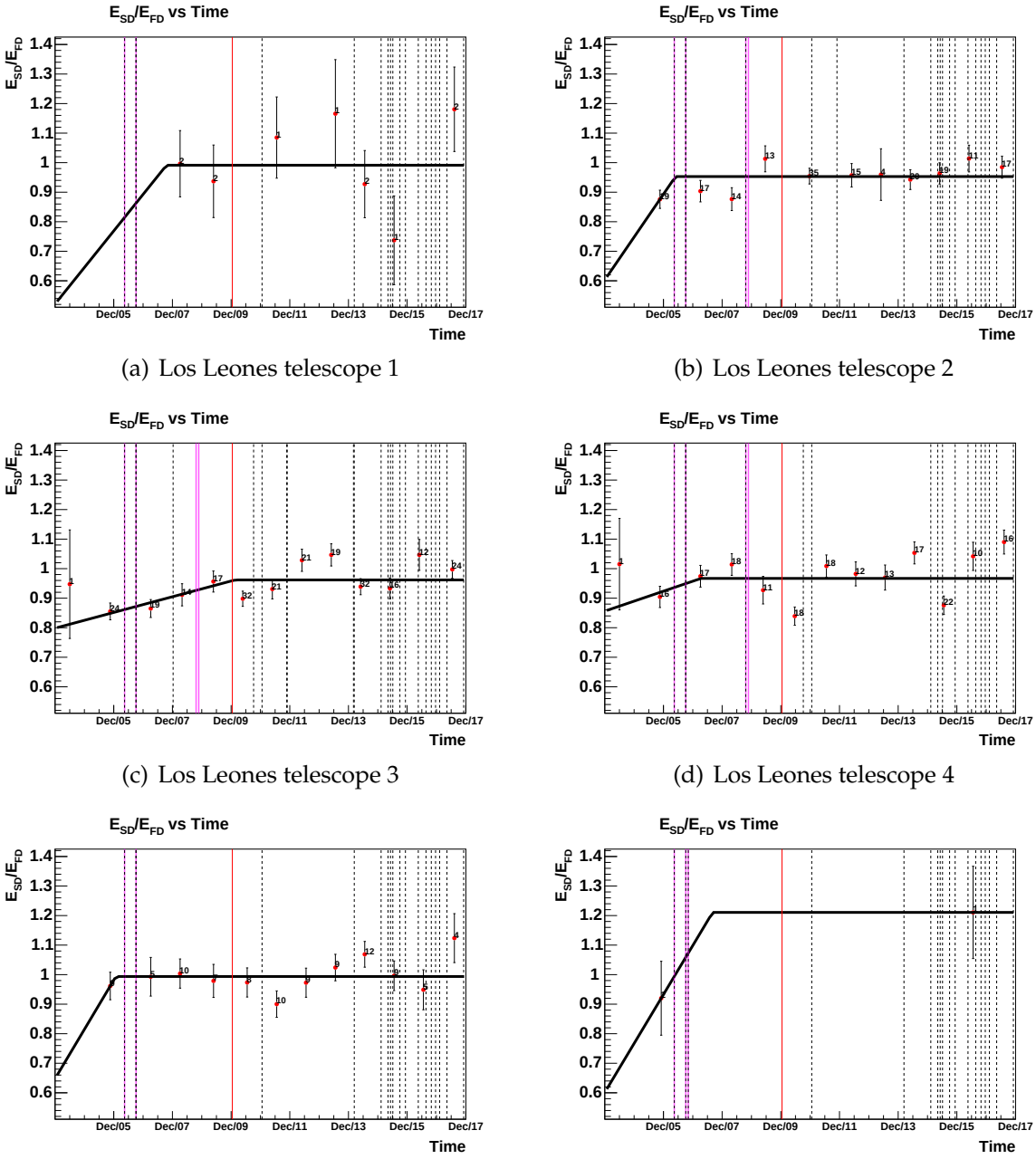
4.3.2.4 Linear-Linear Fit Results

Finally, two linear lines separated by a freely-moving breakpoint were fitted with the following functional form:

$$y(t) = \begin{cases} m_1 \left(\frac{t-t_0}{T}\right) + c_1 & \frac{t-t_0}{T} < b \\ m_2 \left(\frac{t-t_0}{T}\right) + c_2 & \frac{t-t_0}{T} \geq b \end{cases} \tag{4.9}$$

where *y(t)* is the energy ratio as a function of time, m_1 and m_2 are two different slopes both in units of year^{-1}, t is time in seconds since start time, t_0 at 2004, T is the number of seconds in one year, c_1 and c_2 are two distinct y-intercepts and b is the breakpoint as the number of years after 2004. Again, the results shown in Table 4.3 show that a linear-linear fit does not statistically improve the χ^2_{red} for Los Leones telescope 4, which indicates that a more complex function is needed to improve the fitting. This is unlike telescope 3 since the deviation lies between 0-2σ unlike in the linear-horizontal fit. There is also a statistical improvement in the χ^2_{red} for Los Morados telescope 1 and 3, indicating that a linear-linear trend is more favourable. For telescopes 5 and 6 of Loma Amarilla, the fits have improved from the linear-horizontal trend to a linear-linear trend as the deviations change from beyond 3σ to 2σ-3σ and 0-2σ. However, this is not the case for telescopes 1 and 3, since the deviations lie beyond 3σ for all fits as well. Lastly, telescope 2 of Coihueco has a statistical improvement in the fit from linear-horizontal to linear-linear since the deviation changes from 2-3σ to a range between 0-2σ. This indicates that a linear-linear fit is more suited to the behaviour of the energy ratio for all telescopes at Coihueco.

In terms of the fit parameters, Coihueco seems favour a positive slope before the breakpoint (placed at $\sim$ 2010 with the exception of Telescope 1) while after the breakpoint, three out of six of the telescopes exhibit an energy ratio having a negative slope. This positive-negative slope behaviour is the same for Loma Amarilla telescope 4 and 6, but not the case for all the telescopes at Los Morados (positive-positive). Telescopes 1, 2, 3 and 5 of Los Leones also exhibit a positive-positive slope.



(a) Los Leones telescope 1

(b) Los Leones telescope 2

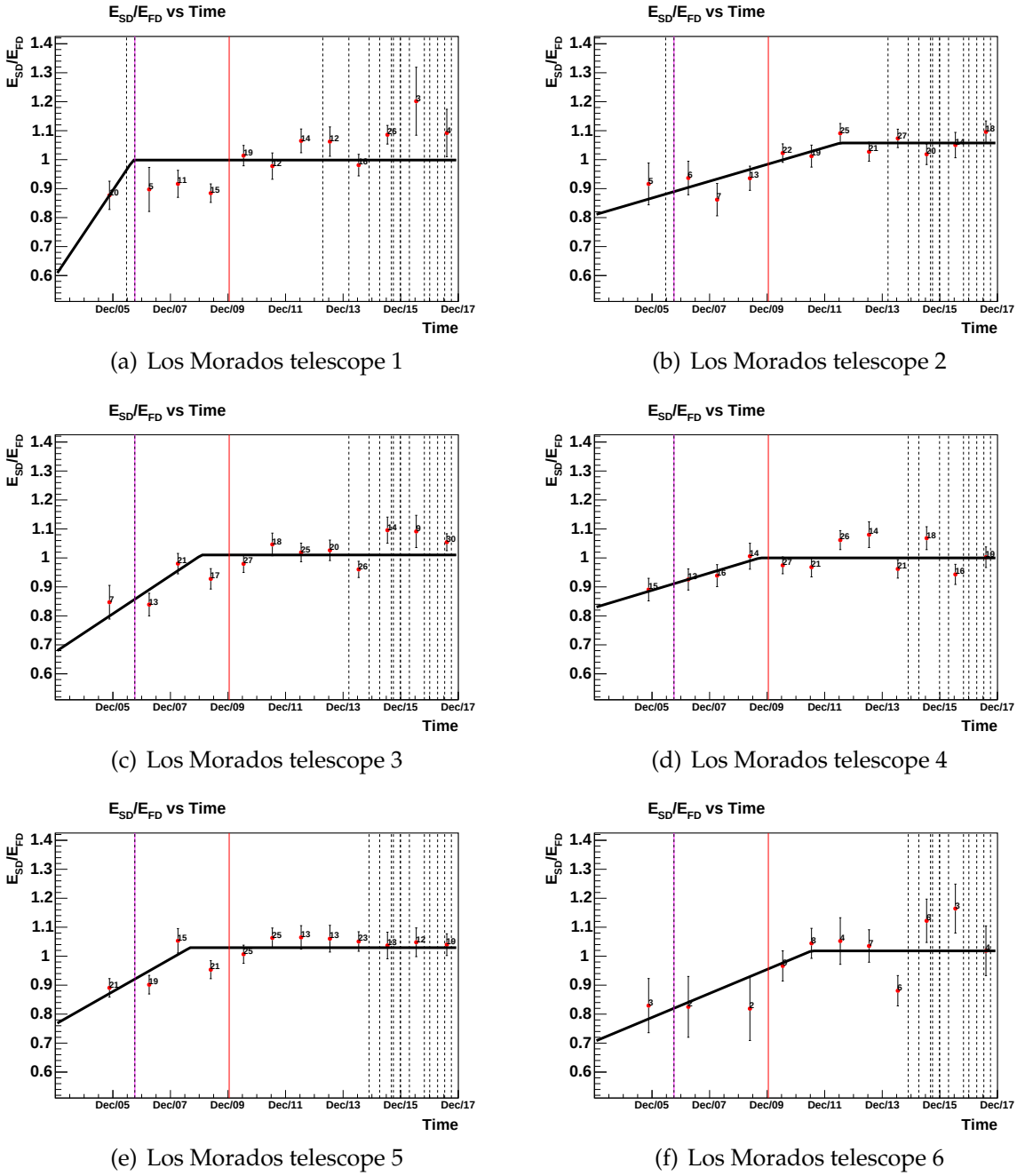
(c) Los Leones telescope 3

(d) Los Leones telescope 4

(e) Los Leones telescope 5

(f) Los Leones telescope 6

Figure 4.23: Linear-horizontal fits on the energy ratio as a function of time for each telescope at Los Leones. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint. The result of the fit in telescope 6 is trivial, since there are only two datapoints and more than 2 parameters making the χ^2_{red} values infinite.



(a) Los Morados telescope 1

(b) Los Morados telescope 2

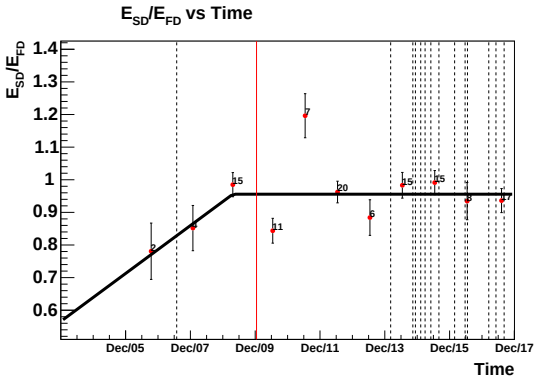
(c) Los Morados telescope 3

(d) Los Morados telescope 4

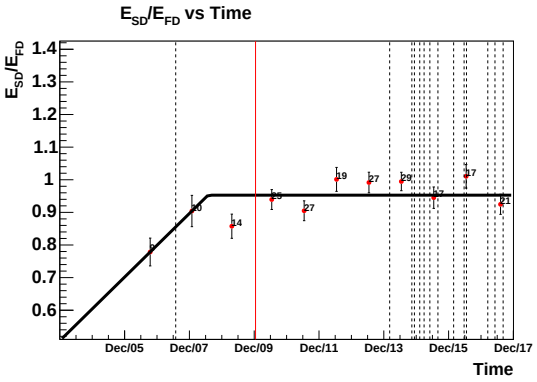
(e) Los Morados telescope 5

(f) Los Morados telescope 6

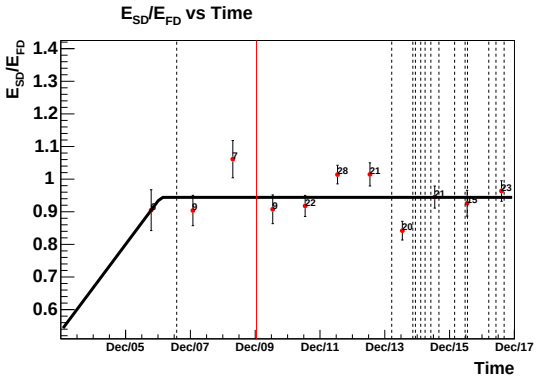
Figure 4.24: Linear-horizontal fits on the energy ratio as a function of time for each telescope at Los Morados. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



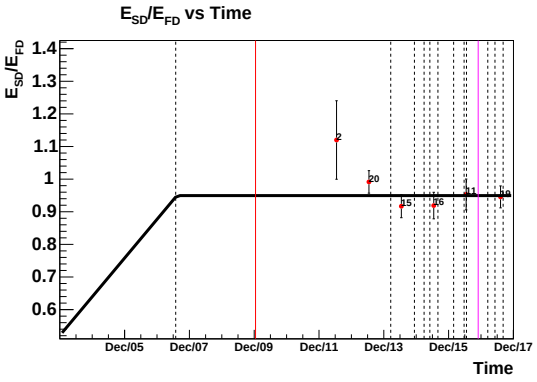
(a) Loma Amarilla telescope 1



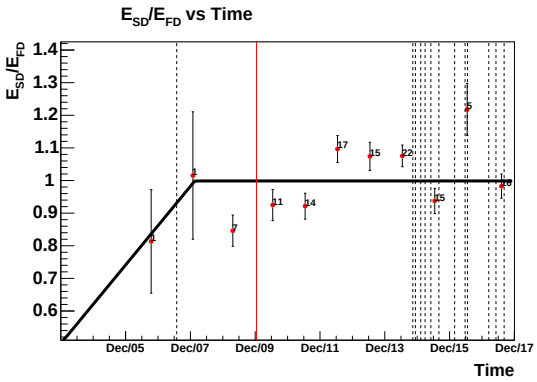
(b) Loma Amarilla telescope 2



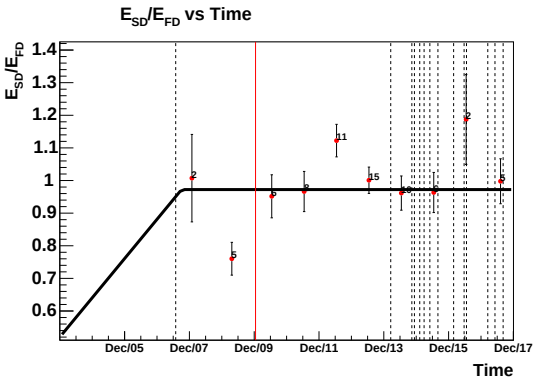
(c) Loma Amarilla telescope 3



(d) Loma Amarilla telescope 4

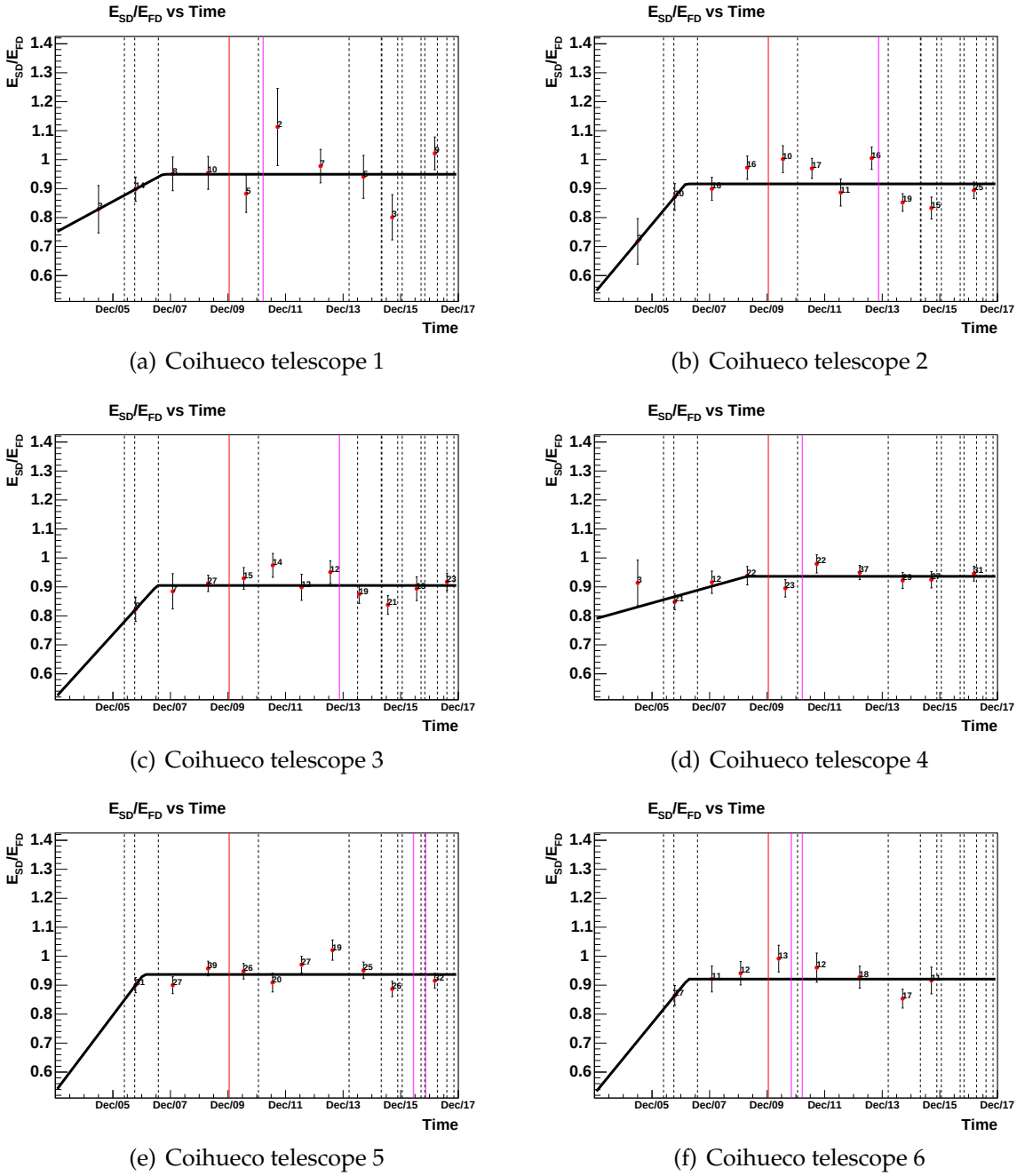


(e) Loma Amarilla telescope 5



(f) Loma Amarilla telescope 6

Figure 4.25: Linear-horizontal fits on the energy ratio as a function of time for each telescope at Loma Amarilla. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



(a) Coihueco telescope 1

(b) Coihueco telescope 2

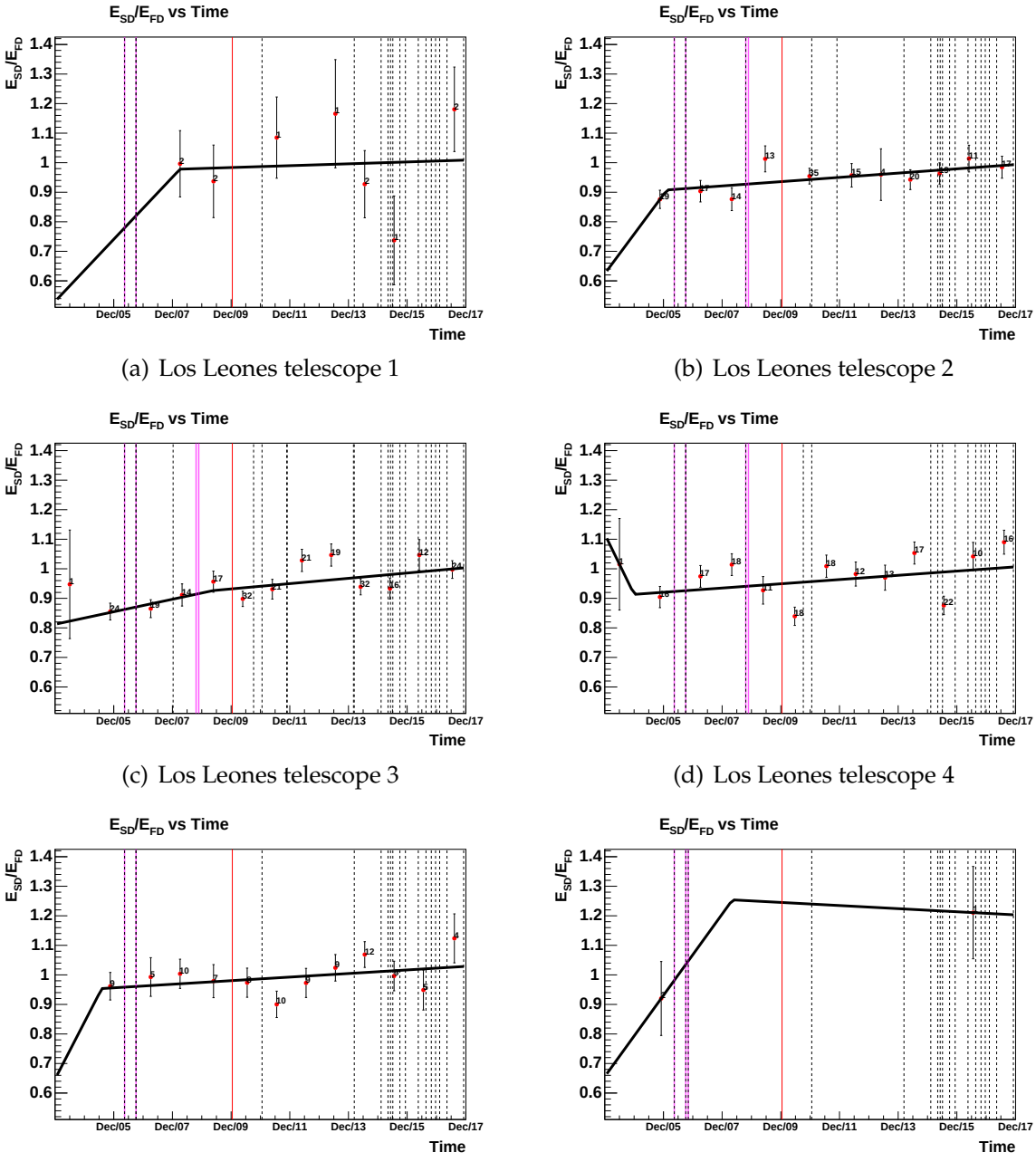
(c) Coihueco telescope 3

(d) Coihueco telescope 4

(e) Coihueco telescope 5

(f) Coihueco telescope 6

Figure 4.26: Linear-horizontal fits on the energy ratio as a function of time for each telescope at Coihueco. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



(a) Los Leones telescope 1

(b) Los Leones telescope 2

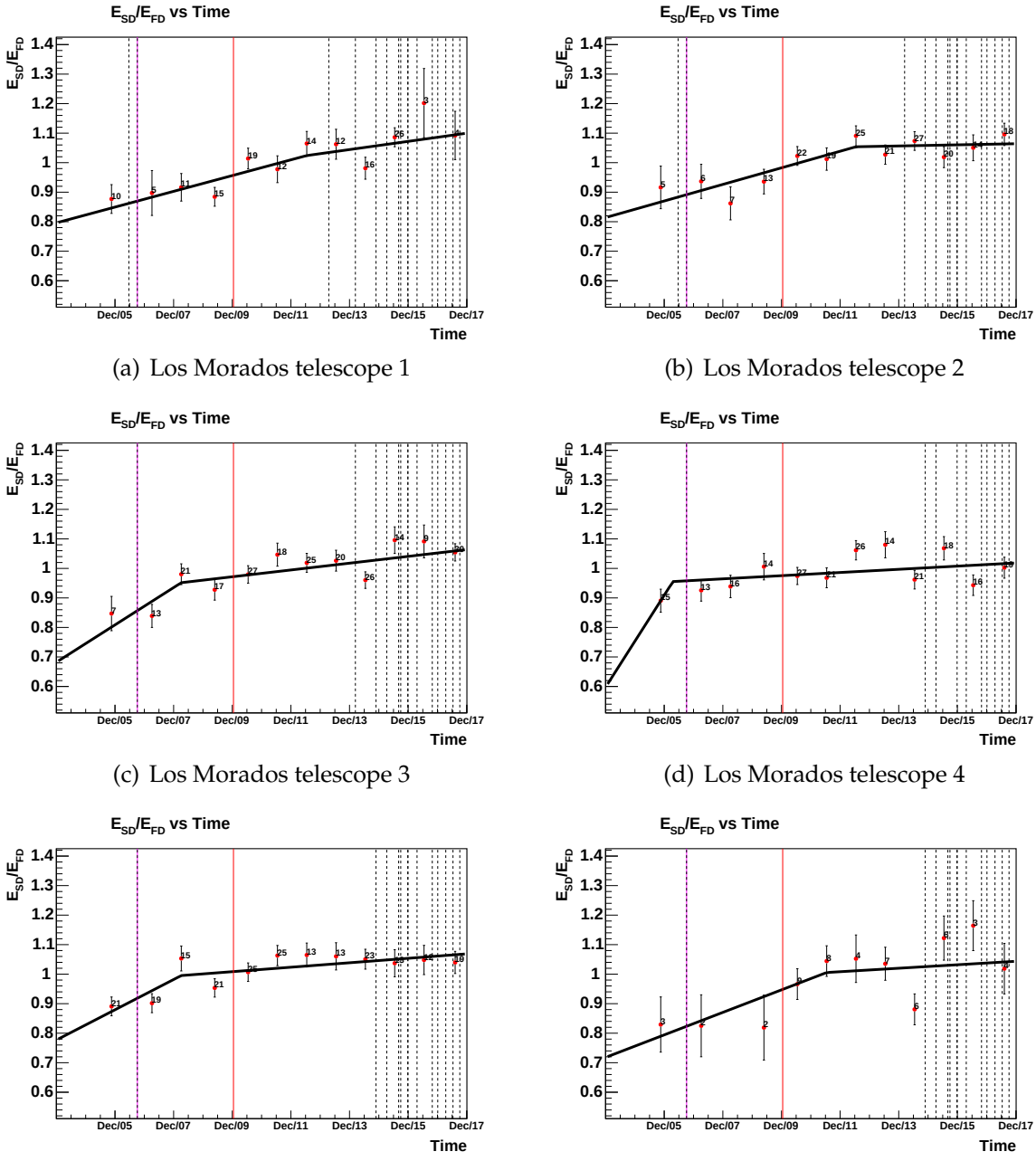
(c) Los Leones telescope 3

(d) Los Leones telescope 4

(e) Los Leones telescope 5

(f) Los Leones telescope 6

Figure 4.27: Linear-linear fits on the energy ratio as a function of time for each telescope at Los Leones. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.The result of the fit in telescope 6 is trivial, since there are only two datapoints and more than 2 parameters making the χ^2_{red} is infinite.



(a) Los Morados telescope 1

(b) Los Morados telescope 2

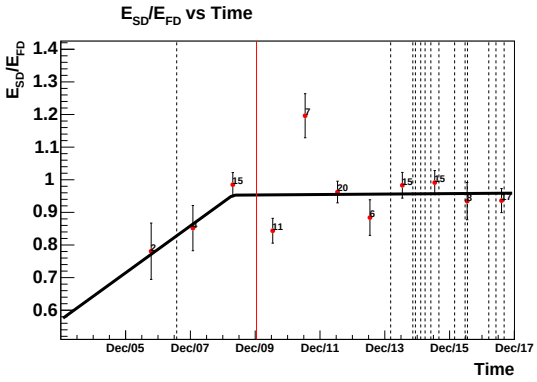
(c) Los Morados telescope 3

(d) Los Morados telescope 4

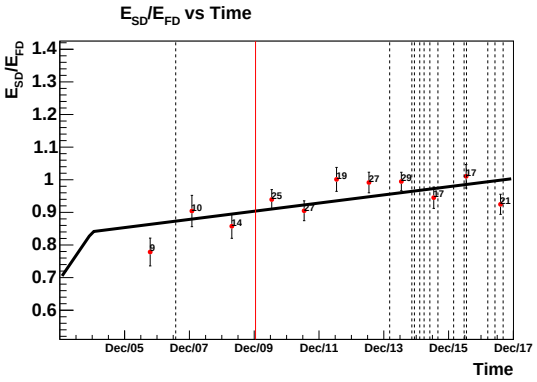
(e) Los Morados telescope 5

(f) Los Morados telescope 6

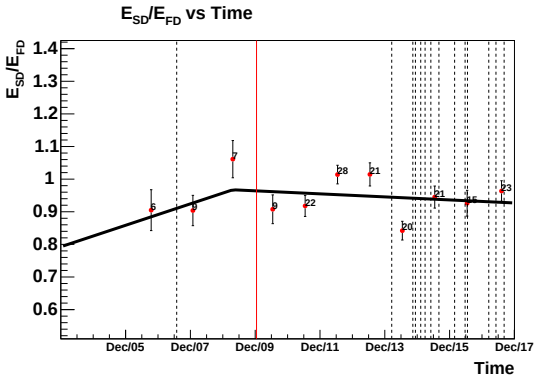
Figure 4.28: Linear-linear fits on the energy ratio as a function of time for each telescope at Los Morados. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



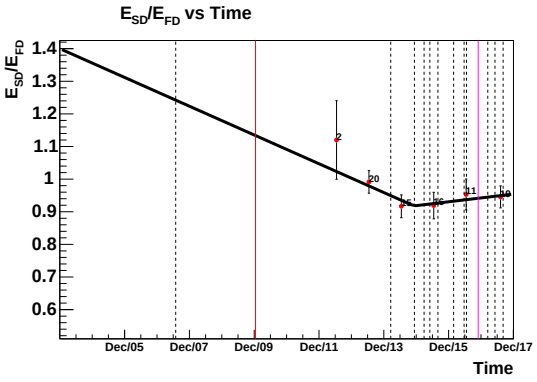
(a) Loma Amarilla telescope 1



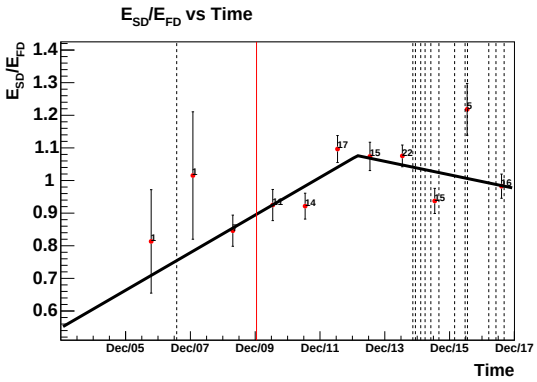
(b) Loma Amarilla telescope 2



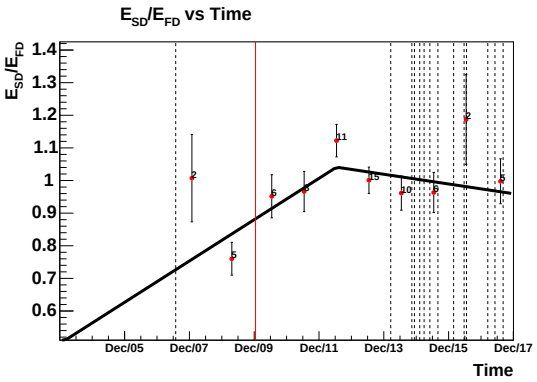
(c) Loma Amarilla telescope 3



(d) Loma Amarilla telescope 4

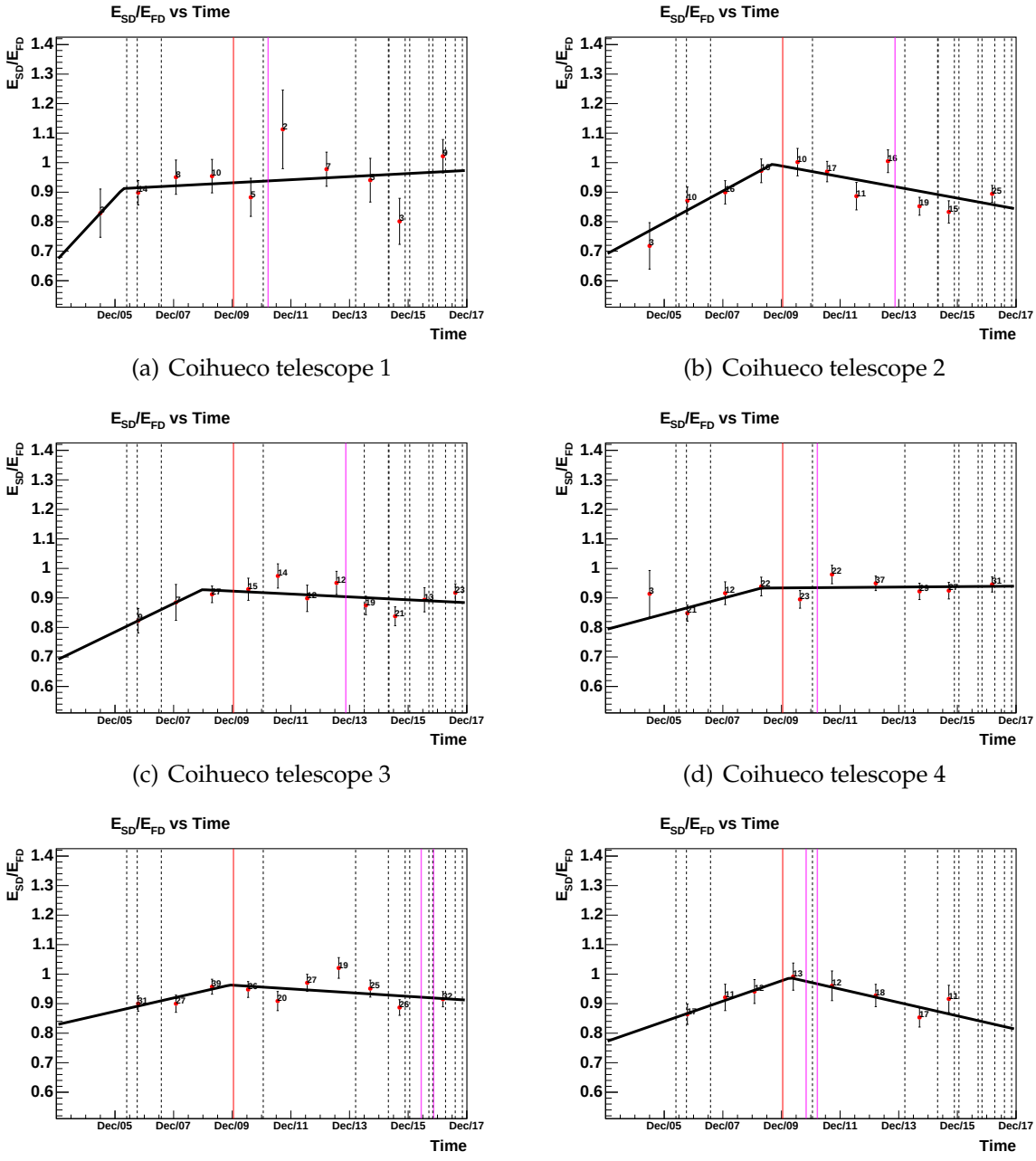


(e) Loma Amarilla telescope 5



(f) Loma Amarilla telescope 6

Figure 4.29: Linear-linear fits on the energy ratio as a function of time for each telescope at Loma Amarilla. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.



(a) Coihueco telescope 1

(b) Coihueco telescope 2

(c) Coihueco telescope 3

(d) Coihueco telescope 4

(e) Coihueco telescope 5

(f) Coihueco telescope 6

Figure 4.30: Linear-linear fits on the energy ratio as a function of time for each telescope at Coihueco. The red vertical line indicates when the drum calibration was done. The black dashed lines indicate dates of filter cleanings while the magenta lines represent the dates of mirror cleanings. The number of events in each bin is labelled next to every datapoint.

Fit type	FD Site	Tel	χ^2	NDF	χ^2_{red}	Probability	# Filter	# Mirror
Horizontal	Los Leones	1	6.5	6	1.08	0.369	18	2
		2	15.8	10	1.58	0.106	20	4
		3	39.9	12	3.32	7.5×10^{-5}	23	4
		4	49.3	12	4.11	1.87×10^{-6}	17	4
		5	11.6	11	1.05	0.396	18	2
		6	2.11	1	2.11	0.147	18	3
	Los Morados	1	40	11	3.64	3.51×10^{-5}	16	1
		2	27.8	11	2.53	0.00347	15	1
		3	43.1	11	3.92	1.05×10^{-5}	14	1
		4	27.3	11	2.48	0.00419	10	1
		5	37.6	11	3.42	9.06×10^{-5}	13	1
		6	22.3	10	2.23	0.0138	13	1
	Loma Amarilla	1	31.6	10	3.16	0.000473	14	0
		2	34.5	10	3.45	0.000152	14	0
		3	29.9	10	2.99	0.000899	14	0
		4	4.94	5	0.987	0.424	12	1
		5	42	10	4.2	7.47×10^{-6}	13	0
		6	30.2	9	3.35	0.00041	14	0
	Coihueco	1	10.9	9	1.22	0.28	15	1
		2	30.4	10	3.04	0.000723	15	1
		3	14.1	10	1.41	0.17	15	1
		4	13.9	9	1.55	0.124	13	1
		5	16.6	9	1.84	0.0558	14	2
		6	9.84	7	1.41	0.198	14	2
Linear	Los Leones	1	6.46	5	1.29	0.264	18	2
		2	6.92	9	0.768	0.646	20	4
		3	18.2	11	1.66	0.0764	23	4
		4	44.1	11	4.01	7.11×10^{-6}	17	4
		5	9.76	10	0.976	0.462	18	2
		6	1.78×10^{-22}	0	inf	0	18	3
	Los Morados	1	12.1	10	1.21	0.276	16	1
		2	11.8	10	1.18	0.297	15	1
		3	18.3	10	1.83	0.0494	14	1
		4	20.6	10	2.06	0.024	10	1
		5	15.6	10	1.56	0.112	13	1
		6	15.3	9	1.7	0.082	13	1
	Loma Amarilla	1	30	9	3.34	0.000433	14	0
		2	20.1	9	2.23	0.0175	14	0
		3	29.8	9	3.32	0.000466	14	0
		4	4.13	4	1.03	0.388	12	1
		5	34.4	9	3.82	7.6×10^{-5}	13	0
		6	22.9	8	2.86	0.00354	14	0
	Coihueco	1	9.15	8	1.14	0.33	15	1
		2	29.8	9	3.31	0.000471	15	1
		3	14.1	9	1.56	0.12	15	1
		4	9.39	8	1.17	0.31	13	1
		5	16.5	8	2.06	0.0356	14	2
		6	9.76	6	1.63	0.135	14	2

Table 4.2: Results from fitting a horizontal and linear line for the energy ratio of all telescopes at four different FD sites. The colour green signifies fit probabilities which have deviations between 0-2σ, orange is for deviations between 2σ-3σ and red means a deviation that lies beyond 3σ. The two fit types are applied to every FD site and telescope. The last two columns are the number of filter and mirror cleanings done at each telescope.

Fit type	FD Site	Tel	χ^2	NDF	χ^2_{red}	Probability	# Filter	# Mirror
Linear-hori	Los Leones	1	6.5	4	1.63	0.165	18	2
		2	10.4	8	1.29	0.241	20	4
		3	21.9	10	2.19	0.0159	23	4
		4	47.4	10	4.74	8.04×10^{-7}	17	4
		5	11.2	9	1.24	0.265	18	2
		6	2.9×10^{-13}	0	inf	0	18	3
	Los Morados	1	34.2	9	3.8	8.14×10^{-5}	16	1
		2	8.03	9	0.893	0.531	15	1
		3	21.7	9	2.41	0.00977	14	1
		4	21.3	9	2.37	0.0112	10	1
		5	13.2	9	1.46	0.155	13	1
		6	13.7	8	1.71	0.091	13	1
	Loma Amarilla	1	25.6	8	3.2	0.00123	14	0
		2	18.2	8	2.27	0.02	14	0
		3	29.5	8	3.69	0.00026	14	0
		4	4.94	3	1.65	0.177	12	1
		5	40.7	8	5.08	2.42×10^{-6}	13	0
		6	30.2	7	4.31	8.83×10^{-5}	14	0
	Coihueco	1	8.1	7	1.16	0.323	15	1
		2	23.5	8	2.94	0.00275	15	1
		3	10.5	8	1.31	0.235	15	1
		4	6.35	7	0.907	0.5	13	1
		5	14.9	7	2.12	0.0379	14	2
		6	7.6	5	1.52	0.179	14	2
Linear-linear	Los Leones	1	6.46	3	2.15	0.0912	18	2
		2	6.37	7	0.91	0.497	20	4
		3	17.2	9	1.91	0.0461	23	4
		4	43.6	9	4.84	1.67×10^{-6}	17	4
		5	9.76	8	1.22	0.282	18	2
		6	1.05×10^{-8}	0	inf	0	18	3
	Los Morados	1	11.6	8	1.44	0.172	16	1
		2	7.97	8	0.997	0.436	15	1
		3	14.2	8	1.77	0.0768	14	1
		4	18.8	8	2.35	0.0161	10	1
		5	9.98	8	1.25	0.267	13	1
		6	13.4	7	1.92	0.0626	13	1
	Loma Amarilla	1	25.6	7	3.66	0.000596	14	0
		2	20.1	7	2.87	0.00541	14	0
		3	27.9	7	3.99	0.000227	14	0
		4	1.2	2	0.6	0.55	12	1
		5	19.4	7	2.76	0.00715	13	0
		6	12.8	6	2.13	0.0467	14	0
	Coihueco	1	8.36	6	1.39	0.213	15	1
		2	13.6	7	1.94	0.0591	15	1
		3	8.89	7	1.27	0.261	15	1
		4	6.31	6	1.05	0.389	13	1
		5	11.5	6	1.92	0.0736	14	2
		6	2.47	4	0.618	0.65	14	2

Table 4.3: Linear-horizontal and linear-linear fit results. Last two columns show the number of filter and mirror cleanings done at each telescope. The colour green signifies fit probabilities which have deviations between 0-2σ, orange is for deviations between 2σ-3σ and red means a deviation that lies beyond 3σ.

Fit type	FD Site	Tel	Slope (Year^{-1})	Intercept
Horizontal	Los Leones	1	N/A	0.99 ± 0.05
		2		0.94 ± 0.01
		3		0.937 ± 0.009
		4		0.96 ± 0.01
		5		0.99 ± 0.01
		6		1.0 ± 0.1
	Los Morados	1	N/A	0.99 ± 0.01
		2		1.03 ± 0.01
		3		0.99 ± 0.01
		4		0.98 ± 0.01
		5		1.00 ± 0.01
		6		0.99 ± 0.02
	Loma Amarilla	1	N/A	0.95 ± 0.01
		2		0.94 ± 0.01
		3		0.94 ± 0.01
		4		0.95 ± 0.02
		5		1.00 ± 0.01
		6		0.97 ± 0.02
	Coihueco	1	N/A	0.93 ± 0.02
		2		0.91 ± 0.01
		3		0.90 ± 0.01
		4		0.92 ± 0.01
		5		0.933 ± 0.009
		6		0.91 ± 0.01
Linear	Los Leones	1	0.00 ± 0.01	1.0 ± 0.1
		2	0.009 ± 0.003	0.88 ± 0.02
		3	0.012 ± 0.003	0.85 ± 0.02
		4	0.007 ± 0.003	0.91 ± 0.03
		5	0.006 ± 0.004	0.94 ± 0.04
		6	0.03 ± 0.02	0.9 ± 0.2
	Los Morados	1	0.022 ± 0.004	0.82 ± 0.03
		2	0.015 ± 0.004	0.89 ± 0.004
		3	0.015 ± 0.003	0.86 ± 0.03
		4	0.008 ± 0.003	0.92 ± 0.03
		5	0.014 ± 0.002	0.90 ± 0.02
		6	0.019 ± 0.007	0.83 ± 0.07
	Loma Amarilla	1	0.006 ± 0.004	0.90 ± 0.04
		2	0.013 ± 0.003	0.83 ± 0.03
		3	-0.001 ± 0.004	0.95 ± 0.04
		4	-0.001 ± 0.01	1.1 ± 0.1
		5	0.015 ± 0.005	0.85 ± 0.05
		6	0.021 ± 0.008	0.78 ± 0.07
	Coihueco	1	0.007 ± 0.005	0.88 ± 0.04
		2	-0.003 ± 0.003	0.93 ± 0.03
		3	0.000 ± 0.003	0.90 ± 0.03
		4	0.006 ± 0.003	0.88 ± 0.02
		5	0.001 ± 0.002	0.93 ± 0.02
		6	-0.001 ± 0.005	0.92 ± 0.04

Table 4.4: Fit parameters obtained from the horizontal and linear fit applied to the energy scale ratio for each telescope. There are no slope values for the horizontal fit-type. Full credit is given to Violet Harvey for providing code which fixes the parameters to be the same number of decimal places as the error which was set to be displayed with one significant figure.

Fit type	FD Site	Tel	Slope (Year^{-1})	Intercept	Breakpoint (Years after'04)	Slope 2 (Year^{-1})
Linear-hori	Los Leones	1	0.1 ± 0.2	0.4 ± 0.7	4 ± 7	N/A
		2	0.13 ± 0.07	0.6 ± 0.1	2.5 ± 0.4	
		3	0.03 ± 0.01	0.80 ± 0.04	6 ± 1	
		4	0.03 ± 0.03	0.86 ± 0.09	3 ± 1	
		5	0.1 ± 0.1	0.7 ± 0.2	2.1 ± 0.4	
		6	0.2 ± 0.5	1 ± 1	4 ± 6	
	Los Morados	1	0.12 ± 0.07	0.7 ± 0.1	2.9 ± 0.7	N/A
		2	0.029 ± 0.007	0.81 ± 0.05	8.541 ± 0.002	
		3	0.07 ± 0.03	0.7 ± 0.1	5.1 ± 0.8	
		4	0.12 ± 0.06	0.7 ± 0.1	2.7 ± 0.5	
		5	0.06 ± 0.02	0.76 ± 0.07	4.7 ± 0.8	
		6	0.04 ± 0.01	0.7 ± 0.1	7.54 ± 0.02	
	Loma Amarilla	1	0.07 ± 0.03	0.6 ± 0.2	5.309 ± 0.008	N/A
		2	0.10 ± 0.05	0.5 ± 0.2	4.6 ± 0.7	
		3	0.14 ± 0.05	0.5 ± 0.1	3.1 ± 0.5	
		4	0.14 ± 0.06	0.3 ± 0.2	4 ± 2	
		5	0.15 ± 0.05	0.4 ± 0.2	4 ± 1	
		6	0.15 ± 0.07	0.4 ± 0.3	4 ± 2	
	Coihueco	1	0.05 ± 0.07	0.7 ± 0.2	4 ± 2	N/A
		2	0.12 ± 0.07	0.5 ± 0.2	3.2 ± 0.6	
		3	0.14 ± 0.05	0.4 ± 0.1	3.4 ± 0.4	
		4	0.03 ± 0.01	0.79 ± 0.06	5.3 ± 0.8	
		5	0.14 ± 0.04	0.5 ± 0.1	3.0 ± 0.2	
		6	0.14 ± 0.06	0.5 ± 0.2	3.2 ± 0.3	
Linear-linear	Los Leones	1	0.1 ± 0.2	0.5 ± 0.8	4 ± 7	0.00 ± 0.02
		2	0 ± 2	0.6 ± 0.7	2 ± 5	0.007 ± 0.004
		3	0.021 ± 0.009	0.81 ± 0.04	5.396 ± 0.008	0.009 ± 0.004
		4	-0.2 ± 0.1	1.1 ± 0.1	1 ± 10	0.007 ± 0.003
		5	0.2 ± 0.1	0.6 ± 0.2	2 ± 1	0.006 ± 0.004
		6	0.1 ± 0.4	0.7 ± 0.6	4 ± 6	0.0 ± 0.2
	Los Morados	1	0.027 ± 0.008	0.80 ± 0.05	8.541 ± 0.009	0.01 ± 0.01
		2	0.028 ± 0.008	0.81 ± 0.05	8.541 ± 0.005	0.002 ± 0.008
		3	0.06 ± 0.02	0.68 ± 0.09	4.261 ± 0.05	0.012 ± 0.004
		4	0.16 ± 0.06	0.6 ± 0.1	2.3 ± 0.3	0.005 ± 0.003
		5	0.05 ± 0.02	0.78 ± 0.06	4.261 ± 0.002	0.008 ± 0.004
		6	0.04 ± 0.02	0.7 ± 0.1	7.541 ± 0.002	0.01 ± 0.01
	Loma Amarilla	1	0.07 ± 0.03	0.6 ± 0.2	5.309 ± 0.003	0.001 ± 0.005
		2	0.1 ± 0.1	0.7 ± 0.1	1 ± 6	0.013 ± 0.003
		3	0.03 ± 0.02	0.8 ± 0.1	5.31 ± 0.01	-0.005 ± 0.005
		4	-0.044 ± 0.002	1.4 ± 0.2	11 ± 2	0.01 ± 0.03
		5	0.06 ± 0.02	0.5 ± 0.1	9.2 ± 0.7	-0.02 ± 0.01
		6	0.063 ± 0.003	0.5 ± 0.2	8.541 ± 0.001	-0.01 ± 0.01
	Coihueco	1	0.1 ± 0.1	0.7 ± 0.2	2 ± 1	0.005 ± 0.005
		2	0.05 ± 0.02	0.69 ± 0.08	5.7 ± 0.7	-0.018 ± 0.006
		3	0.05 ± 0.06	0.7 ± 0.2	5 ± 2	-0.005 ± 0.004
		4	0.03 ± 0.01	0.79 ± 0.07	5.31 ± 0.02	0.001 ± 0.004
		5	0.02 ± 0.01	0.83 ± 0.06	6 ± 1	-0.006 ± 0.005
		6	0.03 ± 0.02	0.77 ± 0.09	6 ± 1	-0.02 ± 0.01

Table 4.5: Linear-horizontal and linear-linear fit parameters. It is noted that the uncertainty in the breakpoint is implicitly constrained to be an integer given that the sampling interval of the energy ratio is at least a year long.

4.4 Conclusions

The long term stability of the ESR known as E_{FD}/S_{38} and E_{SD}/E_{FD} from 2004 to 2017 have been presented using high quality hybrid cosmic ray data. There has been an improvement on the aerosol database since the study done by Nguyen [110] and this was presented. The reduced χ^2 value of the ESR in this work (1.92) is slightly closer to 1 compared to the previous study (2.04) done by Nguyen, but the drift is still present at -0.7% per year post-2014. A comparison between weather-corrected and non-weather corrected data in 2015 exhibited amplitudes of seasonal modulation that were close in proximity which is interesting. The amplitude is the smallest for data calibrated with the running cal A procedure (0.54 %). This indicates that the running calibration process managed to reduce the fluctuations in pixel gain, thus providing a more accurate calibration for energy reconstruction.

The effects of filter and mirror cleanings as well as the drum calibration were observed when analysing the E_{SD}/E_{FD} ratio in the long term. The most notable effect is the filter cleaning at the 2014 discontinuity, which seems to cause a drop in the energy ratio for most telescopes but other than this drop, there is no clear drop after many of the filter cleanings probably due to how often it is done. The effect of the drum calibration is also unclear, since the energy ratio for all telescopes seem to either increase or decrease unpredictably.

Then, four different types of fits were applied to the energy ratio for all telescopes. Overall, it can be said that the energy ratio is unlikely to have a horizontal trend. Los Leones telescope 4 seems to indicate that a more complex structure is preferable. This may be attributed to some irregularity in filter cleanings in years prior to 2015. The other telescopes at Los Leones tend to have one increasing trend. All of the telescopes at Los Morados have the best values in terms of the reduced χ^2 for linear and linear-linear behaviour. Telescopes 1 and 3 at Loma Amarilla also have an energy ratio which most likely contains a more complex structure. However, seeing how these telescopes have the same cleaning dates as all the other telescopes, it is unknown as to why the behaviour differs. As for Coihueco, all telescopes have an energy ratio more suitable for a linear-linear type fit.

In terms of the linear fit parameter values, Los Leones seems to have an increasing energy ratio up to 3 ± 2 % per year. As for Los Morados, the slope varies from 0.8 ± 0.3 % to 2.2 ± 0.4 % among all telescopes. From the linear-linear fit analysis, the energy ratios obtained from Loma Amarilla and Coihueco resulted in first slope values at 3 % to 12 %. After the breakpoint, the slopes are negative for most telescopes with a maximum slope of 0.3% and the minimum at -1.9%. The breakpoint at Coihueco is favoured to be around 2010. Further work on the energy ratio can include applying more sophisticated fits.

Chapter 5

Examining the Stability of Event Rates Recorded by the Surface Detector at the Pierre Auger Observatory

This chapter details the analysis on stability of the event rates recorded by the surface detector (SD). The aim is to monitor the energy assignments by observing rates as a function of position across the array and time. This analysis can show the differences in the behaviour of observed stations which may be correlated with positions on the array, uptime and commission time.

To calculate rates of events, data on the number of events detected by each station was obtained. The SD data consists of the finalised reconstruction (weather-corrected) used for the International Cosmic Ray Conference 2019 (*icrc19*). The analysis only looks at quality showers detected by the surface detector which have energies above $3{\times}10^{18}$ eV, with a maximum zenith angle of 60° since this is the threshold energy at which the SD is fully efficient. Full efficiency indicates that the geometrical acceptance of air showers does not depend on the nature of the primary particle, its energy or arrival direction [103]. It was required that each event detected by an SD station fulfils the 6T5 criteria. This ensures that each event detected has its highest signal station surrounded by 6 other stations which are active, thus avoiding the collection of events where the air showers lie outside the array or where there is any faulty station. A list of quality cuts is provided in Appendix A. Also, only the events of stations in the 1500 m array are included. So, the stations in the 750 m array, as well as doublet and test stations are ignored in the following analysis. The data contains events detected between 2004 and 2017.

5.1 Introduction

Figure 5.1 shows the number of events detected by the SD each year from 2004 to 2017. The number of events are shown to increase from 2004 to 2008. The year

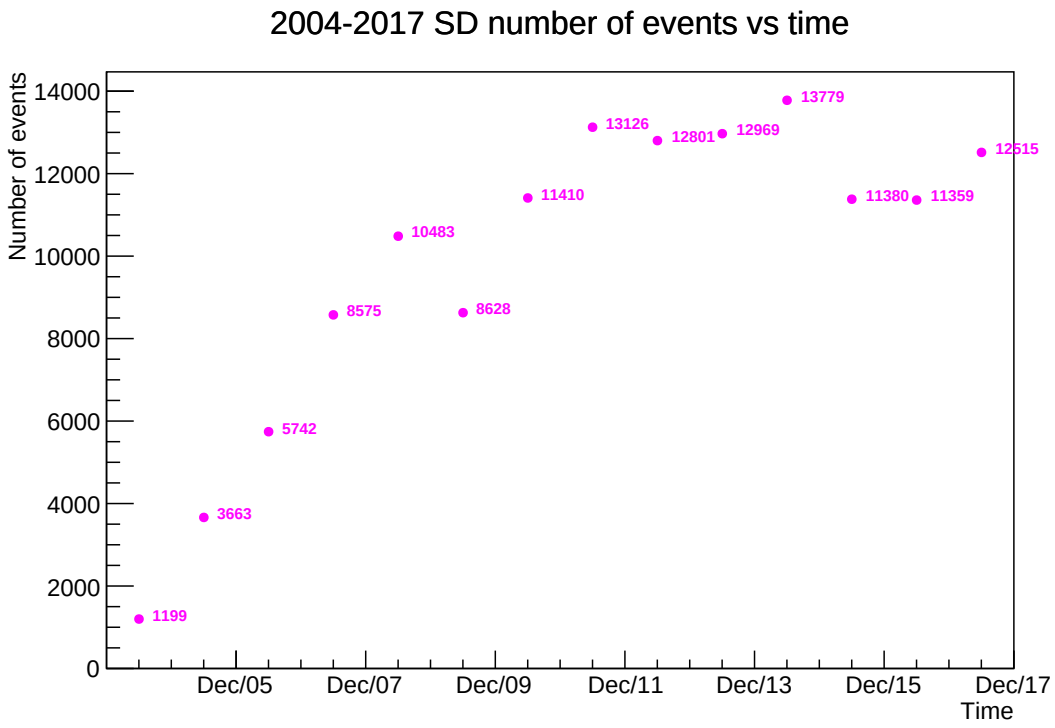


Figure 5.1: Number of events detected each year. All events have energies above 3 EeV and zenith angle below $60°$. The data used is the *icrc19* weather-corrected data. All datapoints are labelled with the number of events for each year. The uncertainty for each data point is calculated to be the square root of the number of events as per Poisson statistics. However, their values are too small to be displayed as error bars in the above graph due to scale.

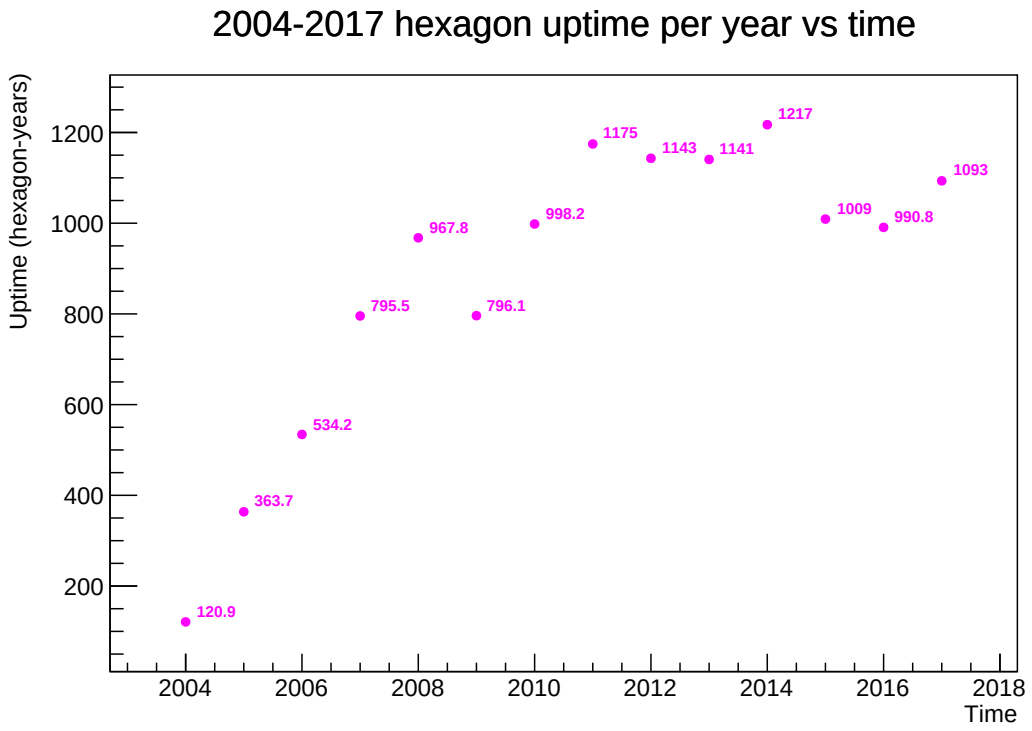


Figure 5.2: The number of hexagon-years for every year from 2004-2017. The hexagon-years are labelled for clarity.

2008 marks the completion of the SD array. There exists a steep decrease in the number of events in 2009 and this is most likely due to the communications crisis which occurred between the months of June to November. This communications crisis had caused some of the data within this period to be unreliable [133]. The number of annual events was fairly stable from 2010-2015, after which there was another steep decrease. This was attributable to technical issues in some stations in that year, and the rate recovered in subsequent years as these issues were addressed.

The uptime of stations are obtained from a dataset called Hexalife files [134]. These files contain information on the amount of time every station in a full hexagon of six active stations is operational and readily taking data. This is referred to as uptime, while the station surrounded by a full active hexagon is referred to as a hexagon. It was ensured that the uptimes during the bad periods and the communication crisis were not included in the analysis, since the data in those time frames are unreliable. The uptimes for each year were calculated by summing up the number of seconds each hexagon was active in that year. Then, this total hexagon uptime in seconds was divided by the total number of seconds in 1 year which is 31,557,600 seconds/year[1]. Doing this results in a new value, *hexagon-years*. Since the hexagon-years were obtained for each year, this gives an

[1]This is a result of the product of 86400 seconds/day $\times$ 365.25 days/year.

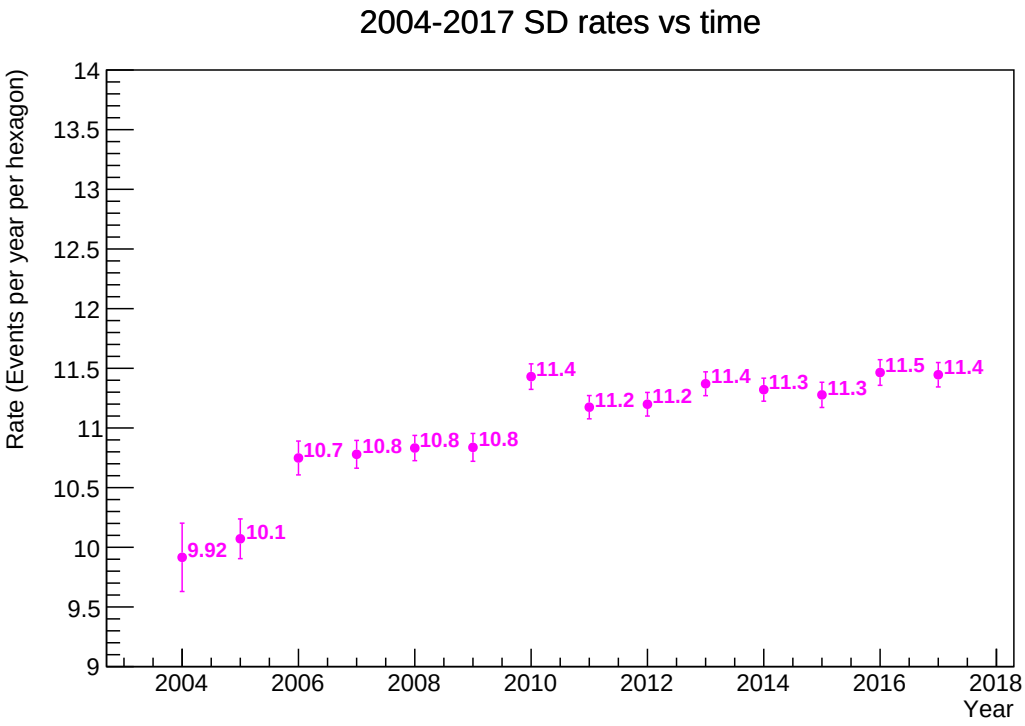


Figure 5.3: Yearly rates as a function of time from 2004 to 2017. The uncertainties as depicted by the error bars were calculated to be the square root of the number of events detected in that year by all stations divided by the total uptime of those stations.

average value for the number of active hexagons for each year. Figure 5.2 shows how this varies from 2004 to 2017 and the shape is very similar to the variation in the number of events observed by the stations as shown in Figure 5.1. Since the number of events and the hexagon years for each year are obtained, the rate, can be obtained as the quotient of the two, resulting in the evolution from 2004 to 2017. This is in units of the number of events/hexagon/year. This is shown in Figure 5.3. This trend is investigated in more detail later in Section 5.4, where several fits were made to analyse the possible trends.

5.2 Maps of Rates as a Function of Position from 2004 to 2013

To determine if the event rates are a function of position, rates for each station are illustrated in a 2D histogram, with the x and y-axis representing positions and the z-axis as rate. All stations on the maps are represented by bins of area 0.244 km^2 with a length of 0.376 km in the x-direction and 0.65 km in the y-direction. To ease the calculation of positions, the coordinates for each station is set to the Site Coordinate System, which means the centre of the array is set to position (0,0) with north and east being positive in y and x respectively. It was also noted

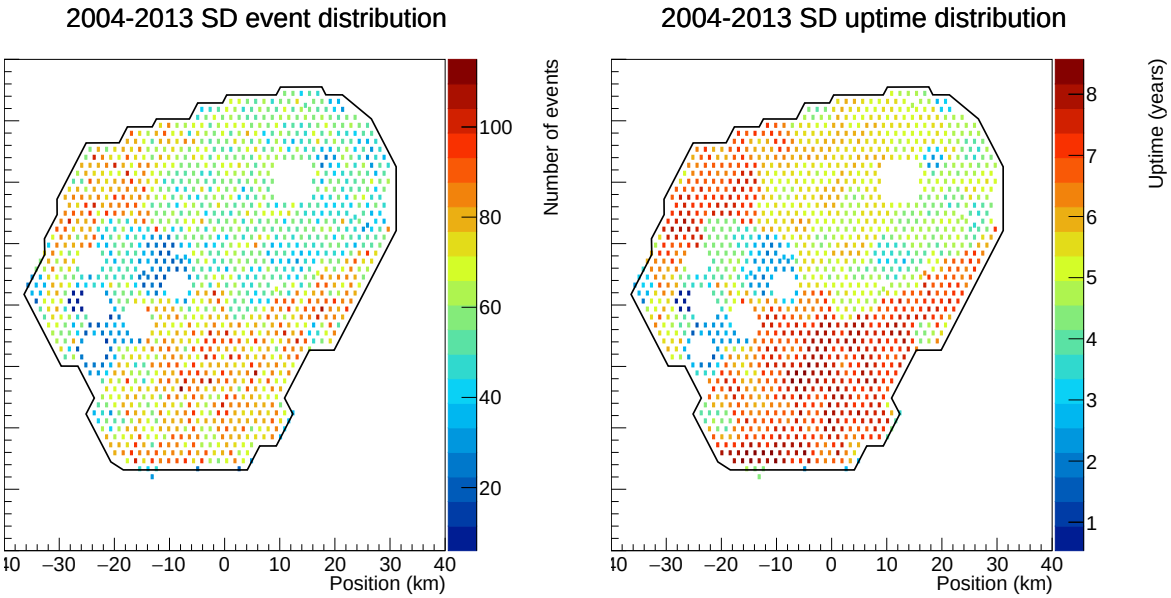
that throughout the history of data-collection conducted by the Pierre Auger observatory, some stations have been removed, and some have been replaced with stations which were assigned new ID numbers. In the following analysis, these removed and replaced stations have been taken into account to ensure that the calculation of rates is done correctly.

Firstly, the rate as a function of position across the array from 2004 to 2013 was calculated. An analysis for the years 2014-2017 is shown later in Section 5.3. This separation of data into two time-frames is due to the introduction of the new Time-over-threshold deconvolved (ToTd) and Multiplicity of Positive Steps (MoPS) station-level triggers at the end of 2013, described in Section 3.1.2.1. This may affect the sensitivity or efficiency of the stations detecting the events. The number of events assigned to a station is obtained by counting the number of times that station received the highest signal. Here, the station is referred to as the *hottest* station. These events are once again subject to the cuts in Appendix A. Then the uptimes of each station are obtained. Figure 5.4 shows two maps, where 5.4(a) depicts the number of events associated with each station between 2004 and 2013 while 5.4(b) depicts the uptime of the stations in years. The maximum and minimum values on the z-axis shown are scaled according to the highest and lowest values obtained from all stations. As expected, the stations which have higher uptimes correspond to the stations which have detected more events. There is a general trend for longer uptimes to the south of the array as a result of the south-to-north sequence of initial deployment.

A cross-check was done for the uptimes shown in Figure 5.4(b), by obtaining the age of all stations, defined as the amount of time since they were deployed. Both the uptimes and station age show a trend as a function of position, since over time, many stations were deployed up until 2008 which marks the completion of the SD array. The station age corresponds to the time between the date of deployment and midnight of the 1st of January 2014. A comparison between the station ages and the integrated uptimes is shown in Figure 5.5. It can be seen that there is a similarity in the structure showing higher uptimes and station age from the South to the East of the array. It is also noted that the maximum station age reaches 12 years because some stations were deployed prior to 2004. The uptimes only reach up to 8.3 years due to accumulated *bad periods* or the *communications crisis* as explained in Appendix A.

The rates for each station were calculated as the ratio of the number of events and the uptimes. Figure 5.6(a) shows a map of rates for every station from 2004-2013. The maximum and minimum rates are scaled according to the highest and lowest rates obtained by all stations. Figure 5.6(b) shows a map of measured uncertainties for each station calculated as:

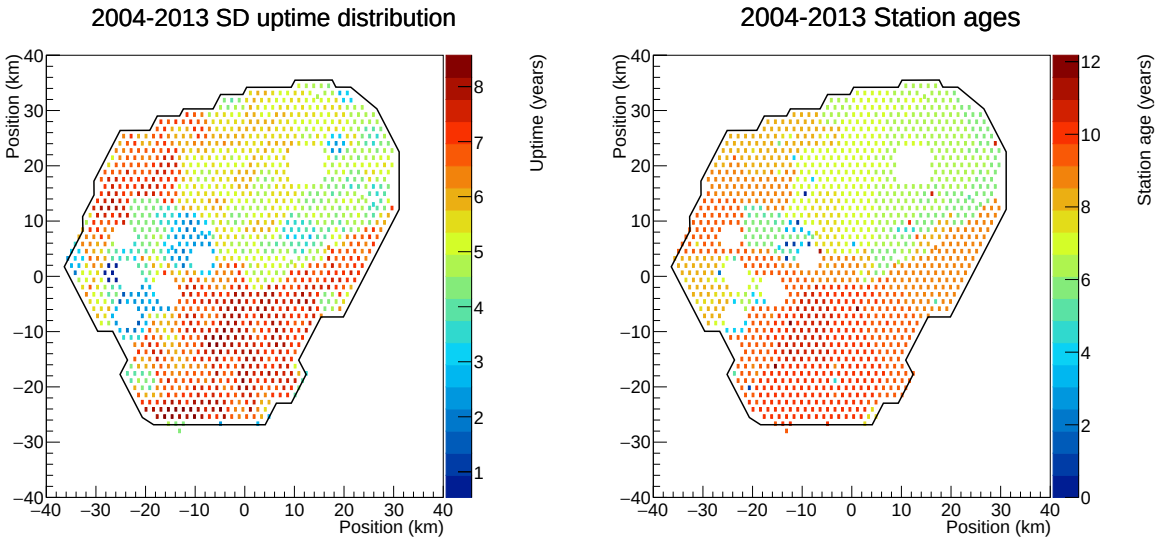
$$\sigma_{\text{measured}} = \frac{\sqrt{\text{Number of events detected by the station}}}{\text{Uptime of the station}} \tag{5.1}$$



(a) Number of events detected by every hottest station across the array.

(b) Integrated uptimes across the array.

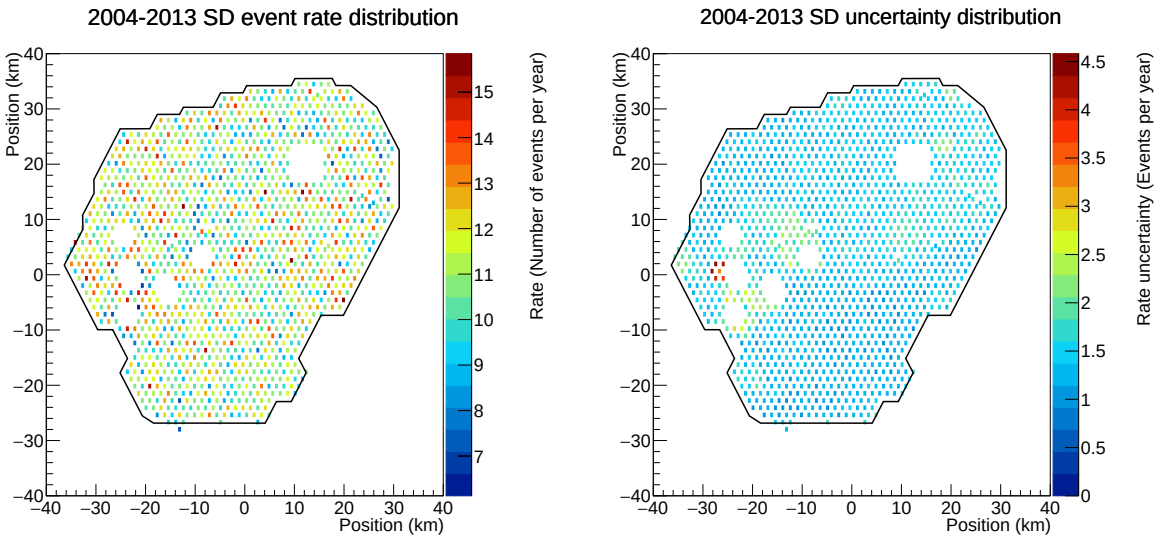
Figure 5.4: Number of events with threshold energy of 3 EeV (weather-corrected), detected by each hottest station accumulated from 2004 to 2013 (left) as well as their respective uptimes (right). The black outline represents the latest update of the array outline in 2020, where stations which lie on the outline represent stations on the border of the array as calculated by V. Harvey. It is noted that at a particular stage in time, some stations located on the current border previously had more stations surrounding them which were located outside the borders. As time progressed, some stations have been removed and replaced.



(a) The uptime of each station across the array, from the start of 2004 to the end of 2013.

(b) The age of all stations defined as the amount of time since each station had been deployed up to the end of 2013.

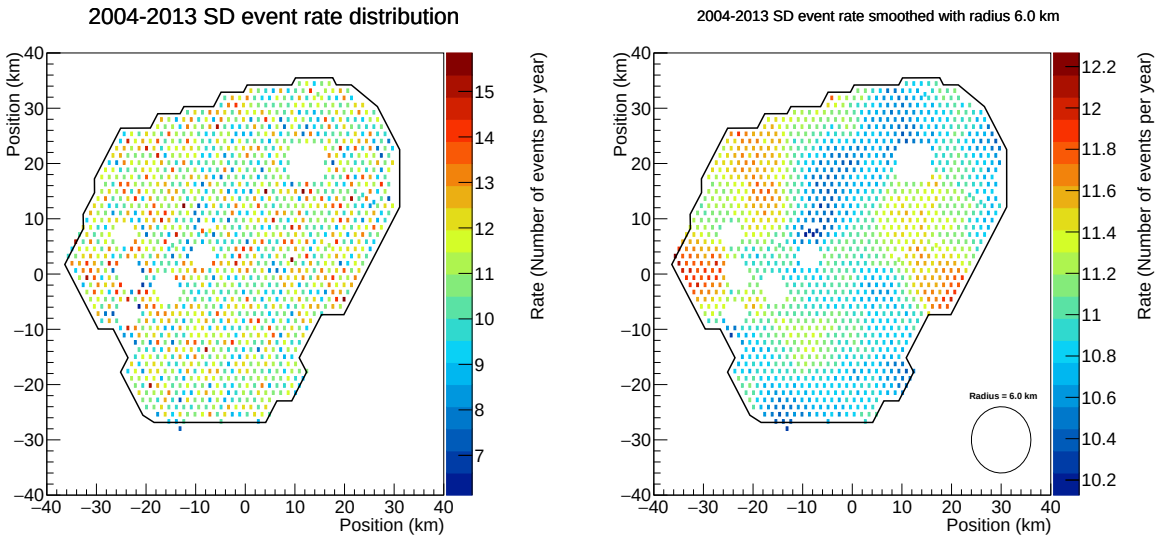
Figure 5.5: The uptimes of every station from 2004-2013 which signify the amount of time the station had been actively taking data (left) and the station age which signifies the amount of time since deployment (right).



(a) Rates across the array.

(b) Measured uncertainties across the array.

Figure 5.6: Map showing rates of events with a threshold energy of 3 EeV (weather-corrected) across the array compared with a map of uncertainties plotted as the square root of the number of events associated with each hottest station, divided by their respective uptime.



(a) Rates of each station across the array from 2004 to 2013.

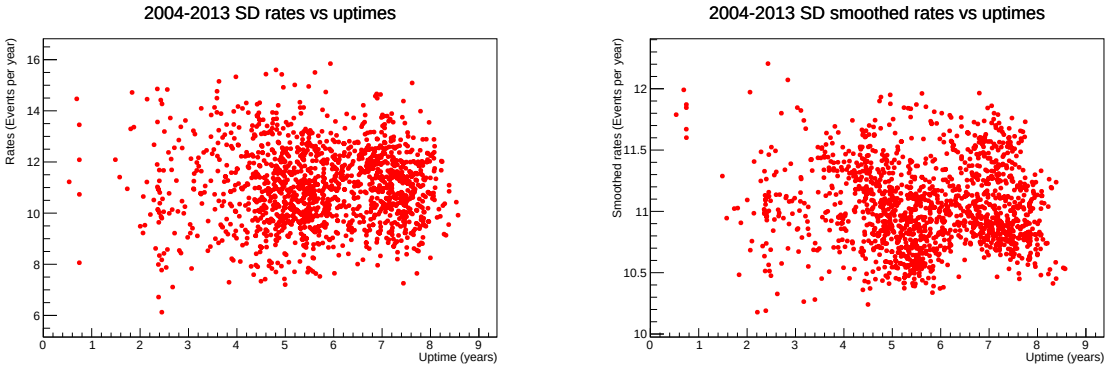
(b) Smoothed rates across the array from 2004 to 2013.

Figure 5.7: Comparison between an unsmoothed and smoothed rate map for data between 2004 to 2013.

There are a few stations with low uptimes, as shown by the high rate uncertainties in the small western area of the array. From Figure 5.6(a), it is difficult to observe whether there are any significantly major hot or cold spots in terms of rates, due to fluctuations between each station. To suppress these fluctuations, a *smoothing* algorithm was applied to the rate map. Firstly, a circle of a chosen radius is drawn with a station at its centre. Then, the number of events detected and the uptimes of all stations enclosed by the circle is summed up. The ratio of the two is then calculated to get an overall value of the rate shown as follows:

$$\text{Overall rate} = \frac{\sum_{i=1}^{N} \text{Number of events detected by station } i}{\sum_{i=1}^{N} \text{Uptime of station } i} \tag{5.2}$$

where N is the number of stations enclosed by the circle. Then, this overall rate is assigned to the station at the centre of the drawn circle. This process is repeated for all other stations to make up a map of averaged overall rates. Figure 5.7 shows a comparison between the map of rates of each station shown previously and a map of *smoothed* or averaged rates using the previously described method. A radius of 6km was chosen as it includes approximately 55 stations in a filled circle of stations which is statistically sufficient in the smoothing method. It can be seen that there are several hotspots on the eastern and western part of the array. The cold spots span across from the north to the south of the array. The scale of the smoothed rates span between 10.2 events per year to 12.2 events per year. The hot spots seem to be located close to the borders of the array, compared to the cold spots.



(a) Rates as a function of uptime for all stations from 2004 to 2013.

(b) Smoothed rates as a function of uptime for all stations from 2004 to 2013.

Figure 5.8: Graphs showing rates(both unsmoothed and smoothed) as a function of uptime from 2004 to 2013.

Two graphs of rates (both unsmoothed and smoothed) as a function of uptime, were produced, analogous to Figure 5.7 but with the removal of position dependence. This is to see if there exists an uptime dependence. Figure 5.8 shows two very faint populations of stations, more prominent particularly in Figure 5.8(b). There seems to be two populations of stations but the trend is difficult to discern. This became a motivation behind the study on possible differences in the behaviour of station rates based on the year of commission. This was investigated in Section 5.5.

5.2.1 Toy Monte Carlo for the Analysis of Rates

The Monte Carlo method was used to create many random rate maps. The aim of this analysis is to confirm whether false details appear due to just a Gaussian distribution of rates and to see how prominent the random fluctuations are after smoothing. Firstly, the distribution of rates from the real map of data from 2004-2013 was obtained. Then, a Gaussian function was fitted to the distribution and the fit parameters represented by the mean value, μ as well as the width, σ were obtained. This is shown in Figure 5.9, where μ and σ are shown as 11 events/year and 1.51 events/year. The process of generating random maps begins with randomly sampling a value of rate from the Gaussian distribution fit done in Figure 5.9 for every station on the real map. This process is repeated to produce many different maps to see how these random maps can vary. However, in order to smooth the maps, it is not possible to obtain the total number of events and uptimes, since the maps were generated from a Gaussian distribution of numbers representing rates. So, the maps are smoothed by calculating the average, or

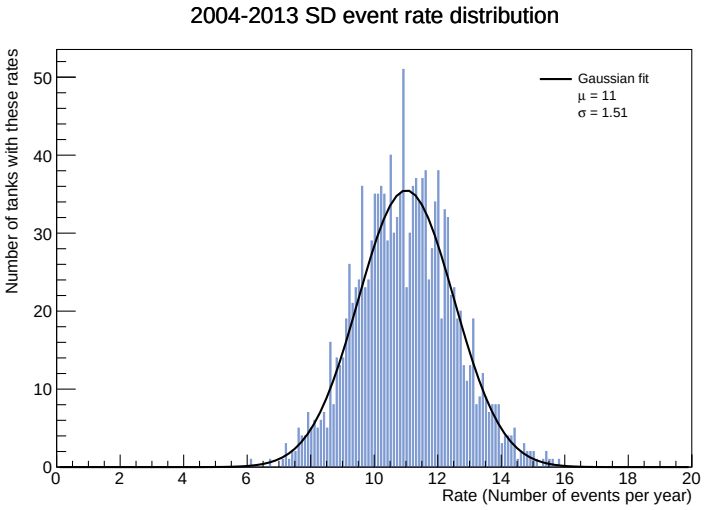


Figure 5.9: A one-dimensional distribution of rates (unsmoothed) for the years 2004-2013.

mean rates shown as the following:

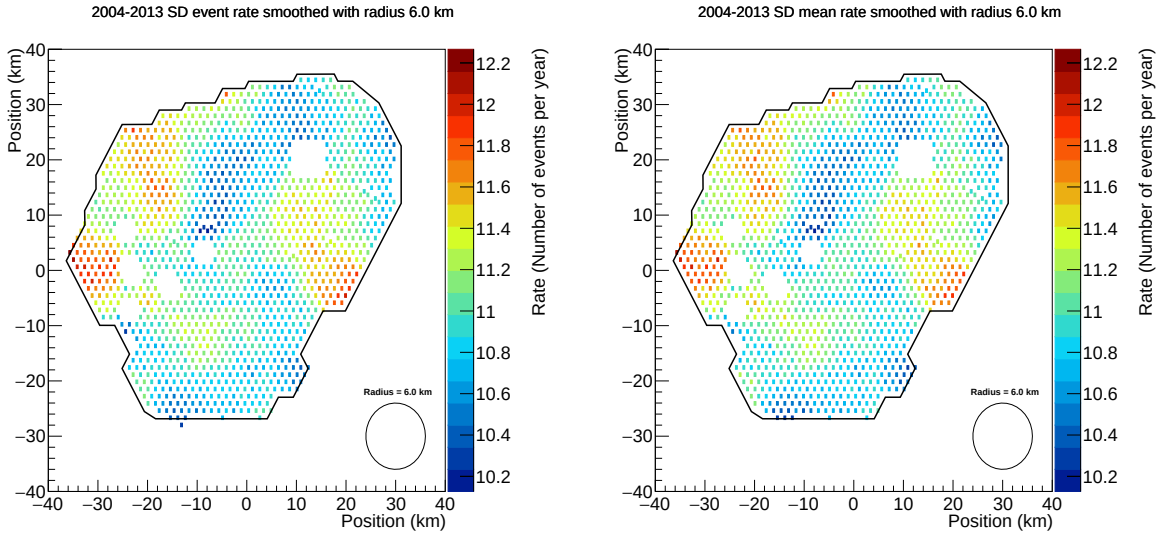
$$\text{Mean rate} = \frac{\sum_{i=1}^{N} \text{Monte Carlo rate of station}}{\text{Number of stations, } N} \tag{5.3}$$

where N is the number of stations in a circle of radius 6 km. To check if smoothing by calculating the mean, differs greatly from the method described by Equation 5.2, maps using the real data from 2004 to 2013 were produced using these two methods for comparison. For the real map data, the smoothed rate is calculated as

$$\text{Overall rate} = \frac{\sum_{i}^{N} \text{Number of events from station } i \text{ / Uptime of station } i}{\text{Number of stations, } N} \tag{5.4}$$

where N is the number of stations in a circle of radius 6km. Figure 5.10 shows two maps with rates smoothed using the overall and mean rate calculation. Figure 5.11 shows the absolute differences in smoothed rate calculations and they have a range from around -0.1 to 0.1 which is 5%. So the overall and mean rates are in good enough agreement with each other. Thus, the mean rate method is used to smooth random maps generated from Monte Carlo methods. The random number generator used is obtained from a library called the C++ BOOST library.

Figure 5.12 shows some examples of Monte Carlo generated smoothed maps while Figure 5.13 shows corresponding rate distributions as one-dimensional histograms with a fitted Gaussian function. From the comparison between Figure 5.13 and Figure 5.14 there is more spread in the real data compared to the Monte Carlo generated data as shown by the values of σ labelled on each plot. This implies that the false details which appear purely from a random Gaussian



(a) Overall rates calculated using Equation 5.2 from 2004 to 2013.

(b) Mean rates calculated using Equation 5.3 from 2010 to 2013.

Figure 5.10: A comparison between the calculation of overall and mean rates.

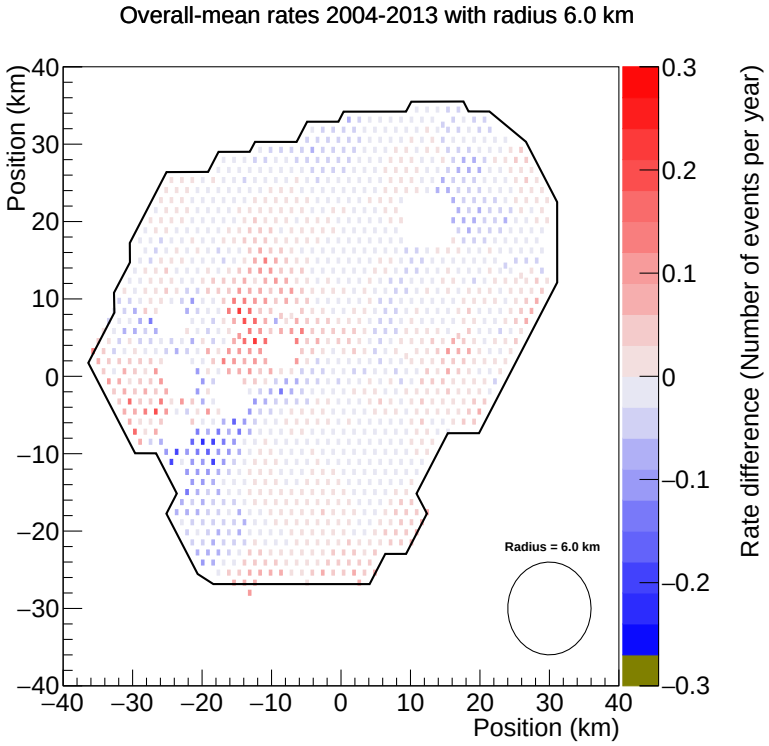
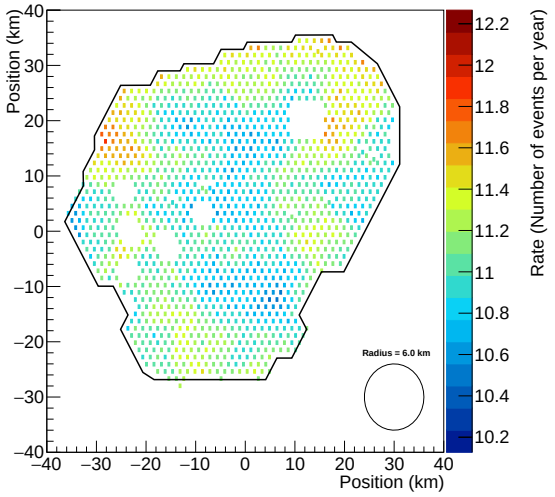
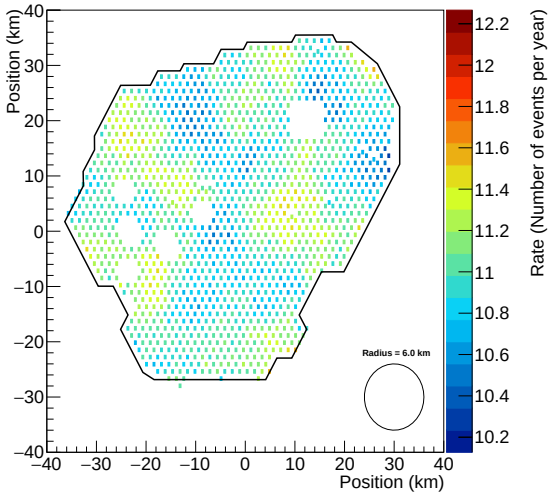


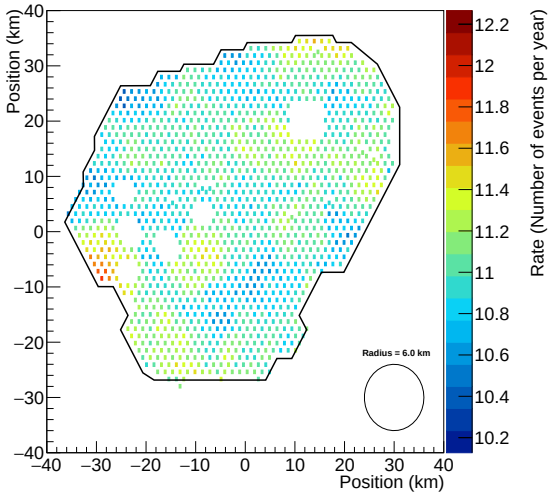
Figure 5.11: Map showing the difference between the overall and mean rates method (overall-mean rate), using red-blue colour scale to ease visualisation, for histograms with negative and positive values in their bin contents.



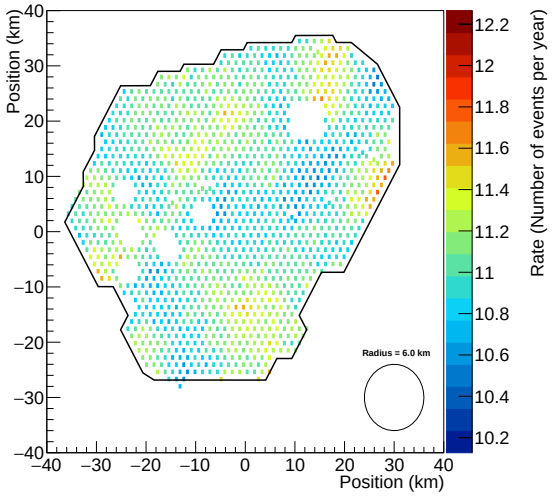
(a) Random smoothed rate map 1.



(b) Random smoothed rate map 2.

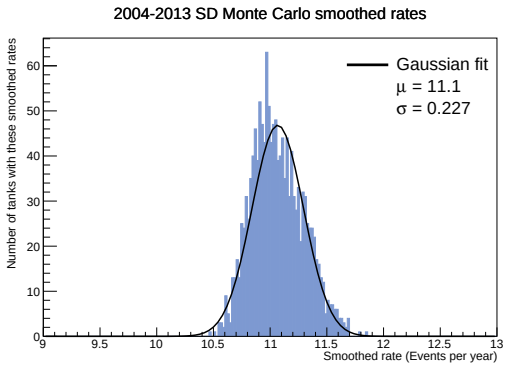


(c) Random smoothed rate map 3.

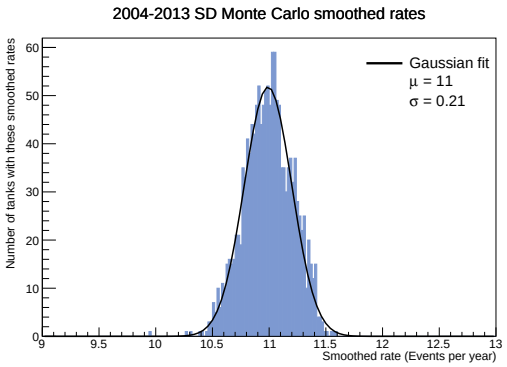


(d) Random smoothed rate map 4.

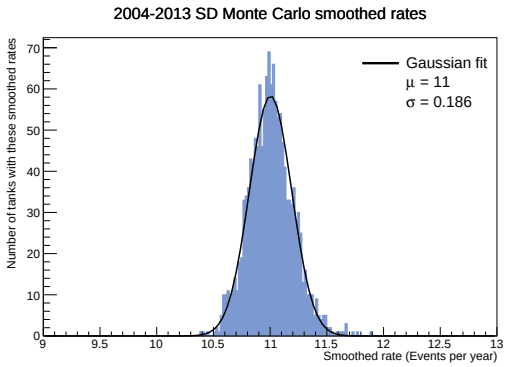
Figure 5.12: Four examples of smoothed random Monte Carlo rate maps for the years 2004-2013.



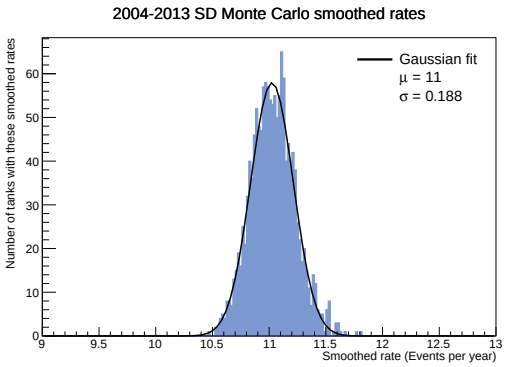
(a) Random smoothed rate distribution 1



(b) Random smoothed rate distribution 2



(c) Random smoothed rate distribution 3



(d) Random smoothed rate distribution 4

Figure 5.13: Four examples of smoothed random Monte Carlo rate distributions for the years 2004-2013.

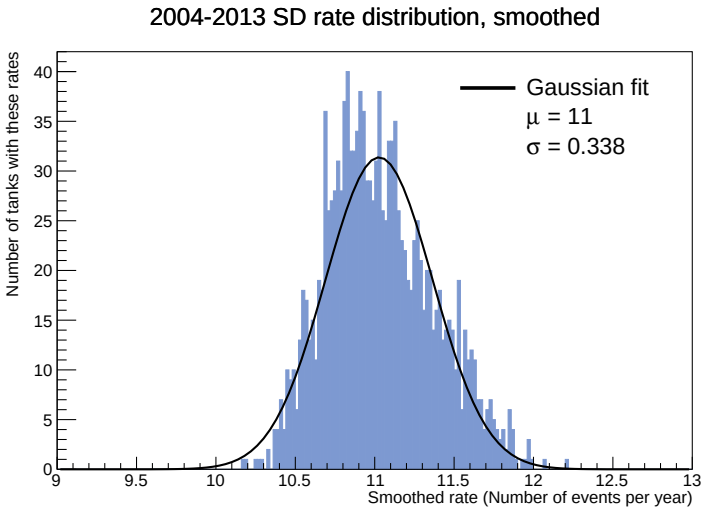
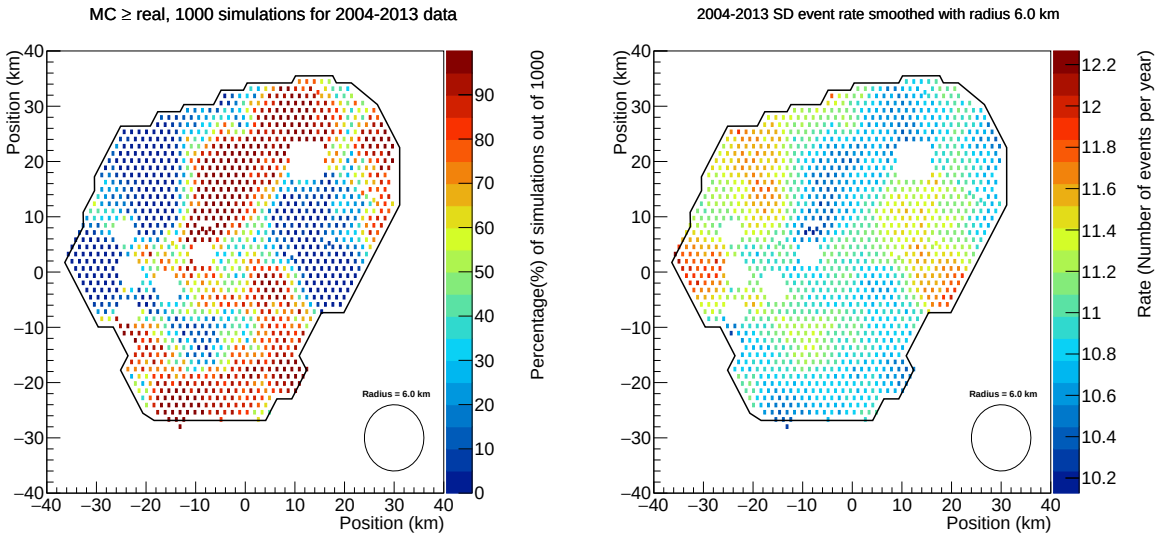


Figure 5.14: A one-dimensional distribution of the true smoothed rates for the years 2004-2013.



(a) Percentage of Monte Carlo simulations ≥ real smoothed rate for 1000 simulations.

(b) Smoothed rates across the array.

Figure 5.15: A comparison between the Monte Carlo check and the real map for 2004-2013 data.

distribution of rates are much less compared to the variation of rates on the real map.

5.2.1.1 Quantifying Significance from Monte Carlo Maps

From the random maps produced by the Monte Carlo method, the significance of rates can be quantified by comparing the Monte Carlo smoothed rates with the real smoothed rates. Firstly, 1000 Monte Carlo smoothed rate maps are simulated. For each station on every map, a check was done to see if the Monte Carlo smoothed rate is greater than the real smoothed rate. The result is plotted as a percentage of Monte Carlo simulations that are greater than the real simulated station rate. Figure 5.15 shows a comparison between the map showing percentages that each station rate is greater than the real station rate and the map showing the real smoothed rates. It can be seen that the stations with lower probabilities correspond to the stations which have higher rates in the real map as expected.

A distribution of these percentages as probabilities were plotted in Figure 5.16 on a one-dimensional histogram to observe the frequency of probabilities. There seems to be an increasing trend towards 100 % and 0 % which implies the presence of real systematics in the smoothed rates across the array. For the interpretation of these probabilities in terms of a significance, they were converted to the number of standard deviations away from the mean represented by σ values in the Gaussian distribution function. Figure 5.17 shows a map of σ values, obtained from the conversion of probabilities using the

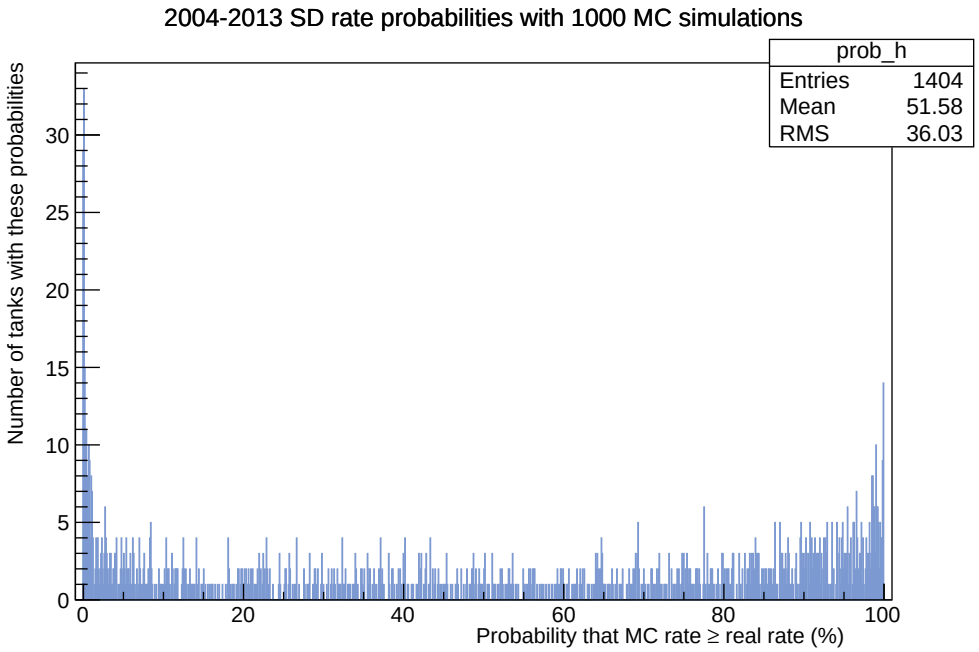


Figure 5.16: A one-dimensional distribution of probabilities for checking if the simulated Monte Carlo rate is greater than the real rate.

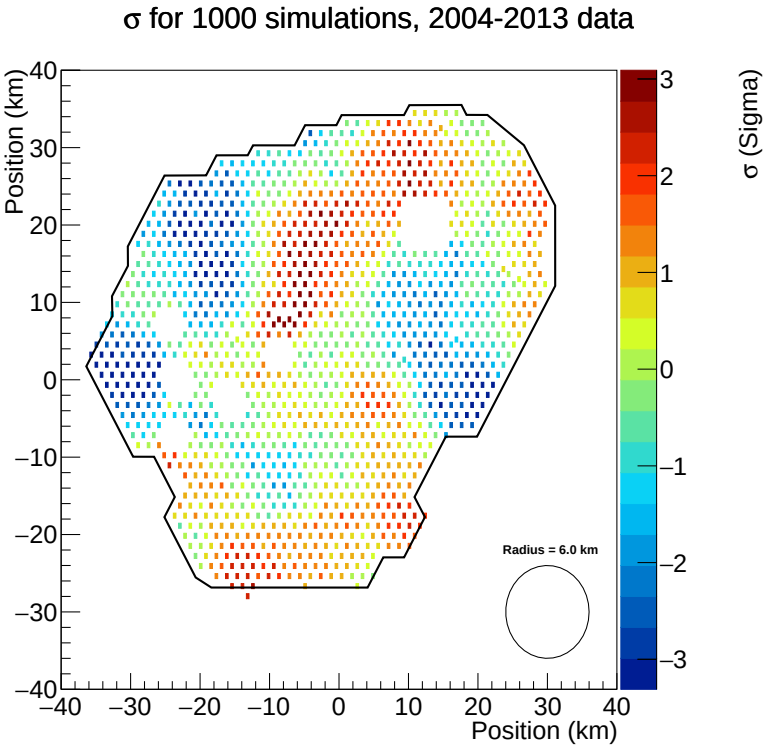


Figure 5.17: A map of the σ values, where 0 % is set to be -3.31σ while 100 % is set to be +3.31σ. A probability of 50 % will correspond to a σ value of 0.

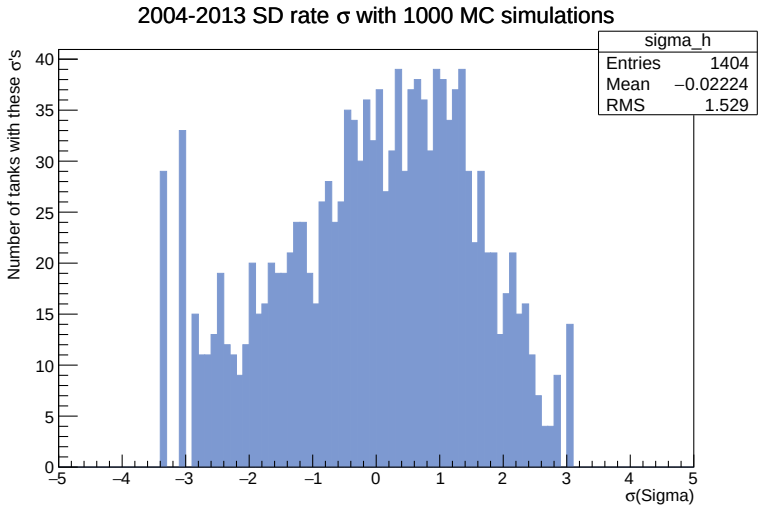
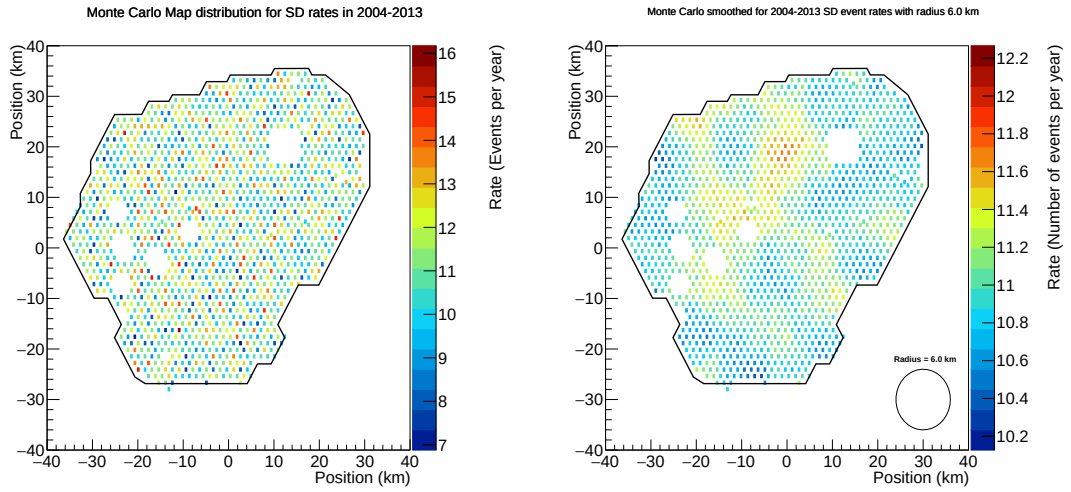


Figure 5.18: A one-dimensional distribution of σ values, where 0 % is set to be -3.31σ while 100 % is set to be +3.31σ. A probability of 50 % will correspond to a σ value of 0.

ROOT::normal_cdf function. Figure 5.18 shows that there seems to be a skew to the right in the distribution of σ, as well as a high frequency near $\sim$-3σ and +3σ, corresponding to the similar behaviour is seen in the distribution of probabilities shown in Figure 5.16.

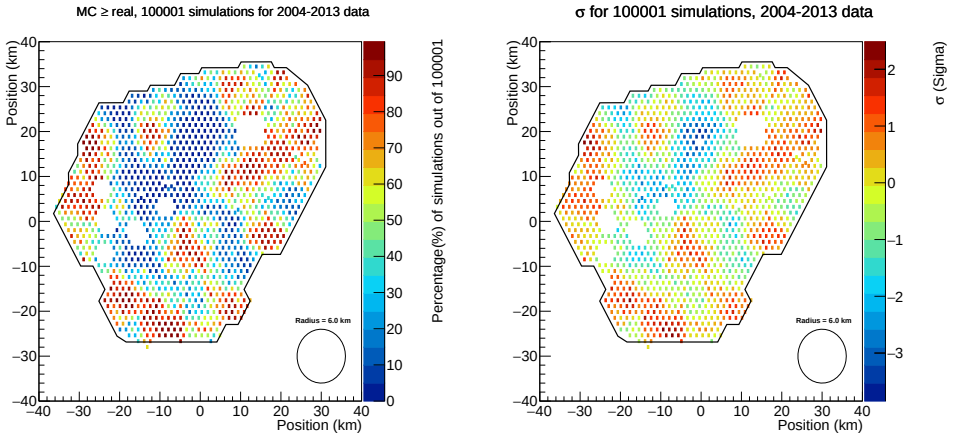
To ensure that the presence of systematics is not due to the smoothing process, but rather due to the true data, a check was done by choosing a random Monte Carlo simulated smoothed rate map and treating it as though it was the real map. The same process is applied and the distribution of probabilities is observed to see if there are areas with systematically high or low rates. The process was repeated but with 100,000 simulations to extend the significance or σ values to $\pm$4σ. Figure 5.19 shows the chosen Monte Carlo rate and smoothed rate map. Figure 5.20 shows the resulting probabilities as well as the σ values. From Figure 5.21 it can be seen that the distribution of probabilities is mostly flat, with the maximum number of stations having any given probability to be 8 and no spike towards 100 % or 0 %. The distribution of σ is shown to not be skewed, with no high frequency towards $\pm$4σ. This is expected since the parent distribution is a Gaussian and this also confirms that the smoothing method does not alter the rate information. There is a longer tail towards -4σ which may be due to the distribution of rates of the chosen Monte Carlo smoothed map.



(a) Chosen map of rates simulated through Monte Carlo methods.

(b) Chosen map of simulated smoothed rates.

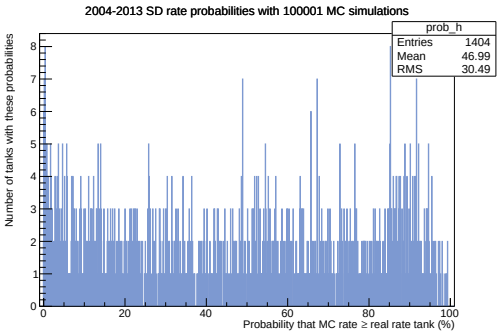
Figure 5.19: An example of a random map generated using Monte Carlo methods.



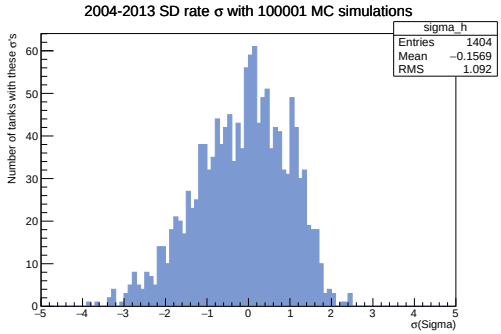
(a) Map of probabilities for 100,000 Monte Carlo simulations.

(b) Significance, σ converted from probability.

Figure 5.20: Two-dimensional histograms of probabilities and significances for the maps in Figure 5.19.



(a) Distribution of probabilities from the Monte Carlo test.



(b) Distribution of significance, σ with an RMS value of 1.09 and a mean of -0.157.

Figure 5.21: Distribution of probabilities and significance as σ values.

5.3 Rates as a Function of Position for Separate Time Periods

To observe the behaviour of rates across the array as time progresses, four maps were produced to show the rates at different periods of time. They are known as the 2004-2006, 2007-2009, 2010-2013 and 2014-2017 time periods. Durations of three and four years in these periods were chosen to maximise the number of events possible for statistical sufficiency. The number of events and uptimes for all stations in these periods were obtained and the rates were calculated and smoothed using the method shown by Equation 5.2.

Figures 5.22, 5.23 and 5.24 show the smoothed rates of all three time periods within the 2004-2013 block. The black square in Figure 5.24 represents a station which has detected events but have no uptime from 2010 to 2013. This was plotted just for reference, and will be of greater significance in Section 5.4. It can be seen that the rates in the 2004-2006 time period for the stations in the southern area are much lower than other stations in the array as well as the other time periods.

Lastly, the rates for the final time period, using events from 2014 to 2017 were calculated. Figure 5.25 shows a smoothed rate map for data from 2014-2017. Compared to the other time periods, the rates for all stations are much higher. This could be attributed to the introduction of new triggers at the end of 2013, making the stations more sensitive in event detection, but it is unclear how large this contribution is. From adjusting the scale, Figure 5.26 shows a prominent hotspot located on the north eastern corner of the array. It is unknown as to why there is a higher rate in this area. This is investigated in more detail in Section 5.4. The small hotspot in the north-western region is similar to the one in the 2004-2013 time period. Overall it has been observed that the rates of all stations seem to increase from the first (2004-2006) to the fourth (2014-2017) time period.

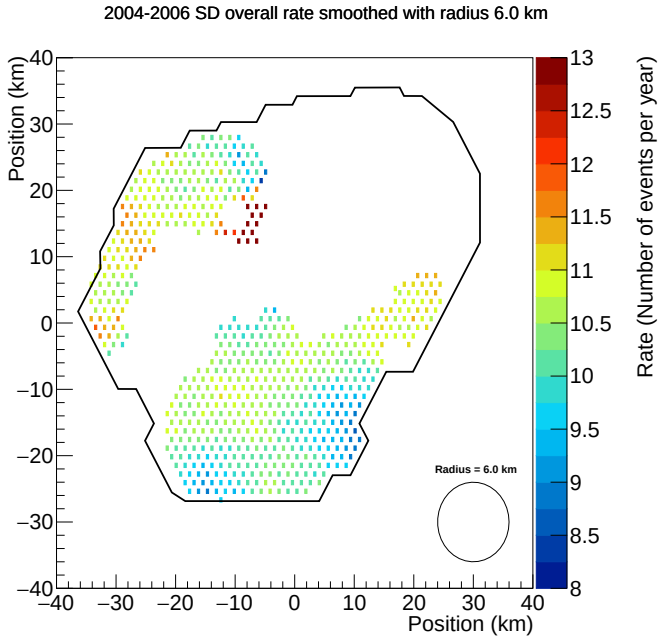


Figure 5.22: Smoothed rates as a function of position across the array for the years 2004 to 2006.

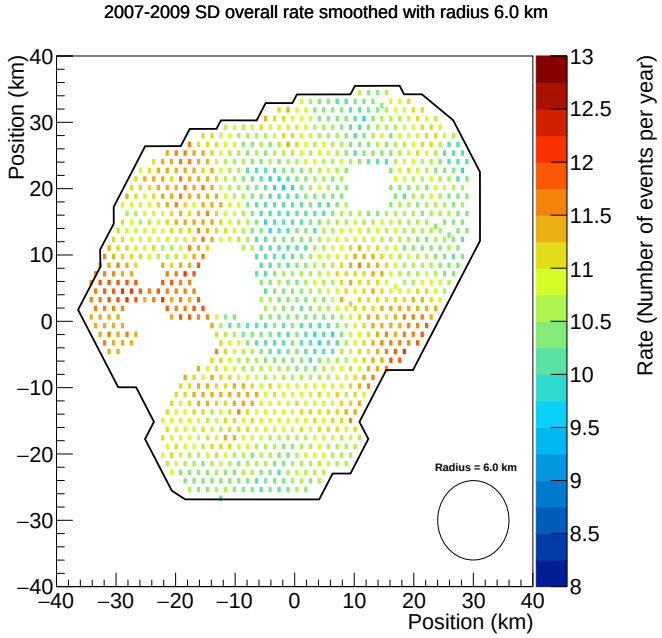


Figure 5.23: Smoothed rates as a function of position across the array for the years 2007 to 2009.

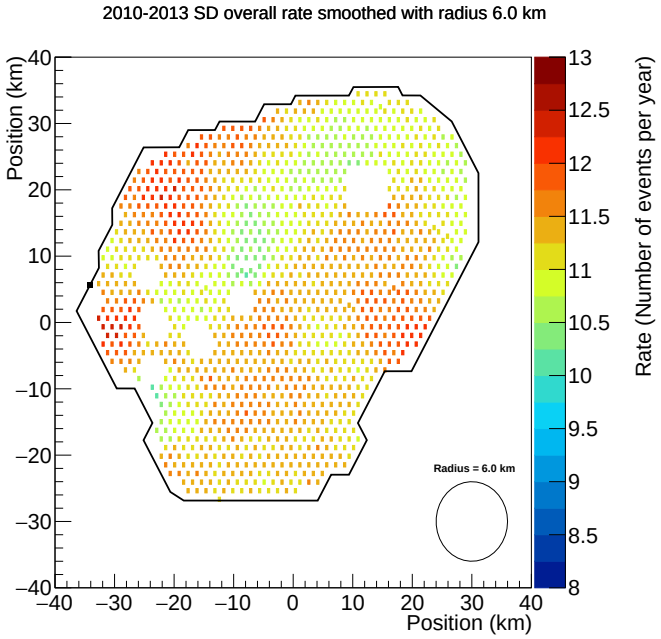


Figure 5.24: Smoothed rates as a function of position across the array for the years 2010 to 2013. The black square representing a station located on the western border of the array had detected some events, but had no uptime accumulated between 2010 and 2013. This station may have either had an issue during data acquisition or was in the process of being assigned as a border station as it should not detect any 6T5 events.

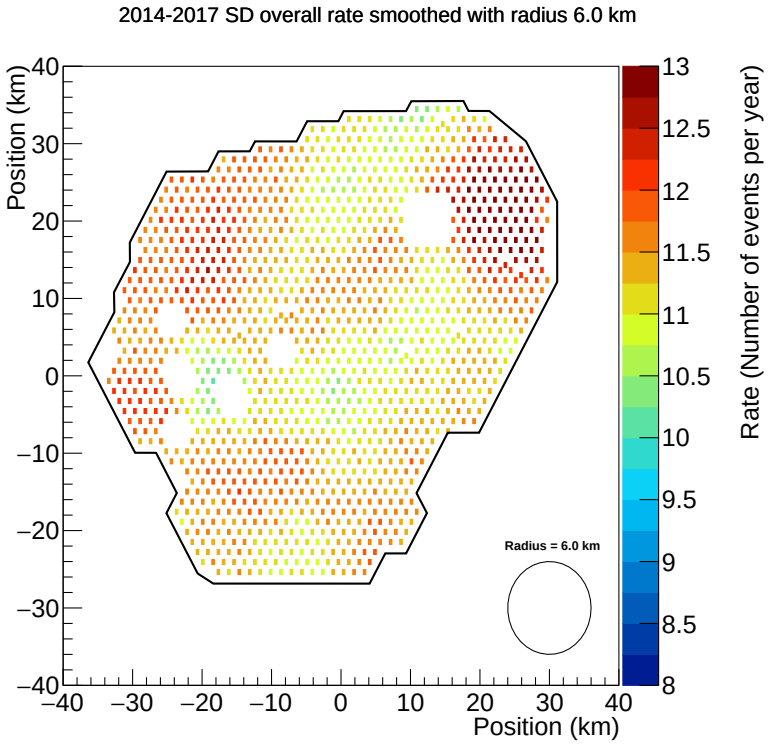


Figure 5.25: A smoothed rate map for data between 2014 and 2017.

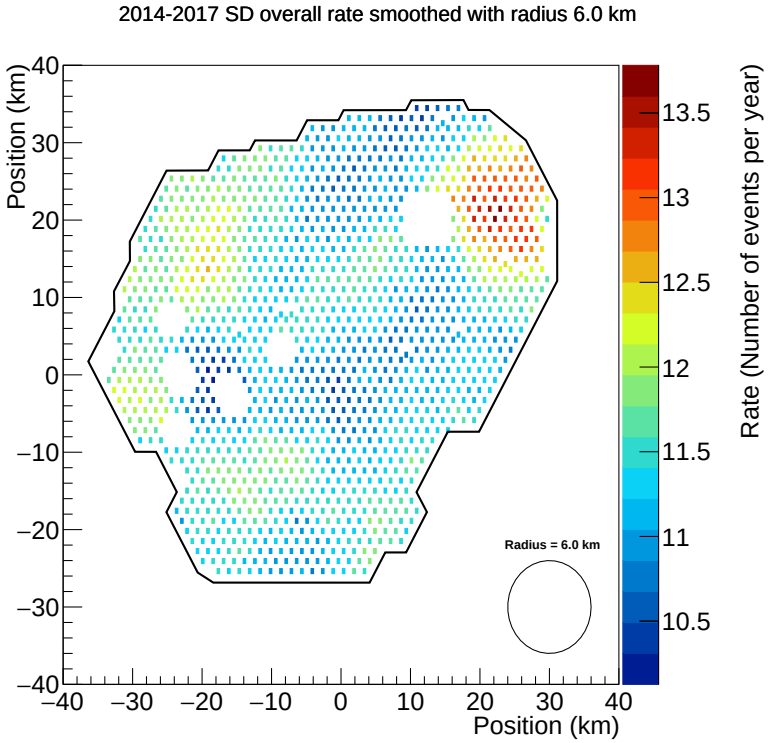


Figure 5.26: A smoothed rate map for data between 2014 and 2017 with scales adjusted to reveal an interesting hotspot shown on the north west region of the array.

5.4 Long Term Analysis of Rates Recorded by the SD from 2004 to 2017

In principle, the rates of all stations should be constant as time progresses. An analysis of rates in the long term was done to see if a trend exists for all stations across the array. Further analysis on Figure 5.3 was done by obtaining the goodness of different types of fits applied.

5.4.1 Fits on Rates as a Function of Time

Four different fits were done onto the graph of rates as a function of time and they are shown in Figure 5.27. The four fits consist of a horizontal line fit, a linear fit, a linear-horizontal fit and a linear-linear fit, similar to the studies done in Sections 4.3.2.1, 4.3.2.2, 4.3.2.3 and 4.3.2.4 regarding the energy ratio as a function of time. Figure 5.27(a) shows the result of the horizontal line fit. As for the linear fit, the functional form used is shown as

$$y(t) = m(t - t_0) + c \tag{5.5}$$

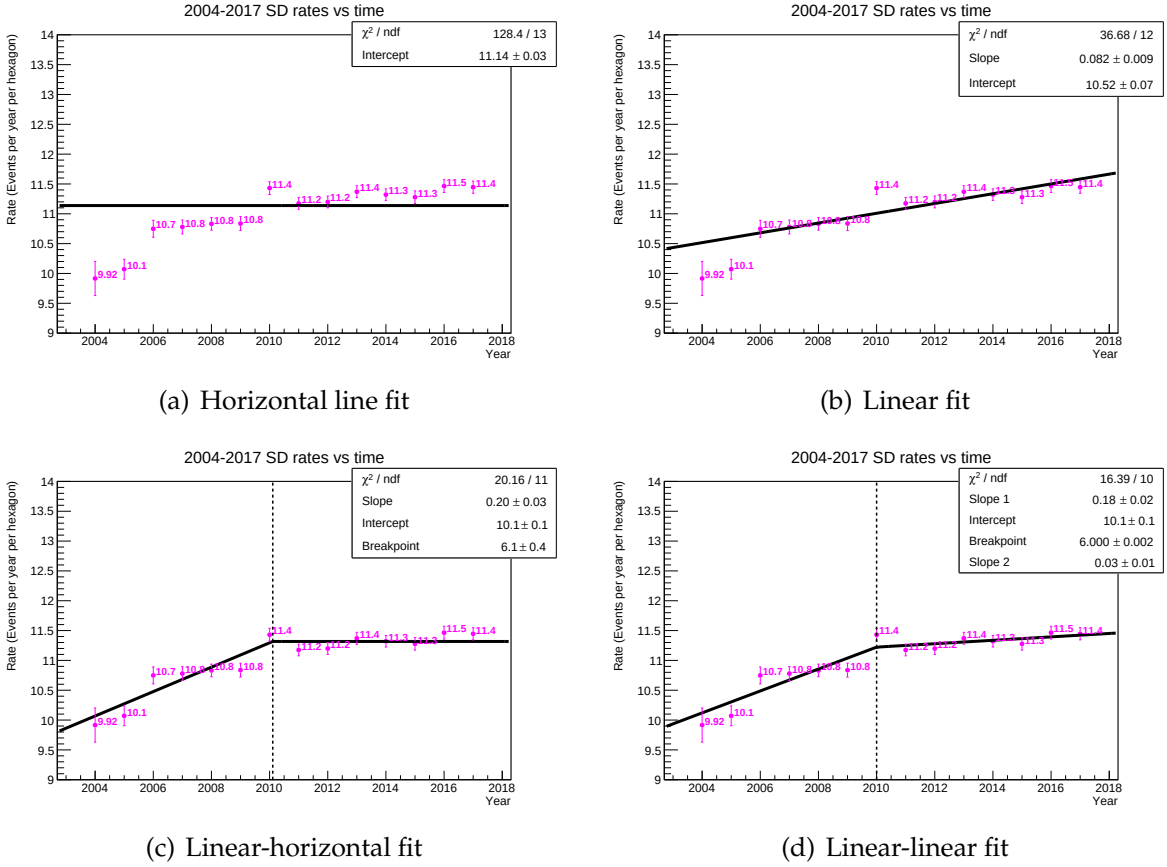
where *y(t)* is the rate as a function of year, m is the slope parameter in events/year^2 /hexagon, t is time in years, t_0 is the year 2004 and c is the y-intercept parameter in events/year/hexagon. Figure 5.27(b) shows the result of the linear fit. Then, the linear-horizontal fit, the functional form is

$$y(t) = \begin{cases} m(t - t_0) + c & t - t_0 < b \\ mb + c & t - t_0 \geq b \end{cases} \tag{5.6}$$

where b is the breakpoint as the number of years after 2004. Figure 5.27(c) shows the result of the linear-horizontal fit. Lastly, for the linear-linear fit, the functional form is shown as

$$y(t) = \begin{cases} m_1(t - t_0) + c_1 & t - t_0 < b \\ m_2(t - t_0) + c_2 & t - t_0 \geq b \end{cases} \tag{5.7}$$

where m_1 and m_2 are two slope parameters, while c_1 and c_2 are two y-intercept parameters. Figure 5.27(d) shows the linear-linear fit result. Using the reduced χ^2 values, the probabilities were calculated. They are shown in Table 5.1. The fit parameters are listed in Table 5.2. From these results, it can be seen that the horizontal fit is statistically unlikely for the behaviour of the rates recorded by the SD as a function of time, since the deviation is beyond 3σ. There is an improvement in the χ^2_{red} for the linear fit. However, its deviation still lies beyond 3σ. This fit in particular is almost comparable to the work of Isabelle Lhenry-Yvon [135] shown in Figure 5.28. The data used is reconstructed using the Central Data Acquisition System (CDAS) reconstruction code with an analysis program called *Herald* which is used specifically for the surface detector. This data spans



(a) Horizontal line fit

(b) Linear fit

(c) Linear-horizontal fit

(d) Linear-linear fit

Figure 5.27: Fits on plots of rates as a function of time for all stations between 2004 to 2017.

Fit type	χ^2	NDF	χ^2_{red}	Probability
Horizontal	128.4	13	9.88	4.36×10^{-21}
Linear	36.7	12	3.06	0.000252
Horizontal (2006 onwards)	66.4	13	5.11	3.66×10^{-9}
Linear (2006 onwards)	18.42	12	1.54	0.104
Linear-horizontal	20.2	11	1.83	0.0432
Linear-linear	16.4	10	1.64	0.0890

Table 5.1: Results from horizontal, linear, linear-horizontal fit for rates detected by all hexagons as a function of time. The colour red represents the probabilities that have a deviation from the hypothesis of more than 3σ, the colour orange represents deviations between 2σ-3σ and green notes a range, 0-2σ.

Fit type	Slope 1 (Events/year^2 /hexagon)	Intercept (Events/year /hexagon)	Breakpoint (Years after 2004)	Slope 2 (Events/year^2 /hexagon)
Horizontal	N/A	11.14 ± 0.03	N/A	N/A
Linear	0.082 ± 0.009	10.52 ± 0.07	N/A	N/A
Horizontal (2006 onwards)	N/A	11.19 ± 0.03	N/A	N/A
Linear (2006 onwards)	10.80 ± 0.06	0.065 ± 0.009	N/A	N/A
Linear-horizontal	0.20 ± 0.03	10.1 ± 0.1	6.1 ± 0.4	N/A
Linear-linear	0.18 ± 0.02	10.1 ± 0.1	6.000 ± 0.002	0.03 ± 0.01

Table 5.2: Fit parameter results from horizontal, linear, linear-horizontal fit for rates of all stations as a function of time.

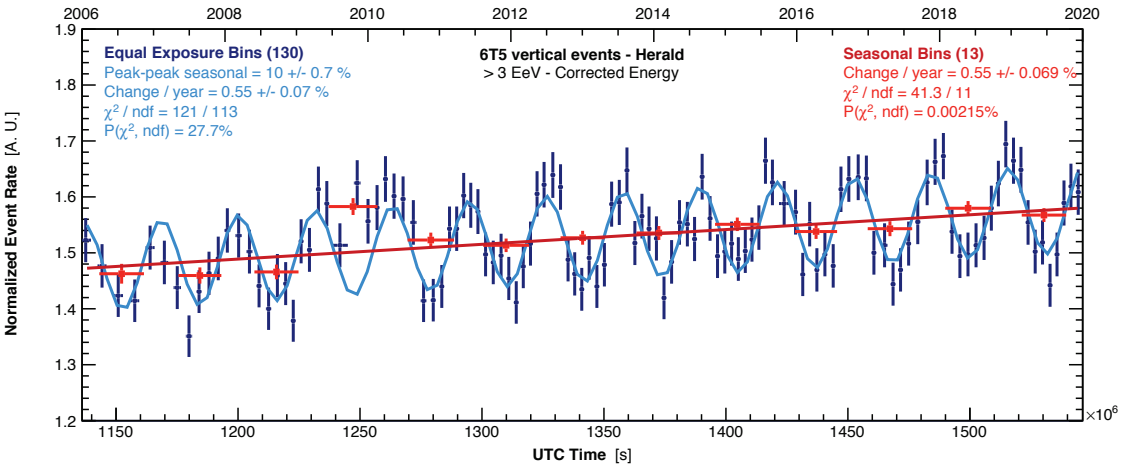


Figure 5.28: Analysis of the SD rate as a function of time for weather-corrected events, with energies above 3 EeV and using 6T5 triggered stations [135]. The data used is reconstructed using the Central Data Acquisition System (CDAS) reconstruction code with an analysis program called *Herald*. The dark blue points represent data in equal exposure bins with a light blue sine wave fitted to it. The red data points are separated into seasonal bins. The red line is a linear fit applied to the red points. The bottom x-axis represents UTC time in seconds while the top x-axis is year. The change per year, χ^2, number of degrees of freedom and probabilities from the fits are shown for the blue and red data points.

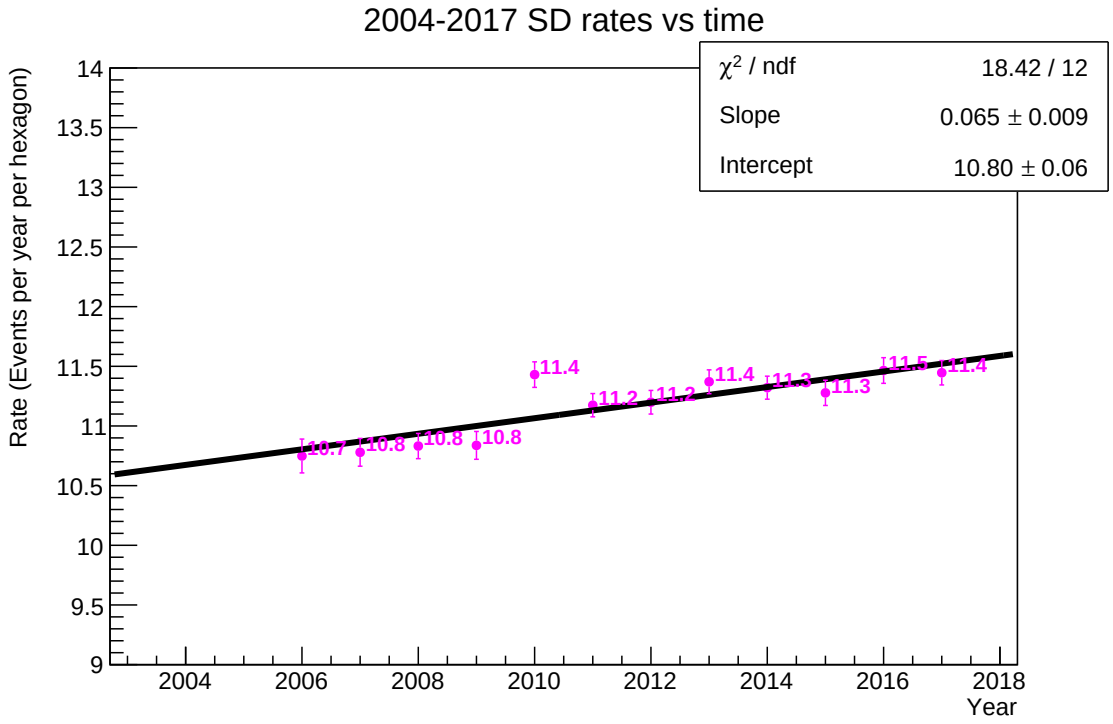


Figure 5.29: A linear fit applied to only the data points post-2016 (inclusive) for a comparison with the work done by Lhenry-Yvon shown as the red fit in Figure 5.28.

from 2006 to 2019. Looking in particular at the red data points and linear fit, the rates recorded by the SD are plotted in approximately yearly bins. The rates have also been normalised to obtain arbitrary units in the y-axis. It can also be seen that there is a jump in the normalised event rate in 2010 as shown prominently by the red data point. The red line representing the linear fit gives a change per year of 0.55 ± 0.07 %. To compare with this value, a modified version of the fit previous shown in Figure 5.27(b) is generated. The first two points representing the rate in 2004 and 2005 were removed and a new linear fit was done. The result is shown in Figure 5.29. The χ^2_{red} value shown gives a probability of ~0.104, which is within the 0-2σ deviation. By taking the value of the slope from the linear fit in this result and an average rate obtained as the y-intercept parameter from a horizontal line fit of the rates as a function of time, from 2006 onwards, the change per year is calculated as

$$\text{Change per year} = \frac{\text{Slope}}{\text{Average rate}} \times 100\,\% \tag{5.8}$$

The result of the horizontal line fit from 2006 onwards is shown in Figure 5.30. Taking the value of the slope which is 0.065 events/year^2/hexagon and the average rate, which is 11.19 ± 0.03 events/year/hexagon, the change per year is calculated as

$$\text{Change per year} = \frac{0.065}{11.19} \times 100\,\% \simeq 0.58\,\% \tag{5.9}$$

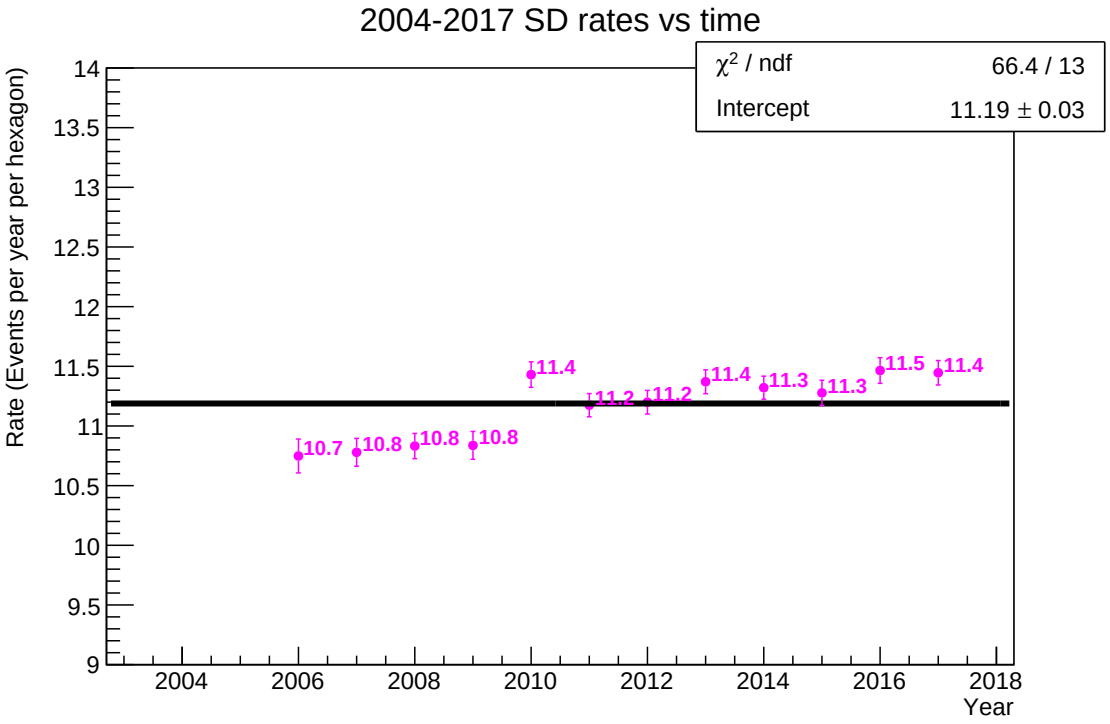


Figure 5.30: Modified fit done to compare with the work done by Lhenry-Yvon.

The uncertainty in the change per year is calculated as

$$\sigma_{\text{Change per year}} = \frac{\sigma_{\text{Slope}}}{\text{Average rate}} \times 100\ \% \tag{5.10}$$

and since the uncertainty in the slope, $\sigma_{\text{Slope}} = 0.009$,

$$\sigma_{\text{Change per year}} = \frac{0.009}{11.19} \times 100\ \% = 0.08\ \% \tag{5.11}$$

Thus, it was found that the change per year of the rates recorded by the SD found in this work, which is 0.58 ± 0.08 % is within the uncertainty of the value found in the work shown in Figure 5.28 which is 0.55 ± 0.07 %. This is obtained despite the difference in the reconstruction methods and the data timespan. The data used in this work is reconstructed using $\overline{\text{Off}}\underline{\text{line}}$, which is software developed to reconstruct data from the SD as well as many other detectors at the Pierre Auger Observatory [98]. Also, only data from 2006-2017 is displayed. In contrast, the work shown in Figure 5.28 has data reconstructed using the CDAS Herald code and has seasonal binning spanning from 2006 to 2019. Not only that, the increase in the rate from 2009 to 2010 as shown in Figure 5.29 seems to match the one in Figure 5.28. In spite of this, it is noted that the probability obtained from the fit in Figure 5.28 has a deviation which lies beyond 3σ, as opposed to that of this work which is between 0-2σ. A plot of monthly rates as a function of time was

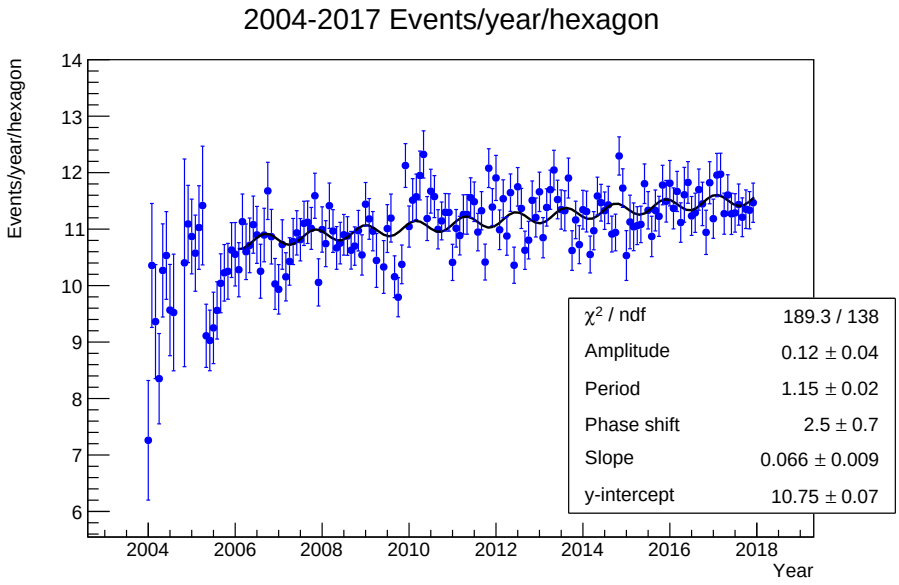


Figure 5.31: Evolution of monthly rates. The black line is a sinusoidal fit shown in Equation 5.12 and only fits the data from 2006 to the end of 2017. The probability for this fitted sinusoidal behaviour is 0.2 %.

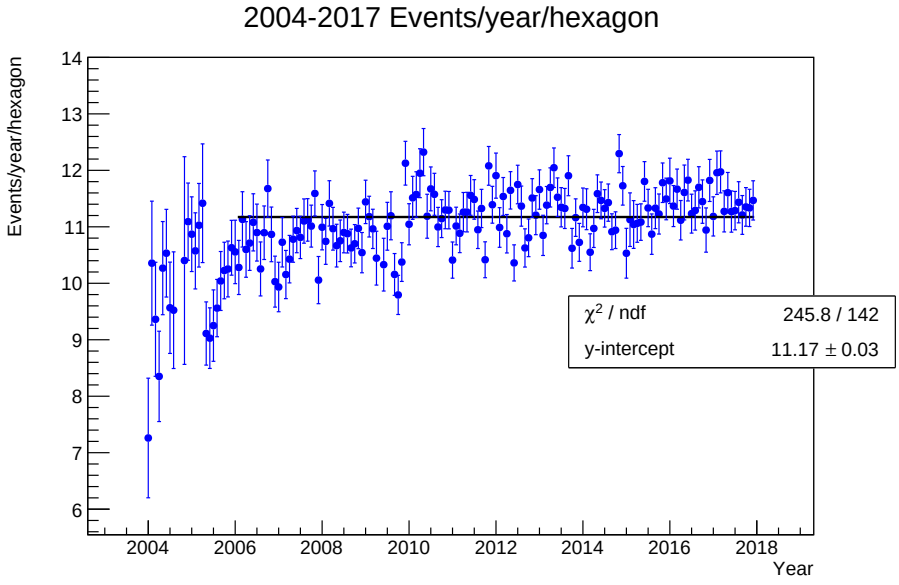


Figure 5.32: Evolution of monthly rates. The black horizontal line is a fit to find the average rate over 2006 to the end of 2017.

generated, to see if there is a modulation similar to that shown in Figure 5.28. The result is shown in Figure 5.31 with a sinusoidal fit applied to the plot only in the range between 2006 and the end of 2017. Each month was calculated to have 2,629,800 seconds.[1]. The fit function is shown as

$$\text{Monthly rate (t)} = A\cos\left(2\pi\left(\frac{t - t_{\text{shift}}}{T}\right) + \phi\right) + m(t - t_{\text{shift}}) + \textit{y-int} \tag{5.12}$$

where A is the amplitude, t_{shift} is the start of 2006, T is the period, ϕ is the phase shift, m is the slope and *y-int* is the y-intercept. The result of the fit is shown to have an amplitude of 0.12 $\pm$ 0.04 events/year/hexagon, a period of 1.15 $\pm$ 0.02 years, a phase shift of 2.5 $\pm$ 0.7 radians, a drift of 0.066 $\pm$ 0.009 events/year^2/hexagon and a y-intercept of 10.75 $\pm$ 0.07 events/year/hexagon. To calculate the peak-to-peak value as a percentage, the amplitude was doubled and converted shown as

$$\text{Peak-to-peak (\%)} = \frac{\text{Amplitude} \times 2}{\text{Average rate}} \times 100\,\% \tag{5.13}$$

The average rate is obtained as the fitted parameter from a horizontal line fit applied to the monthly rates and this is shown in Figure 5.32. Thus, using the values of 0.12 $\pm$ 0.04 events/year/hexagon and 11.17 $\pm$ 0.03 events/year/hexagon, the peak-to-peak (%) is calculated as

$$\text{Peak-to-peak (\%)} = \frac{0.12 \times 2}{11.17} \times 100\,\% = 2.2\,\% \tag{5.14}$$

The calculation of the uncertainty in this value is shown as

$$\sigma_{\text{Peak-to-peak}} = \frac{\sigma_{\text{Amplitude}} \times 2}{\text{Average rate}} \times 100\,\% \tag{5.15}$$

$$\sigma_{\text{Peak-to-peak}} = \frac{0.04 \times 2}{11.17} \times 100\,\% = 0.7\,\% \tag{5.16}$$

The peak-to-peak (%) is calculated as 2.2 $\pm$ 0.7 % which is much smaller compared to the value shown in Figure 5.28 i.e. 10 $\pm$ 0.7 %. The fit probability of 0.2 % which depicts a deviation beyond 3σ could indicate that the fit function is not suitable for the behaviour of monthly rates. It seems that the data has a lot more noise and it may be more difficult for the fitter to discern a more obvious pattern and amplitude. This is in contrast to the sinusoidal function shown in Figure 5.28 which is shown to have a probability of 27.7%. The source of this large difference is unknown.

Then, a linear-horizontal fit was applied to the yearly rates as shown in Figure 5.27(c) rates and it was found that the χ^2_{red} improved significantly from the linear fit (over 2004-2017). There seems to be a prominent increasing rate trend from

[1]This is the result of 86400 seconds/day $\times$ 365.25 days/year divided by 12 months/year.

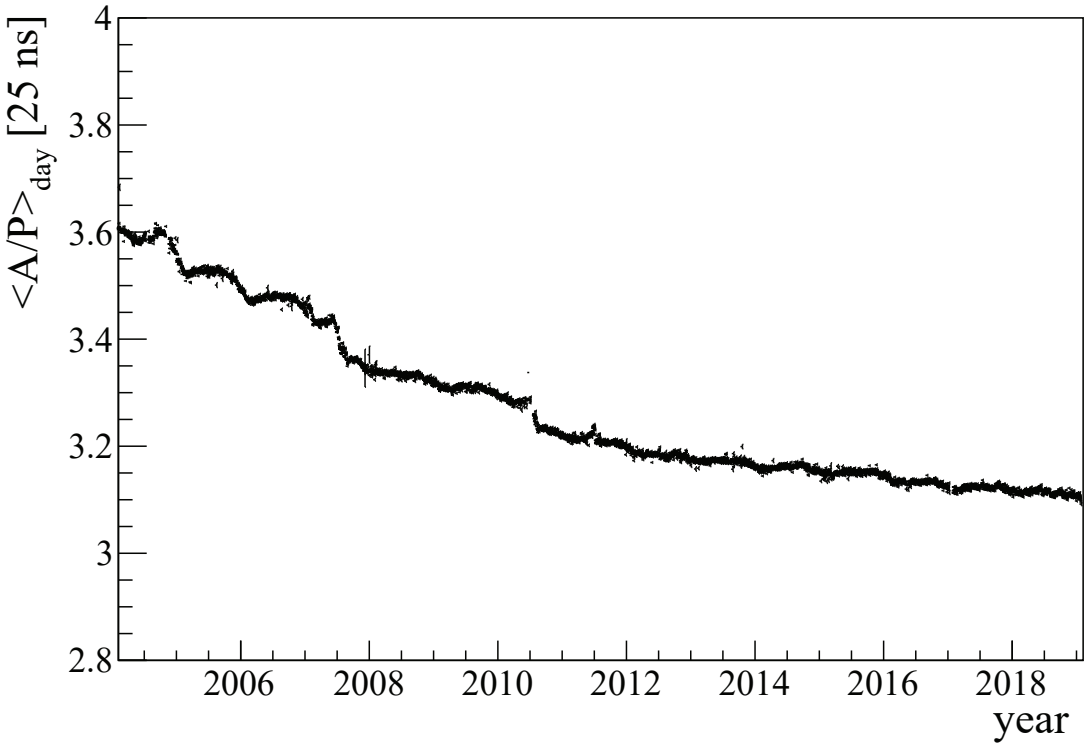


Figure 5.33: The area over peak as a function of time [136]. The data is obtained from one of the three PMTs in a station named Rocio with station ID 270.

2004 to 2010, where the breakpoint is located. It is unknown why there is a spike in rates in 2010 and this is further investigated in chapter 6, using weather station data. Finally, a linear-linear fit was applied. It was found that the second slope is much smaller than the first slope and it is unclear as to why this is so.

It can be seen that the behaviour of the rates as a function of time is similar to the *area over peak* evolution. The area over peak is a ratio that is constantly monitored by the SD calibration and monitoring system. It provides the basis for calibration for each station [98, 137]. The analog signal from the PMT of a station is digitised at a 40 MHz rate which means one sample is obtained every 25 ns. This sample is a measure of the width of a muon pulse in a station. This 40 MHz digitization rate is fast enough to obtain the muon signal shape, where A is the sum of digitised information around the region where the main signal appears and P is the maximum value of this signal [137]. Therefore, the area over peak value (A/P) is obtained in units of 25 ns. The A/P has a strong dependence on the reflection coefficient in a station, determined by the Tyvek liner and the water absorption factor. Figure 5.33 shows the evolution of A/P obtained from Station 270. The data spans from 2004 to the end of 2018 [136]. There seems to be a rapid decrease in the area over peak value shown as <A/P>, up to the start of 2010, where there is a slight gap in the data. A very steep slope prior to this gap, is similar to the behaviour shown in Figure 5.27(d), where before the breakpoint located in 2010, the slope is steeper. The decay in A/P indicates that there could

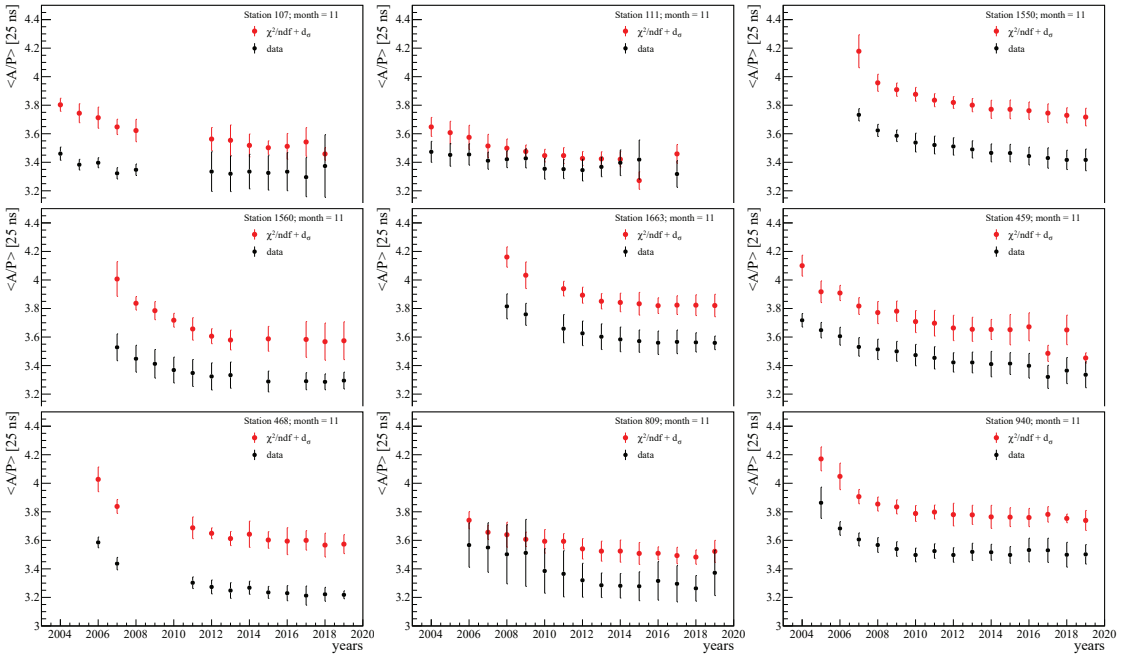


Figure 5.34: The area over peak evolution for 9 different stations using data obtained in the last week of November from 2006 to 2019 [138]. The black points represent real data while the red points are simulations.

be a combined effect from the reduction in reflectivity of the Tyvek liner and the water absorption.

Figure 5.34 shows the area over peak evolution for 9 different stations in the last week of November. The red points represent simulations done to observe the signal response to air showers and calibration procedure with a changing Tyvek liner reflectivity [138]. It can be seen that the area over peak has a rapid decrease over time for most stations, up to $\sim$ 2010, after which the trend becomes flatter and smoother. This trend has been attributed to a strong dependency on the Tyvek liner reflectivity of the station, through preliminary analysis done in [138]. Although there is a still a difference in the area over peak of 0.3 [25 ns] between the simulations and the data as shown in Figure 5.34, the overall decreasing trend has been reproduced through the decrease of the liner reflectivity.

Another very interesting study was done by Isabelle Lhenry-Yvon and this is shown in Figure 5.35 where there is a direct correlation between the decay in the A/P and the rates of events at all energies (no threshold energy cuts). This correlation indicates that there is a reduction in the trigger efficiency attributed to the degradation of the Tyvek liner reflectivity. Since the station calibration histograms only keep track of the response to muons, this indicates a need for further study into whether calibration using muons every minute is not keeping track of the drift in sensitivity. There is also a possibility that the response of stations to the electromagnetic component is not tracked perfectly and is detailed briefly in Section 5.6.

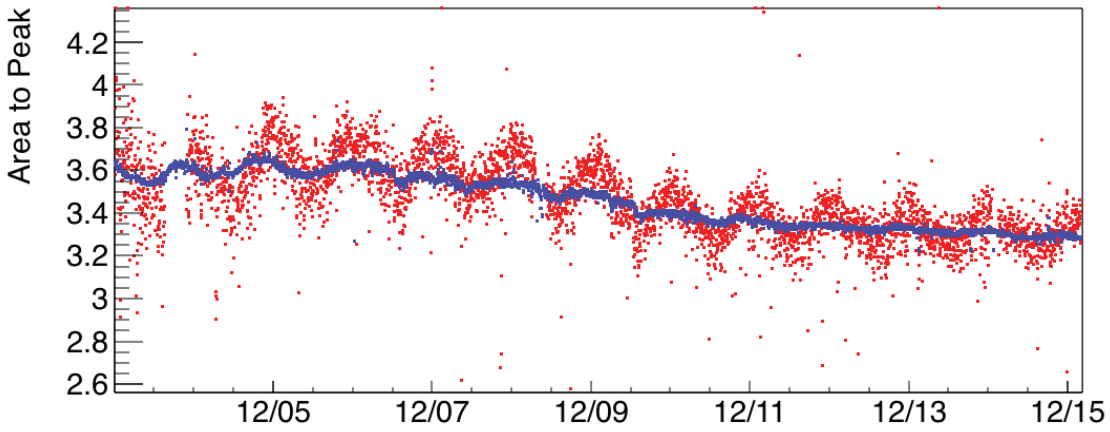


Figure 5.35: Evolution of area over peak (blue points) averaged for all stations in the 1500 m array compared with re-scaled event rates (red points) [139].

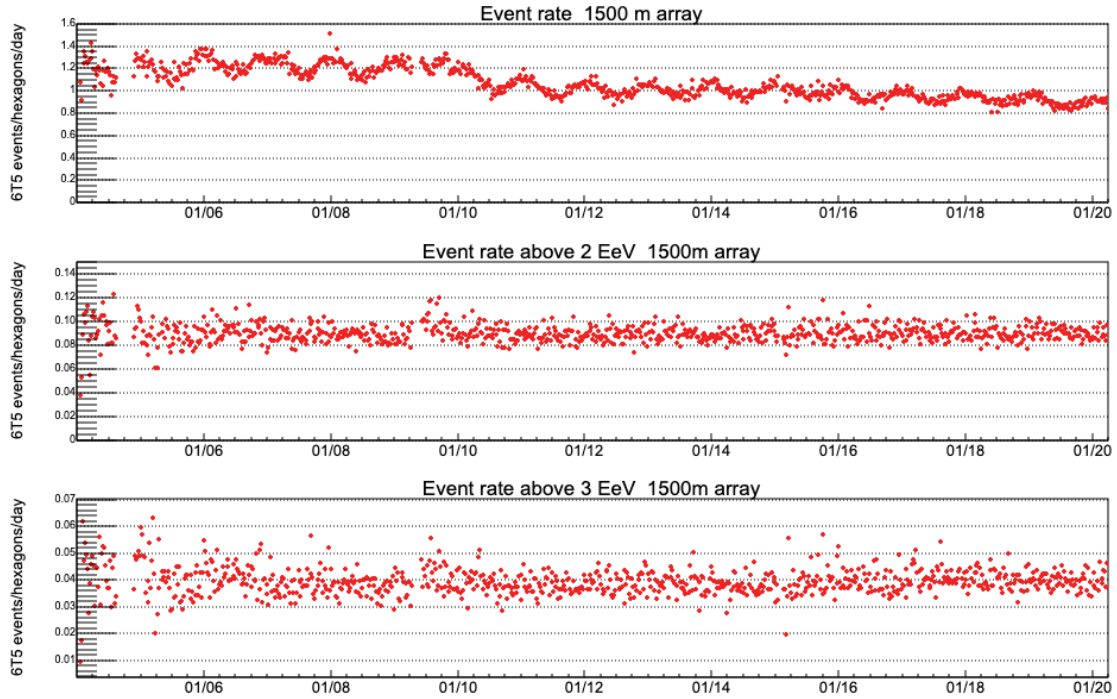
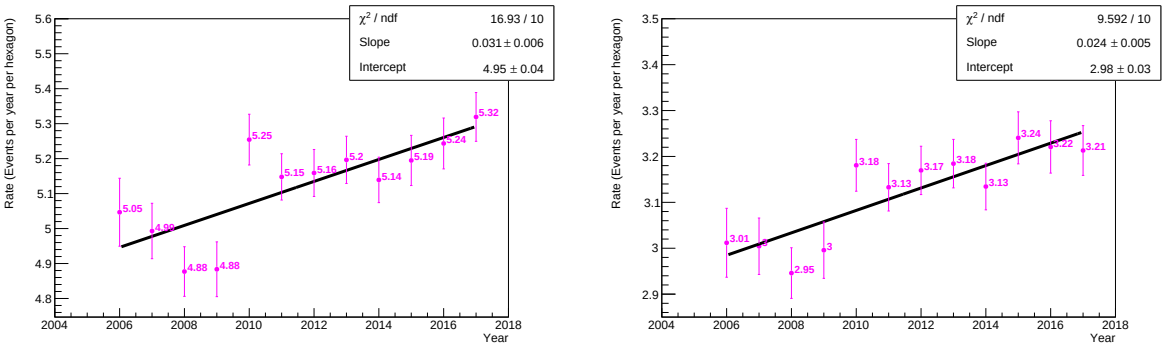


Figure 5.36: Three panels showing weather-corrected daily event rates between 2004 to 2020 for the 1500 m array [135]. Top panel shows event rates for all energies, middle panel shows rates for events with energies above 2 EeV and the bottom panel shows event rates with energies above 3 EeV.



(a) Rates above 4.5 EeV

(b) Rates above 6 EeV

Figure 5.37: Rates as a function of time with a threshold energy of (a) 4.5 EeV and (b) 6 EeV.

Lhenry-Yvon's work also shows a comparison between the event rates (with no threshold) as shown in Figure 5.35, with two other plots of rates with threshold energies. Figure 5.36 shows weather-corrected daily event rates at three thresholds; for all energies, for energies above 2 EeV and energies above 3 EeV. The slope is seen to be negative for event rates at all energies (with a sharp discontinuity mid-2010) and the slope is $\sim$ zero for energies above 2 EeV. It is not directly obvious that the slope of rates above 3 EeV is positive. The flat-looking trend may be due to the very low increase in rate day after day. Since the rate increases by 0.58 % in 12 years as shown in this work in Equation 5.9, it can be said that at least in first order, the rate increases by 0.048 % each year, implying that the rate increase per day is around 0.00013 % per day. In any case, the much more obvious increasing yearly event rates above 3 EeV could be due to the drift in the calibration of SD energy.

It was then checked whether the weather-corrected yearly rates above higher threshold energies exhibit a decreasing, flat or increasing trend. Figure 5.37 shows the rates of events with energies above 4.5 EeV and 6 EeV. The rate change per year (%) was obtained to be 0.6 $\pm$ 0.1 % and 0.8 $\pm$ 0.2 % respectively. These values are within the uncertainty of the change per year in rates above 3 EeV which is 0.58 $\pm$ 0.08 %. The reason behind this agreement may be because these thresholds are far enough above the full efficiency threshold and with a steep energy spectrum (i.e. very few events at such high energies), the slope in the rates as a function of time do not increase high enough.

Previous analysis was also done on the event rates per unit exposure (km^{-1} sr^{-1} yr^{-1}) observed by the SD above a higher threshold energy as a function of time and station age. It was found that with a threshold energy above 5 EeV and 10 EeV, the event rates per unit exposure increase as function of time and station age as shown in Figures 5.38 and 5.39. Additionally, it was found that the increase in event rates may be related to two periods of freezing events. During the winter months of 2007 and 2010, there were some stations that experienced

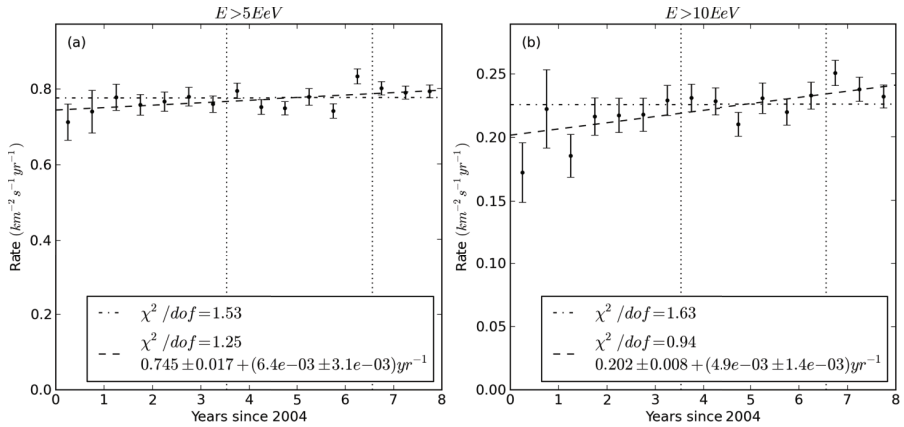


Figure 5.38: Event rate per unit exposure for events with energies above 5 EeV and 10 EeV as a function of time in 6 month bins [140]. The two vertical lines indicate freezing events in 2007 and 2010.

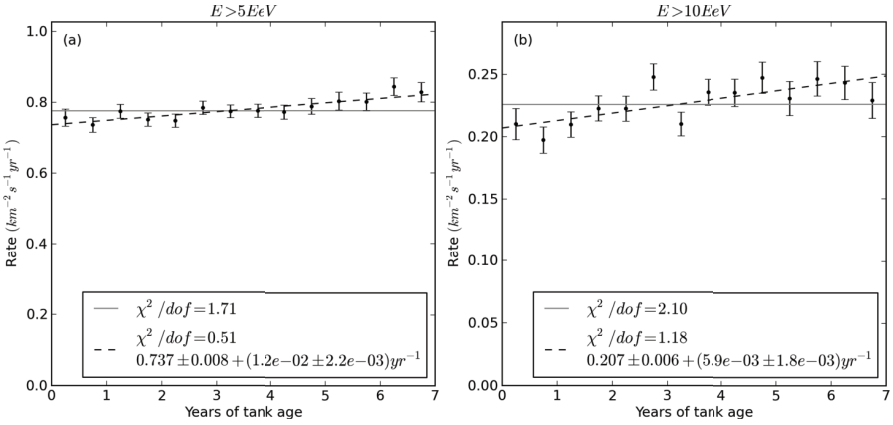


Figure 5.39: Event rate per unit exposure for events with energies above of 5 EeV and 10 EeV as a function of station age in 6 months periods [140].

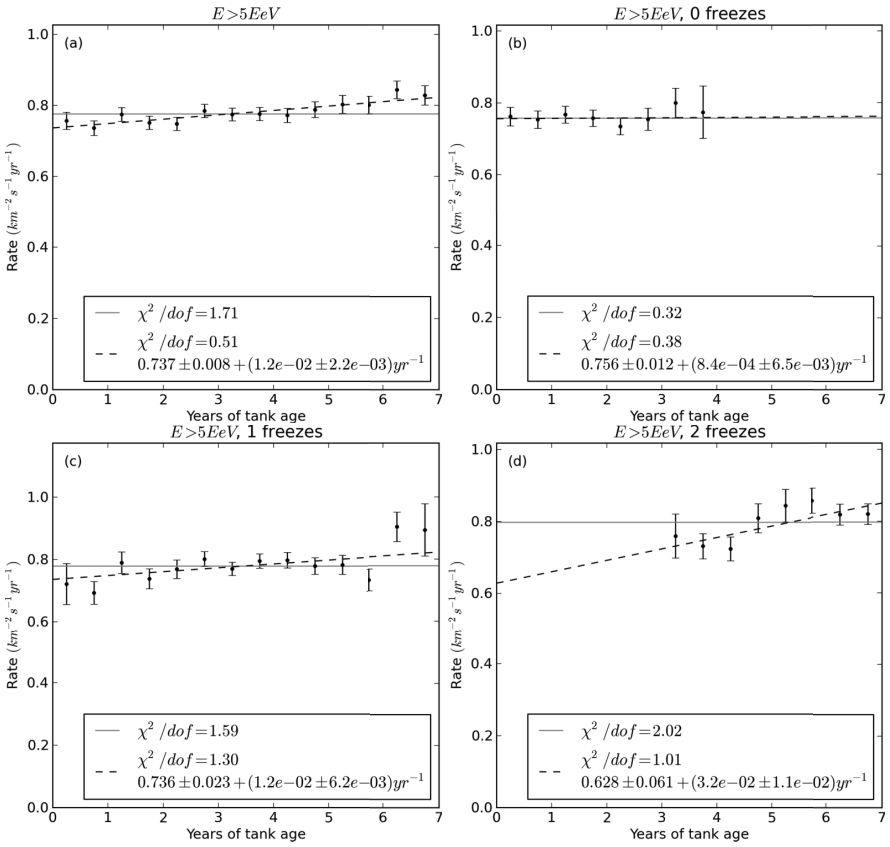


Figure 5.40: Comparison of rates above 5 EeV for (a) all stations (b) only unfrozen stations, (c) stations that froze once and (d) stations that froze twice [140].

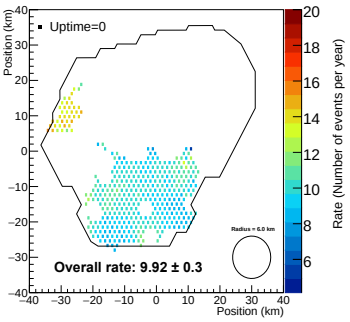
freezing and this was visible in the area over peak data obtained from the PMTs of those stations [140]. These events appeared to have taken place from the 6th to the 21st of July in 2007 and from the 18th to the 24th of July in 2010 [140]. When the analysis was done on stations which did not experience freezing during these periods, the rates are seen to be constant. This is shown in Figure 5.40 showing a comparison between the rates observed by all stations, rates of stations that have experienced no freezing, 1 freeze and 2 freezes. The result suggests that the increasing rate is related to freezing but may not be confined to two isolated events in 2007 and 2010 [140]. Other potential contributions to the increasing rates were hypothesised and detailed in Section 5.6.

5.4.2 Rates as a Function of Time and Position

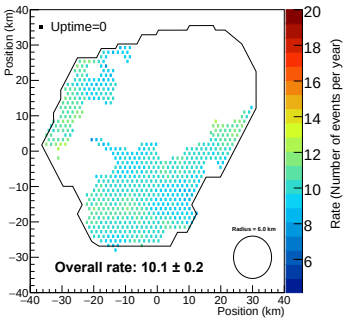
The variation in rates as a function of both time and position were observed. Maps of smoothed rates were plotted for every year and every station. It was noted that there are some stations which have detected events, but provide no uptimes for certain years. The lack of uptimes may be due to some technical issues with the stations, which may have still recorded data even though it could have been under maintenance.

Figures 5.41 and 5.42 show how the rates vary as a function of position on the array and time, as they show the rates of all hexagons in each particular year. The overall rates are calculated as the total number of events divided by the total available uptime of all stations in every year. The black squares represent the stations which have recorded events, but no available uptime. These are noted on the maps just for completeness and may be able to explain the hotspot seen in the rate calculation from 2014-2017 shown in Figure 5.25.

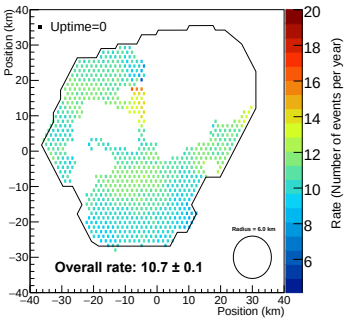
Specifically, there is one station which has recorded events but has no uptime from 2010 to 2013. There are also 37 stations which fulfil this criterion in 2016 and 12 stations in 2017. This makes a total of 50 stations with this criterion. There is a warm spot which appears in 2015 on the north-eastern corner of the array. Then in 2016, many stations are missing and the stations in left in the area have high rates. The situation stabilises again in 2017. The 37 stations with no uptime may have been under maintenance for a certain period of time, affecting the hexagons around them. For example, if there are 5 stations that are not active, this eliminates up to 30 stations, due to the 6T5 trigger criteria. Then in 2017, the stations recovered. However, there exists two extra holes in the array. The one station which does not have uptime from 2010 to 2013, is later removed from 2014 onwards and this may have been due to the removal process of the station from the 6T5 trigger criteria since it was located at the border of the array. As for the 12 stations in 2017, the stations may have been under maintenance creating two extra holes in the array



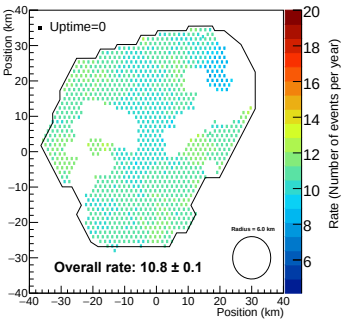
(a) Smoothed rates 2004



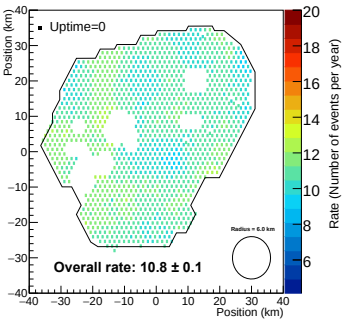
(b) Smoothed rates 2005



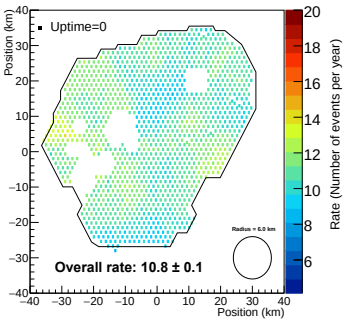
(c) Smoothed rates 2006



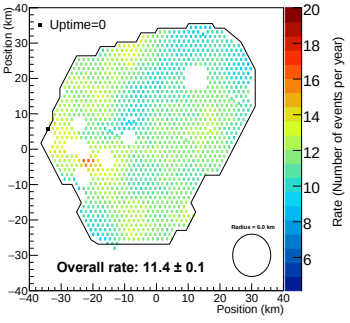
(d) Smoothed rates 2007



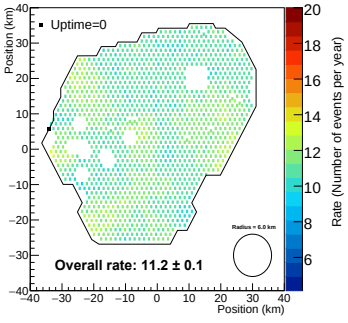
(e) Smoothed rates 2008



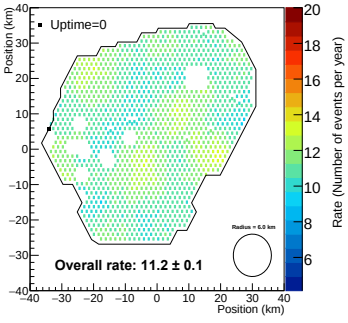
(f) Smoothed rates 2009



(g) Smoothed rate 2010

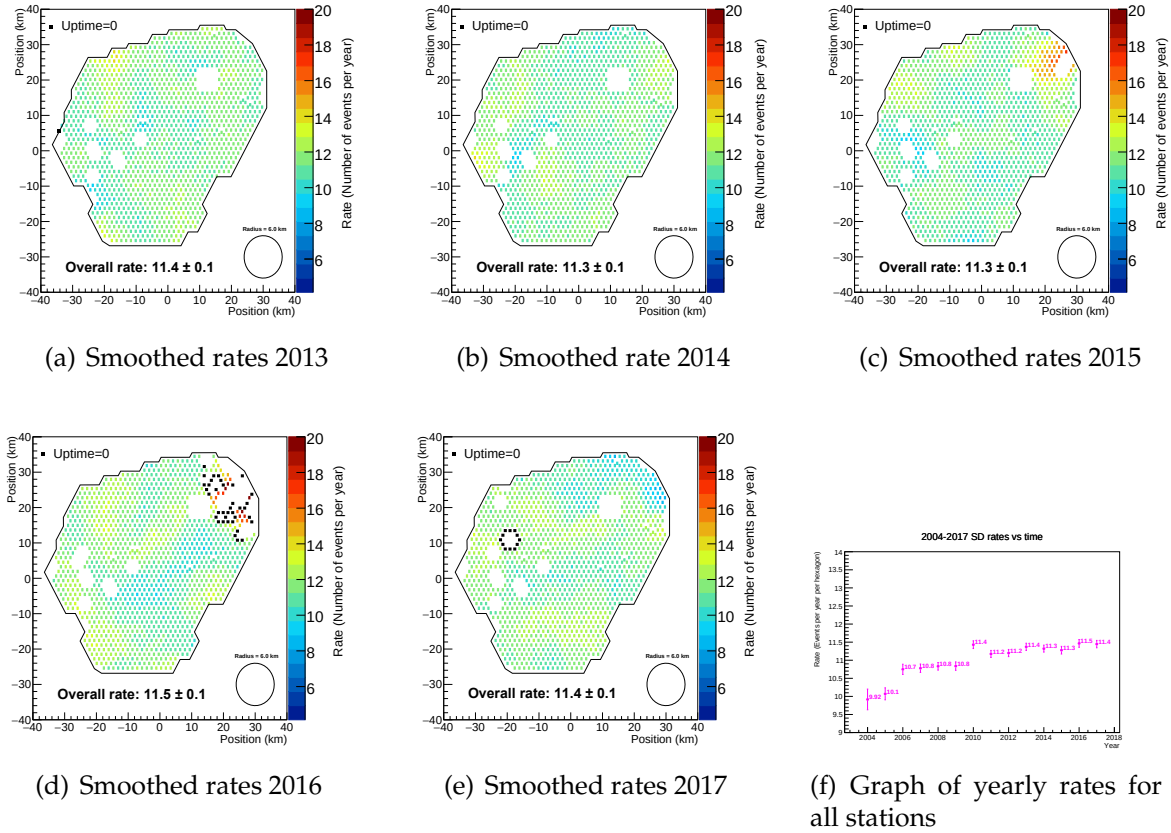


(h) Smoothed rates 2011



(i) Smoothed rates 2012

Figure 5.41: Smoothed rates for each year from 2004 to 2012. The overall rate is calculated as the total number of events divided by the total available uptime of all stations in every year. These values match with the data points in Figure 5.42(f). There is one black square representing a station which has recorded events, but has no available uptime, appearing on the border of the western side of the array, from 2010 to 2012.



(a) Smoothed rates 2013 (b) Smoothed rate 2014 (c) Smoothed rates 2015

(d) Smoothed rates 2016 (e) Smoothed rates 2017 (f) Graph of yearly rates for all stations

Figure 5.42: Smoothed rates for each year from 2013 to 2017. The overall rate is calculated as the total number of events divided by the total available uptime of all stations in every year. These values match with the data points in Figure 5.42(f). The black squares represent the stations which have recorded events, but no available uptime.

5.4.3 Rate of Events and Threshold Energy

The rate of events above a threshold energy as detected by the SD can be an indicator of its energy measurement stability. The relationship between event rate and threshold energy (E_{th}) of detectors can be obtained from the differential energy spectrum equation [110]. Firstly, it is known that

$$\frac{dN}{dE} \propto E^{-\gamma} = kE^{-3} \qquad \mathrm{m^{-2}s^{-1}sr^{-1}eV^{-1}} \quad (5.17)$$

where *k* is a proportionality constant and we take a spectral index of γ = 3 as an approximation of the true spectrum around the threshold energy of 3 EeV. During event detection, cosmic rays of different energies are measured. Therefore, the differential energy spectrum above a certain threshold energy can be integrated with respect to energy resulting in a value that represents event rate. This gives

$$\int_{E_{th}}^{\infty} \frac{dN}{dE} dE = \textit{Event rate} \qquad \mathrm{m^{-2}s^{-1}sr^{-1}} \quad (5.18)$$

Equation 5.18 represents the event rate above the threshold energy E_{th}. We can then substitute Equation 5.17 into 5.18 to obtain an event rate shown as

$$\textit{Event rate} = \int_{E_{th}}^{\infty} kE^{-3} dE = \left[-\frac{k}{2E_{th}^2} \right]_{E_{th}}^{\infty} = \frac{k}{2E_{th}^2} \quad (5.19)$$

Then, differentiating with respect to E_{th} on both sides results in

$$\frac{d(\textit{Event rate})}{dE_{th}} = \frac{-k}{E_{th}^3} \quad (5.20)$$

Equation 5.20 can be rearranged to obtain

$$d(\textit{Event rate}) = \frac{-kdE_{th}}{E_{th}^3} \quad (5.21)$$

and then divided by event rate shown in Equation 5.19 to obtain

$$\frac{d(\textit{Event rate})}{\textit{Event rate}} = \frac{-2dE_{th}}{E_{th}} \quad (5.22)$$

Equation 5.22 shows a direct relationship between event rate and threshold energy [110]. For example, if the event rate increases by 1 %, the threshold energy must decrease by half a percent. Since, the varying rates may be due to the varying threshold energy of stations in the array, using this relationship, the implied threshold energy can be obtained for all rate maps from 2004 to 2017. From Equation 5.22, the implied fractional change in threshold energy exhibited by every station can be calculated as a negative half of the fractional change in

rates. Using this relationship, the fractional changes in threshold energy for every station on the rate map as a function of year were plotted in Figures B.6-B.11 in Appendix B.

A brief comparison was done with the study detailed in Section 4.3.2.2 of Chapter 4, where it was shown that for most telescopes, E_{SD}/E_{FD} increase with time. The behaviour of this ratio may be partially caused by changes in the threshold energy of the SD. From Figure 5.27(b), the change per year of rates is calculated to be

$$\text{Change per year} = \frac{0.082}{11.14} \times 100\,\% \simeq 0.74\,\% \tag{5.23}$$

with uncertainties calculated as

$$\sigma_{\text{Change per year}} = \frac{\sigma_{\text{Slope}}}{\text{Average rate}} \times 100\,\% \tag{5.24}$$

$$\sigma_{\text{Change per year}} = \frac{0.009}{11.14} \times 100\,\% \simeq 0.08\,\% \tag{5.25}$$

As the rates recorded by the SD are increasing with time with a change per year of 0.74 $\pm$ 0.08 % as obtained from Figure 5.27(b), the true threshold energy is expected to decrease by 0.37 $\pm$ 0.04 % per year according to Equation 5.22. If the rates above an assigned energy of 3 EeV are going up with time, it indicates that the energies assigned to air showers are also going up with time. This indicates that the assigned SD energies are increasing by 0.37 % every year. If there was no contribution from the drift in E_{FD}, this agrees with the behaviour of the increasing trend of E_{SD}/E_{FD} at each telescope (shown in Chapter 4) due to the increasing assigned SD energies, E_{SD}. However, it was hypothesised that the FD energy does depend on the aerosol database and can have an effect due to UV filter and mirror cleanings as shown in Chapter 4. Since it was found that the energy scale ratio has a drift of 0.7 % post-2014, this would mean that after correcting the SD energies (by applying a reduction of 0.37 %), the remaining drift of 0.33 % will be due to a drift in FD energies. Horizontal, linear, linear-horizontal and linear-linear fits were also done on the energy ratio for each FD site(rather than each telescope), but the χ^2 values obtained were significantly poor. For reference, they are shown in Figures B.2, B.3 B.4 and B.5 in Appendix B. This indicates that the trend in the FD energy is compounded when all telescopes are taken into account in the plot of energy ratio against time at each FD site. This signifies a need for more complex fits in future work.

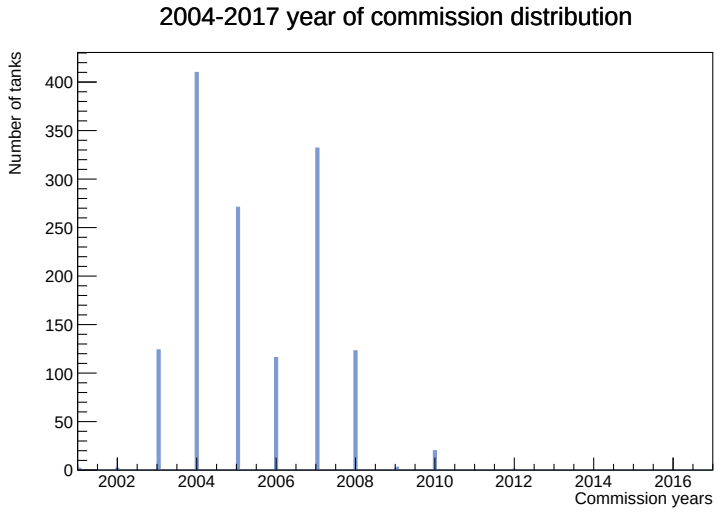


Figure 5.43: The number of stations commissioned between 2001 and 2016.

5.5 Rates of Stations Commissioned in Different Years

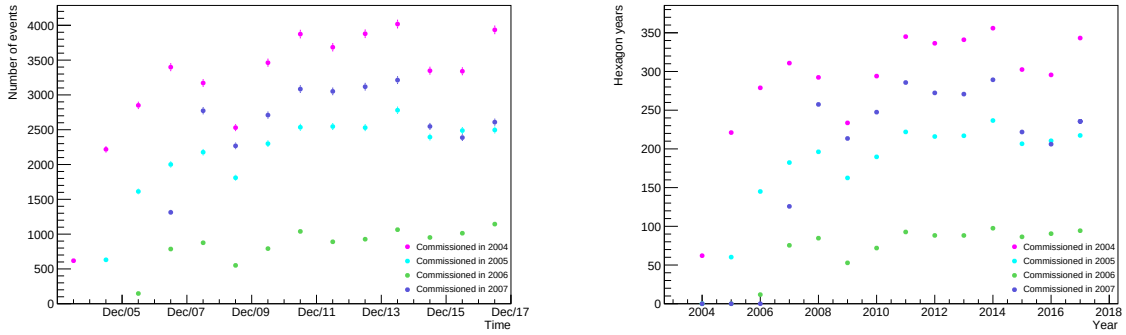
This section looks at the rates of stations based on their commission years. The aim is to observe whether the increasing rates shown previously in Section 5.4 is shown for stations commissioned in different years, and if they exhibit the same behaviour but at different start times. It was then checked whether these four station cohorts were the cause of hot/cold spots on the maps showing rates as a function of time and position across the array. Firstly, the number of stations commissioned in each year was obtained.

Figure 5.43 shows the number of stations commissioned between 2001 to 2016. Four of these commission years were chosen to see if the stations exhibit the same behaviour in the overall rates as a function of time as shown in Figure 5.3. The number of events and uptime of stations commissioned in 2004, 2005, 2006 and 2007 were obtained and shown in Figures 5.5 and 5.5. It can be seen that both plots exhibit similar behaviour which is expected. The higher the number of stations commissioned in any particular year, the higher the uptime and number of detected events.

Then the rates as a function of time were calculated as the ratio of the two, for each cohort. Figure 5.45 shows this result. Horizontal line, linear, linear-horizontal and linear-linear fits were applied to the varying rates from 2004 to 2017 for each station cohort to quantify their trends.

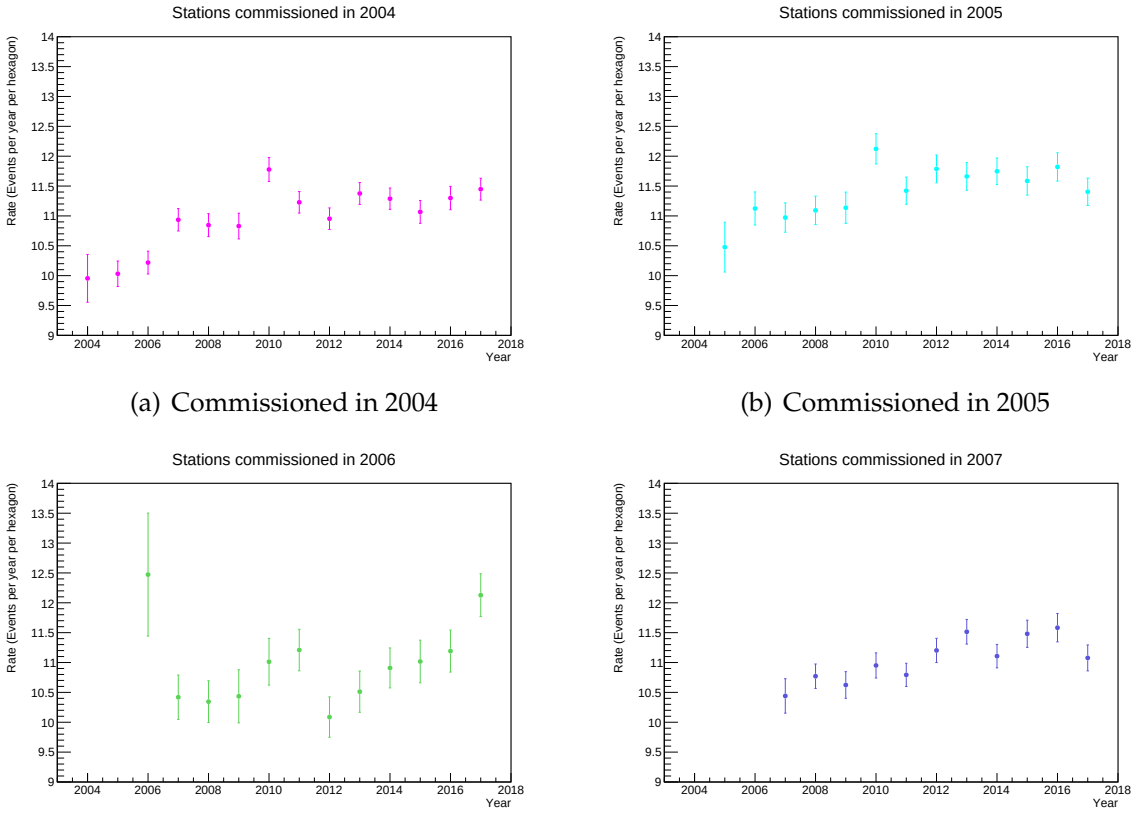
5.5.1 Horizontal Line Fits

Figure 5.46 shows the fit results for a horizontal line. The goodness-of-fit values are listed in Table 5.3. It can be seen that the fit for the 2004 cohort has the lowest



(a) (b)

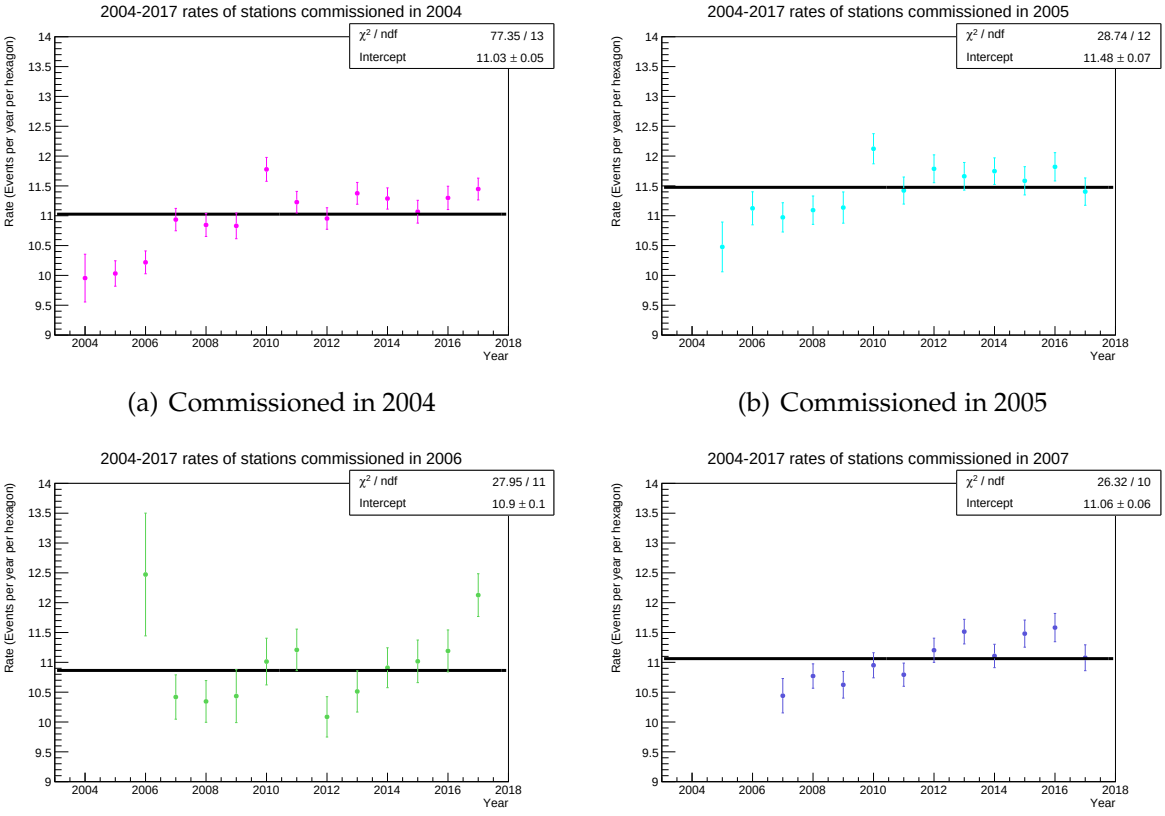
Figure 5.44: Plots showing (a) number of events as a function of time for stations commissioned in four different years and (b) uptimes as a function of time for stations commissioned in four different years. The vertical error bars represent the square root of the number of events. Pink represents stations commissioned in 2004, cyan is for 2005, green is for 2006 and purple for 2007.



(a) Commissioned in 2004 (b) Commissioned in 2005

(c) Commissioned in 2006 (d) Commissioned in 2007

Figure 5.45: Rates of events with a threshold energy of 3 EeV (weather-corrected) seen by four cohorts of stations as a function of time.



(a) Commissioned in 2004

(b) Commissioned in 2005

(c) Commissioned in 2006

(d) Commissioned in 2007

Figure 5.46: Stations commissioned in 2004, 2005, 2006 and 2007 with a horizontal line fit.

probability. The deviation of this fit also lies beyond 3σ. Due to this, the rates for the stations commissioned in 2004 is very unlikely to have a non-varying behaviour as a function of time. As for the 2005, 2006 and 2007 cohorts, the deviations from the fit lie between 2-3σ. The fit parameters as listed in Table 5.4 also show that out of all the y-intercept (normalisation) parameters, the 2005 cohort has the highest value among the others, with a value of 11.48 $\pm$ 0.07 events/year/hexagon. It is also noted that since the reduced χ^2 values (χ^2_{red}) are poor, the uncertainties in the parameters are not reliable.

5.5.2 Linear Fits

A linear fit was applied to all station cohort rate plots to test for statistically significant improvement of goodness-of-fit values. The following fit equation was used to fit to the data:

$$y(t) = m(t - t_{com}) + c \tag{5.26}$$

Fit type	Commission year	χ^2	NDF	χ^2_{red}	Probability
Horizontal	2004	77.3	13	5.95	3.47×10^{-11}
	2005	28.7	12	2.4	0.0043
	2006	28	11	2.54	0.00329
	2007	26.3	10	2.63	0.00333
Linear	2004	36.1	12	3.01	0.00031
	2005	17.8	11	1.62	0.0861
	2006	19.2	10	1.92	0.0377
	2007	10.9	9	1.21	0.283
Linear-horizontal	2004	17.1	11	1.55	0.105
	2005	9.86	10	0.986	0.453
	2006	25.3	9	2.81	0.00263
	2007	7.65	8	0.956	0.468
Linear-linear	2004	17	10	1.7	0.0751
	2005	9.7	9	1.08	0.375
	2006	12.3	8	1.54	0.139
	2007	7.61	7	1.09	0.369

Table 5.3: Summary of results for all fit-types. The colour green signifies deviations between 0-2σ, orange signifies deviations between 2σ-3σ and red represents a deviation beyond 3σ.

Fit type	Comm. year	Slope (Events/year^2 /hexagon)	Intercept (Events/year /hexagon)	Breakpoint (Years after '04)	Slope 2 (Events/year^2 /hexagon)
Horizontal	2004	N/A	11.03 ± 0.05	N/A	N/A
	2005		11.48 ± 0.07		
	2006		10.9 ± 0.1		
	2007		11.06 ± 0.06		
Linear	2004	0.09 ± 0.01	10.4 ± 0.1	N/A	N/A
	2005	0.06 ± 0.02	11.1 ± 0.1		
	2006	0.10 ± 0.03	10.2 ± 0.2		
	2007	0.09 ± 0.02	10.6 ± 0.1		
Linear-horizontal	2004	0.23 ± 0.06	9.9 ± 0.2	6.0 ± 0.9	N/A
	2005	0.21 ± 0.05	10.6 ± 0.2	5.000 ± 0.002	
	2006	0.12 ± 0.08	10.4 ± 0.3	5.00 ± 0.01	
	2007	0.14 ± 0.03	10.5 ± 0.2	6.000 ± 0.004	
Linear-linear	2004	0.24 ± 0.04	9.9 ± 0.2	6.000 ± 0.004	-0.01 ± 0.03
	2005	0.22 ± 0.06	10.6 ± 0.2	5.000 ± 0.004	-0.01 ± 0.03
	2006	-0.02 ± 0.07	10.7 ± 0.3	7 ± 1	0.4 ± 0.2
	2007	0.14 ± 0.04	10.5 ± 0.2	6.00 ± 0.03	-0.01 ± 0.06

Table 5.4: Summary of resulting fit parameters for all fit-types. The uncertainties in the breakpoints are implicitly constrained to an accuracy of 1 or more years, since the data points have a yearly binning. For example, the uncertainty in the breakpoint for the linear-linear fit on rates of stations commissioned in 2007 is 1 year, while for all others, the uncertainty is zero.

Fit type	Commission year	Normalised rate (Events/year /hexagon)	Slope (Events/year^2 /hexagon)	Change/year (%)
Linear	2004	10.70 ± 0.05	0.09 ± 0.01	0.84 ± 0.09
	2005	11.13 ± 0.07	0.06 ± 0.02	0.5 ± 0.2
	2006	10.5 ± 0.1	0.10 ± 0.03	1.0 ± 0.3
	2007	10.73 ± 0.06	0.08 ± 0.02	0.8 ± 0.2

Table 5.5: The rate changes per year (%) obtained from the linear fits. The normalisation is obtained from the y-intercept parameters in the horizontal line fits.

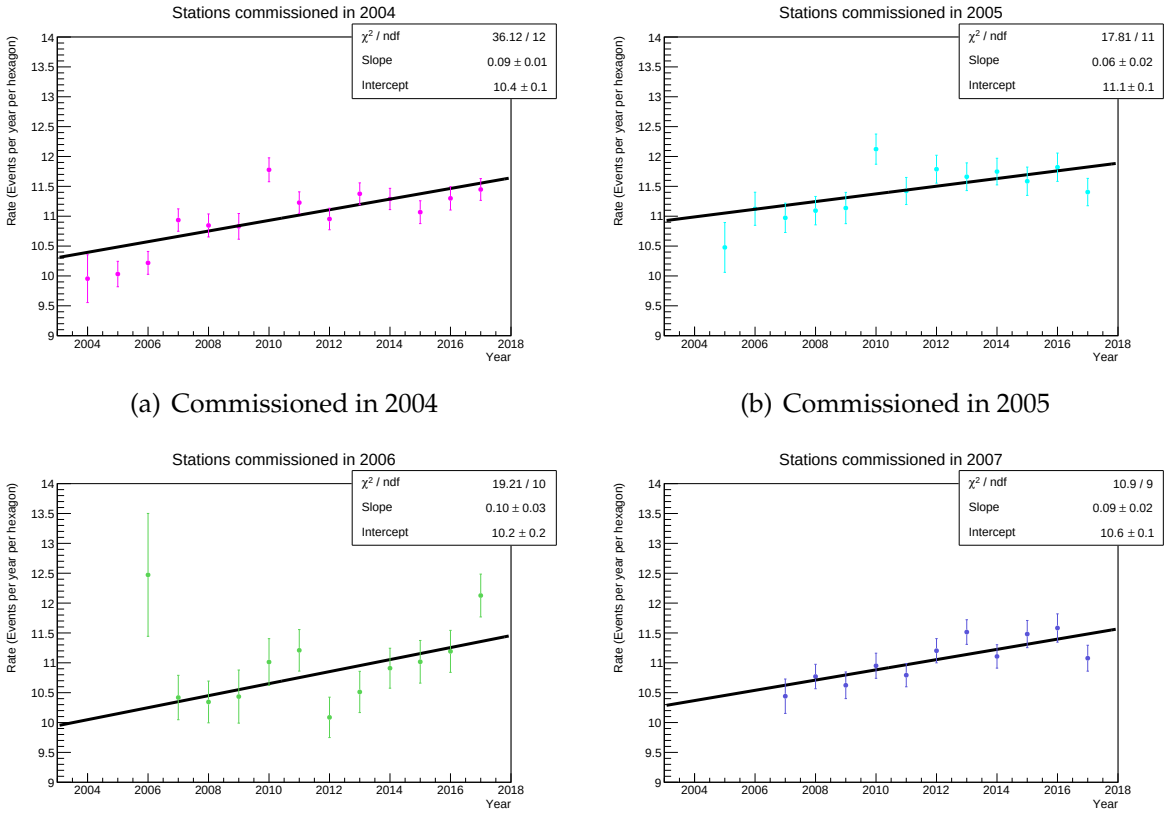
where *y(t)* is the rate as function of year, m is the freely-varying slope parameter, t is time in years, t_{com} is the year of commission of stations, and c is the freely-varying y-intercept. Figure 5.47 shows the linear fits for all cohorts. From Table 5.3 it is shown that the 2004 cohort has a probability which has a deviation of more than 3σ, deeming the linear fit once again unlikely for the varying rates. On the other hand, the other cohorts have linear fits with deviations lying between 0-2σ which is a clear statistical improvement from the horizontal line fit. It can be seen that a linear fit is more likely than a horizontal trend for stations commissioned in 2005, 2006 and 2007. The low probability in the linear fit for the 2004 cohort indicates a presence of a more complex structure in the trend. In terms of the slope parameters shown in Table 5.4, all cohorts have positive slope. Above all, the rates of stations commissioned in 2006 have the biggest slope. The percentage in the change per year of rates for each cohort were calculated and listed in Table 5.5. It is noted that the stations commissioned in 2005 and 2007 have a change per year which lies within the uncertainty of the change per year in rates obtained by Lhenry-Yvon [135] in Figure 5.28. Another notable feature is the spike in rates in the year 2010, for stations commissioned in 2004 and 2005. The highest rate value is measured in this year. This is similar to the rate spike in 2010 shown in Figure 5.28.

5.5.3 Linear-horizontal Fits

Another fit that was applied to the rates of these four station cohorts is the linear-horizontal fit, where a freely-varying breakpoint separates a linear and horizontal line. The equation used in this fit is shown as

$$y(t) = \begin{cases} m(t - t_{com}) + c & t - t_{com} < b \\ mb + c & t - t_{com} \geq b \end{cases} \tag{5.27}$$

where *y(t)* is the rate as a function of time, m is the slope parameter in units of events/year^2/hexagon, t is time in years, t_{com} is the year of commission of stations, c is the y-intercept parameter and b is the breakpoint parameter as the number of years after the year of commission. Figure 5.48 shows the fit results



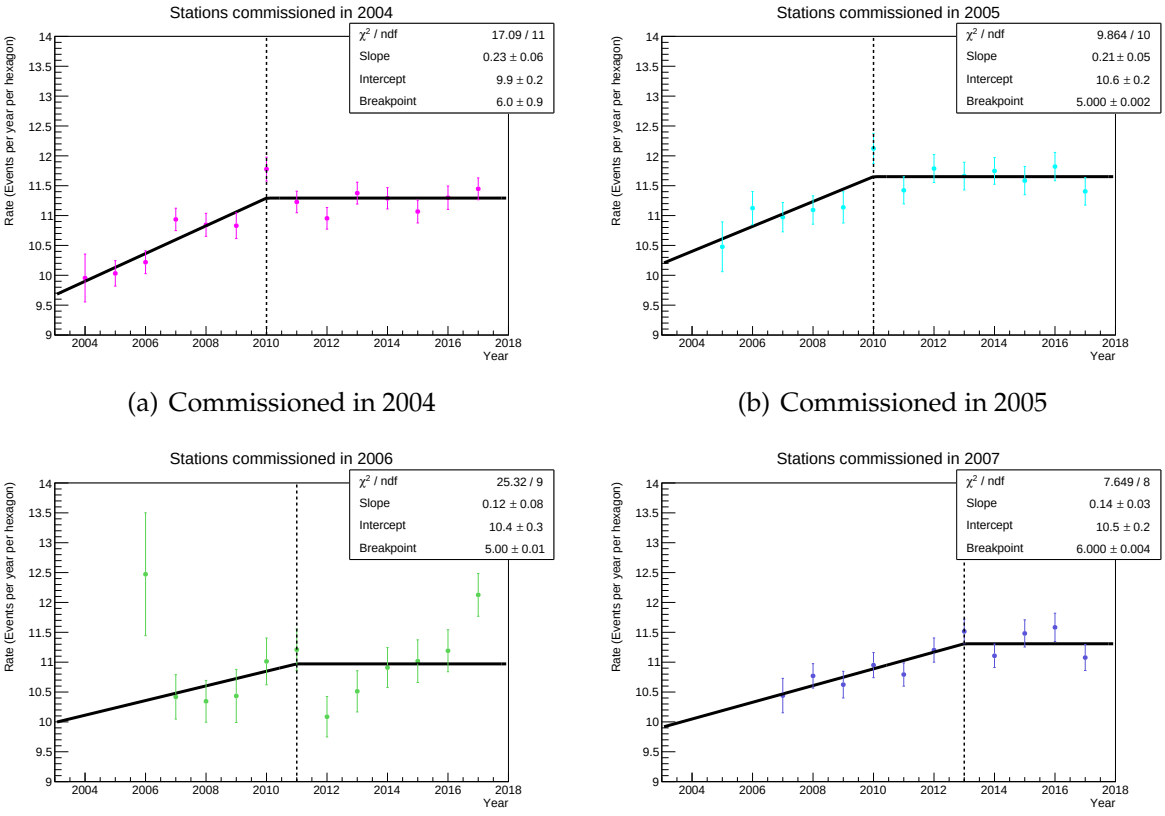
(a) Commissioned in 2004

(b) Commissioned in 2005

(c) Commissioned in 2006

(d) Commissioned in 2007

Figure 5.47: Stations commissioned in 2004, 2005, 2006 and 2007 with a linear fit.



(a) Commissioned in 2004

(b) Commissioned in 2005

(c) Commissioned in 2006

(d) Commissioned in 2007

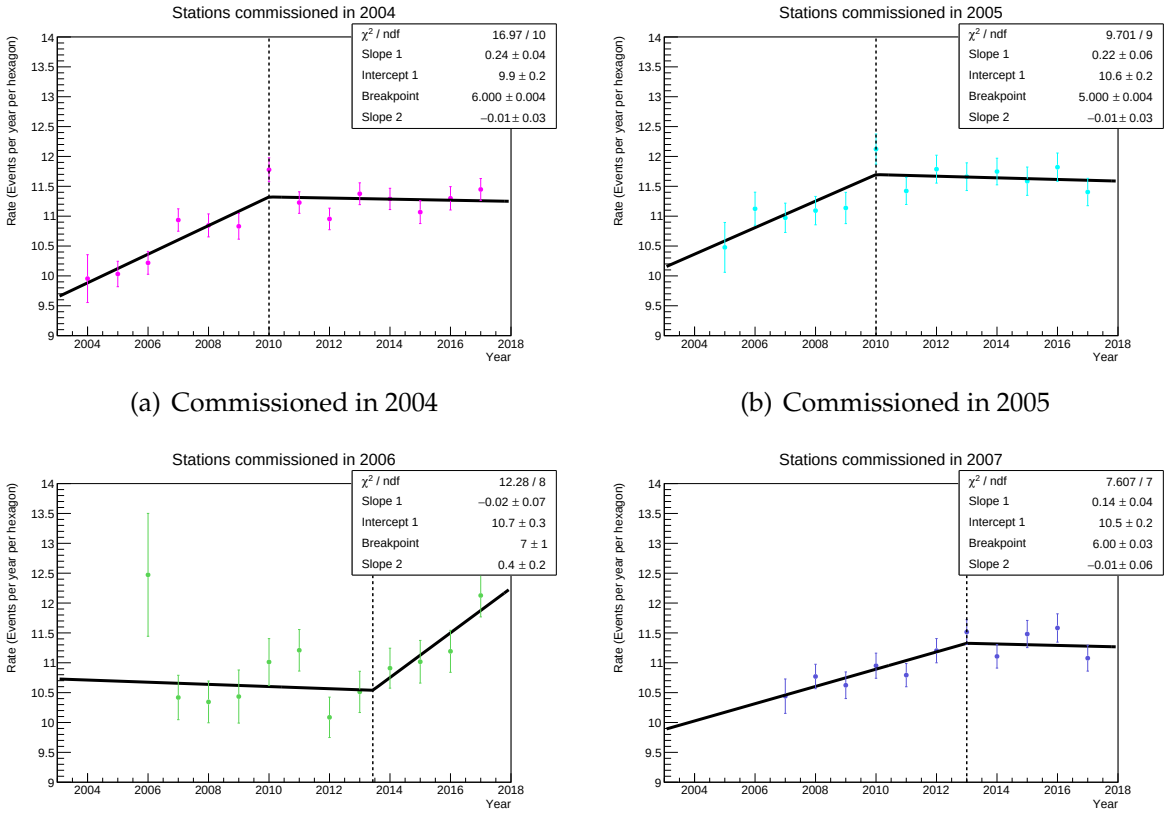
Figure 5.48: Stations commissioned in 2004, 2005, 2006 and 2007 with a linear-horizontal line fit.

while the goodness-of-fit values and parameters were calculated and listed in Tables 5.3 and 5.4. There is a statistically significant improvement in the reduced χ^2 for the 2004 station cohort from the linear fit to the linear-horizontal fit. This is in contrast to the fit for the 2006 cohort. It can be said that the behaviour of the rates of the 2006 station cohort is very unlikely to be possessing a linear-horizontal behaviour, shown prominently by the other three cohorts. A notable feature is that the breakpoint for both the 2004 and 2005 cohort lies in 2010, which is where the spike in rates occur. For the 2007 cohort, the breakpoint lies in 2013, probably due to the absence of the sudden spike in rates in 2010.

5.5.4 Linear-linear Fits

Finally, two linear lines separated by a freely varying breakpoint were fitted with the following functional form:

$$y(t) = \begin{cases} m_1(t - t_{com}) + c_1 & t - t_{com} < b \\ m_2(t - t_{com}) + c_2 & t - t_{com} \geq b \end{cases} \tag{5.28}$$



(a) Commissioned in 2004

(b) Commissioned in 2005

(c) Commissioned in 2006

(d) Commissioned in 2007

Figure 5.49: Stations commissioned in 2004, 2005, 2006 and 2007 with two fits separated by a freely varying breakpoint. Each region has one freely-varying slope to minimise the χ^2 value in the overall fit. It can be seen that the 2006 cohort shows a trend which differs from the rest, where there is a negative then highly positive slope.

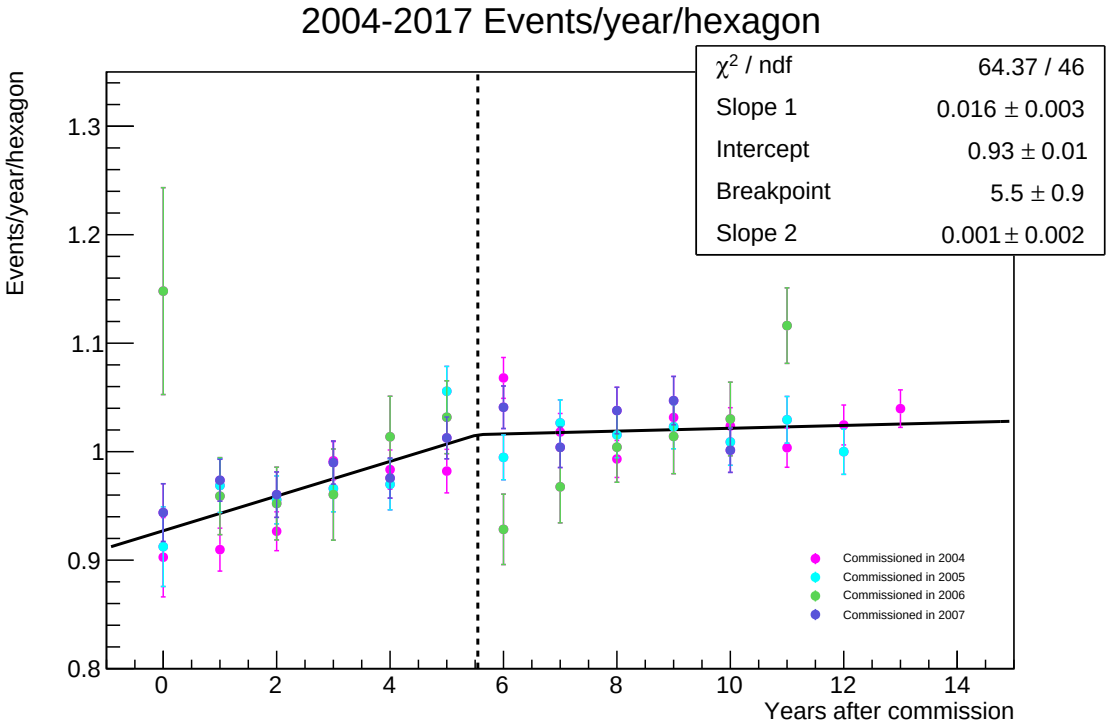


Figure 5.50: Normalised rates as a function of time with an applied linear-linear fit for stations commissioned in 2004, 2005, 2006 and 2007.

where m_1 and m_2 are two freely-varying slopes while c_1 and c_2 are two y-intercepts. Figure 5.49 shows the linear-linear fits for all four station cohorts. Table 5.3 shows that there is a statistically significant improvement in the 2006 cohort from the linear-horizontal to the linear-linear fit, making the probabilities for all the fits have deviations within 0-2σ. It seems that a linear-linear fit is more suited to the behaviour of the rates of all station cohorts. For the 2006 cohort, the breakpoint now lies in mid-2013 as shown in Figure 5.49(c) , while the breakpoint in 2007 remains in 2013 just like in the linear-horizontal fit. Also, the 2006 station cohort is the only one with a slope that changes from negative to positive after the breakpoint occurs. It is also noted that since slope 2 of the 2006 cohort is 40 times larger than the slopes of the others, it could be responsible in making the slope of the rates of all stations after the breakpoint shown in Figure 5.27(d) to be slightly positive.

5.5.5 Overall Behaviour of Rates of All Four Station Cohorts

Finally, it was checked whether the rates as seen by stations commissioned in 2004, 2005, 2006 and 2007 exhibit the same behaviour with a shift in time. The rates as a function of time for these four cohorts were normalised and plotted in Figure 5.50. A linear-linear fit was applied to this to see if this behaviour agrees with the rates for all four cohorts. The rates were normalised by using the

weighted mean; calculated in the horizontal fits applied to the plots in Section 5.5.1. The calculation of the normalised points for each cohort are shown below as

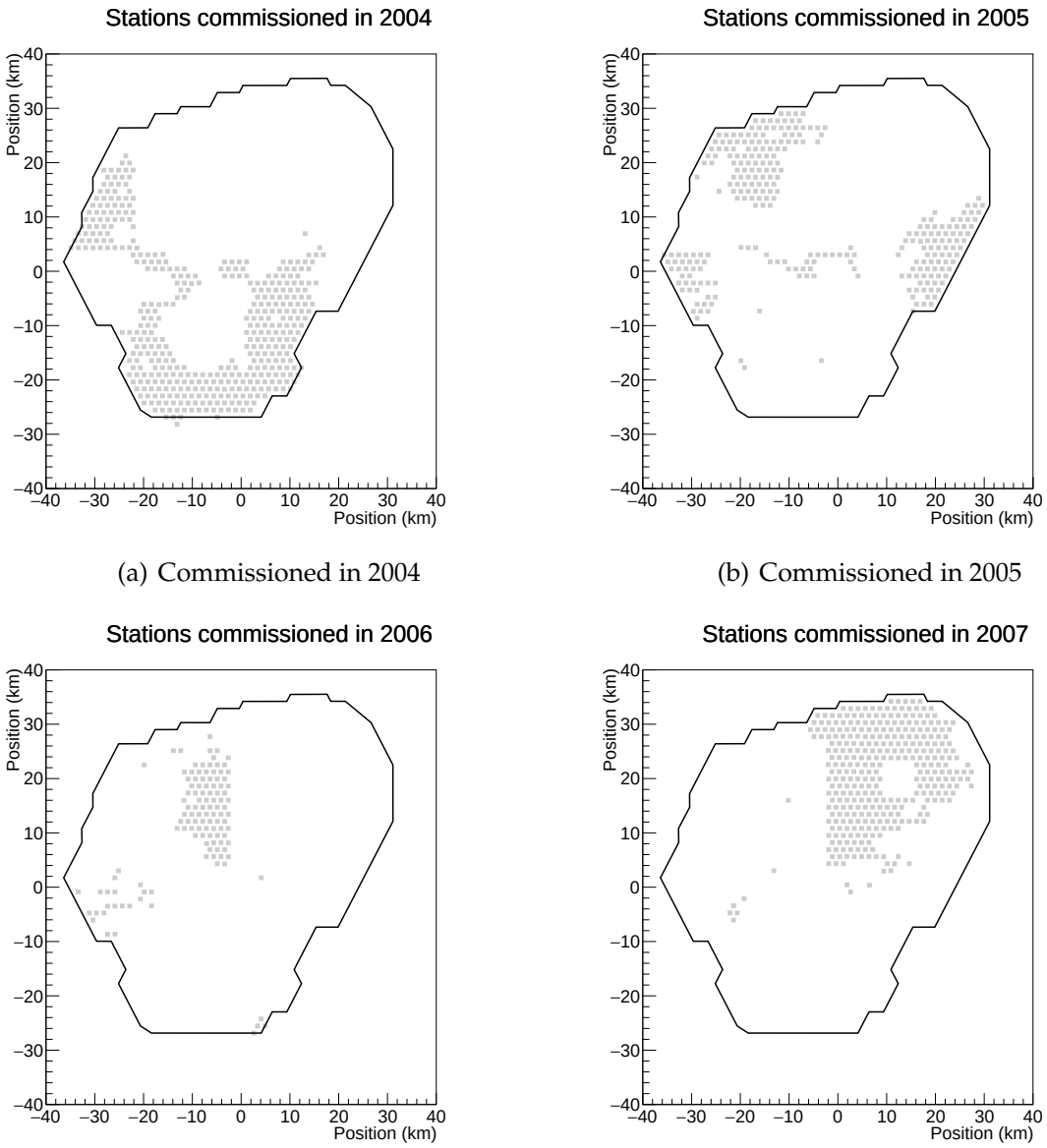
$$\text{Rate for each year} = \frac{x_i}{x_\mu} \tag{5.29}$$

where x_i is the number of events divided by the uptime for each year and x_μ is the weighted mean, which is the y-intercept parameter fitting for a constant rate, obtained in Table 5.4. The uncertainties of these points are calculated as

$$\sigma_{\text{Rate for each year}} = \frac{x_i}{x_\mu} \times \sqrt{\frac{\sigma_i^2}{x_i^2} + \frac{\sigma_{x_\mu}^2}{x_\mu^2}} \tag{5.30}$$

where σ_i is the uncertainty in the rate for each year known as the square root of the number of events divided by the uptime, and σ_{x_μ} is the uncertainty in the y-intercept (constant rate) parameter shown in Table 5.4. The fit probability obtained in Figure 5.50 is 1.3%. This lies in a deviation range between 2-3σ. From this, it is not unlikely that the rates have a linear-linear behaviour. The breakpoint is shown to lie 5.5 years from the respective years of commission for each cohort and the slope after this breakpoint is lower in magnitude compared to the slope before the breakpoint. This is similar to the behaviour shown in Figure 5.27(d), indicating that the rates of stations for each of these four cohorts have rates which could be increasing drastically until it reaches its highest average rate, then increases again but more slowly. This is a strong indication that the PMTs of all stations are aging in a similar manner, despite the minute-by-minute calibration process using charge and amplitude histograms generated by muons in air showers.

Then, the rates of stations as a function of position across the array are compared with the positions of stations commissioned in different years. Figure 5.51 show four maps which depict the positions of stations commissioned in 2004, 2005, 2006 and 2007. The positions of these stations on the array were then compared with the respective rates plotted in Section 5.3. Out of all the rate maps, there seems to be only two examples of any obvious correlation. One of them includes the high rates of stations commissioned in 2006, as some are positioned in the hotspot shown in Figure 5.52(a). This is also the case for stations commissioned in 2007, for the hotspot shown in Figure 5.52(b). It is difficult to say whether the hot/cold spots are caused purely by stations commissioned in certain years.



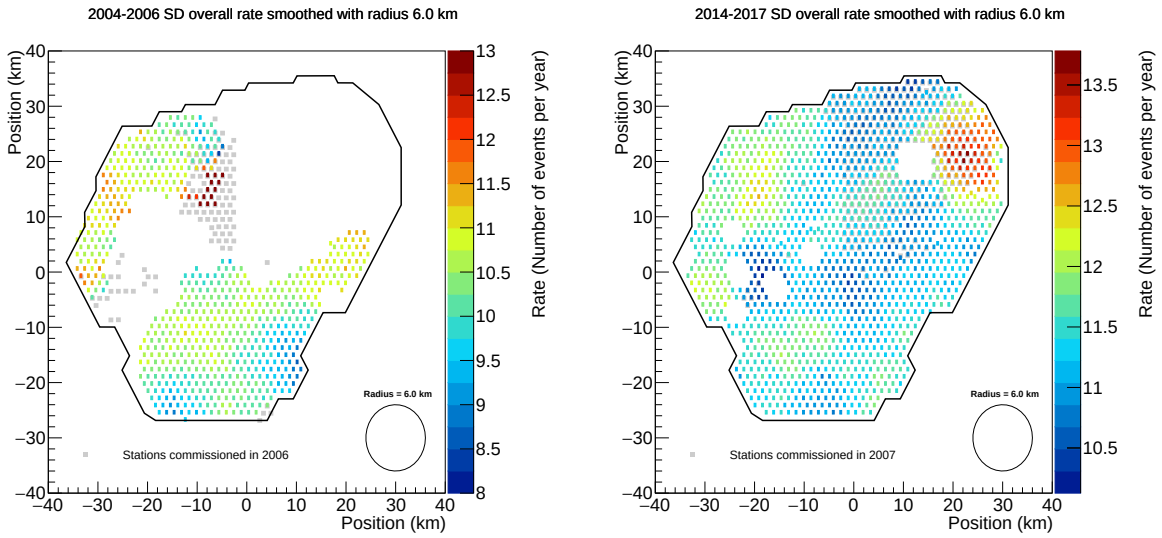
(a) Commissioned in 2004

(b) Commissioned in 2005

(c) Commissioned in 2006

(d) Commissioned in 2007

Figure 5.51: Positions across the array of stations commissioned in 2004, 2005, 2006 and 2007.



(a) (b)

Figure 5.52: Map of rates from 2004-2006 with (a) stations commissioned in 2006 and (b) with stations commissioned in 2007 both plotted in grey.

5.6 Possible Causes of the Increasing SD Event Rates

One hypothesis for the cause of the increasing rates is that the SD energy resolution is getting poorer with time. Since it is known that the energy spectrum is very steep, there are many more low energy events than there are high energy events. A worsening SD energy resolution can cause more low energy events to be reconstructed as higher energy events (above full efficiency threshold) [140]. To investigate this, a study was done to see how much the SD energy resolution needs to worsen to obtain the increasing rate observed. The event rate is obtained as the integral of the cosmic-ray flux above the full efficiency threshold. So, in order to obtain the rate, the functional form of the cosmic-ray flux must be obtained. The cosmic-ray flux in the energy spectrum is described by a function

Parameter	Value $\pm \sigma_{\text{stat.}} \pm \sigma_{\text{sys.}}$
J_0 [km^{-2} sr^{-1} yr^{-1} eV^{-1}]	$(1.315 \pm 0.004 \pm 0.400) \times 10^{-18}$
γ_1	$3.29 \pm 0.002 \pm 0.10$
γ_2	$2.51 \pm 0.03 \pm 0.05$
γ_3	$3.05 \pm 0.05 \pm 0.10$
γ_4	$5.1 \pm 0.3 \pm 0.1$
E_{12} [eV] (ankle)	$(5.0 \pm 0.1 \pm 0.8) \times 10^{18}$
E_{23} [eV]	$(13 \pm 1 \pm 2) \times 10^{18}$
E_{34} [eV] (suppression)	$(46 \pm 3 \pm 6) \times 10^{18}$

Table 5.6: Energy spectrum best-fit parameters with statistical and systematic uncertainties for completeness [141].

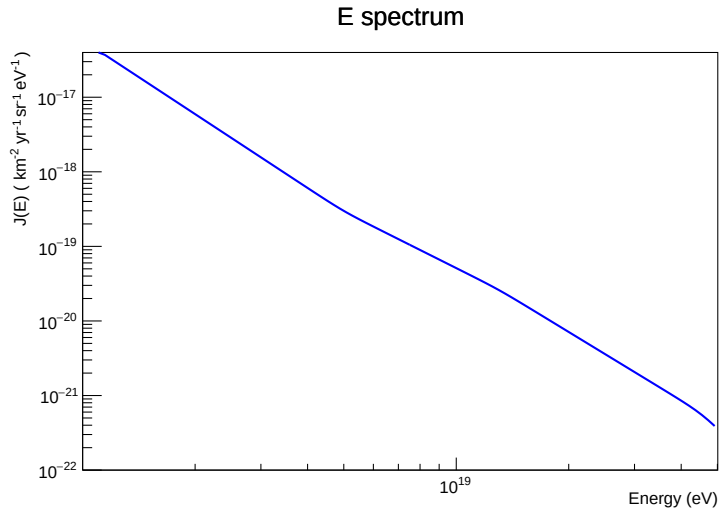


Figure 5.53: Energy spectrum as calculated from Equation 5.31 and the parameters in Table 5.6.

shown as

$$J(E) = J_0 \left(\frac{E}{E_0}\right)^{\gamma_1} \prod_{i=1}^{3} \left[1 + \left(\frac{E}{E_{ij}}\right)^{\frac{1}{\omega_{ij}}}\right]^{(\gamma_i - \gamma_j)\omega_{ij}} \tag{5.31}$$

where E_0 is fixed at $10^{18.5}$ eV, ω_{ij} is fixed at 0.05 for all values of i and j and the other best-fit parameters for the energy spectrum in this equation are shown in Table 5.6 [141]. The role of the indices i and j is to control the steepness of the spectrum at different energy ranges. This spectrum is plotted in Figure 5.53. It is the final result obtained after removing the effects of the SD energy resolution, and it is known as the *true* or *corrected* spectrum produced by Nature. To investigate the effects of a changing SD energy resolution upon observed event rates, this spectrum is convolved with a probability distribution function (in this case; a Gaussian) with a width determined by the SD energy resolution function, to obtain a 'forward-folded' spectrum. This is called the *raw* or *observed* spectrum which has distortions in its shape caused by the finite energy resolution [141]. The SD energy resolution is shown as

$$\frac{\sigma_{\mathrm{SD}}(E)}{E} = \sigma_0 + \sigma_1 \exp(-\frac{E}{E_\sigma}) \tag{5.32}$$

where the values of the parameters are obtained as: σ_0 = 0.078, σ_1 = 0.16 and E_σ = 6.6 $\times$ 10^{18} eV [141]. This function is plotted in Figure 5.54. It can be seen that the resolution goes from a maximum of $\sim$ 22 % to a minimum of $\sim$ 8 %. In the numerical convolution process, the SD energy resolution represents the width of the Gaussian convolved with the energy spectrum. So, depending on the energy

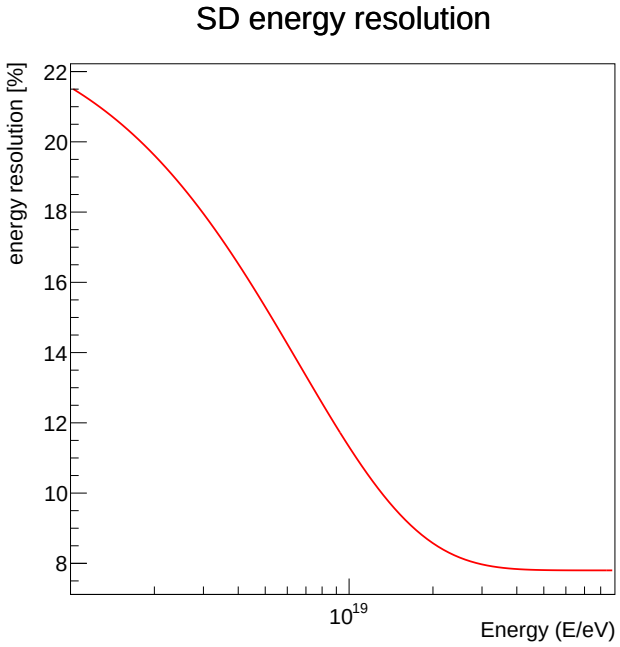


Figure 5.54: SD energy resolution as calculated from Equation 5.32.

bin, the width of the Gaussian changes according to the SD energy resolution function.

Another factor which must be taken into account is a bias in the energy reconstruction, below the full efficiency threshold. This bias arises due to the SD having a higher efficiency of shower detection for heavier primary nuclei [141]. Hence, the distribution of S_{38} below the full efficiency threshold may no longer be fairly averaged over the underlying mass distribution. This may result in a bias on E_{SD} as extrapolated from the calibration procedure and trigger effects [141]. From this, positive fluctuations in S_{38} and hence, in E_{SD} are favoured over negative ones. The bias is encapsulated in an equation shown as

$$b_{\mathrm{SD}}(E,\theta) = (b_0 + b_1 \exp(-\lambda_b(\cos\theta - 0.5)))\log_{10}\left(\frac{E_*}{E}\right) \tag{5.33}$$

for $\log_{10}(\mathrm{E/eV}) \leq \log_{10}(\mathrm{E_*/eV}) = 18.4$ and $b_{SD} = 0$ otherwise [141]. The parameters are obtained to be $b_0 = 0.20$, $b_1 = 0.59$ and $\lambda_b = 10.0$ [141]. The variable θ is the zenith angle. Figure 5.55 shows the relative bias values of the SD as a function of $\cos\theta$ for four different energy ranges. It can be seen that as the energy approaches the full energy efficiency threshold, the bias tends towards zero.

It is also noted that the shape of the spectrum is affected by the detection efficiency curve in the range where the array is unable to detect events at 100 % efficiency, i.e. there exists a dependence on the nature of the primary particle, energy or arrival direction. The efficiency is represented by an error function (erf)

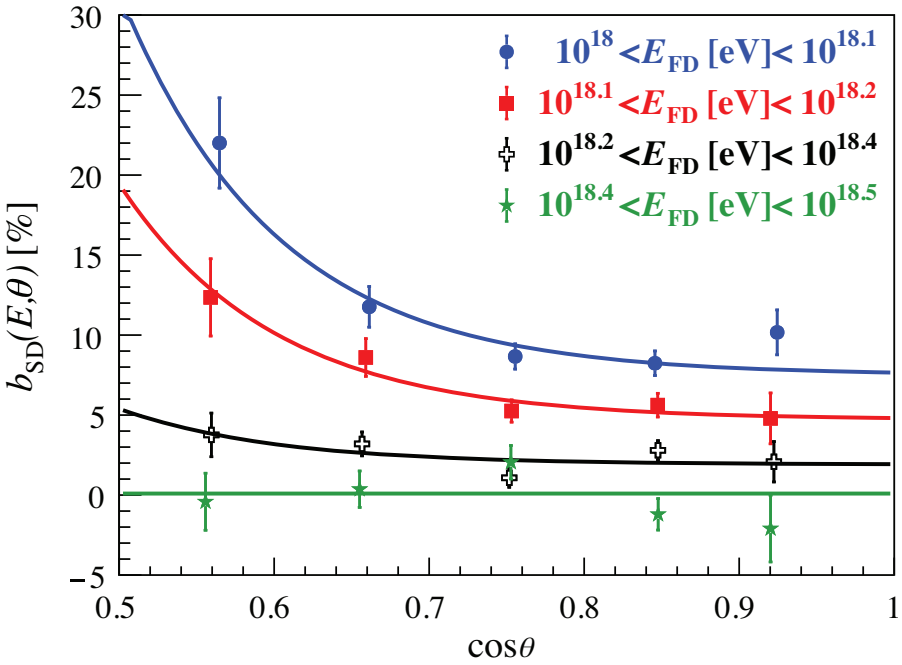


Figure 5.55: The bias (%) of the SD as a function of the cosine of the zenith angle for four different energy ranges, according to Equation 5.33 [141].

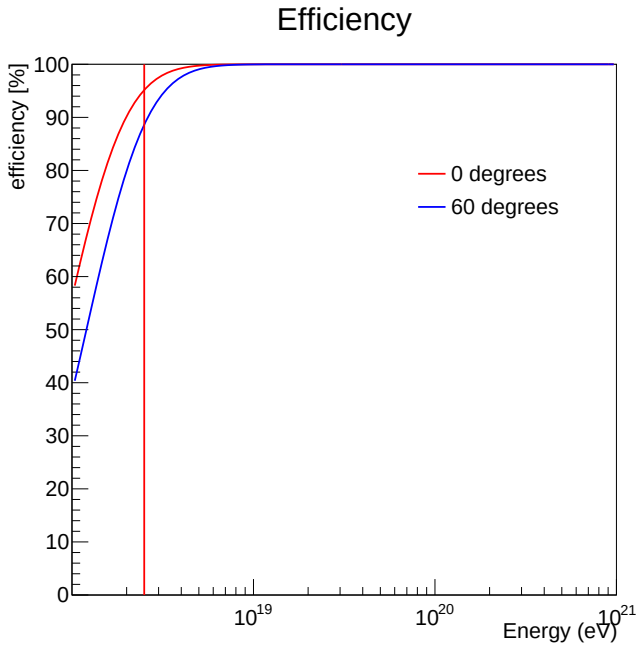


Figure 5.56: The efficiency function at two zenith angles; 0° and 60° as described by Equation 5.34. The vertical red line marks an energy of 2.5 EeV.

shown as

$$\epsilon(E,\theta) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\log_{10}(E/eV) - p_0(\theta)}{p_1}\right)\right] \tag{5.34}$$

where θ is the zenith angle of the shower, p_1 = 0.373 and $p_0(\theta)$ = 18.63 - 3.18 $\cos^2\theta$ + 4.38 $\cos^2\theta$ -1.87 $\cos^6\theta$. Figure 5.56 shows a plot of the efficiency at two zenith angle extremes; at 0° and 60°. It can be seen that above 2.5 EeV, the efficiency is over 85 %.

To perform the forward folding process, a numerical convolution between a Gaussian and the corrected spectrum is performed. This is done by dividing the energy range into a finite, but sufficiently large number of slices, then drawing a Gaussian in each bin. It is ensured that the bin slices are much smaller than the width of the Gaussian, so that the forward-folding process is performed correctly. The Gaussian is a probability distribution function, incorporating the SD energy resolution as the width of the Gaussian and the bias function. The functional form of this Gaussian is shown as

$$\kappa(E_{\text{SD}}|E;\theta) = \frac{1}{\sqrt{2\pi}\sigma_{\text{SD}}(E,\theta)} \times \exp\left[-\frac{(E_{\text{SD}} - E(1 + b_{\text{SD}}(E,\theta)))^2}{2\sigma^2_{\text{SD}}(E,\theta)}\right] \tag{5.35}$$

where E is the true energy (the middle of each bin) and E_{SD} is the measured energy. Due to the non-zero bias below $10^{18.4}$ eV, the Gaussian is centred at a value of $E(1+b_{SD}(E,\theta))$. To mitigate the zenith angle dependence, it is first noted that the number of arriving cosmic rays is distributed as

$$\frac{dN}{d\theta} \propto \cos\theta \sin\theta \tag{5.36}$$

where the cosine term accounts for the decreasing projected area of the array with increasing zenith angle and the sine term describes the differential spherical segment due to the changing solid angle [106]. The zenith angle, θ is randomly sampled from a $\sin 2\theta$ distribution. This originates from the trigonometric identity showing that $\sin 2\theta = 2\sin\theta\cos\theta$ and the factor of 2 is included as a normalisation factor, ensuring that this distribution maximises at a value of 1. After a random angle is obtained, it is substituted into the equation for bias and efficiency. Then, a fixed number of shower energies are randomly sampled from the Gaussian at every bin, which have tails that overlap neighbouring bins creating bin-to-bin event migrations. All of the random sampling performed here is done using the random number generator in the C++ BOOST library. These new shower energies are the *measured* energies, E_{SD}. These values are used to fill a set number of bins in a new histogram. As the histogram is filled, the contents of each bin are scaled by the efficiency.

In this work, 10,000 shower energies are randomly sampled and used to fill a new histogram with 6000 bins, ranging from 10^{18} eV to 10^{21} eV to ensure no edge effects are seen especially at the full efficiency threshold. The energy range is divided equally in the linear scale, resulting in a bin width of $\sim 1.67 \times 10^{17}$ eV.

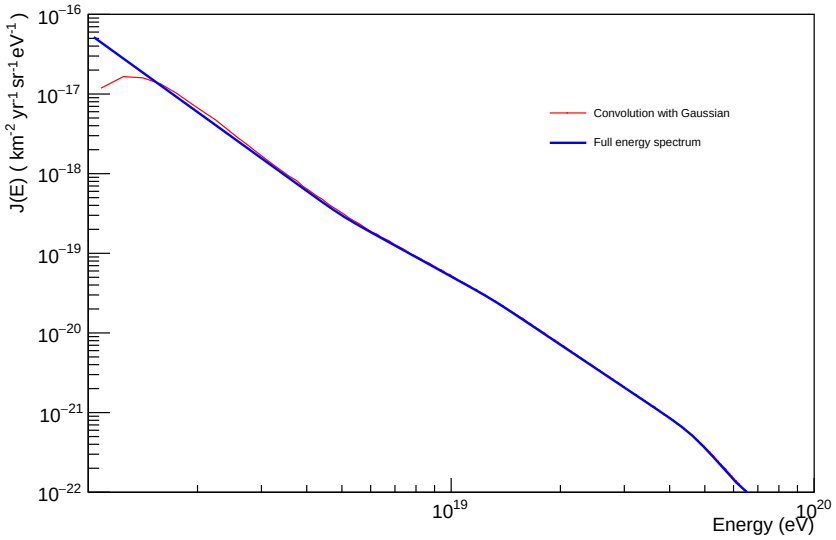


Figure 5.57: The corrected spectrum in blue and the forward-folded, observed spectrum in red as a result of a numerical convolution between a Gaussian with a width of the SD energy resolution function, shifted by a bias and with the efficiency taken into account.

This is smaller than the width of the Gaussian at the lowest energy range, which is $\sim 2.2 \times 10^{17}$ eV (22 % of 10^{18} eV), indicating a sufficient number of bins for the convolution process to be done correctly. The resulting forward-folded spectrum, with all of these factors taken into account is shown in Figure 5.57.

A cross-check was done to ensure that the observed spectrum obtained is correct. The ratio of the *true* spectrum to the *measured* spectrum, known as the *correction* coefficients was calculated. This curve was compared with published work shown in Figure 5.58. It can be seen that the correction factor from this work is consistent in terms of shape and normalisation with the published work shown indicating that the observed spectrum obtained is correct. This is emphasized in Figure 5.59.

Following on from this is the calculation of event rate above the full efficiency threshold. As mentioned, this is done as the integral of the flux above 3 EeV. This was done for both the corrected and observed spectrum in Figure 5.57. The value obtained directly from the integration process for the corrected spectrum is 2.32 km^{-2} sr^{-1} yr^{-1}. For the observed spectrum, the value is 2.44 km^{-2} sr^{-1} yr^{-1}. In order to obtain the event rate which is the number of events per year, this value must be multiplied by the aperture, which is obtained as the effective area under zenith angle θ, integrated over the solid angle Ω within which the showers are observed [141]. The aperture for each station was found to be 4.59 km^2 sr [141]. Therefore, the measured rate is effectively 2.44 events km^{-2} sr^{-1} yr^{-1} $\times$ 4.59 km^2 sr $\simeq$ 11.2 events yr^{-1}. This matches with the average rate obtained in this work shown in Section 5.4.1 which is 11.19 $\pm$ 0.03 events per year.

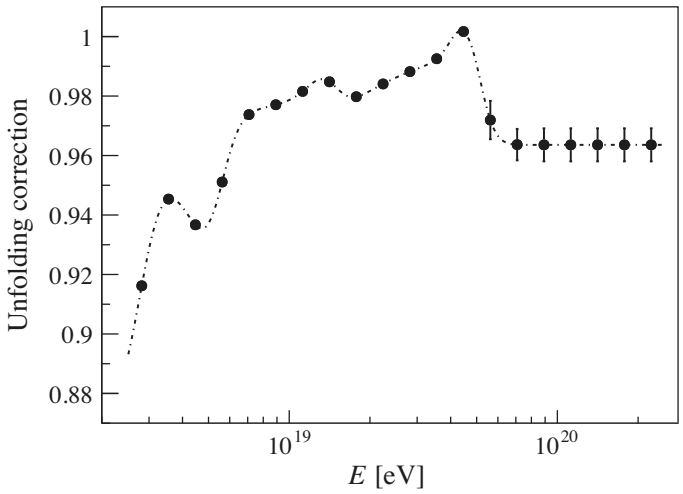


Figure 5.58: Unfolding correction factor applied to the observed spectrum to account for the detector effects, as a function of cosmic-ray energy [141].

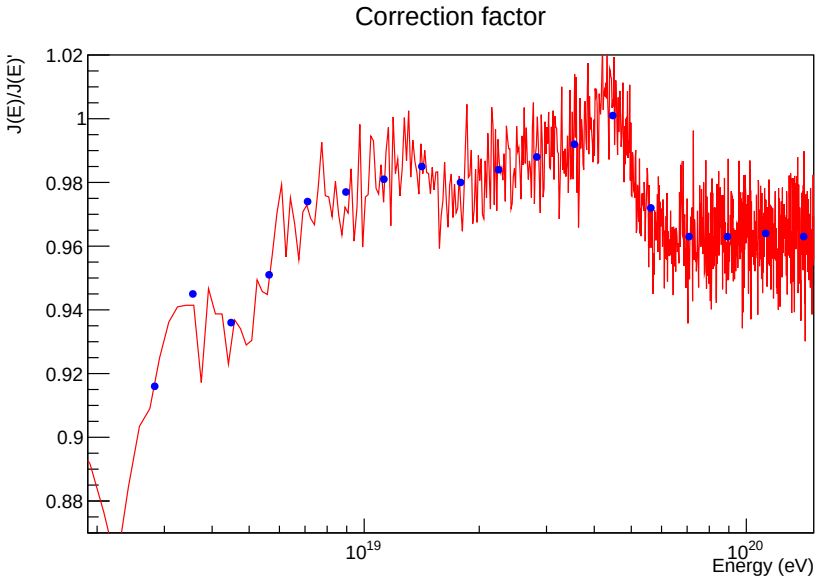


Figure 5.59: Unfolding correction factor obtained in this work (in red) and the correction factor from a published work as shown in Figure 5.58 [141] (in blue).

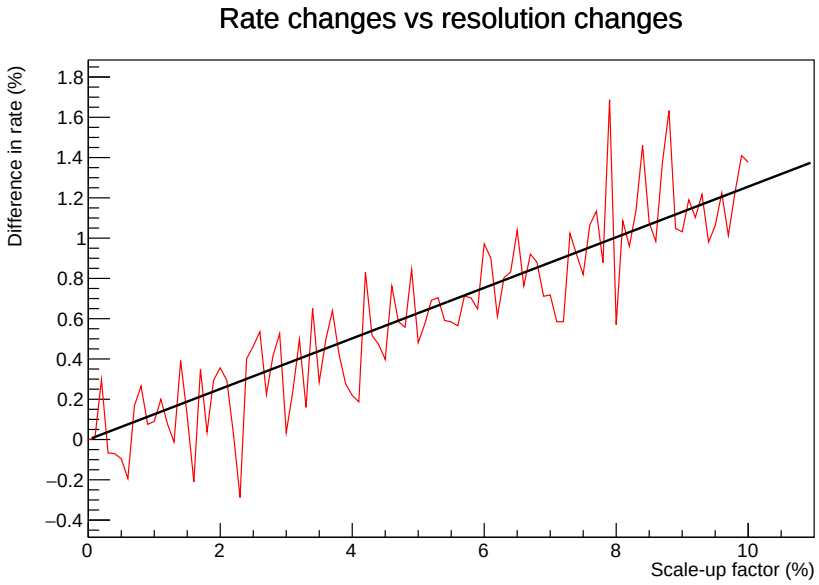


Figure 5.60: Rate changes (%) as a function of the multiplicative factor in the SD energy resolution function (%). The black line represents a fit to the data points obtained from the simulation, with a freely-varying slope. The slope obtained is 0.13.

As stated in the beginning of this section, the suggested hypothesis to explain the increasing rates is the worsening SD energy resolution with time. As obtained from Section 5.4.1, the rates from 2006 onwards are increasing by 0.58 $\pm$ 0.08 % every year. It was checked to see how much the SD energy resolution had to worsen, to obtain an increase in rate of 0.58 %, by repeating the numerical convolution process but with a multiplicative factor in front of the SD energy resolution function shown in Equation 5.32. As the factor is changed, the measured rate changes. This rate change can be calculated as a percentage. Firstly, it is noted that there is a small difference in the event rates between the corrected and observed spectrum shown as

$$\text{Base difference in rate (\%)} = \frac{2.44 - 2.32}{2.32} \times 100\,\% = 5.17\,\% \tag{5.37}$$

As the SD energy resolution worsens by a few percent, the observed rate increases. This is because if a shower was measured to have a higher energy than the true energy, the effect would be obvious due to an extremely steep spectrum, inducing a slightly higher normalisation and hence, a slightly higher rate. This effect becomes more and more prominent as the SD energy resolution worsens. In this analysis, the multiplicative factor was increased up to 10 % with 0.1 % increments resulting in 100 simulations. Throughout this process, the change in rate was recorded. A plot of the rate change (%) as a function of this multiplicative factor is generated and shown in Figure 5.60, where the base difference in rate calculated in Equation 5.37 is calibrated as 0 % on the plot. A linear fit was done,

with a resulting slope of 0.13. It is concluded that for a change in rate of ~ 0.58 % per year, the SD energy resolution must worsen by ~ 4.5 % every year, which is unreasonably large. This suggests that the worsening of the SD energy resolution is most likely not a large contribution to the increasing rates.

It was also found from current work done by Ioana Mariş that there could be non-triggered stations further away from the shower axis, causing the reduction in accuracy of the lateral distribution function (LDF) of a given shower [142]. These non-triggering stations are described as stations which are set as 'active' in the T2 file information, have distances to the shower axis that are less than a few hundred metres, but obtain a zero signal, despite having a higher expected signal higher than the signal for full efficiency. An inaccurate LDF affects the S(1000) signal, worsening its resolution, which could be the cause of the increasing rates above 3 EeV. Further investigation entails an analysis on the number of non-triggering stations as a function of time, as it could be correlated with the increasing rates.

Another possible cause for the increasing rates as seen by the SD is that the station calibration histograms using muons (every minute) are not keeping track of the drift in sensitivity of the stations. Since the histograms are really only tracking the muon response, it is possible that the electromagnetic component is not tracked just as well. In addition, the Tyvek liner reflectivity of the stations largely affects the number of reflections of Cherenkov photons produced from the particles travelling through the water. It was mentioned in Section 5.4.1 that the Tyvek liner reflectivity is attributed to the decreasing sensitivity of the stations, causing a decay in the area over peak. A preliminary analysis was done by Zapparrata where 1000 air showers were simulated with two extreme liner reflectivity values (0.962 and 0.922). The aim was to simulate both the signal response to air showers and calibration procedure with changing reflectivity. It was found that the difference in S(1000) obtained (which determines E_{SD}), between the two liner reflectivity values is less than 1.5 % which is not substantial considering the large difference in reflectivity over 14 years [143]. Current work is still being done to refine the results as the simulations were performed using FastTankSimulatorOG which is a module in $\overline{\text{Off}}\underline{\text{line}}$ that is less complete than G4TankSimulatorOG available in GEANT4 (a toolkit for the simulation of particles going through a detector) [144].

5.7 Conclusions

An analysis of the rates recorded by the SD as a function of position and time at the Pierre Auger Observatory has been presented using SD data and Hexalife files, over a span of 14 years. Firstly, the rates as a function of position were observed, particularly the rates of all stations detecting events from 2004-2013. It can be seen that in this time period there are several hotspots towards the border of the array while the cold spot spans from the north to the south particularly in

the inner region of the array. It was checked whether the hot and cold spots are the result of random fluctuations using Monte Carlo analysis. Consequently, it was shown that these hot or cold spots are mostly due to systematic effects of the stations. Then, the rates of stations were observed for separate time periods; for 2004-2006, 2007-2009, 2010-2013 and 2014-2017. The lowest rates were shown in 2004-2006. Moving forward through the time periods, the average rates tend to increase. The highest average rate was shown in the final time period, as well as an interesting hotspot located in the north-west region of the array.

Then, the rates as a function of time were investigated. Several fits were applied to the rate trend and a comparison was done with the work of Lhenry-Yvon [135], Choi [136] and Zapparrata [138]. There seems to be similarity in the behaviour of the rates particularly in the increasing trend. The change per year obtained in this work which is 0.58 ± 0.08 % is within the uncertainty of the change per year obtained by Lhenry-Yvon which is 0.55 ± 0.07 %. Despite this, it was noted that there were discrepancies in how the data was analysed, such as the difference in data reconstruction methods and the timeline of the data, being 2006 to 2019 in comparison with data from 2006 to 2017. Another interesting feature is the spike in rates in 2010 which is very prominent in Lhenry-Yvon's work and also in this work. Monthly binnings were applied to the data to see if there is a similar modulation in rates shown by the work of Lhenry-Yvon. The obtained peak-to-peak value is 2.2 ± 0.7 %; five times smaller than the peak-to-peak obtained in [135] which is 10 ± 0.7 %. However, it was noted that this large discrepancy could be due to the equal exposure binning which was done in the work of Lhenry-Yvon but not in this work.

The trend of rates as a function of time was investigated in more detail by applying a linear-horizontal and linear-linear fit. The probability for the linear-horizontal fit had a deviation between 0-2σ, with a 'settling-down' behaviour after 2010. This is similar to the work shown by Choi as well as Zapparrata in the evolution of the station area over peak values from 2004 to 2019. The area over peak tends to decrease intensely from 2004 to around 2010, then it continues to decrease at a lower rate. It was then found in an analysis done by Lhenry-Yvon that the evolution of rates of events above all energies directly correlates with the area over peak evolution. This indicates a degradation in the sensitivity of the stations which had been attributed to the Tyvek liner reflectivity. It was also found in this work that when a threshold of 3 EeV or higher was applied, the rates are seen to be increasing, which indicates a drift in the energy calibration. Previous work also showed an increasing rate as a function of time and station age. It was shown that this may be partially related to two freezing events which happened in 2007 and 2010.

Following this, rates as a function of time and position were investigated in more detail by generating maps of rates for each year. The rates generally increased from 2004 onwards, but there were some missing stations in the years of 2015, 2016 and 2017, causing holes in the array. This could be due to maintenance

being done on stations or issues with data acquisition. Particularly in 2016, there were some stations located at the north-eastern corner of the array, with rates abnormally higher than other years. The stations with these rates, accompanied with some missing stations with technical issues in data acquisition, are most likely the cause of the hotspot in the 2014-2017 map. The stations then recovered in 2017.

In addition to this, the fractional changes in the threshold energy were calculated. It was noted that the change in rate could be attributed to a change in the energy calibration of the SD. This was compared with the E_{SD}/E_{FD} study done in Chapter 4 as the energy ratio gives an indication of the drifting energy scale of the SD. It was found that the energy ratio overall is increasing as a function of time, indicating an increase in the SD energy. This implies a drifting energy scale for the SD and agrees with the increasing rate as a function of time as seen by the SD.

Lastly, the rates of stations commissioned in four different years were investigated to see if different cohorts show similar evolutions shifted in time. The rates of stations commissioned in 2004, 2005, 2006 and 2007 were plotted and applied with four different fits. The linear-linear fit was shown to be the most likely behaviour for all station cohorts. The rates of all four cohorts were plotted and it was shown that a linear-linear fit gave a reasonable result, with a probability of 1.3 %. The breakpoint is shown to lie 5.5 years from the year of commission, with a steep slope prior and a less steep slope after the breakpoint. It was then checked if there was any correlation that can be seen between the stations commissioned in specific years to the high and low rate regions in the maps of rate as a function position across the array. Maps of the station positions according to cohort were plotted and overlaid with the rate maps. However, only two examples of correlations can be seen, where stations commissioned in 2004 and 2006 were involved in the high rate regions.

Finally, a few possible causes of the increasing rates were detailed. One of which is the worsening SD energy resolution. This was investigated by performing a forward-folding process onto the energy spectrum, to obtain an observed spectrum which has its shape distorted by a given finite SD energy resolution function. A multiplicative factor was applied to the SD energy resolution function to see the extent of its increase needing to satisfy a rise in rate of 0.58 % per year. It was found that the SD energy resolution must increase by 4.5 % each year to contribute to a rate increase of 0.58 % per year. This large result suggests that the worsening of the SD energy resolution could not possibly be a large contributor. Non-triggering stations further from the shower axis during the time of an event may also contribute to a worsening SD energy resolution. Current work is being done to see if the average number of stations involved in detecting events above the full efficiency threshold is constant or not. Another cause is the degradation of the Tyvek liner reflectivity causing a decay in the sensitivity of stations. Current work is being done to simulate response

to air showers and the calibration procedure under two different liner reflectivity values. From preliminary results, it was found that the S(1000) signal differs up to only 1.5 % which is not substantial, given that the liner reflectivity values reflect the changes in a period of 14 years.

Chapter 6

Assigning Hybrid Event Rates to Each Fluorescence Detector Through the Study of SD Event Rates and Cross-checks

This chapter details an assignment of expected hybrid event rates to each of the four FD sites; Los Leones, Los Morados, Loma Amarilla and Coihueco, using SD and hybrid events. An analysis of the SD rates also seen by each FD site as a function of time is done to observe a potential presence of more or less prominent time dependence using hybrid events. The structures observed are then compared with the rate increase of as a function of time obtained in Chapter 5 as well as the energy ratio increase seen from Chapter 4. Fractional changes in the threshold energy as a function of time are then calculated using the event rate and threshold relationship. This is then followed by an analysis of rates as a function of distance between the FD site and the hottest station to investigate the possibility of significant discrepancies due to distance from air showers. The possible structures seen are compared with the structures of rates as a function of position on the array as well. Then, further investigations were done on the high rate observed in the year 2010 as seen in Section 5.4.1 of Chapter 5.

The hybrid data used is the finalised reconstruction of weather-corrected ICRC 2019 data with applied cuts as described in Appendix A, resulting in 8586 events from 2004 to 2017. For every hybrid event detected from each FD site, the hottest station is used to assign a yearly rate from the previous analysis outlined in Section 5.4.2. Additionally, for every event, the distance between a hottest station and an FD site associated with that event was calculated and assigned.

6.1 Introduction

The number of hybrid events for each station between the year 2004 and 2017 was calculated and plotted on Figure 6.1. There seems to be a high number of hybrid events detected in the northern region between Loma Amarilla and Coihueco.

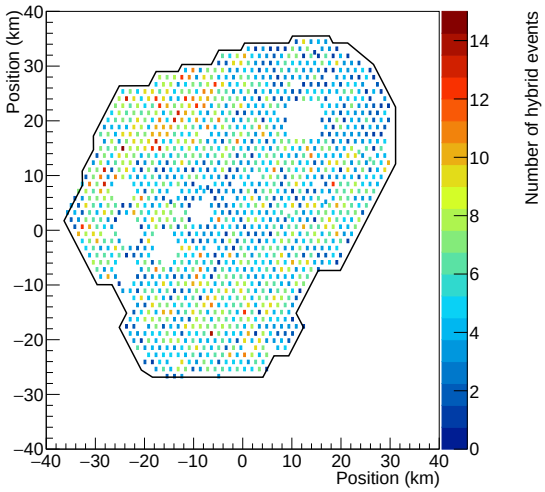


Figure 6.1: The number of hybrid events with energies above 3×10^{18} eV and zenith angle $< 60°$ detected by each station from 2004 to 2017.

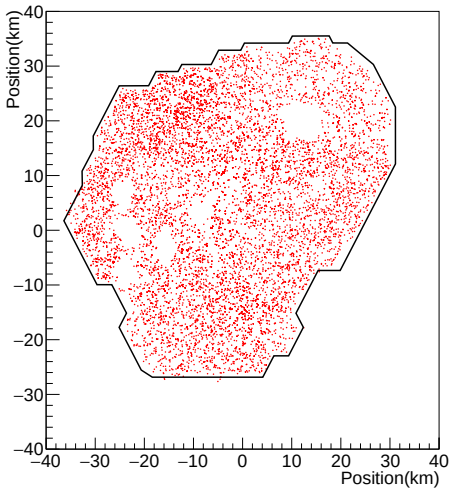


Figure 6.2: Core positions for every hybrid event with energies above 3×10^{18} eV and zenith angle $< 60°$ detected from 2004 to 2017.

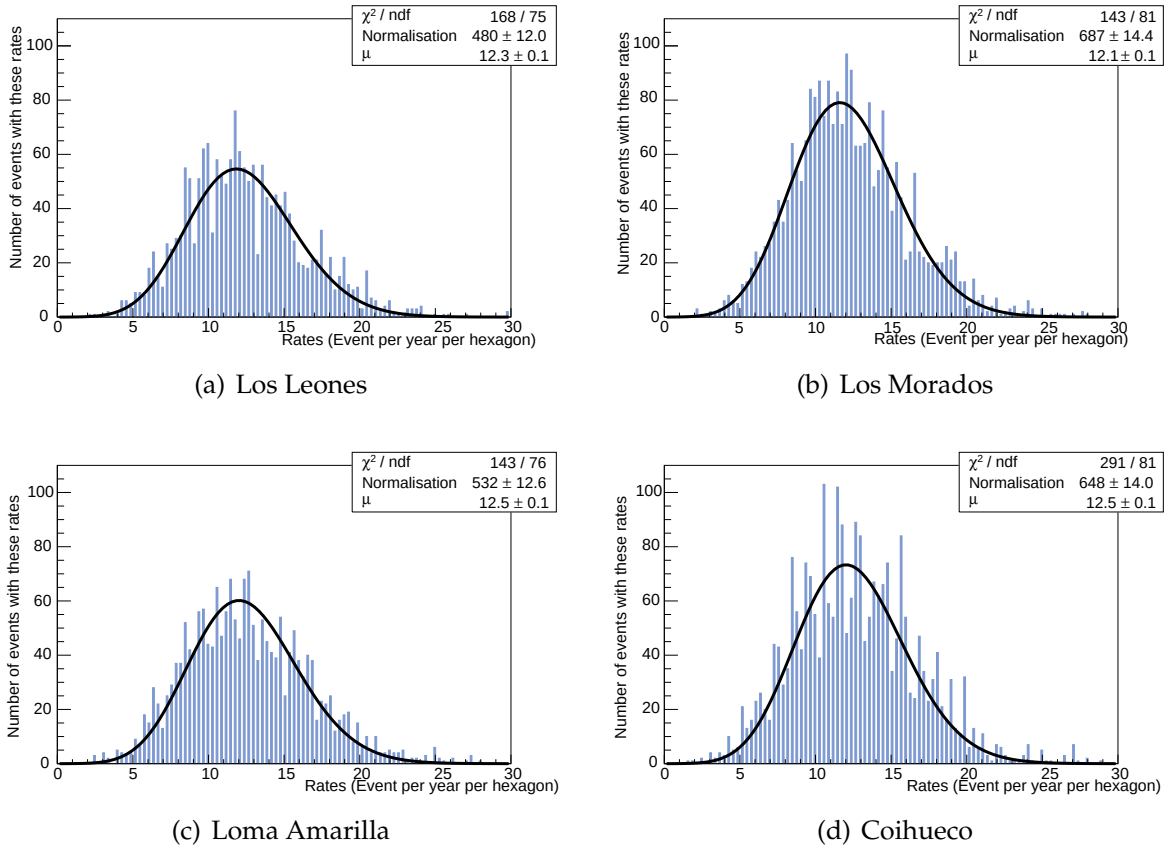
Then, the reconstructed shower cores (intersection between the shower axis and the ground), for all hybrid events were plotted in Figure 6.2 to compare with Figure 6.1 and a faint clump is shown in the north-western region. This could be due to the two FD sites (Loma Amarilla and Coihueco) being close together in this region, causing some overlap in the number of hybrid events detected. The following analyses may be able to shed light on any prominent structures in the hybrid event rates seen from each of the four FD sites.

6.2 Rates as a Function of Time As Seen from Each of the Four FD sites

The aim of investigating the hybrid event rates (as calculated form the SD event rates) as a function of time expected for each of the FD sites is to see if there is any time dependence arising from SD variations. It will be compared whether the trends in this analysis are consistent with structures in the rate maps shown by Figures 5.22 - 5.26 and 5.41 - 5.42 from Chapter 5. This is followed by a comparison to the increasing rates, seen by the SD (0.58 ± 0.08 % per year from 2006-2017) as well as a comparison to the increasing energy ratio. Finally, the fractional change in threshold energy for each FD site was then done to potentially be used in future work as corrections to E_{SD}.

The analysis begins with the assignment of yearly rates obtained by the hottest station[1] in each hybrid event. The rate distributions are shown in Figure 6.3. A

[1]The hottest station is defined as the station which receives the highest amount of signal in an event and is surrounded by 6 active stations



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.3: Distributions of yearly event rates for the hottest stations involved in hybrid events seen at Los Leones, Los Morados, Loma Amarilla and Coihueco. A Poisson function is fitted onto each of these distributions, with a normalisation and mean value, μ.

Poisson function was fitted onto these distributions to obtain a mean rate value. It can be seen that the mean rate value for Los Morados is 4σ away from Loma Amarilla and Coihueco. This might be due to the hotspots seen in the rate maps shown in Figure 5.7(b) of Chapter 5. Despite the hotspots seen as a function of position across the array, the mean rates seen from each FD site do not deviate significantly from each other.

Then, plots showing yearly rates from 2004 to 2017 are generated. Sections 6.2.2, 6.2.3, 6.2.4 and 6.2.5 show an analysis of the trends present in the plots. This analysis can determine if there is a strong dependence of rate as a function of time at each of the four FD sites. It was also noted that there were no hybrid events (subject to the applied cuts in Appendix A) detected by Los Morados in 2004, none seen by Loma Amarilla in 2004, 2005 and 2006 and only one event seen in Coihueco in 2004. Figure B.12 in Appendix B shows this for reference. An important point to note is that the mean rate at each of the FD sites is slightly biased upwards compared with the mean rate measured from SD events, due to a hybrid selection bias explained in the following section.

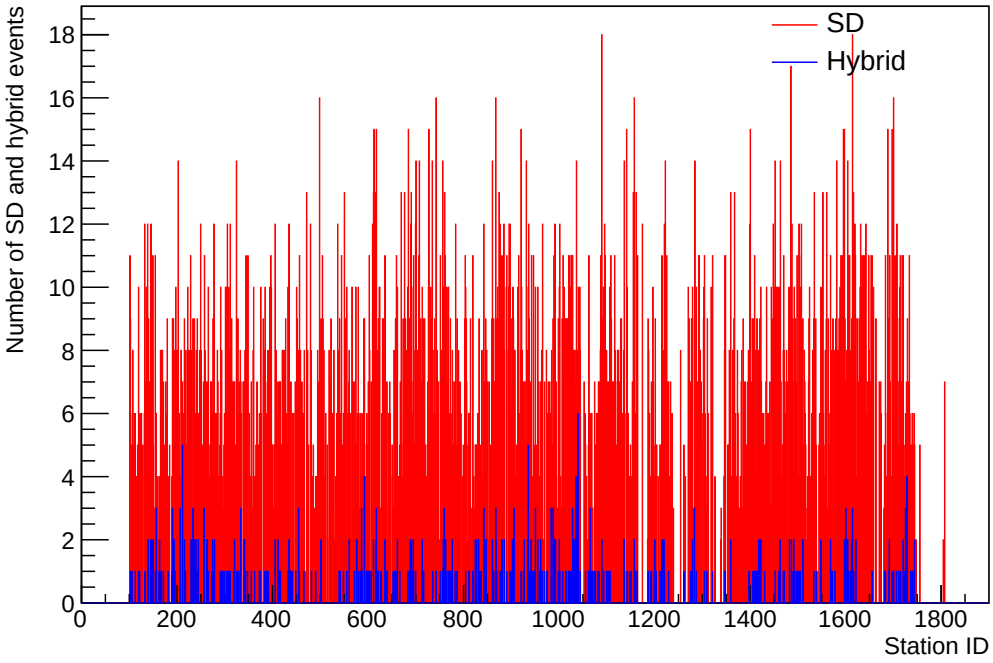


Figure 6.4: The number of hybrid and SD events with energies above 3 EeV (weather-corrected) and $\theta < 60^{\circ}$ detected by every hottest station in the year 2009. The blue bars are the hybrid events while the red bars are the SD events. All the stations are identified with a station ID.

6.2.1 Hybrid Selection Bias

In the process of assigning yearly rates to each hybrid event, there is a slight bias upwards in the mean rate. To illustrate this, distributions of both hybrid and SD events for all stations are shown in Figure 6.4 only for the year 2009 as an example. The average rate of a station in this year (calculated as the total number of SD events for all stations divided by the number of stations) is 6.61 events/year. But, the average rate of a hybrid-triggered station (calculated as the number of SD events of only hybrid-triggered stations divided by the number of stations) is 7.61 events/year. So, all hybrid-triggered stations incidentally possess a higher than average (by about 15 %) yearly rate. Profile plots of these assigned yearly rates to hybrid events seen at each FD site are plotted against time and shown in Figures 6.5-6.8 where several fits are applied. These plots are made by taking the average of all the assigned yearly rates of hybrid events in each bin (which has a length of one year) and plotting it with an error bar representing the standard error on the mean.

6.2.2 Horizontal Line Fits

Horizontal line fits for the rates as a function of time as seen by Los Leones, Los Morados, Loma Amarilla and Coihueco are shown in Figure 6.5. The χ^2

values were obtained and the probabilities of each fit were calculated and shown in Table 6.1. The fit parameters are listed in Table 6.2. It can be seen that the four FD sites have fits with deviations from the hypothesis either beyond 3σ or between 2σ-3σ. The former is shown by Los Leones and Coihueco while the latter is for Los Morados and Loma Amarilla. Firstly, it is noted that the intercepts are systematically higher than the average rate obtained in the SD-only analysis in Figure 5.27(a) of Chapter 5 where it shows an average rate of around 11.14 ± 0.03 events per year per hexagon. This greatly shows the upward bias as described in 6.2.1. Los Leones and Los Morados have their intercepts systematically lower than that of Loma Amarilla and Coihueco. From these intercepts, the fractional change in rates and (implicit) threshold energies were calculated and plotted in Figures 6.9 and 6.10 of Appendix B. This is done by using this Equation mentioned in Section 5.4.3 repeated here for clarity:

$$\frac{d(\textit{Event rate})}{\textit{Event rate}} = \frac{-2dE_{th}}{E_{th}} \tag{6.1}$$

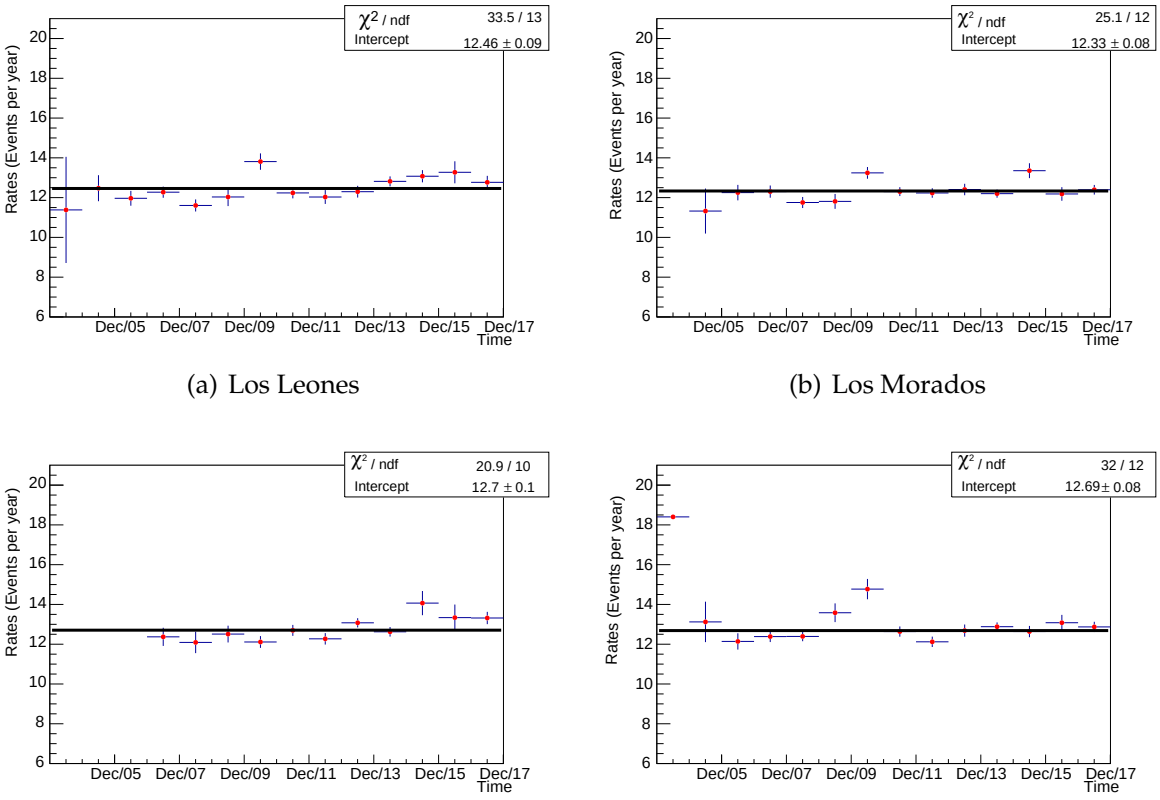
where *Event rate* is the y-intercept at each FD site and *d(Event rate)* is the absolute difference between each data point in Figure 6.5 and the y-intercept. The fractional change in threshold energy is obtained to be negative two times the fractional change in rates. So as rate above 3 EeV goes up by a 2 % for example, the implicit threshold energy goes down by 1 %, indicating that more events with energies less than 3 EeV are being reconstructed as having higher energies.

6.2.3 Linear Fits

A linear fit was applied to the rate trends as a function of time for the four FD sites. The following fit equation was used to fit the data:

$$y(t) = m\left(\frac{t - t_0}{T}\right) + c \tag{6.2}$$

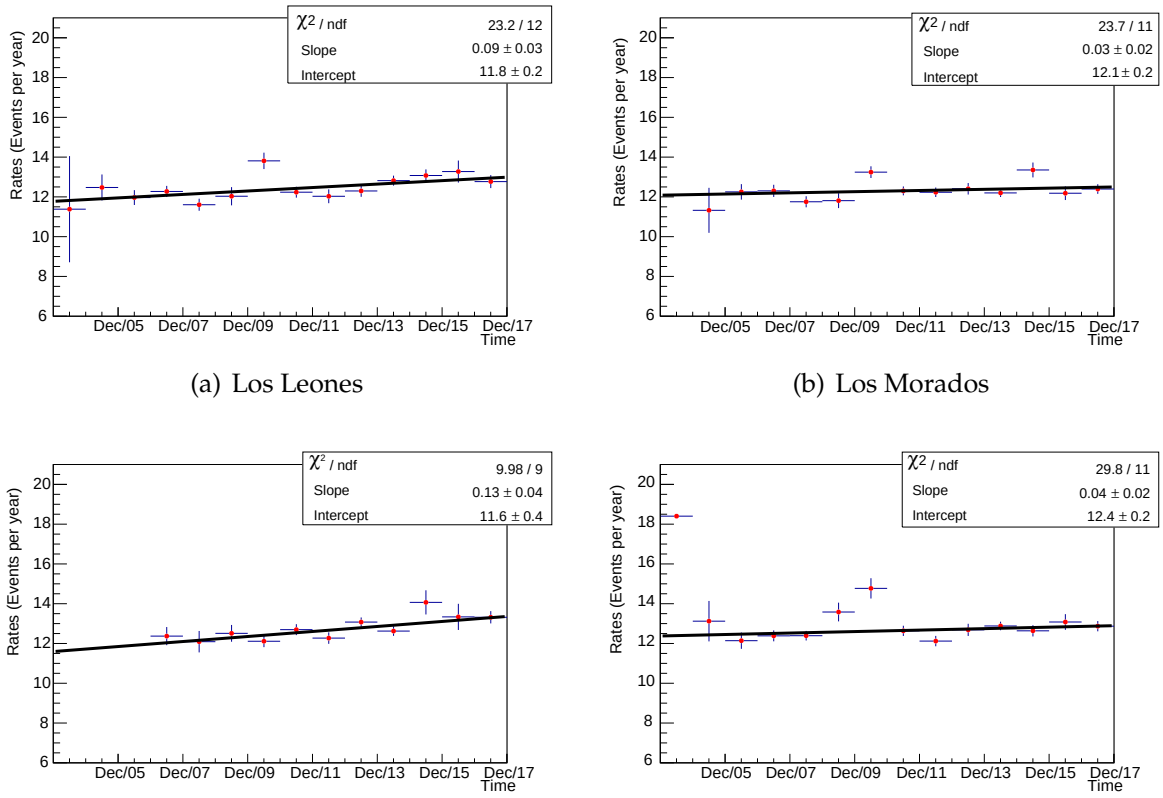
where *y(t)* is the rate as a function of time, m is the slope parameter in events/year^2 /hexagon, t is time in seconds since t_0 which is the start of 2004, T is the number of seconds in a year and c is the y-intercept parameter in events/year/ hexagon. Figure 6.6 show the fit results along with the probabilities and parameters respectively listed in Tables 6.1 and 6.2. It is shown that out of all linear fits for the four FD sites, Loma Amarilla has the highest probability of exhibiting a linear trend. Los Leones and Los Morados have deviations from the hypothesis which lies between 2σ-3σ, while the fit for Coihueco has a deviation beyond 3σ. To get a change per year (%), the slope parameters as well as their uncertainties were divided by the average rate represented by the horizontal fit intercepts and these are listed in Table 6.3. It is noted that the change per year in rates seen by just the SD from 2006 to 2017 obtained in Section 5.4.1, which is 0.58 $\pm$ 0.08 % (from 2006-2017) is at most only 1σ away from the change per year (%)



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.5: Horizontal line fits applied on rates of weather-corrected events with a threshold energy of 3 EeV, as a function of time seen by each FD site. All error bars represent the standard error on the mean for each yearly bin.



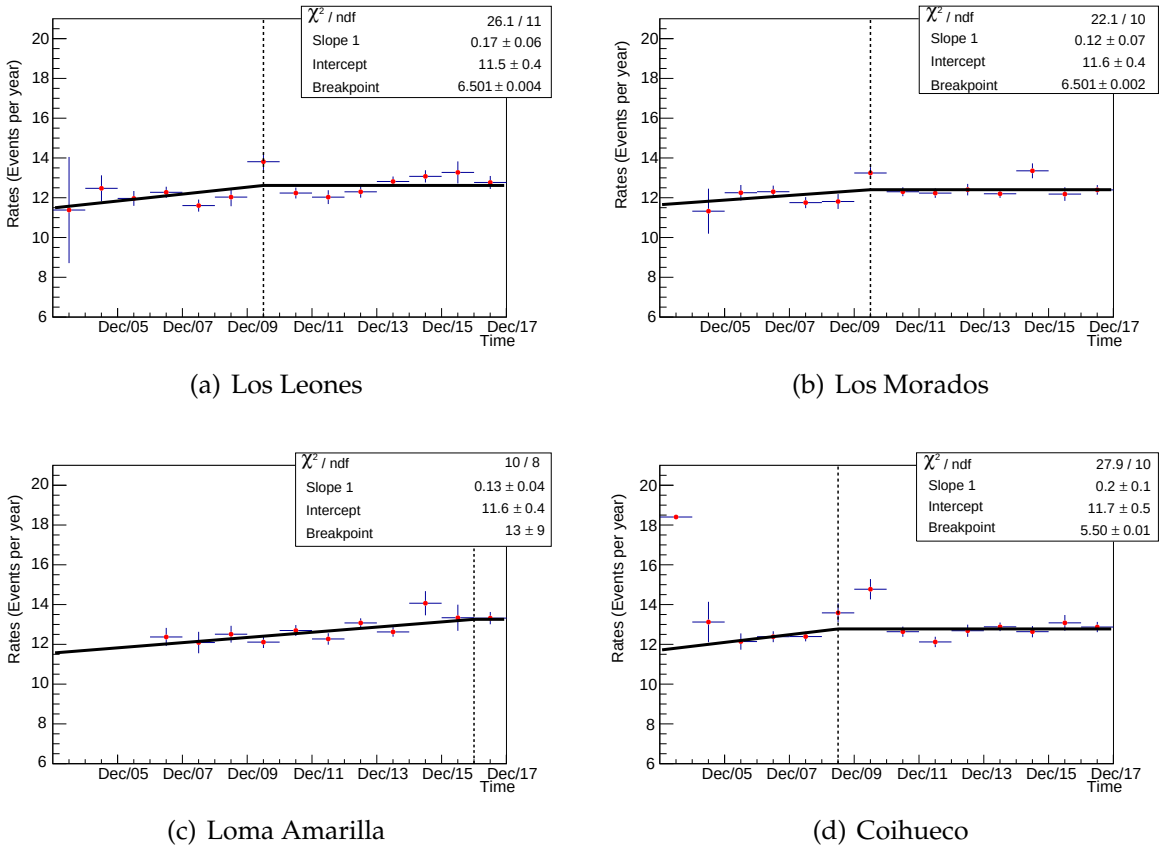
(a) Los Leones

(b) Los Morados

(c) Loma Amarilla

(d) Coihueco

Figure 6.6: Linear fits for rates as a function of time where each plot represents a single FD site. All error bars represent the standard error on the mean for each yearly bin.



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.7: Linear-horizontal fits for rates as a function of time where each plot represents each FD site. All error bars represent the standard error on the mean for each yearly bin. The vertical dashed line represent the fitted breakpoint in time at each site.

obtained by all four sites from 2004-2017 (noting that some sites have no detected events in 2004).

6.2.4 Linear-Horizontal Fits

A linear-horizontal fit was applied to the rates as a function of time for the four FD sites. They consist of two lines separated by a breakpoint where one has a free slope and one is horizontal. The fit function is shown as:

$$y(t) = \begin{cases} m\left(\frac{t-t_0}{T}\right) + c & \frac{t-t_0}{T} < b \\ mb + c & \frac{t-t_0}{T} \geq b \end{cases} \tag{6.3}$$

where *y(t)* is the rate as a function of time, t is the time in seconds since the start time t_0 which is the start of 2004, m is the slope parameter, T is the number of seconds in a year, c is the y-intercept and b is the breakpoint as the number of years after 2004. Figure 6.7 shows the fit results. The goodness-of-fit values and parameters are listed in Tables 6.1 and 6.2 respectively. It is shown in Table 6.1

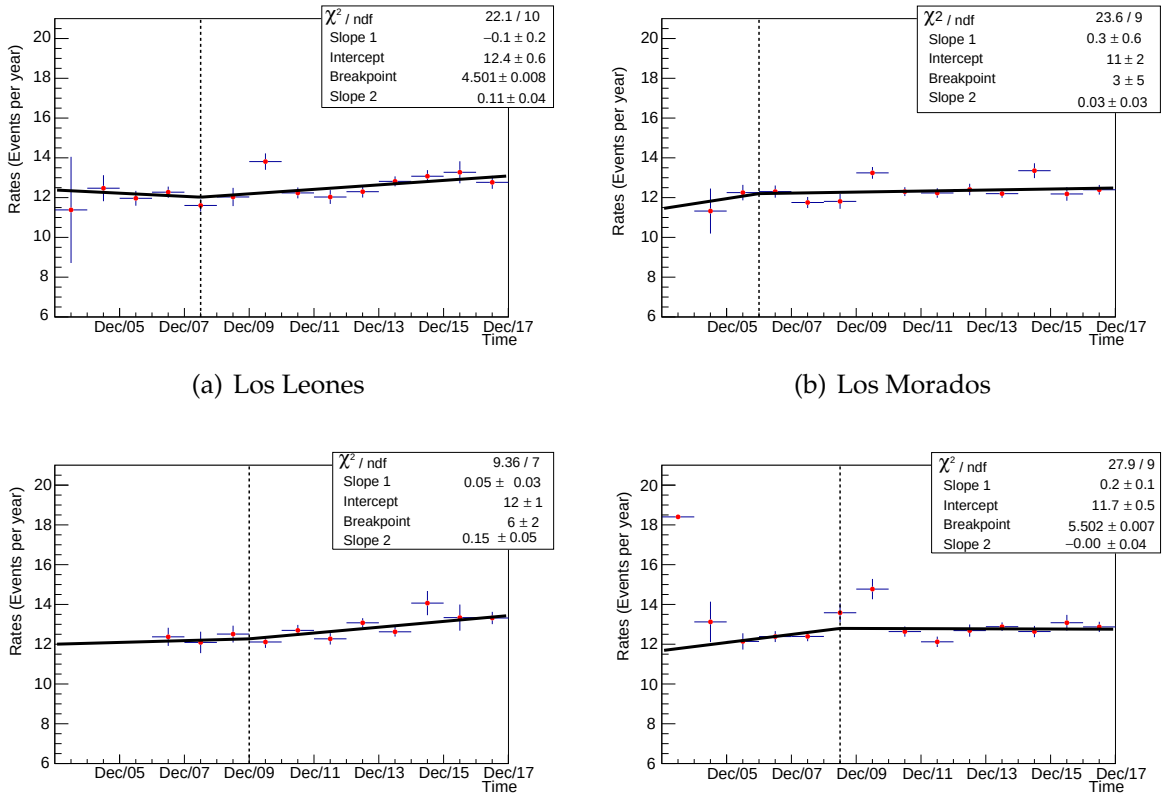
that the probability values for the fits are essentially unchanged from the linear fits, so neither the linear nor linear-horizontal behaviour is preferred over the other. The breakpoints as listed in Table 6.2 are shown to be in mid-2010 for Los Leones and Los Morados. As for Loma Amarilla, the breakpoint seems to be very close to the end of 2017 and the uncertainty on this is large (13 ± 9). For Coihueco, the breakpoint lies in mid-2009. Coihueco is also the only FD site which exhibits a breakpoint position that is not within the uncertainty of the other FD sites.

6.2.5 Linear-Linear Fits

Finally, a linear-linear type fit was applied to the rates as a function of time for the four FD sites. This fit consists of two lines with free slopes separated by a freely-moving breakpoint. The functional form of this fit is shown as

$$y(t) = \begin{cases} m_1 \left(\frac{t-t_0}{T}\right) + c_1 & \frac{t-t_0}{T} < b \\ m_2 \left(\frac{t-t_0}{T}\right) + c_2 & \frac{t-t_0}{T} \geq b \end{cases} \tag{6.4}$$

where *y(t)* is the rate as a function of time, t is the time in seconds since the start time t_0 which is the start of 2004, m_1 is the first slope parameter, m_2 is the second slope parameter, T is the number of seconds in a year, c_1 is the first y-intercept, c_2 is the second y-intercept and b is the breakpoint as the number of years after 2004. Figure 6.8 show the results of this fit. From Table 6.1, it is shown that the fit probabilities are essentially unchanged from the linear-horizontal fit. It is concluded that a more sophisticated functional form must be used to fit to the trend shown in the rates as a function of time as seen from Coihueco, since the deviations for all fits lie beyond 3σ. From Table 6.2 it can be seen that Los Leones and Coihueco have breakpoint positions which are not within statistical uncertainties of each other, but are within the uncertainties of the breakpoint obtained in the fits for Los Morados and Loma Amarilla. The positions of the breakpoints obtained by all four fits lie between 2008 and the start of 2010.



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.8: Linear-linear fits for rates as a function of time where each plot represents each FD site. All error bars represent the standard error on the mean for each yearly bin. The vertical dashed line represent the fitted breakpoint in time at each site.

Fit type	FD Site	χ^2	NDF	χ^2_{red}	Probability
Horizontal	Los Leones	33.5	13	2.58	0.00144
	Los Morados	25.1	12	2.09	0.0143
	Loma Amarilla	20.9	10	2.09	0.022
	Coihueco	32	12	2.67	0.00137
Linear	Los Leones	23.2	12	1.93	0.026
	Los Morados	23.7	11	2.16	0.014
	Loma Amarilla	9.98	9	1.11	0.352
	Coihueco	29.8	11	2.71	0.00169
Linear-horizontal	Los Leones	26.1	11	2.37	0.00633
	Los Morados	22.1	10	2.21	0.0145
	Loma Amarilla	10	8	1.25	0.263
	Coihueco	27.9	10	2.79	0.00184
Linear-linear	Los Leones	22.1	10	2.21	0.0147
	Los Morados	23.6	9	2.62	0.00505
	Loma Amarilla	9.36	7	1.34	0.228
	Coihueco	27.9	9	3.1	0.000981

Table 6.1: Results from horizontal, linear, linear-horizontal and linear-linear fit for the rates as a function of time at each FD site. The colour red represents the probabilities which have a deviation of more than 3σ, the colour orange represents probabilities that have a deviation between 2σ-3σ and green notes probabilities with a deviation between 0-2σ.

Fit type	FD Site	Slope (Events/yr^2/hex)	Intercept (Events/yr/hex)	Breakpoint (Yrs after '04)	Slope 2 (Events/yr^2/hex)
Horizontal	Los Leones	N/A	12.46 ± 0.09	N/A	N/A
	Los Morados		12.33 ± 0.08		
	Loma Amarilla		12.7 ± 0.1		
	Coihueco		12.69 ± 0.08		
Linear	Los Leones	0.09 ± 0.03	11.8 ± 0.2	N/A	N/A
	Los Morados	0.03 ± 0.02	12.1 ± 0.2		
	Loma Amarilla	0.13 ± 0.04	11.6 ± 0.4		
	Coihueco	0.04 ± 0.02	12.4 ± 0.2		
Linear-horizontal	Los Leones	0.17 ± 0.06	11.5 ± 0.4	6.501 ± 0.004	N/A
	Los Morados	0.12 ± 0.07	11.6 ± 0.4	6.501 ± 0.002	
	Loma Amarilla	0.13 ± 0.04	11.6 ± 0.4	13 ± 9	
	Coihueco	0.2 ± 0.1	11.7 ± 0.5	5.50 ± 0.01	
Linear-linear	Los Leones	-0.1 ± 0.2	12.4 ± 0.6	4.501 ± 0.008	0.11 ± 0.04
	Los Morados	0.3 ± 0.6	11 ± 2	3 ± 5	0.03 ± 0.03
	Loma Amarilla	0.05 ± 0.03	12 ± 1	6 ± 2	0.15 ± 0.05
	Coihueco	0.2 ± 0.1	11.7 ± 0.5	5.502 ± 0.007	-0.00 ± 0.04

Table 6.2: Summary of resulting fit parameters for all fit-types in rates as a function of time at each FD site. The uncertainties in the breakpoint positions are implicitly constrained to be an integer.

Fit type	FD Site	Slope (Events/yr^2/hex)	Average rate (Events/yr/hex)	Change/yr (%)
Linear	Los Leones	0.09 ± 0.03	11.8 ± 0.2	0.8 ± 0.3
	Los Morados	0.03 ± 0.02	12.1 ± 0.2	0.3 ± 0.2
	Loma Amarilla	0.13 ± 0.04	11.6 ± 0.4	1.1 ± 0.4
	Coihueco	0.04 ± 0.02	12.4 ± 0.2	0.3 ± 0.2

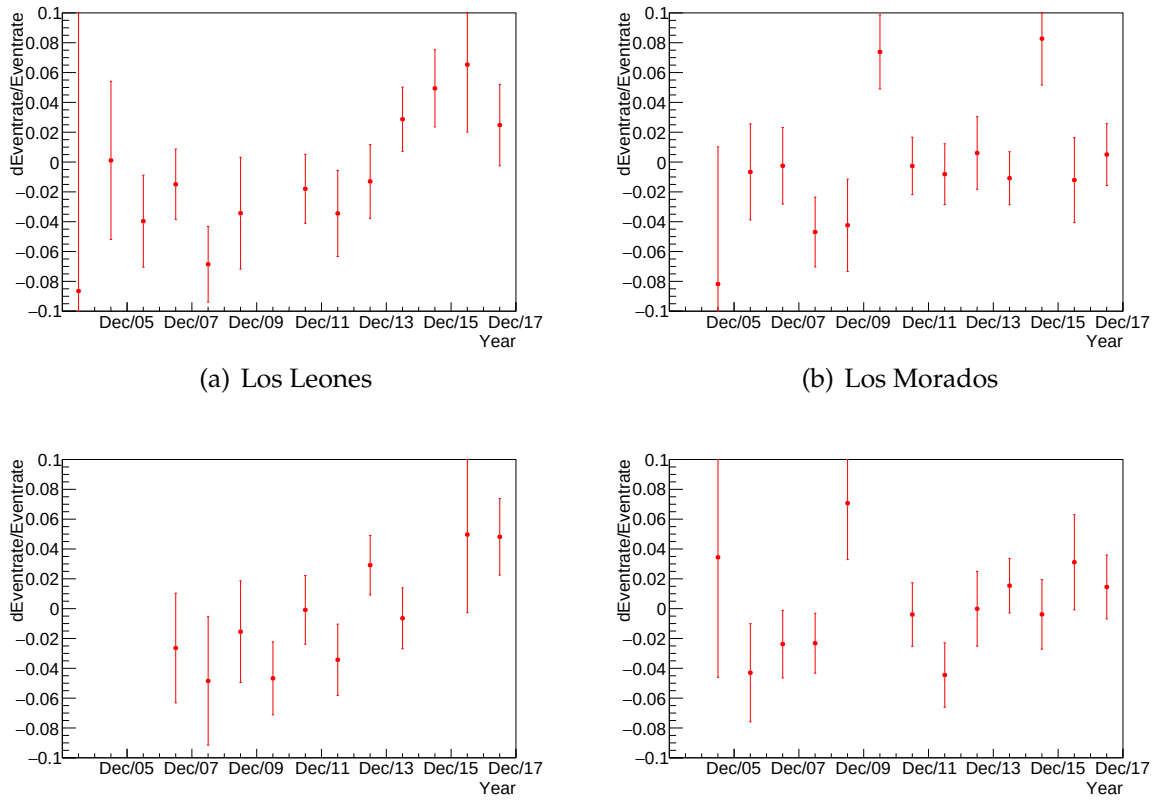
Table 6.3: Change per year (%) of rates as a function of time seen at Los Leones, Los Morados, Loma Amarilla and Coihueco.

6.3 Fractional Changes in Threshold Energy at Each FD Site

Figures 6.9 and 6.10 show the fractional changes in rates and threshold energy as a function of time from Los Leones, Los Morados, Loma Amarilla and Coihueco. This was calculated from the event rate and threshold energy relationship equation shown in Equation 6.1, where the event rate is represented by the normalisations (y-intercepts) from the horizontal line fits listed in Table 6.2. From the analysis on the absolute rates as a function of time, it was shown that only Loma Amarilla's event rate follows a linearly increasing trend. It cannot be said whether Los Leones and Los Morados prefers any of the fits while none of the fits were suited to the trend shown in Coihueco. So, from Equation 6.1, it can be said that as the rates are shown to be increasing (for Loma Amarilla in this case), the threshold energy is decreasing by half the amount. Since these are rates above an assigned energy (which is the full efficiency threshold) of 3 EeV, the increasing trend indicates an increase in the assigned energies. *True* lower energy events are assigned higher energies which means the true threshold energy is decreasing. This is exactly what is shown the trend for Loma Amarilla in Figure 6.10(c), indicating that corrections applied would reduce E_{SD} as a function of time. The trends at the other sites are more complex, which may be a contribution from the clear dips in the threshold energy, shown at Los Leones, Los Morados and Coihueco in 2010. Also, it was concluded previously that the E_{SD}/E_{FD} ratio for most telescopes detailed in Chapter 4 exhibit an increasing linear-linear trend. After corrections are made, depending on the FD site, any remaining drifts in the E_{SD}/E_{FD} trend will be due to a drift in FD energies.

6.3.1 Summary

It can be seen that the rates as a function of time as seen by every FD site is not constant and has distinct behaviour. The rates seen from Loma Amarilla seem to be well-explained by either a linear, linear-horizontal or linear-linear trend. This is in contrast to Los Leones and Los Morados since neither of these trends are significantly preferred. None of the fits were suited to the trend seen at Coihueco, but this could be due to a contribution from the high rate spike in 2010. This high



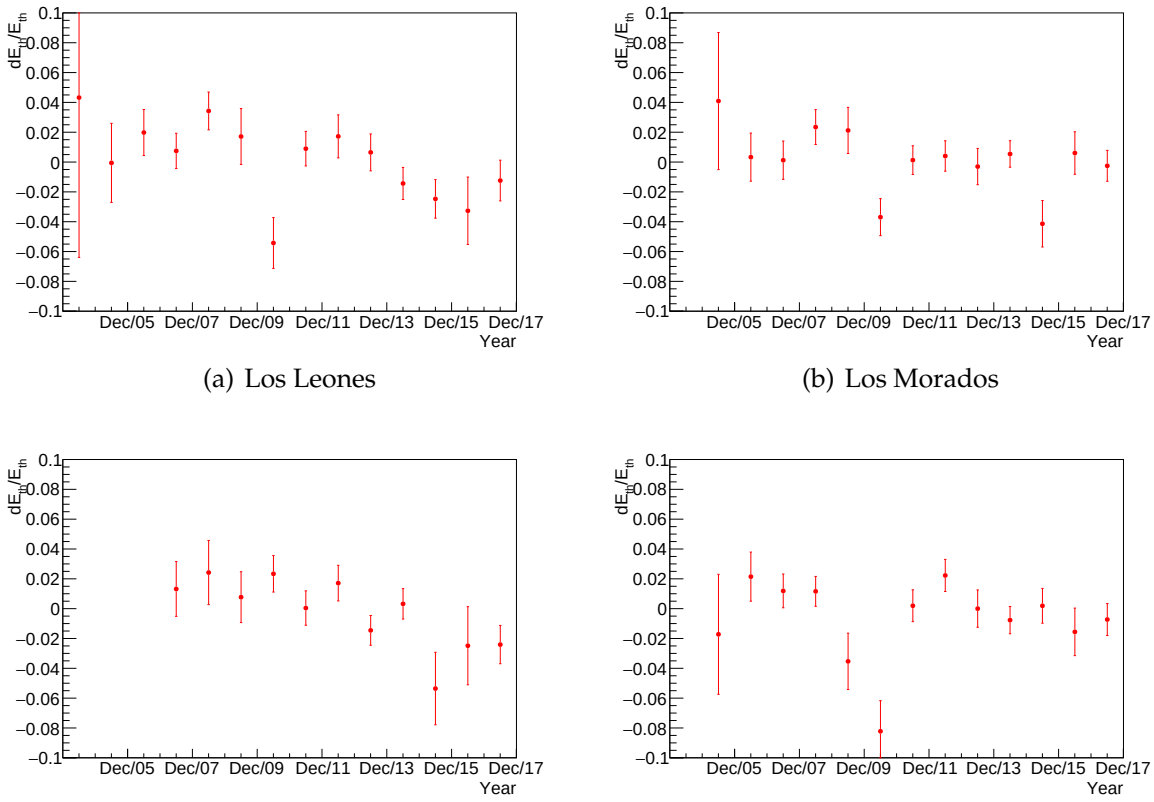
(a) Los Leones

(b) Los Morados

(c) Loma Amarilla

(d) Coihueco

Figure 6.9: Fractional changes in rates as a function of time for each FD site with reference rates listed in Table 6.2 as the horizontal line fit y-intercepts.



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.10: Fractional changes in threshold energy as a function of time for each FD site with reference rates listed in Table 6.2 as horizontal line fit y-intercepts.

rate in 2010 is also seen at Los Leones and Los Morados. In the rate map for 2010 shown in Figure 5.41(g) of Chapter 5, a small hotspot is shown in between Los Leones and Coihueco which may be a contribution to this rate spike. However, the average rates seen in this 2010 rate map is not distinguishably higher than other years. As for the hotspot seen in Figure 5.26 (and more specifically just in 2015 shown in Figure 5.4.2), this seems to be in the vicinity of Loma Amarilla, which may explain the high rate point seen in 2015. However, it is noted that there is no rate spike in 2010 for Loma Amarilla, compared to the other FD sites, which may explain why the time dependence at Loma Amarilla behaves the least erratically.

It is also shown that the breakpoint positions lie between 2008 and 2010. The rate spike in 2010 is slightly odd. Although this rate spike seems to correlate with the freezing events described in Section 5.1, as an additional check, an investigation was performed in Section 6.5 to check for any anomalies in this year. Another function is needed to fit the behaviour of rates seen by Los Leones and Los Morados; to improve the 2σ-3σ deviations to a range between 0-2σ. However, the rates of the stations in the field of view of Los Leones seem to be the most uniform, with few hotspots very near the site. As for Los Morados, there is a rate spike in 2010 and 2015, which is interesting as this is not seen near the FD site on the rate maps. The change per year (%) obtained from the linear fits for all FD sites are not significantly deviating (at most 1σ away) from the change per year obtained from the SD event rate analysis, from 2006 to 2017 (0.58 ± 0.08 %).

It was noted that since the rates above 3 EeV are found to be increasing, this means the assigned energies are also increasing. From comparing the trends in rates as a function of time to the evolving energy ratio at each FD site (E_{SD}/E_{FD} from Chapter 4), it was found that the best fit-type (having the most telescope fit deviations between 0-2σ) is linear-linear. However, the rates seen by Coihueco do not seem to statistically agree with this, due to the 3σ deviation. The linear-linear fit for Los Leones and Los Morados is not strongly preferred like in the energy ratio plots. This may be due to the spike in rates shown in 2010 for the three sites. In contrast, the 0-2σ deviation for the linear-linear fit to the evolving rates seen from Loma Amarilla tends to agree with energy ratio as well, since only two telescopes have a trend with a 3σ deviation. Overall, it can be seen that the FD sites see an overall increasing rate, indicating that the assigned energies are increasing which is shown fairly significantly in most of the energy ratio plots for each telescope.

6.4 Rates as a Function Distance Between FD Sites and Hottest Stations

The following analysis entails an investigation on the possible dependence of SD event rates on distance between FD sites and hottest stations for every

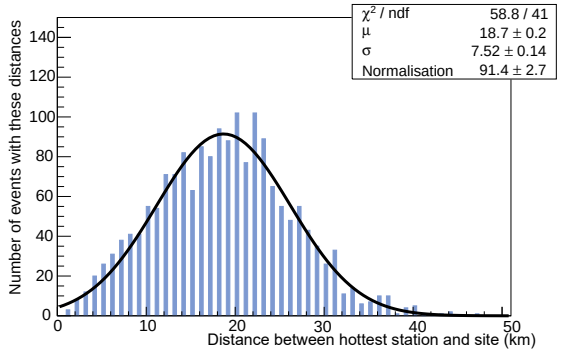
hybrid event. The rates above 3 EeV as seen by each FD site determines the threshold energy, which then dictates S_{38} and finally E_{SD}. From this, the behaviour of these rates as a function of distance can dictate the dependence of energy ratio (E_{SD}/E_{FD}) on aerosol distribution or distance to the shower. The aerosol transmission from the point of shower maximum (X_{max}, indicating the point in the atmosphere at which the size of the shower is largest), denoted $T_{aer,X_{max}}$ is a parameter which encapsulates this. If there is a structure in the rates as a function of distance from each FD site, this will create structures in the energy ratio as a function of $T_{aer,X_{max}}$.

Due to the uniformity in the distribution of stations across the array, the SD energy is used as a reliable normalisation for the FD energy. If a structure exists in the energy ratio as a function of $T_{aer,X_{max}}$, this would be interpreted as either an underestimation or overestimation in the amount of aerosols in the atmosphere as distance to the shower changes. The aerosol distribution dictates the energy deposited in the shower, contributing to the calculation of EFD. However, if E_{SD} in itself also has a dependence as a function of distance, this can affect the interpretation of the aerosol distribution. Since E_{SD} is determined by the rates above 3 EeV (which should be constant across time and position), an analysis of the rates as a function of distance can determine if there would be any contribution to the structure in E_{SD}/E_{FD} versus $T_{aer,X_{max}}$.

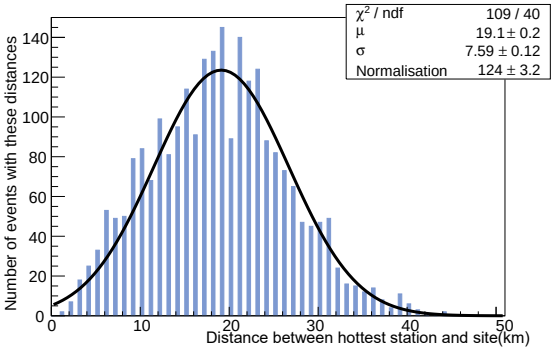
The analysis begins with generating distributions of distances between the FD site and hottest stations for Los Leones, Los Morados, Loma Amarilla and Coihueco. These distributions are shown in Figure 6.11. Unfortunately, it is difficult to predict exactly how much the distributions shown by each of the FD sites should deviate from one another, or at all. This is because the field of views of some of the telescopes at particular FD sites lack an area that is populated with stations (Los Leones telescope 1 and 6 in Figure 3.1 for instance). A Gaussian function is fitted to obtain the values of μ, σ and normalisations. It can be seen that the only FD site which has a mean slightly higher than the other sites is Loma Amarilla. This may be due to the high likelihood of stations detecting hybrid events in the northern region of the array as shown in Figure 6.1. Then, plots of rates as a function of distance from each FD site were generated (scatter plots shown in Figures B.13 of Appendix B for reference) and four fits were applied.

6.4.1 Horizontal Line Fits

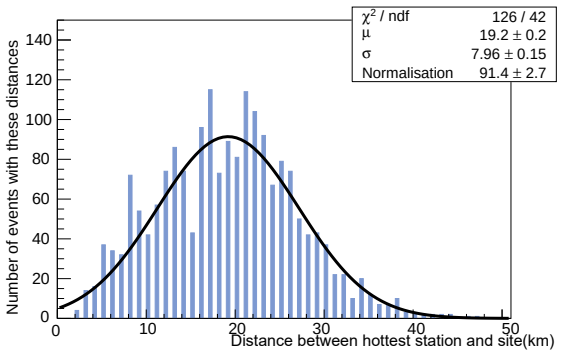
Horizontal line fits for the rates as a function of distance between a hottest station and Los Leones, Los Morados, Loma Amarilla and Coihueco are shown in Figure 6.12. The χ^2 values were obtained and the probabilities of each fit were calculated and shown in Table 6.4. The y-intercepts were also calculated and listed in Table 6.5. It can be seen that only Loma Amarilla has a horizontal fit with a deviation between 0-2σ. The rates as a function of distance between the hottest stations and Coihueco seem to have a behaviour which is the least likely to be a horizontal fit out of the four sites. It is inconclusive whether Los Leones and



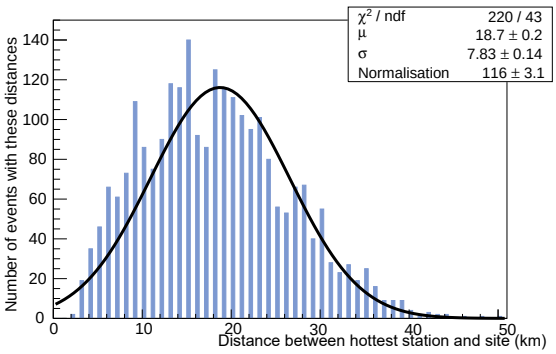
(a) Los Leones



(b) Los Morados

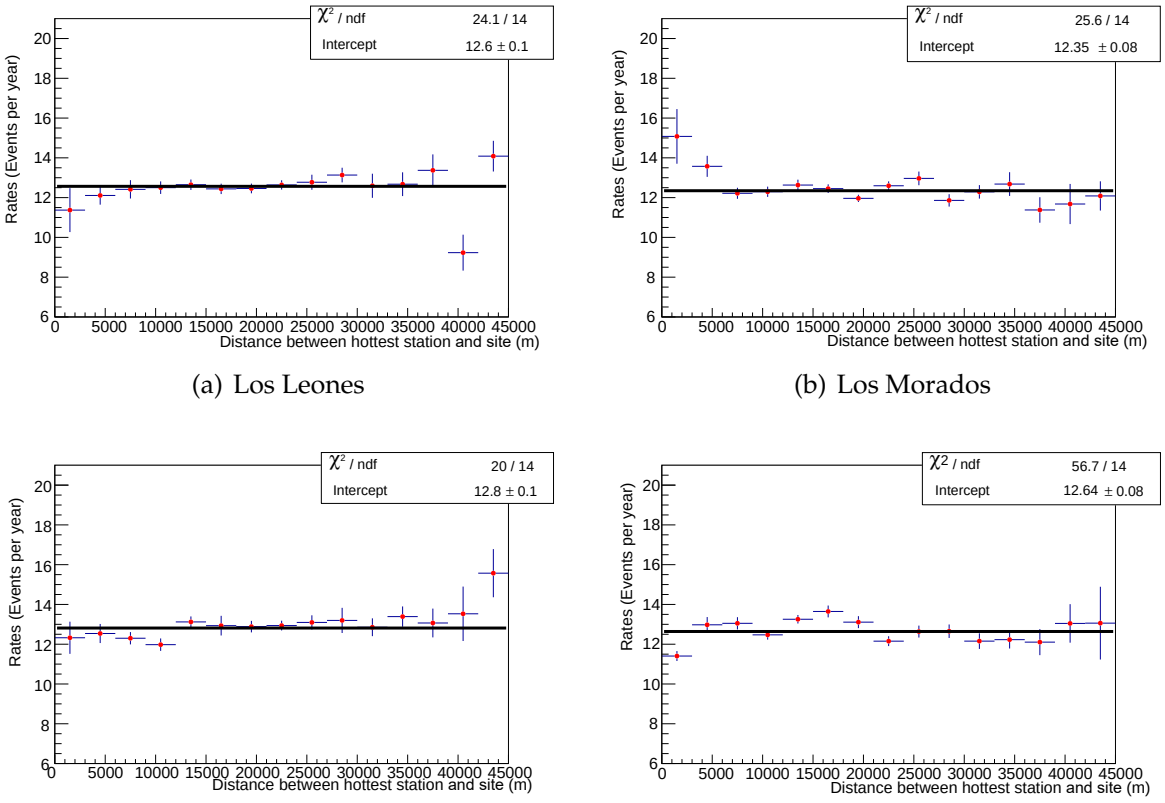


(c) Loma Amarilla



(d) Coihueco

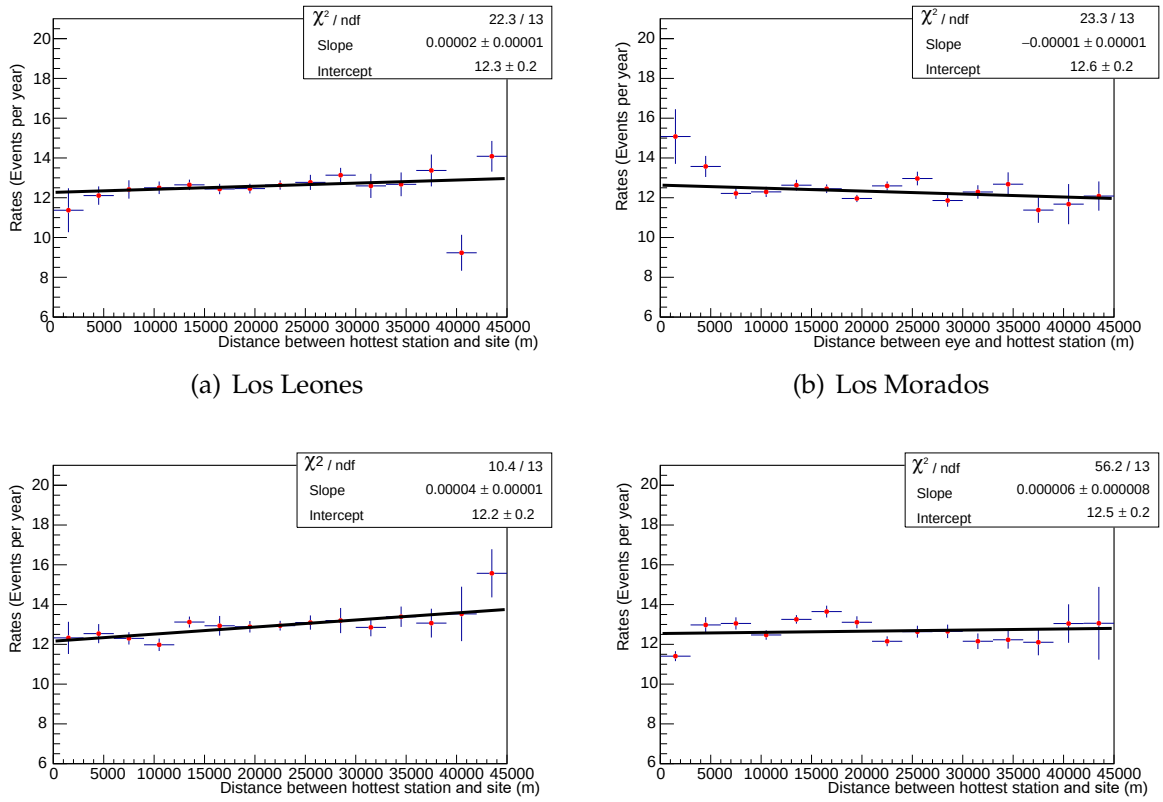
Figure 6.11: Distributions of distances between the hottest station and an FD site. A Gaussian was fitted to each distribution, with the mean μ, deviation σ and normalisation listed.



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.12: Horizontal line fits onto plots of rates as a function of distance between hottest station and each FD site. All error bars represent the standard error on the mean for each yearly bin.



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.13: Plots of rates as a function of distance between hottest station and each FD site. All error bars represent the standard error on the mean for each yearly bin.

Los Morados exhibit a constant rate for all hottest station distances since their deviations lie between 2-3σ. In terms of the y-intercepts, Los Leones, Loma Amarilla and Coihueco have values that are within the statistical uncertainty, which is around 12.7 events/year/hexagon. The normalisation for Los Morados was found to be systematically lower, at 12.35 $\pm$ 0.08 events/year/hexagon. However, the uncertainties in the normalisations are underestimated since the reduced χ^2 values are larger than one.

6.4.2 Linear Fits

A linear fit was applied to the rate trends as a function of distance between the hottest station and FD sites. The following fit equation was used to fit the data:

$$y(x) = m(x) + c \tag{6.5}$$

where $y(x)$ is the rate as a function of distance, x is the distance in metres, m is the slope parameter in events/year/hexagon/m and c is the y-intercept in events/year/hexagon. Figure 6.13 show the fit results along with the

probabilities and parameters respectively listed in Table 6.4 and Table 6.5. It can be seen that a linear fit for Los Leones has improved in terms of deviation from a horizontal to a linear trend, as it now lies between 0-2σ. The fit probabilities are not substantially improved for the other three sites. The slope parameter obtained for Loma Amarilla is systematically higher than the those of the other three sites, while the y-intercepts for all four sites agree within their statistical uncertainties.

6.4.3 Linear-Horizontal Fits

Linear-horizontal fits were then applied to the plots of rate as a function of distance for each FD site to see if the χ^2 values improve. The fit function is shown as follows:

$$y(x) = \begin{cases} m\,(x) + c & x < b \\ mb + c & x \geq b \end{cases} \tag{6.6}$$

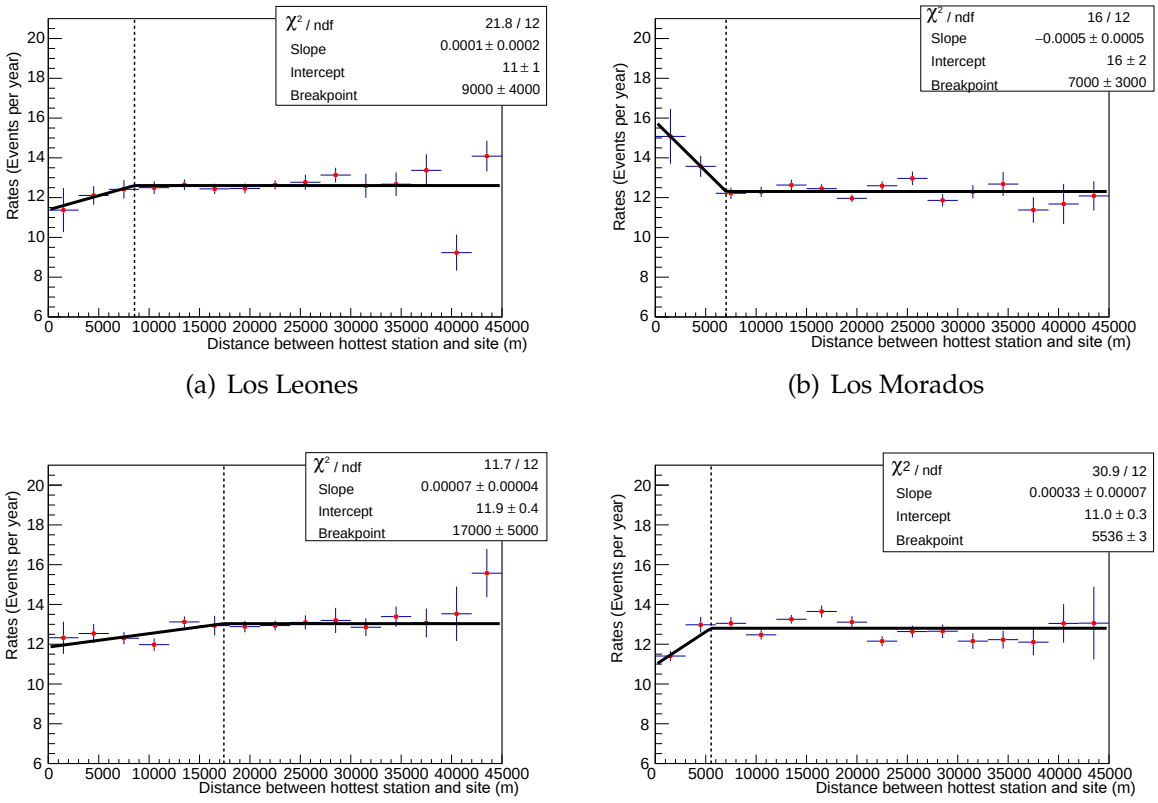
where *y(x)* is the rate as a function of distance, *x* is the distance in metres, *m* is the slope parameter in events/year/hexagon/m, *c* is the y-intercept in events/year/hexagon and *b* is the breakpoint as the distance in metres at which the rate starts to follows a constant behaviour. Figure 6.14 show the fit results. It can be seen that Los Morados is the only site which exhibits a negative slope before the breakpoint. From Table 6.4 it is shown that there is an improvement in the deviation of this fit type for Los Morados, but there is regression for that of Los Leones. A linear-horizontal fit type is most likely suited for Los Morados and Loma Amarilla but not for Los Leones and Coihueco. The y-intercept obtained for Los Morados is substantially higher than the other sites while the breakpoint at Loma Amarilla is significantly further, obtained as 17000 $\pm$ 5000 m.

6.4.4 Linear-Linear Fits

Finally, a linear-linear fit was applied. This fit consists of two lines with free slopes separated by a freely-moving breakpoint. The functional form is shown as

$$y(x) = \begin{cases} m_1(x) + c_1 & x < b \\ m_2(x) + c_2 & x \geq b \end{cases} \tag{6.7}$$

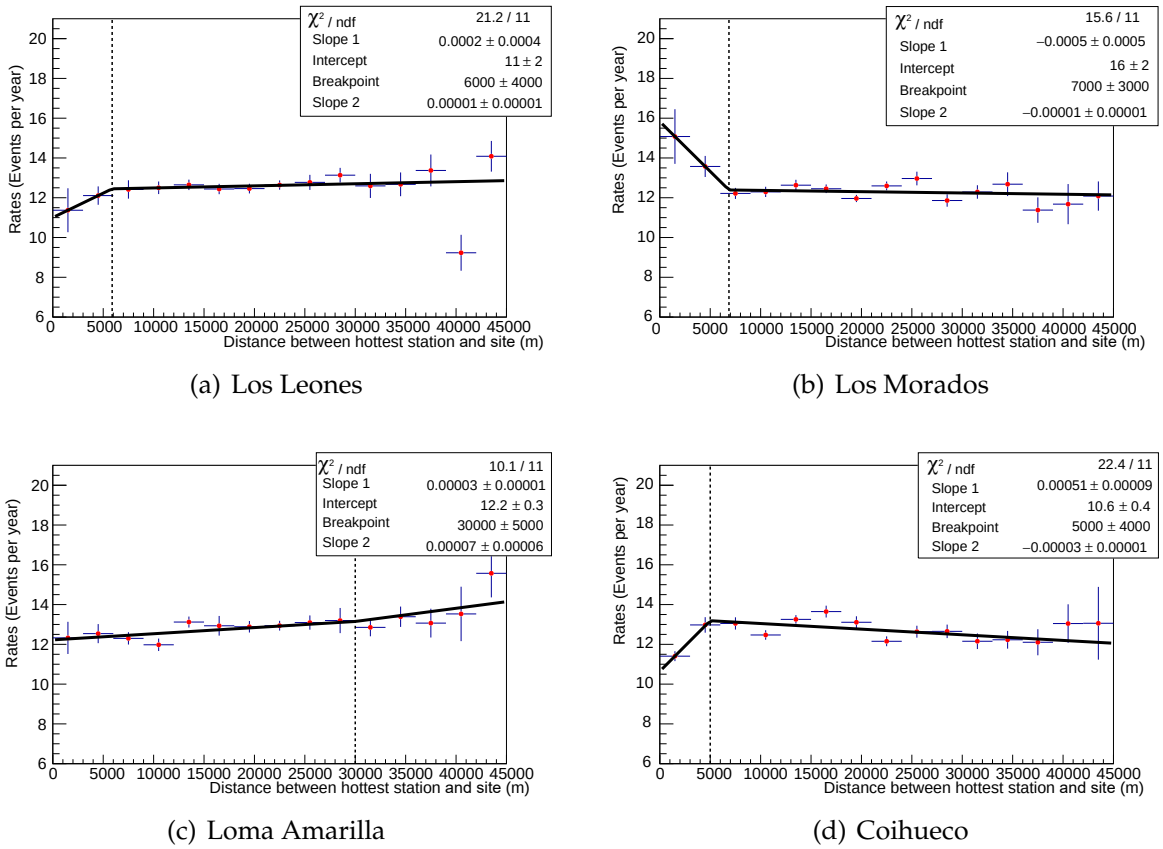
where $y(x)$ is the rate as a function of distance, *x* is the distance in metres, m_1 is the first slope parameter in events/year/hexagon/m, m_2 is the second slope parameter, c_1 is the first y-intercept in events/year/hexagon, c_2 is the second y-intercept and *b* is the breakpoint in metres. Figure 6.15 show the resulting fits. The deviations remain comparable to the previous linear-horizontal type fit, except for Coihueco where there is an improvement. However, since the deviation from the hypothesis for this fit for Coihueco still lies between 2-3σ, it is not conclusive as to whether a linear-linear fit is appropriate. The first slopes for Loma Amarilla and Coihueco are constrained to be positive, while the slopes in Los Leones



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.14: Linear-horizontal fits done on rates as a function of distance between the hottest station and each FD site. All error bars represent the standard error on the mean for each yearly bin.



(a) Los Leones (b) Los Morados

(c) Loma Amarilla (d) Coihueco

Figure 6.15: Linear-linear fits done on plots of rates as a function of distance between hottest station and each FD site. All error bars represent the standard error on the mean for each yearly bin.

and Los Morados agree within the uncertainties. Once again, the intercept for Los Morados is systematically higher; at 16 ± 2 events/year/hexagon while the breakpoint for Loma Amarilla is also even further than the linear-horizontal fit; at 30000 ± 5000 m. The second slope in both Los Morados and Coihueco are shown to be negative and are not within the statistical uncertainties of the second slopes for Los Leones and Loma Amarilla.

6.4.5 Summary

Overall, all four fits for Loma Amarilla have a deviation between 0-2σ while the trend for Coihueco is the most sophisticated. The rates as a function of distance seen at Los Leones is most fitting for a linear fit while for Los Morados, it can either be a linear-horizontal or linear-linear fit.

The linear increase found in Los Leones may be due to the uniform rates across position in the vicinity of the FD site as shown in the rate maps in the previous chapter, with a hotspot appearing at the furthest distance away. There is a hotspot present in the vicinity of Los Morados as shown in Figure 5.7(b) which explains

Fit type	FD Site	χ^2	NDF	χ^2_{red}	Probability
Horizontal	Los Leones	24.1	14	1.72	0.0442
	Los Morados	25.6	14	1.83	0.0294
	Loma Amarilla	20	14	1.43	0.129
	Coihueco	56.7	14	4.05	4.42×10^{-7}
Linear	Los Leones	22.3	13	1.72	0.0506
	Los Morados	23.3	13	1.79	0.0384
	Loma Amarilla	10.4	13	0.8	0.662
	Coihueco	56.2	13	4.32	2.46×10^{-7}
Linear-horizontal	Los Leones	21.8	12	1.82	0.0399
	Los Morados	16	12	1.33	0.191
	Loma Amarilla	11.7	12	0.975	0.468
	Coihueco	30.9	12	2.58	0.00205
Linear-linear	Los Leones	21.2	11	1.93	0.0315
	Los Morados	15.6	11	1.42	0.156
	Loma Amarilla	10.1	11	0.92	0.52
	Coihueco	22.4	11	2.04	0.0215

Table 6.4: Results from applying horizontal, linear, linear-horizontal and linear-linear fits for rates as a function of distance between an FD site and hottest station.The color red represents the probabilities that have a deviation beyond 3σ, the colour orange represents probabilities that have a deviation between 2σ-3σ and green notes a deviation between 0-2σ.

Fit type	FD Site	Slope 1 (Events/yr/hex /m)	Intercept (Events/yr/hex)	Breakpoint (m)	Slope 2 (Events/yr/hex /m)
Horizontal	Los Leones Los Morados Loma Amarilla Coihueco	N/A	12.6 ± 0.1 12.35 ± 0.08 12.8 ± 0.1 12.64 ± 0.08	N/A	N/A
Linear	Los Leones Los Morados Loma Amarilla Coihueco	$(2 \pm 1) \times 10^{-5}$ $(-1 \pm 1) \times 10^{-5}$ $(4 \pm 1) \times 10^{-5}$ $(6 \pm 8) \times 10^{-6}$	12.3 ± 0.2 12.6 ± 0.2 12.2 ± 0.2 12.5 ± 2	N/A	N/A
Linear-horizontal	Los Leones Los Morados Loma Amarilla Coihueco	$(1 \pm 2) \times 10^{-4}$ $(-5 \pm 5) \times 10^{-4}$ $(0.7 \pm 0.4) \times 10^{-4}$ $(3 \pm 7) \times 10^{-4}$	11 ± 1 16 ± 2 11.9 ± 0.4 11.0 ± 0.3	9000 ± 4000 7000 ± 3000 17000 ± 5000 5536 ± 3	N/A
Linear-linear	Los Leones Los Morados Loma Amarilla Coihueco	$(2 \pm 4) \times 10^{-4}$ $(-5 \pm 5) \times 10^{-4}$ $(0.3 \pm 0.1) \times 10^{-4}$ $(5.1 \pm 0.9) \times 10^{-4}$	11 ± 2 16 ± 2 12.2 ± 0.3 10.6 ± 0.4	6000 ± 4000 7000 ± 3000 30000 ± 5000 5000 ± 4000	$(1 \pm 1) \times 10^{-5}$ $(-1 \pm 1) \times 10^{-5}$ $(7 \pm 6) \times 10^{-5}$ $(-3 \pm 1) \times 10^{-5}$

Table 6.5: Summary of resulting fit parameters for all fit-types in rates as a function of distance between a hottest station and each FD site.

high rates very near the site, then a breakpoint, indicating where the rates start to exhibit a negative drift. One of the prominent features at Loma Amarilla are the breakpoint positions, placed at 17000 $\pm$ 5000 m for the linear-horizontal fit and at 30000 $\pm$ 5000 m for the linear-linear fit; both at which are slightly further than the breakpoint positions for the other sites. These are where the rates observed start to exhibit a rising behaviour as the hottest stations are further away. The hotspot seen in the 2014-2017 rate map may be partly responsible for the breakpoint position being placed at these distances; since the hotspots are between 17000 m and 30000 m away from the FD site. Another type of fit is needed for the rates seen by Coihueco as a function of distance for the fit improve to a 0-2σ deviation. This could be correlated with two adjacent hotspots shown in the 2004-2013 rate map (in Figure 5.7(b)) most prominently in the field of view of telescopes 1 and 6.

From these results, it can be said that the rates as a function of distance are linearly increasing for Los Leones. Loma Amarilla fits well with a constant trend and as for Los Morados, there is a sharp decrease, then either a constant or decreasing trend after a breakpoint at around 7000 m. Coihueco could possibly have an increasing then decreasing trend or something more complex. These structures could contribute to a compounding effect in the energy ratio as a function of $T_{aer,X_{max}}$. From this, appropriate corrections can be done to the SD energy, removing such structures to obtain an accurate interpretation of the energy ratio as a function of $T_{aer,X_{max}}$.

6.5 Investigation of the High Rate in 2010

As seen in Figure 5.27 and 5.28 of Section 5.4.1, as well as in Figure 6.5 for Los Leones, Los Morados and Coihueco, there is a jump in the rate observed in 2010. Several checks were done to see if the weather corrections were applied appropriately in 2010 compared to other years. These weather corrections, as described in Section 4.1 of Chapter 4 correct the signal detected by using measurements of variations in density, pressure and geomagnetic field due to their impacts on the air shower components. There can be a variation in the density within a day of about $\pm$ 3 %, corresponding to a temperature change of $\pm$ 8 °C [114]. If weather corrections are not done effectively, this can impact the shower signal detected which in turn affects the energy, thus introducing a systematic error in the rate above the full efficiency threshold. The following analysis checks this by plotting the number of events detected every hour throughout the day. To observe if there is a systematic in the diurnal measurements, the number of events detected throughout the day is accumulatively obtained for all days each year from 2004 to 2017. Then, a sinusoidal fit was applied to these diurnal plots to obtain the amplitude and phase shift. This is to see if there is an outlier in the amplitude and phase for 2010 compared to other years. Then, the air density and pressure measurements were obtained directly from the weather stations at Los Leones, Los Morados, Loma

Amarilla, Coihueco and the Central Laser Facility (CLF). These measurements were plotted to see if the behaviour in 2010 differs significantly from other years.

6.5.1 Diurnal Plots

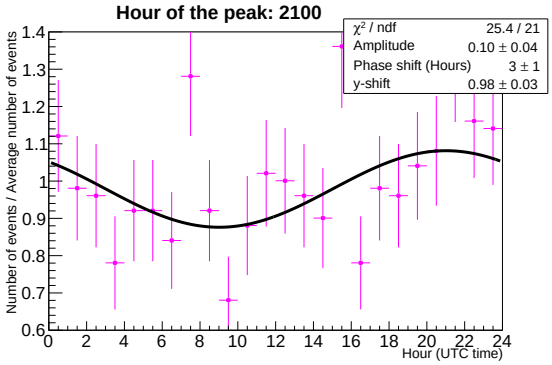
Figures 6.16, 6.17 and 6.18 show the number of events detected every hour in UTC time throughout each year, from 2004 to 2017. The data points were normalised by dividing each bin content with the average value (sum of all events in every hourly bin divided by 24 hours). The fits applied consist of a cosine function shown as

$$y(t) = A\cos\left(2\pi\frac{(t + t_{shift})}{24}\right) + y_{shift} \tag{6.8}$$

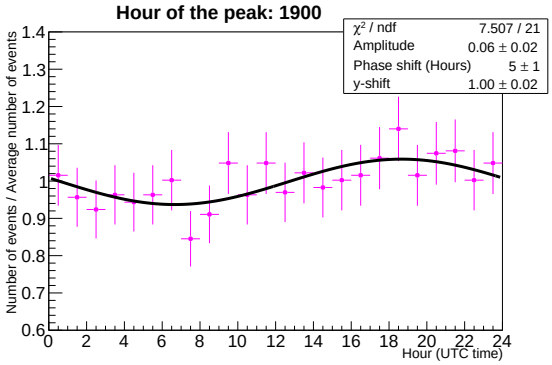
where *y(t)* is the number of events as a function time, A is the amplitude parameter, t is the time in hours, t_{shift} is the phase shift parameter in hours and y_{shift} is the shift upwards in the y-axis. The period was set to be 24 hours. The hour of the peak in the function was reported in every annual plot to show when there is a peak in the number of events in UTC hours. It is also noted that the phase shift represents the shift of the peak of the function to the left in UTC hours. For example, the phase shift obtained for the year 2004 in Figure 6.16(a) is 3 hours, indicating that the hour of the peak is at 2100 hours, which is 3 hours before midnight.

Then, the amplitudes were extracted and plotted in Figure 6.19 showing the evolution. It can be seen from Figure 6.20 that applying weather corrections to the hourly rates of events above 2 EeV, causes a reduction in amplitude. A rough estimation of the uncorrected amplitude is ~ 3 % while the corrected amplitude is ~ 2 %, noting that this is from an analysis of hourly rates with units of day^{-1} km^{-2} rather than a normalised hourly number of events. In this work, the biggest amplitude in the normalised hourly number of events was found to be 0.10 ± 0.04 and this was in 2004. However, it is noted that this might be due to the size of the array being the smallest in this year. A horizontal line fit was performed and it was found that the best constant amplitude was obtained to be 3.7 ± 0.4 %. This amplitude does not agree with the amplitude obtained from published work shown Figure 6.20 and it is unknown as to why this is the case. This horizontal line fit gave a reduced χ^2 value of 1.2 and a probability of 0.303. This probability lies in the deviation value between 0 and 2σ, deeming this fit to be suitable, since the error bars are large enough. An interesting feature is that the year 2010 has one of the highest amplitudes in the number of events, with the exception of 2004 and 2005, which were notably due to the small array size.

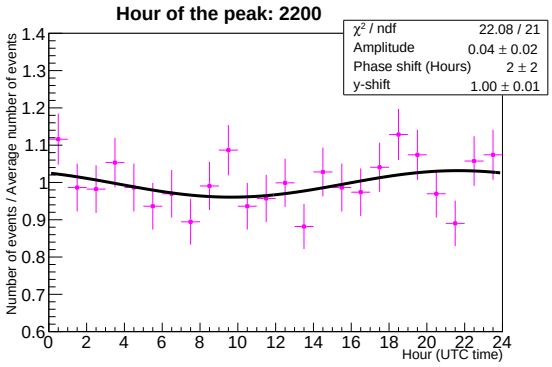
The same process was done with the phase shift in hours and plotted in Figure 6.21. However, since the phase shift is between 0000-2359 hours, to avoid discontinuity between the points, the phase shifts were allowed to oscillate around 2400 hours with a leeway given up to 2600 hours (0200 hours). The fitted constant phase was found to be 23.3 ± 0.4 hours; when converted to the 24 hour



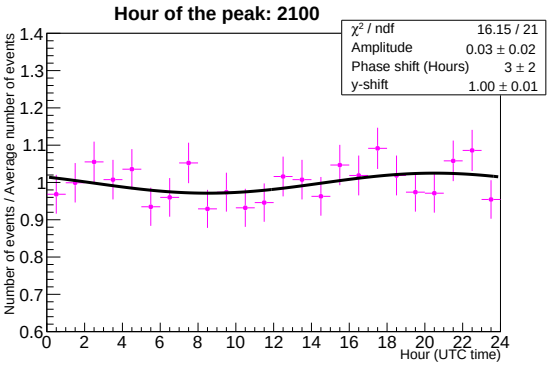
(a) Number of events every hour in 2004



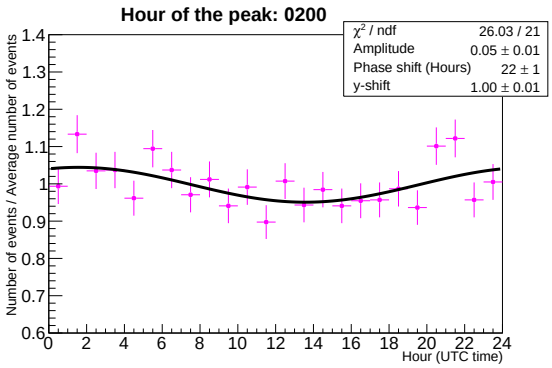
(b) Number of events every hour in 2005



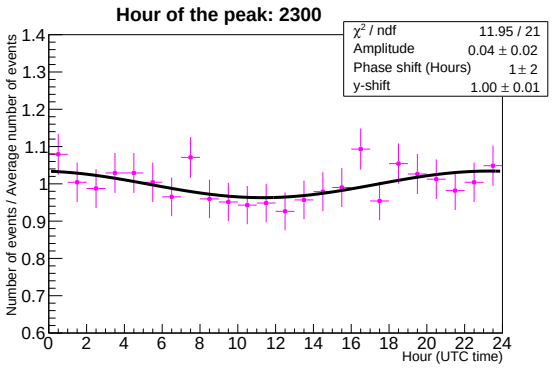
(c) Number of events every hour in 2006



(d) Number of events every hour in 2007

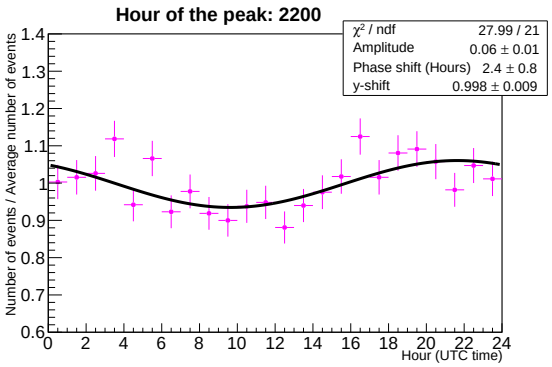


(e) Number of events every hour in 2008

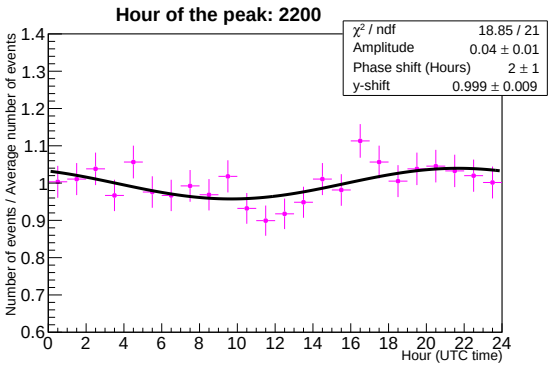


(f) Number of events every hour in 2009

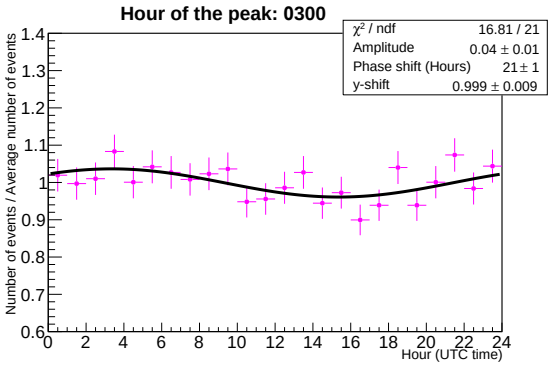
Figure 6.16: Normalised number of events every hour from 2004 to 2009. The hour of the peak of the function is reported in every plot title in UTC time.



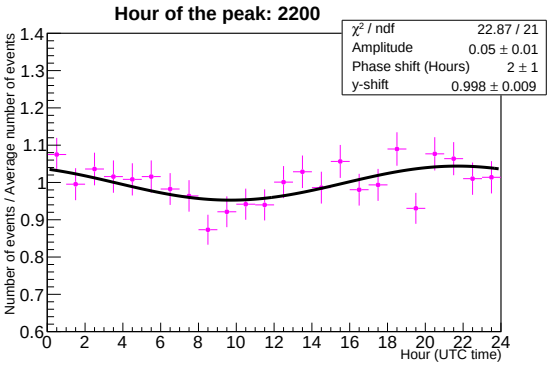
(a) Number of events every hour in 2010



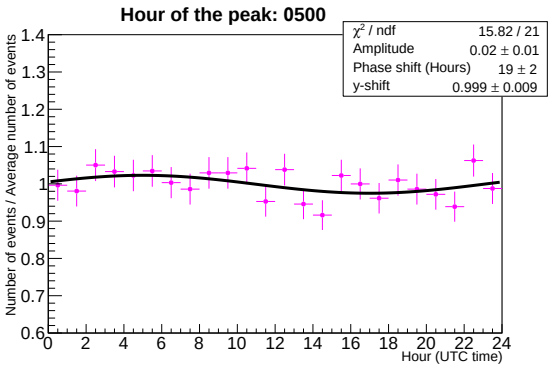
(b) Number of events every hour in 2011



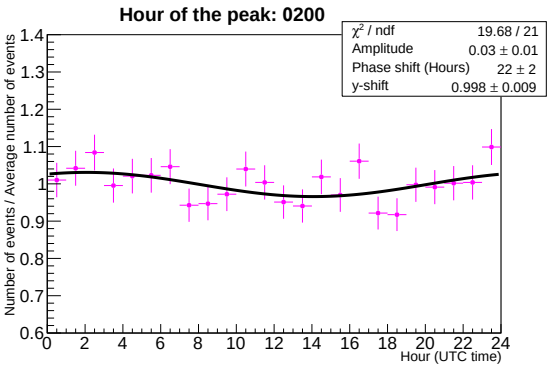
(c) Number of events every hour in 2012



(d) Number of events every hour in 2013

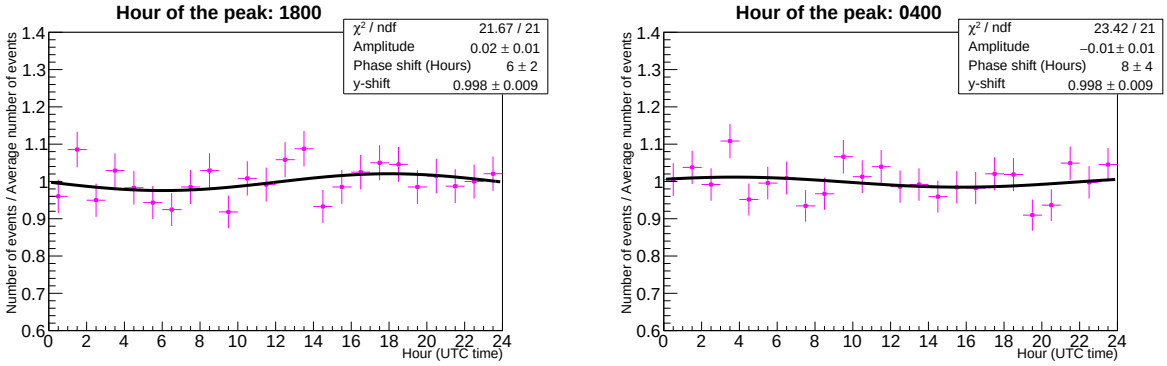


(e) Number of events every hour in 2014



(f) Number of events every hour in 2015

Figure 6.17: Normalised number of events every hour from 2012 to 2017. The hour of the peak of the function is reported in every plot title in UTC time.



(a) Number of events every hour in 2016 (b) Number of events every hour in 2017

Figure 6.18: Normalised number of events every hour in 2016 and 2017. The hour of the peak of the function is reported in every plot title in UTC time.

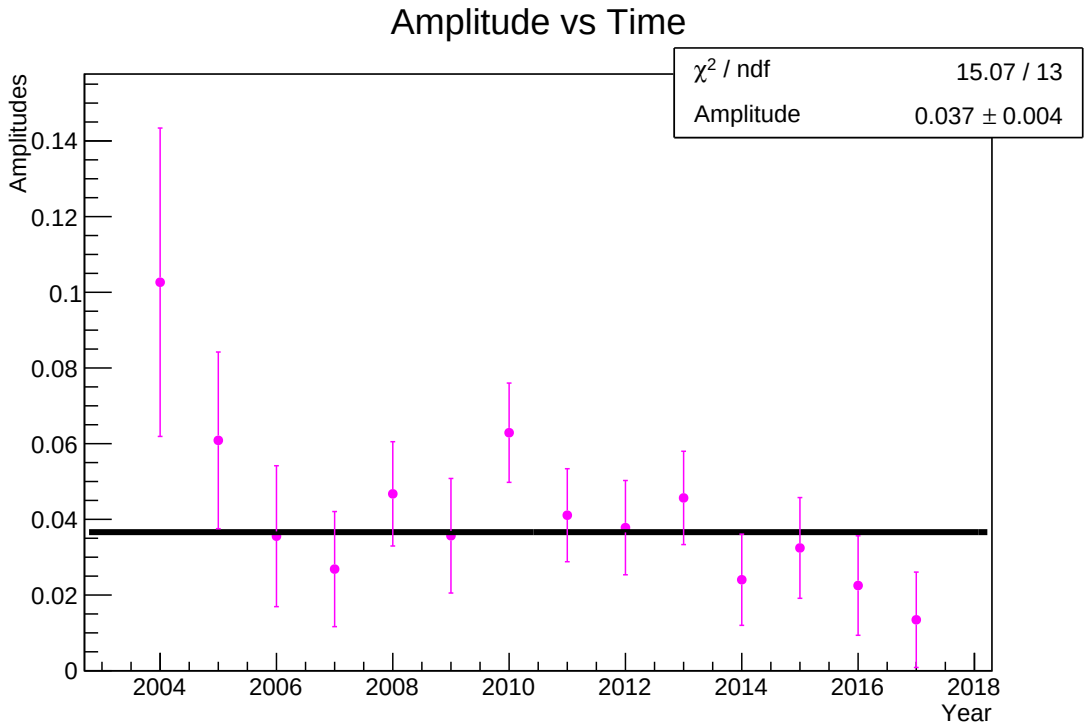


Figure 6.19: Amplitudes for each year between 2004 and 2017. The probability obtained in the horizontal line fit is 30 %.

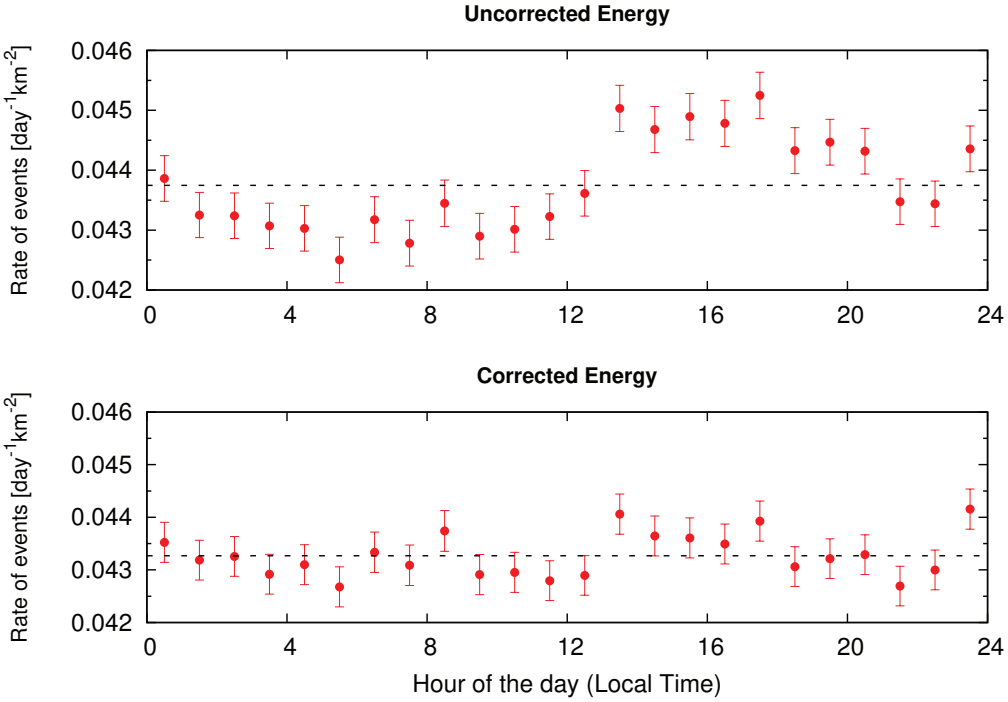


Figure 6.20: Hourly rates before (top) and after (bottom) weather corrections for a threshold of 2 EeV for the 1500 m array [114]. Horizontal dashed lines indicate the average values.

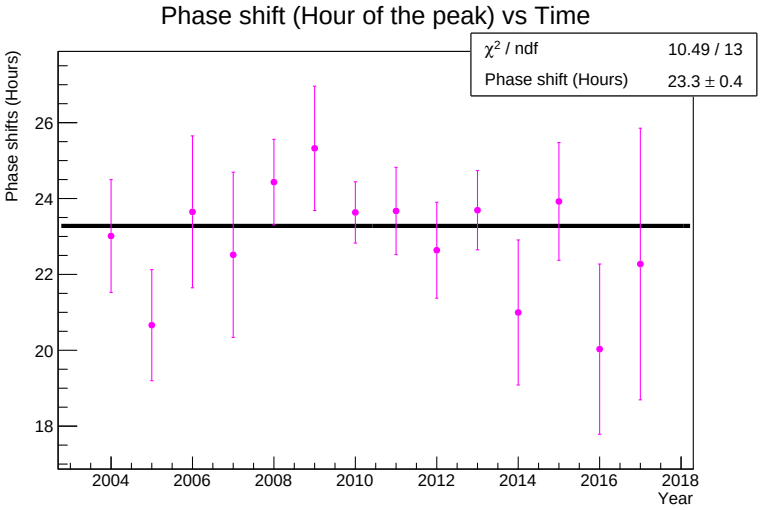


Figure 6.21: Hour of the peak from diurnal plots between 2004 to 2017. The probability obtained in the horizontal line fit is 65 %.

system, the phase was found to be 2318 hours $\pm$ 24 minutes. The probability of this fit is obtained to be 0.65 which indicates that the phase shifts from year to year are not at all different to one another.

6.5.2 Other Checks - Weather Station Data and Daily Rates

For completeness, it was checked to see if there is an anomaly in the pressure and density measured for all years at Los Leones, Los Morados, Loma Amarilla, Coihueco and the Central Laser Facility (CLF). Figures B.14-B.23 in Appendix B show the air density as a function of time each year using data available from 2005 to 2017. Following this are Figures B.24-B.33 showing the variations in pressure as a function of time. From these plots, it can be seen that there were no anomalies found for air density and pressure in the year 2010 for all the sites with weather stations. Plots of daily rates as a function of time for each year were also generated to see if the behaviour in 2010 was significantly different to other years. The number of events detected each day and the number of hexagons were obtained to calculate the ratio as a daily rate. These are shown in Figures B.34 and B.35. The only point notable is that the year 2010 was the only year where there were no days with zero rates. Other than this, the normalisation did not differ from other years and there were no significant outliers.

6.6 Conclusions

Analyses of rates as a function of time seen by Los Leones, Los Morados, Loma Amarilla and Coihueco were done using hybrid events. Each FD site was seen to exhibit a distinct behaviour in the observed rates as a function of time. It was concluded that Loma Amarilla observed a linearly increasing rate. As for Los Leones and Los Morados, another functional form is needed for a better fit. Coihueco sees the most complex evolution of rates. There is a notable rate spike in 2010 for Los Leones, Los Morados and Coihueco. Loma Amarilla does face a rate spike in 2015 which correlates with a hotspot in its vicinity as shown in the rate maps but there is no rate spike seen in 2010 which is slightly odd. As for Los Leones and Los Morados, the rates seem uniform near the site position so it is unknown as to what kind of function is needed to improve the deviation. Most notably, the change per year (%) in rates as seen by the SD (0.58 $\pm$ 0.08 % from 2006-2017), agrees with the change per year in rates as seen from the FD sites (at most only 1σ away). Finally, plots of the fractional changes in rates and threshold energy were generated. These showed that the evolution of true energies with time are reducing as seen by Loma Amarilla while at the other sites, the changes in energies are more complex. This may be a partial contribution from the drop in true energy in 2010.

This was followed by an analysis of the rates as a function of distance between the hottest stations and the FD sites. This was to check if there are any structures

present which could affect the energy ratio as a function of aerosol transmission to the shower maximum. The rates seen from Loma Amarilla is versatile to all four fit types while at Coihueco, the behaviour is most complex. The rates seen from Los Leones tend to increase linearly as a function of distance while at Los Morados it either decreases then stays constant or decreases again after a breakpoint. These trends seem to agree reasonably well with what is seen in the rate maps. Coihueco faces two hotspots adjacent to one another in quite close vicinity, while Loma Amarilla sees a sudden hotspot a lot further away, as seen from the breakpoint position obtained, which is the furthest out of all the sites. From the rate maps, the regions in the vicinity of Los Leones is the most uniform. Despite this, there are hotspots which can be seen, but quite far away, matching the linear increase in its trend with distance. There also exists a high rate region very close to Los Morados, which matches the trend in the rate versus distance plot, then the rates start to stabilise with distance. Further work ensues for appropriate corrections to be applied so this distance dependence can be eliminated and the energy ratio as a function of transmission to shower maximum can be interpreted accurately.

Lastly, the high rate in 2010 was investigated further through a series of diurnal plots from the year 2004 to 2017. These diurnal plots shows an oscillating behaviour in the number of events detected at every hour, which was fitted with a cosine function. The amplitudes and phase shifts from these fits were obtained and plotted separately to see if there is a significant variation throughout the 14-year long data set. Horizontal lines were fitted onto these amplitude and phase plots. A comparison of the average amplitude in this work was done to previous published work and it was found that the result deviates by $\sim 3\sigma$. It is unknown as to why this is the case. Additionally, it was noted that 2010 had the second largest amplitude. The horizontal line fits to the amplitudes and phase shifts resulted in a probability with a deviation well within 0-2σ indicating a preferred constant trend as a function of time. The phase shifts seem to be constant around 2300 hours.

It was also checked if weather corrections were done properly throughout the dataset. Plots of air density and pressure changes with time as well as daily rates were generated to see if the year 2010 had a behaviour that was significantly different to other years. However, no notable features were found that distinguished 2010 from other years indicating that there were no systematic changes in the weather corrections applied.

Chapter 7

Conclusions and Future Work

The primary focus of this thesis is the stability of energy assignments done concurrently by the surface detector and fluorescence detector at the Pierre Auger Observatory. The aim was to provide an overview of the long term trends of the assigned FD and SD energies, checking for a dependence in time and detector location, then quantifying this dependence, subject to many conditions.

Analysis begins with a slight extension to the work done by Phong Nguyen, on an energy scale ratio E_{FD}/S_{38}, defined as the ratio of two independent measurements obtained by the FD and SD. Using high quality hybrid data from 2004 to the end of 2017, the stability of FD energies was checked by investigating the long term trend of this ratio. A quantitative analysis was done to observe the possibility of any correlation with calibration changes, weather and geomagnetic field corrections, as well as an updated aerosol database. It was found that even after all applied corrections, there is still a downwards drift of 0.7 % per year post-2014 and a seasonal modulation with amplitude 0.54 % in the energy scale ratio. The FD was also subject to multiple UV filter and mirror cleaning campaigns as well as a drum calibration done in 2010, which could affect the FD energy. The direct ratio; E_{SD}/E_{FD}, was plotted as a function of time for the 24 telescopes. This analysis found that:

- The E_{SD}/E_{FD} ratio for most telescopes fit best with a linear-linear type fit, with some exceptions shown by telescopes 1 and 3 at Loma Amarilla and telescope 4 at Los Leones. This ratio tends to increase steeply, then exhibit a breakpoint after which it continues to increase less steeply.

- The E_{SD}/E_{FD} ratio had breakpoints in time at around 2010 for most telescopes at Coihueco (Figure 4.30), but varied greatly for others.

- Due to the high frequency of cleaning dates, most of the data did not exhibit a discontinuity after most cleans, with the exception of one done at the start of 2014; where the ratio drops.

Then, to monitor the stability of E_{SD}, the rate of events detected above the full efficiency threshold as a function of time and station location across the array was investigated. Using SD data from 2004-2017 and Hexalife files, the number of events detected and uptimes of the hottest stations involved with those events

were obtained to calculate these rates. This was also done for four separate cohorts of stations; commissioned in 2004, 2005, 2006 and 2007. The results of this analysis are summarised as follows:

- Average rates for all stations are increasing with time. There is a hotspot of rates in the north eastern corner of the SD array near a gap in the detector coverage in the 2014-2017 time block (Figure 5.26).

- The change per year in rates from 2006 to 2017 is 0.58 ± 0.08 % which is in agreement with the work done by Lhenry-Yvon. There is a rate jump in 2010 which is prominent in both works. Rates increase with a steep slope, then exhibit a breakpoint in 2010, after which it continues to increase, but with a less steep slope (Figure 5.27(d)) These features are similar to those of the E_{SD}/E_{FD} ratio, indicating the this ratio is driven by changes in the SD energy scale.

- All four cohorts of stations agree with the linear-linear fit behaviour of rates overall (Figure 5.50). There are only two examples of strong correlations to high rate regions on the maps. This applies for stations commissioned in 2004 and 2006.

- A short study on the rates above 4.5 EeV and 6 EeV had a change per year which was indistinguishable for that of the rates above 3 EeV. Another brief study showed that the SD energy resolution had to worsen by about 4.5 % per year (indicated by Figure 5.60) to generate an increase in rate of 0.58 % per year. This means it is quite likely that the rate increase is not holistically due a worsening SD energy resolution.

Then, the SD rates were assigned to quality hybrid events to study the impact of the SD energy stability on the E_{SD}/E_{FD} ratio as a function of time for each FD site. It was found that:

- A constant evolution of rates with time at Los Leones can be excluded but a linear fit could not be (Figure 6.6(a)). None of the fits can be excluded for Los Morados, but there is no strongly indicated trend either. A linearly increasing rate was seen from Loma Amarilla (Figure 6.6(c)) with a rate spike in 2015, correlating with a hotspot in the rate map. None of the fits done were suited to the trend seen from Coihueco.

- A rate spike was seen in 2010 from Los Leones, Los Morados and Coihueco and not from Loma Amarilla (Figure 6.5). This high rate in 2010 is not shown prominently in the 2010 rate map, compared to other years. From the event rate and threshold energy relationship, true energies seen from Loma Amarilla are decreasing as a function of time. The other sites show a drop in true energies in 2010.

- The rate changes in stations associated with hybrid events for all FD sites, agree with the average rate change over the whole SD array.

Rates as a function of distance between the FD sites and hottest stations were also investigated since trends present could affect interpretation of the energy ratio as a function of aerosol transmission to the shower maximum. Results showed that:

- A constant trend can be excluded from the rates seen from all FD sites except for Loma Amarilla (Figure 6.12(c)). A linearly increasing trend fits best to the rates seen from Los Leones (Figure 6.13(a)). From Los Morados, the rates tend to decrease steeply, then less steeply (Figure 6.15(b)). These trends agree with what is observed on the rate maps.

- None of the fits were suited to the trend seen from Coihueco. The rate map (Figure 5.7(b)) shows two hotspots seen from telescope 1 and 6 which could contribute to the complexity seen in the trend.

The last study was an investigation of the high rate seen in 2010. It was found that the trend in the accumulated number of events every hour in 2010 was comparable to the trends in all other years (Figure 6.17(a)) The only notable feature is how 2010 has the second largest fitted amplitude to the number of events every hour (Figure 6.19).

There are a few suspected causes of the rate trends seen as a function of position and time. One of which is the degradation of station sensitivity seen by the area over peak decay as a function of time. This could be due to a decrease in the Tyvek liner reflectivity of the stations. Another cause is a worsening SD energy resolution. This may be due to an increase in the number of non-triggered stations further away from the shower axis as a function of time. From the analysis presented in this thesis, further work can be done to investigate these possible causes and the drifting energy ratio. These investigations may entail:

- Further collection of hybrid events and calibration data to see if a residual amplitude and drift remains.

- A similar rate map analysis to be done on stations commissioned at different times, including commission years beyond 2007.

- Obtaining information on stations which have experienced freezing (as shown in a previous work [140]), then excluding these stations to see a potential effect on the change per year in rates and energy scale ratio.

- An analysis on the number of stations involved in SD events as a function of time and its effect on the SD energy resolution.

- A study on the effect of Tyvek liner degradation on S(1000) signal and its compounding effect on SD energy as well as rates above 3 EeV as a function of time.

- Using the fractional changes in rates to correct assigned SD energies and seeing how this correction improves the energy scale ratio as a function of time and aerosol transmission to the shower maximum.

Appendix A

Event Quality Cuts

Listed below are cuts used to determine the SD events, FD events and hybrid events that are suitable to be used for the analysis in Chapters 4, 5, and 6. There were also additional cuts applied to the geometry for hybrid events.

Cuts on reconstructed SD data

- In each event, the station which detected the highest amount of signal is surrounded by by 6 working stations. This is also known as the 6T5 trigger (Section 3.1.2.4).
- The energy reconstructed by the SD is more than 3 EeV.
- Each shower in that event has a zenith angle less than 60°.
- The shower was not detected during a *bad period* which is known as the time periods where there existed problems with the stations. Applying this cut allowed the acquisition only of reliable data.
- The shower was not detected during the *communications crisis* which is an extended period of time where there were fast and frequent disconnections between the stations and CDAS. This occurred in 2009 between the months of June and November.

Cuts on reconstructed FD data

- The energy reconstructed by the FD is greater than 3 EeV.
- The maximum relative uncertainty on the energy is 0.2.
- Each shower in that event has a zenith angle less than 60°.
- The shower was not detected during a *bad period* where events had bad calibration constants, GPS glitches and during an aerosol bad period caused by a high amount of backscattered light interpreted as aerosols due to clouds.
- The maximum value of the vertical aerosol optical depth (VAOD) which is the measure of the attenuation of light in the atmosphere, is 0.1.

- Information on the Mie scattering of light in the Mie database, exists for the event.

Additional Figures

B.1 Energy Ratio for Each Telescope

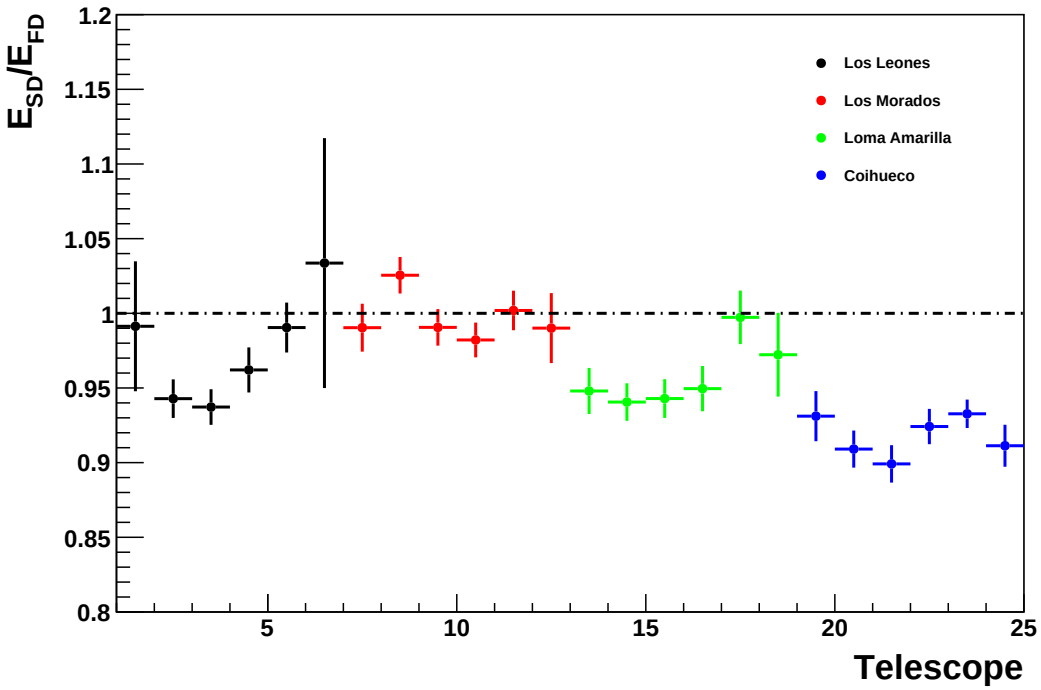
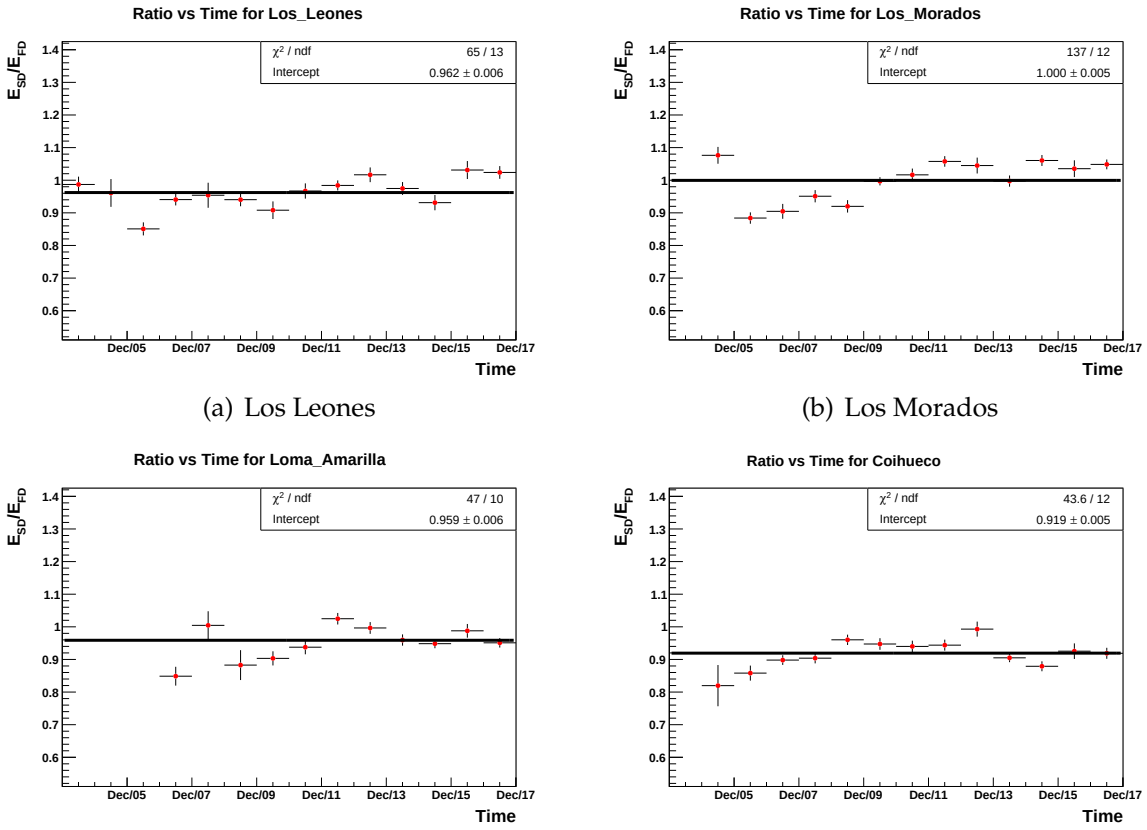


Figure B.1: A weighted plot of E_{SD}/E_{FD} as a function of telescope. The weighted means and error on the means are calculated as shown in Equations 4.4 and 4.6. The energy ratio of telescopes at Los Leones are black, while for Los Morados they are red, Loma Amarilla are green and Coihueco are blue. The x-axis represents the telescope numbers 1 to 6 for each FD site, where telescope 1 of Los Leones is at x = 1.5 and telescope 6 at Coihueco is at x = 24.5. A black horizontal dashed line at $E_{SD}/E_{FD} = 1$ is shown for display.

Figure B.1 shows the weighted values of the energy ratio as a function of telescope. These values depend on the energy resolution of the SD and FD described in Section 4.3.2. Comparing this plot with the non-weighted version (Figure 4.9), most of these values are weighted down with an exception to the point with the largest error bar (Los Leones telescope 6). It is also noted that the energy ratio for each telescope matches the intercepts obtained from the horizontal line fits (shown in Table 4.4).

B.2 Energy Ratio for Each FD Site



(a) Los Leones

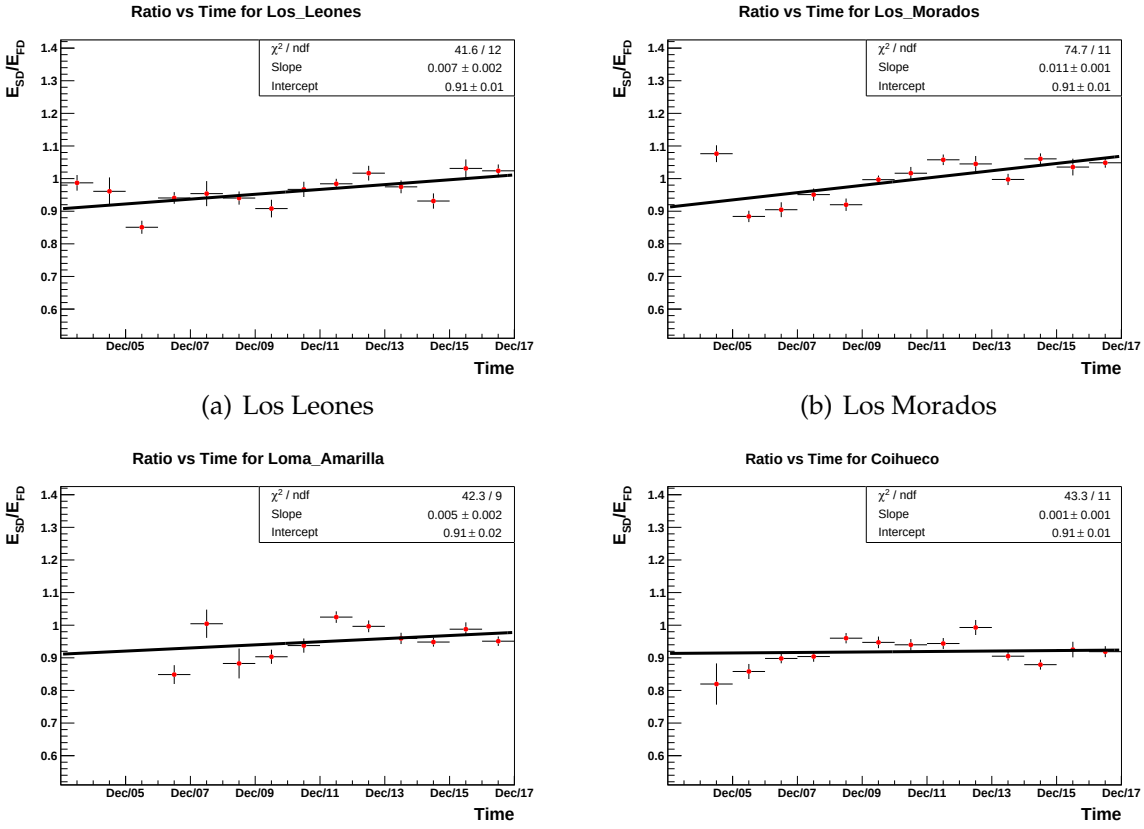
(b) Los Morados

(c) Loma Amarilla

(d) Coihueco

Figure B.2: Ratio of SD energy to FD energy as a function of time for each FD site fitted with a horizontal line. The ratios are weighted according to the energy resolution functions for both SD and FD. It is also noted that since this illustrates the ratio of two energies for the same events, in theory it should yield 1 as a constant value throughout time.

Figures B.2 - B.5 show four different fits applied to the energy ratio for each FD site. The goodness-of-fit results are very poor, with values of reduced χ^2 beyond 3, indicating the need for a non-linear function.



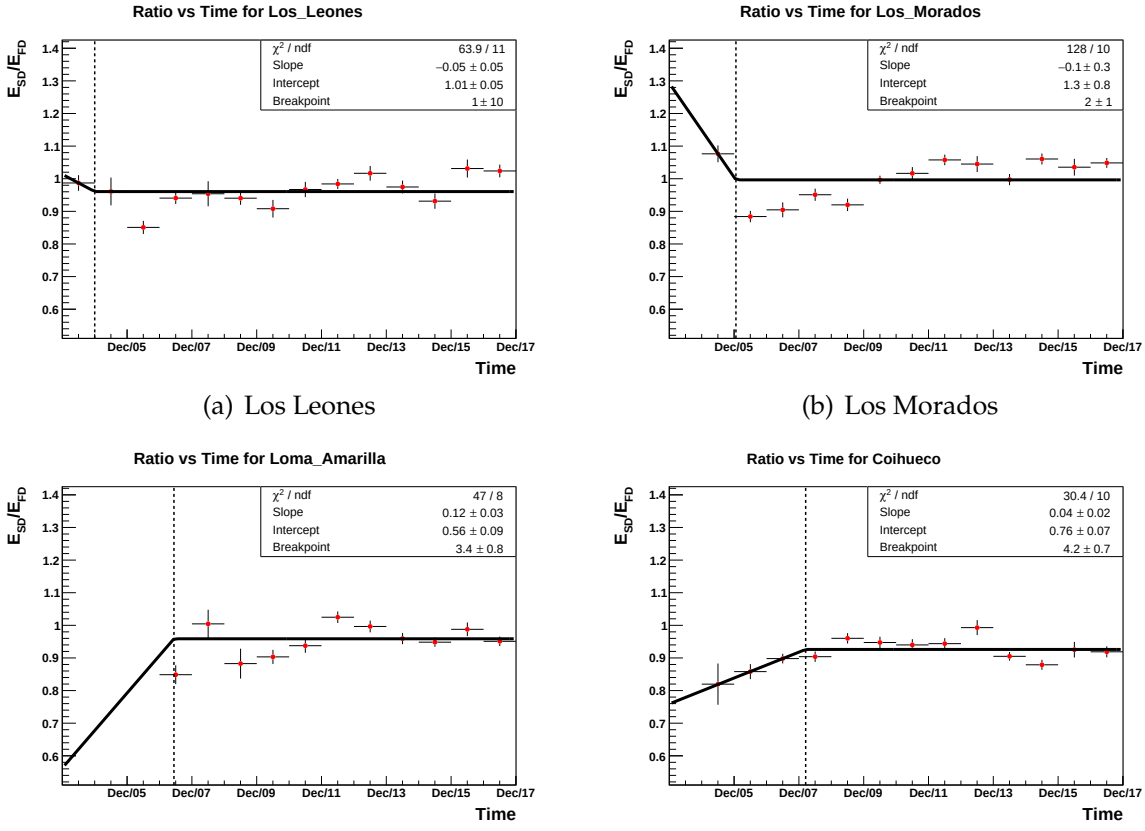
(a) Los Leones

(b) Los Morados

(c) Loma Amarilla

(d) Coihueco

Figure B.3: Ratio of SD energy to FD energy as a function of time for each FD site fitted with a line of non-zero slope. The ratios are weighted according to the energy resolution functions for both SD and FD. It is also noted that since this illustrates the ratio of two energies for the same events, in theory it should yield 1 as a constant value throughout time.



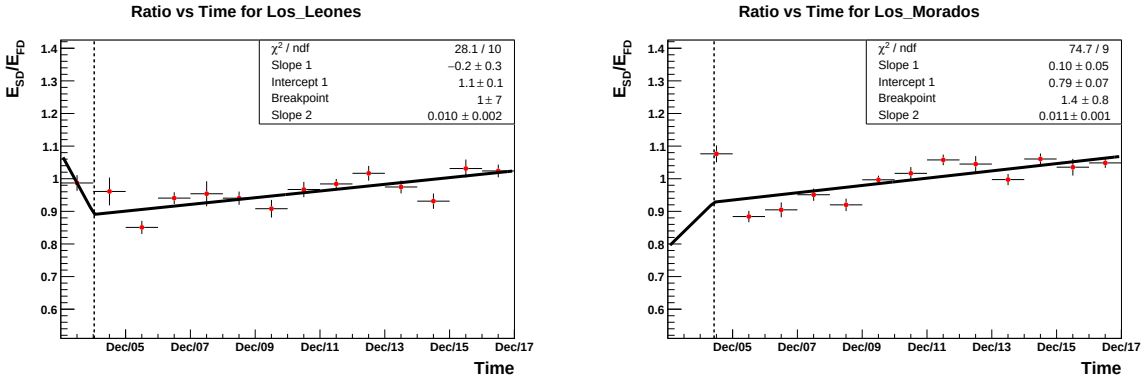
(a) Los Leones

(b) Los Morados

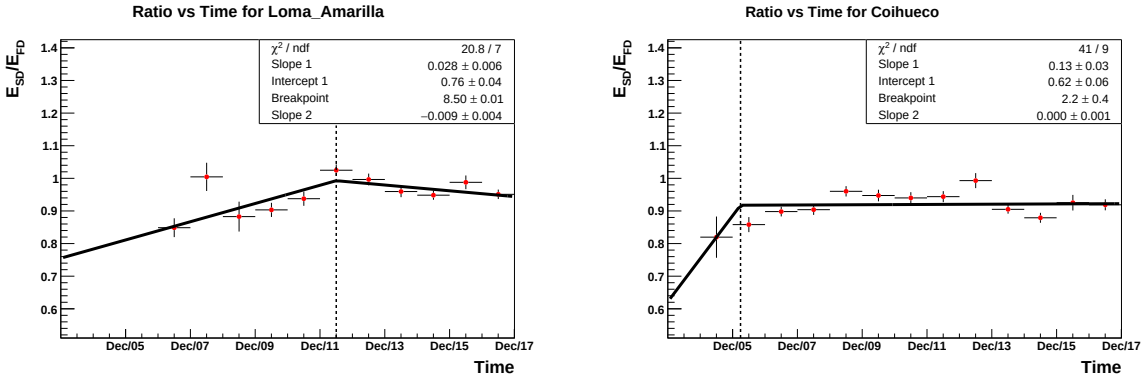
(c) Loma Amarilla

(d) Coihueco

Figure B.4: Ratio of SD energy to FD energy as a function of time for each FD site fitted with a line with non-zero slope and a horizontal line separated by a breakpoint. The ratios are weighted according to the energy resolution functions for both SD and FD. It is also noted that since this illustrates the ratio of two energies for the same events, in theory it should yield 1 as a constant value throughout time.



(a) Los Leones (b) Los Morados

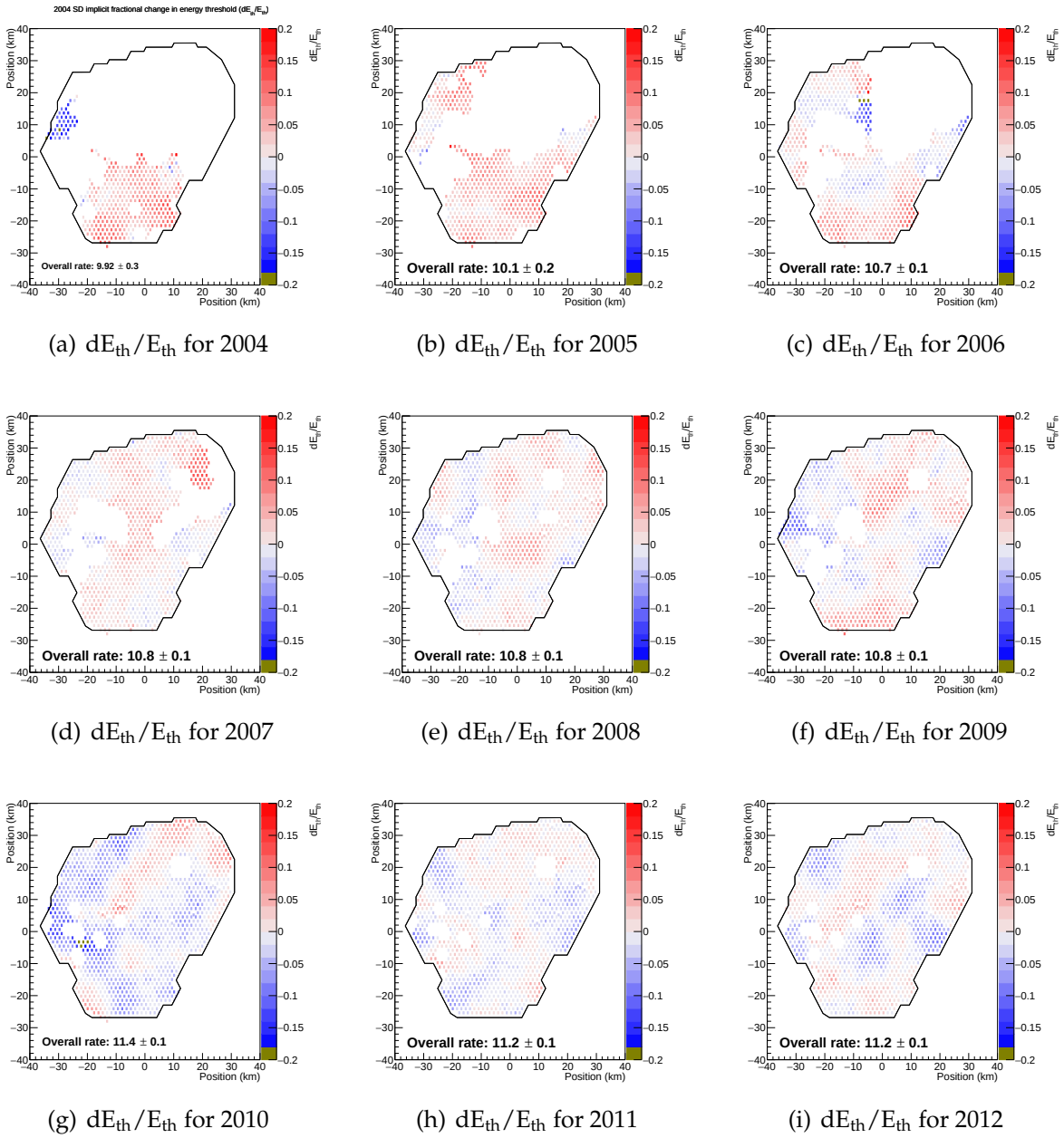


(c) Loma Amarilla (d) Coihueco

Figure B.5: Ratio of SD energy to FD energy as a function of time for each FD site fitted with a two lines of non-zero slope separated by a breakpoint. The ratios are weighted according to the energy resolution functions for both SD and FD. It is also noted that since this illustrates the ratio of two energies for the same events, in theory it should yield 1 as a constant value throughout time.

B.3 Fractional Change in Event Rates and Threshold Energy On Yearly Maps

B.3.1 Using Yearly Rates



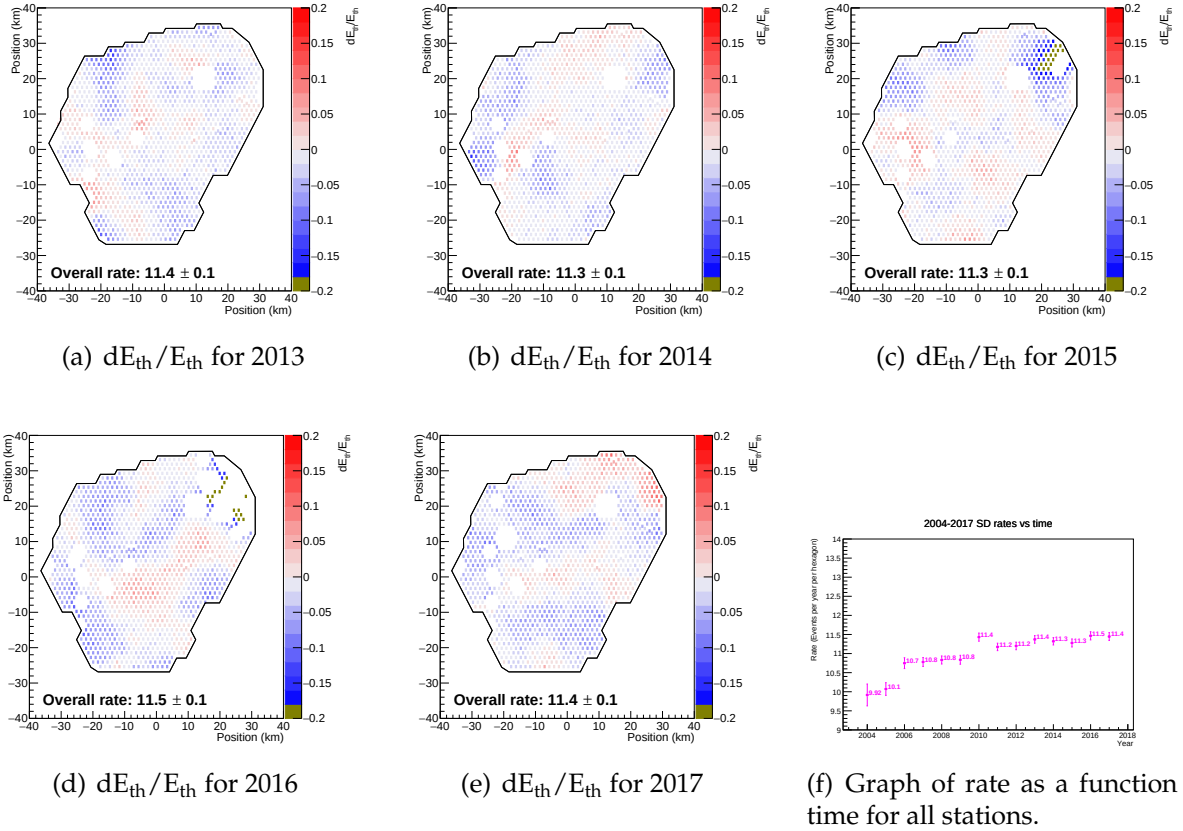
(a) dE_{th}/E_{th} for 2004 (b) dE_{th}/E_{th} for 2005 (c) dE_{th}/E_{th} for 2006

(d) dE_{th}/E_{th} for 2007 (e) dE_{th}/E_{th} for 2008 (f) dE_{th}/E_{th} for 2009

(g) dE_{th}/E_{th} for 2010 (h) dE_{th}/E_{th} for 2011 (i) dE_{th}/E_{th} for 2012

Figure B.6: Fractional difference in threshold energy, with overall rates listed at the bottom left from 2004 to 2012.

Figures B.6 and B.7 show how the fractional threshold energy change varies from 2004 to 2017. The fractional change in event rates are calculated using the yearly rate values shown in Figure B.7(f). These results show a variation in the



(a) dE_{th}/E_{th} for 2013

(b) dE_{th}/E_{th} for 2014

(c) dE_{th}/E_{th} for 2015

(d) dE_{th}/E_{th} for 2016

(e) dE_{th}/E_{th} for 2017

(f) Graph of rate as a function time for all stations.

Figure B.7: Fractional difference in threshold energy, with overall rates listed at the bottom left from 2013 to 2017. Figure B.7(f) is shown as reference to the overall rates.

rates across the array in each year. In 2004, there seems to be a strong negative fractional energy change in the western region of the array.

B.3.2 Using One Average Rate

To show the variation in rates across the array as a function of year, a fixed reference rate for all years was used and this was obtained as the y-intercept obtained from the horizontal line fit done in Figure 5.27(a) of Section 5.4. This rate value is obtained as 11.14 ± 0.03 events/year/hexagon. As a cross-check, $\frac{\text{d(Event rate)}}{\text{Event rate}}$ is first plotted on maps and on a graph of rates as a function of time. The final result is shown in Figures B.8 and B.9. From the graph shown in Figure B.9(f), $\frac{\text{d(Event rate)}}{\text{Event rate}}$ can be calculated as follows:

$$\frac{\text{d(Event rate)}}{\text{Event rate}} = \frac{\text{Rate for all stations - Reference rate}}{\text{Reference rate}} \tag{B.1}$$

or simply as

$$\frac{\text{d(Event rate)}}{\text{Event rate}} = \frac{\text{Rate for all stations}}{\text{Reference rate}} - 1 \tag{B.2}$$

where the rate for all stations is obtained from Figure B.7(f). The uncertainties in B.9(f) are calculated as shown in the following:

$$\sigma^2_{\text{d(Event rate)/Event rate}} = \sigma^2_{\text{Reference rate}} \times \left(\frac{\partial(\text{d(Event rate)/ Event rate})}{\partial \text{Reference rate}} \right)^2 + \sigma^2_{\text{Rate for all stations}} \times \left(\frac{\partial(\text{d(Event rate)/ Event rate})}{\partial \text{Rate for all stations}} \right)^2 \tag{B.3}$$

So, since

$$\left(\frac{\partial(\text{d(Event rate)/ Event rate})}{\partial \text{Reference rate}} \right) = \frac{-\text{Rate for all stations}}{\text{Reference rate}^2} \tag{B.4}$$

and

$$\left(\frac{\partial(\text{d(Event rate)/ Event rate})}{\partial \text{Rate for all stations}} \right) = \frac{1}{\text{Reference rate}} \tag{B.5}$$

Then

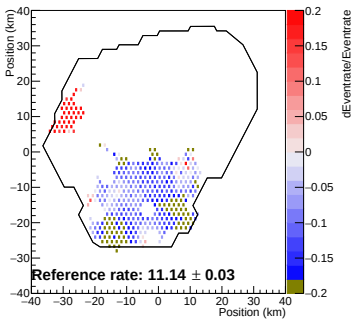
$$\sigma^2_{\text{d(Event rate)/Event rate}} = \sigma^2_{\text{Reference rate}} \times \left(\frac{\text{-Rate for all stations}}{\text{Reference rate}^2} \right)^2 + \sigma^2_{\text{Rate of stations}} \times \left(\frac{1}{\text{Reference rate}} \right)^2 \tag{B.6}$$

As expected, the fractional change in event rate is more negative in the earlier years as shown by the maps in Figure B.8. Then from 2008 onwards, there is more of a balance between the positive and negative fractional event rates across the array. It is also noted that in the north-eastern corner of the array, the fractional change in event rates are high (and positive) in 2015 and 2016. This is followed by a negative fractional change in event rates in that same area in 2017. This corresponds to the high number of stations in this area which were not sending data to the the Central Data Acquisition System, also known as ”black tanks”.

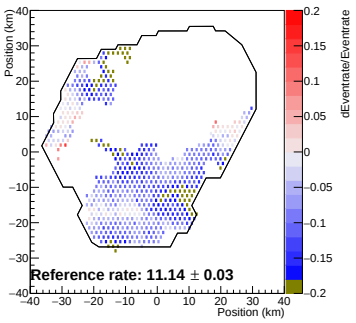
Plots showing the fractional changes in threshold energy are generated by taking the fractional event rate changes in Figures B.8 and B.9 and multiplying each z-bin value by -1/2, as stated by Equation 5.22. The results are shown in Figures B.10 and B.11 with a graph of the fractional energy changes as a function of time shown in Figure B.11(f). The uncertainties for the fractional changes in threshold energy represented by the error bars in Figure B.11(f), were calculated by propagating errors from $\sigma_{\text{dEvent rate/Event rate}}$ to obtain

$$\sigma_{\text{dE}_{\text{th}}/\text{E}_{\text{th}}} = \frac{1}{2} \times \sigma_{\text{d(Event rate)/Event rate}} \tag{B.7}$$

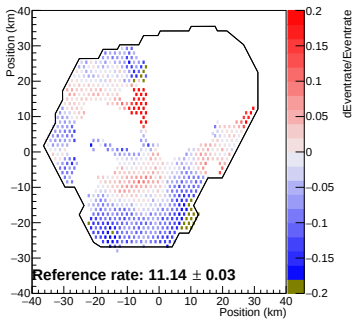
It can be seen from Figure B.11(f) that as time progresses, the fractional change in energy decreases from 0.06 to -0.01, since there are as many stations with a negative fractional change in threshold energy as there are positive ones.



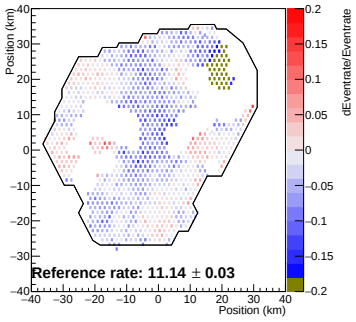
(a) d(Event rate)/Event rate 2004



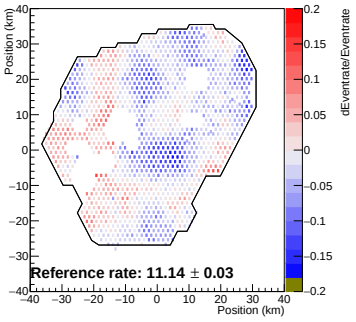
(b) d(Event rate)/Event rate 2005



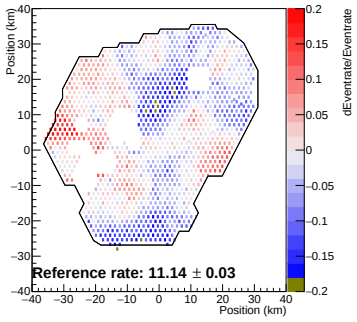
(c) d(Event rate)/Event rate 2006



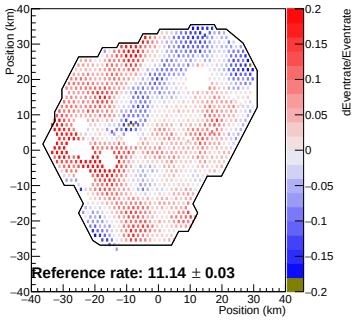
(d) d(Event rate)/Event rate 2007



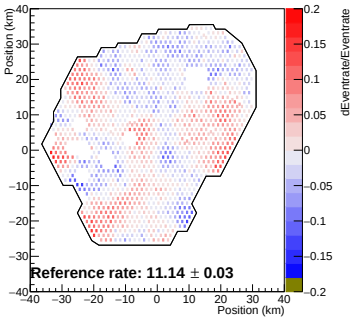
(e) d(Event rate)/Event rate 2008



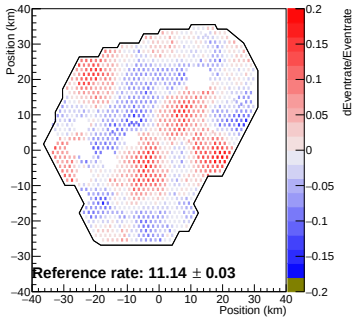
(f) d(Event rate)/Event rate 2009



(g) d(Event rate)/Event rate 2010

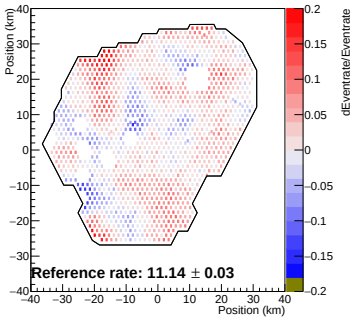


(h) d(Event rate)/Event rate 2011

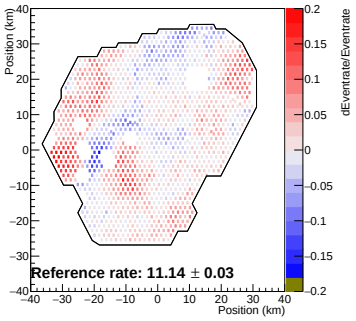


(i) d(Event rate)/Event rate 2012

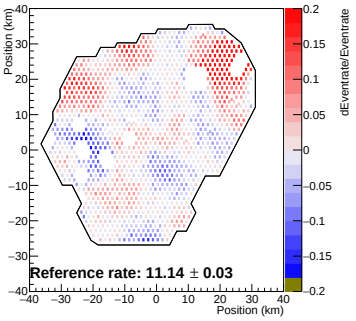
Figure B.8: Fractional difference in event rates, with a reference rate of 11.14 ± 0.03 events/year/hexagon.



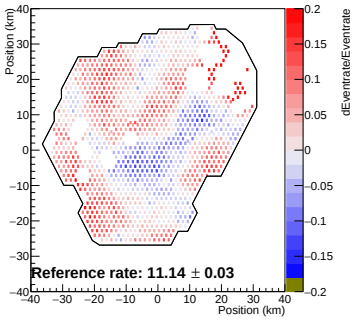
(a) d(Event rate)/Event rate 2013



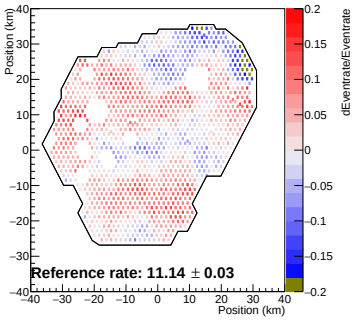
(b) d(Event rate)/Event rate 2014



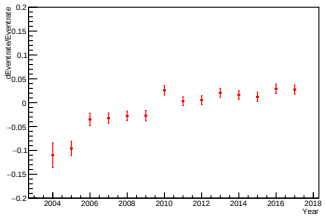
(c) d(Event rate)/Event rate 2015



(d) d(Event rate)/Event rate 2016

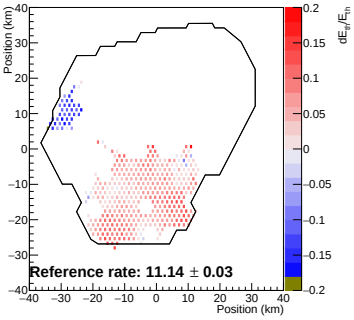


(e) d(Event rate)/Event rate 2017

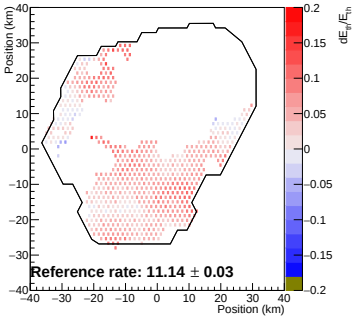


(f) d(Event rate)/Event rate as a function of time

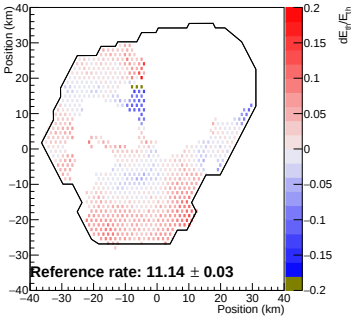
Figure B.9: Fractional difference in event rates, with a reference rate of 11.14 ± 0.03 events/year/hexagon.



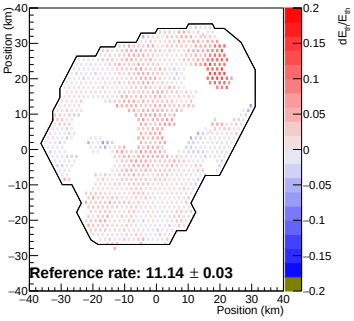
(a) dE_{th}/E_{th} for 2004



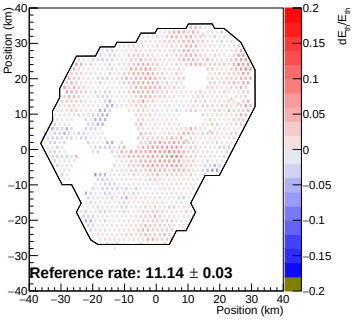
(b) dE_{th}/E_{th} for 2005



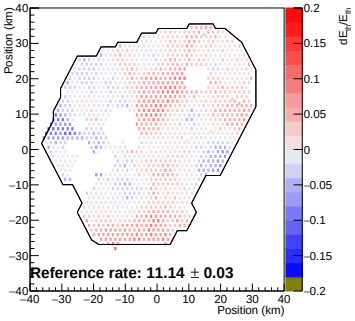
(c) dE_{th}/E_{th} for 2006



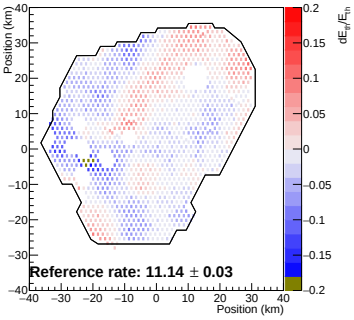
(d) dE_{th}/E_{th} for 2007



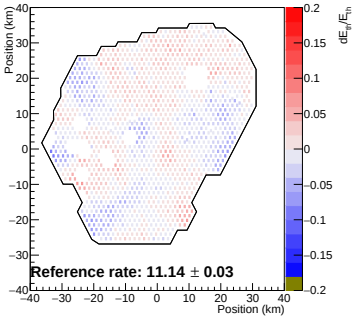
(e) dE_{th}/E_{th} for 2008



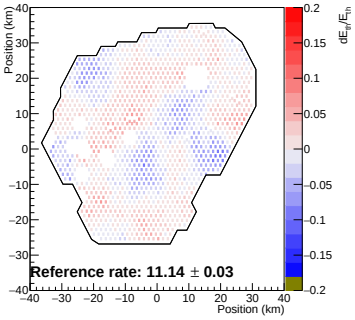
(f) dE_{th}/E_{th} for 2009



(g) dE_{th}/E_{th} for 2010

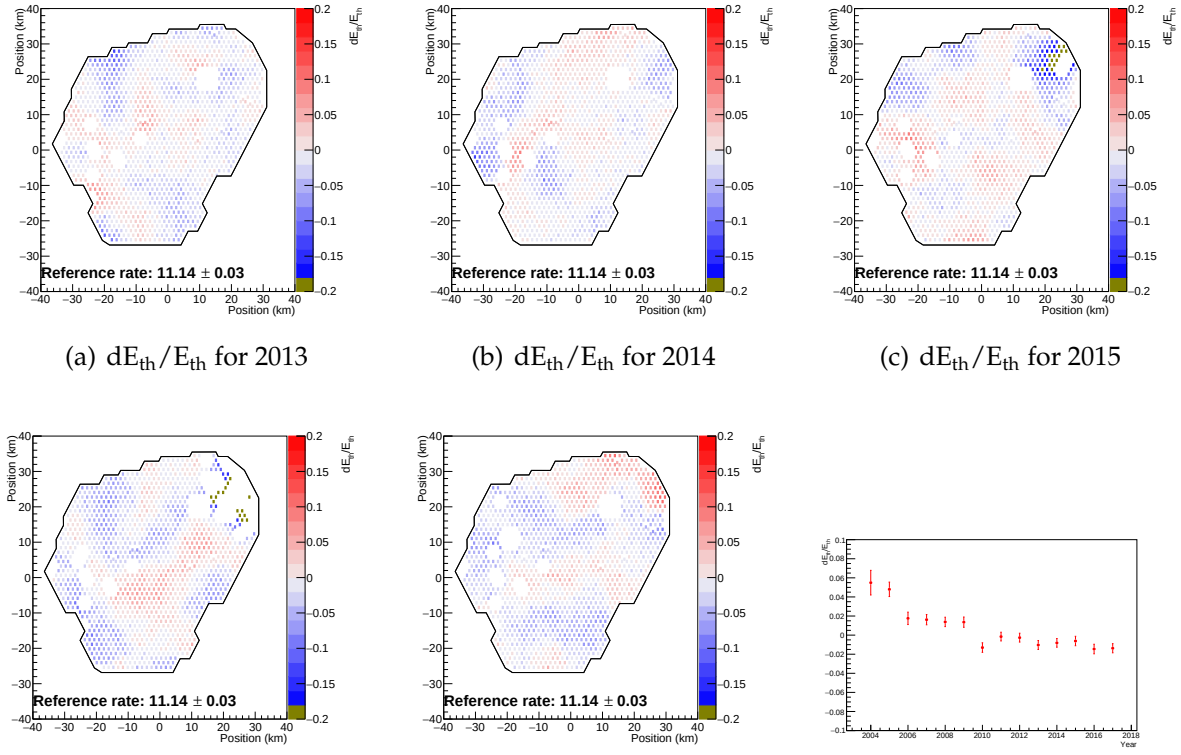


(h) dE_{th}/E_{th} for 2011



(i) dE_{th}/E_{th} for 2012

Figure B.10: Fractional difference in threshold energy, with a reference rate of 11.14 ± 0.03 events/year/hexagon.



(a) dE_{th}/E_{th} for 2013

(b) dE_{th}/E_{th} for 2014

(c) dE_{th}/E_{th} for 2015

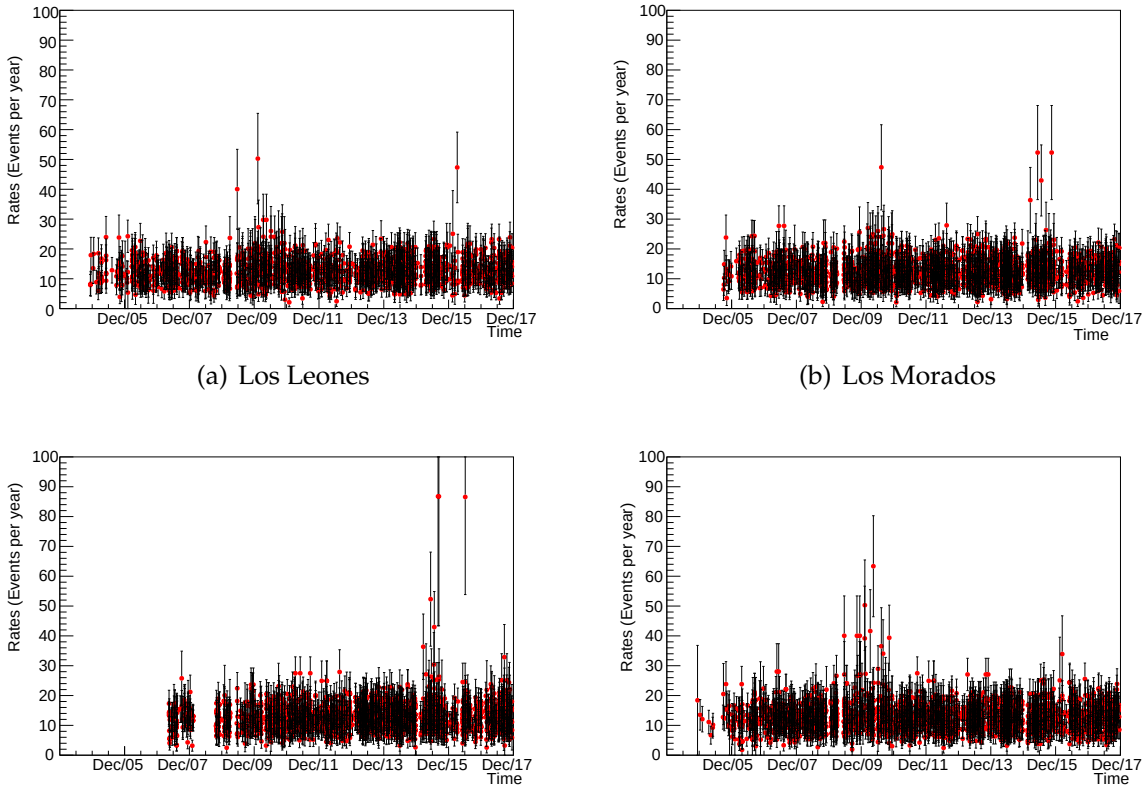
(d) dE_{th}/E_{th} for 2016

(e) dE_{th}/E_{th} for 2017

(f) dE_{th}/E_{th} as a function of time

Figure B.11: Fractional difference in threshold energy, with a reference rate of 11.14 ± 0.03 events/year/hexagon. This is accompanied by a graph showing the fractional threshold energy changes as a function of time.

B.4 Scatter Plots of Rates as a Function of Time and Distance From Each of the FD Sites



(a) Los Leones

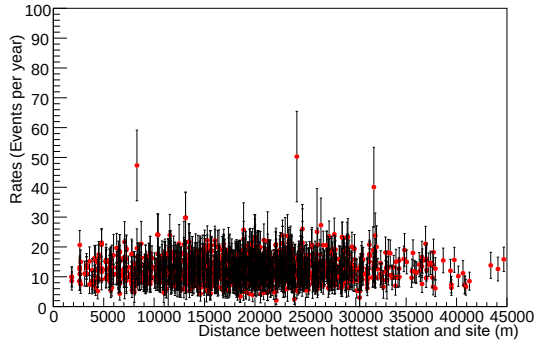
(b) Los Morados

(c) Loma Amarilla

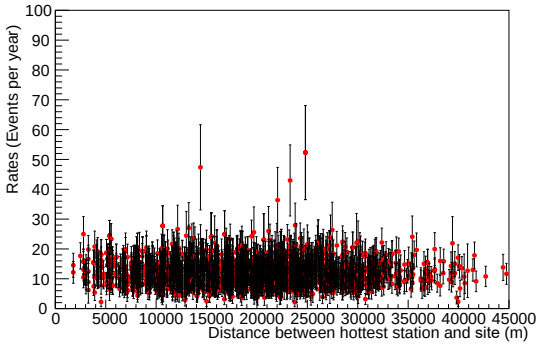
(d) Coihueco

Figure B.12: Scatter plots of rates as a function of time for Los Leones, Los Morados, Loma Amarilla and Coihueco. All error bars represent the statistical uncertainty of each rate which is the square root of the number of events divided by the uptime of each station. There were no hybrid events (subject to the applied cuts in Appendix A) detected by Los Morados in 2004, none seen by Loma Amarilla in 2004, 2005 and 2006 and only one event seen in Coihueco in 2004.

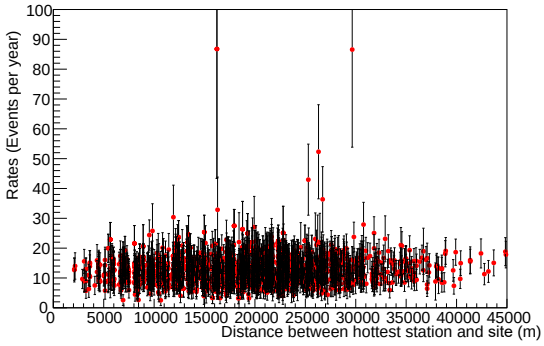
Scatter plots showing yearly rates as a function of time from each FD site are shown in Figure B.12. It is noted that there were no hybrid events (subject to the applied cuts in Appendix A) detected by Los Morados in 2004, none seen by Loma Amarilla in 2004, 2005 and 2006 and only one event seen in Coihueco in 2004. Similarly, scatter plots showing yearly rates as a function of distance from each FD site are shown in Figure B.13.



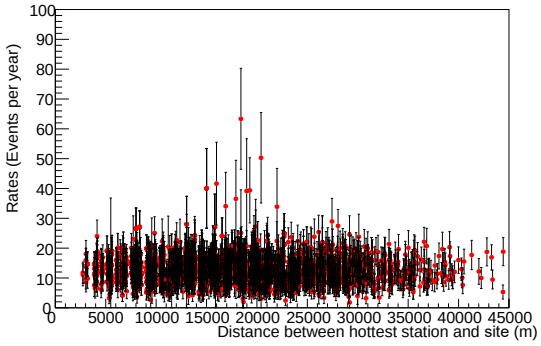
(a) Rates as a function of distance between hottest station and Los Leones.



(b) Rates as a function of distance between hottest station and Los Morados.



(c) Rates as a function of distance between hottest station and Loma Amarilla.

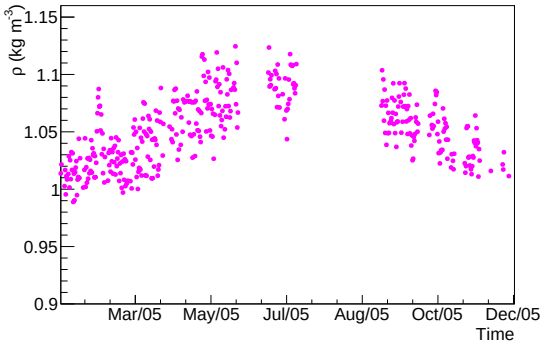


(d) Rates as a function of distance between hottest station and Coihueco.

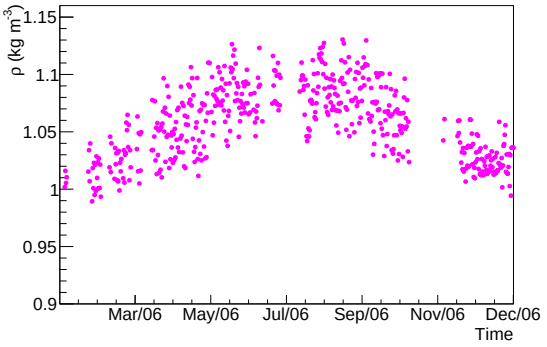
Figure B.13: Scatter plots of rates as a function of distance between hottest station and each FD site. All error bars represent the standard error on the mean for each yearly bin.

B.5 Air Density as a Function of Time

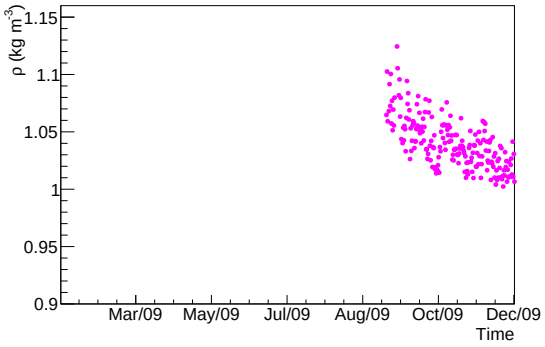
Figures B.14-B.23 show plots of the air density as a function of time as recorded by the weather stations at Los Leones, Los Morados, Loma Amarilla, Coihueco and the Central Laser Facility. It was noted that no data was taken in 2007 and 2008 at Los Leones.



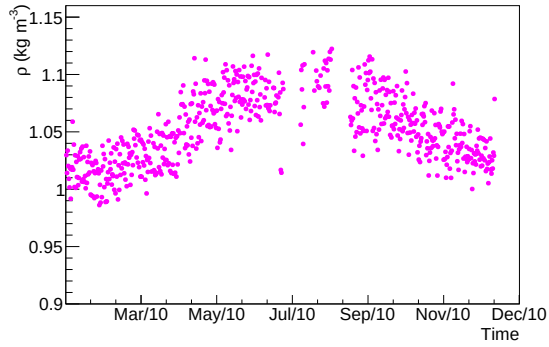
(a) Air density variation in 2005



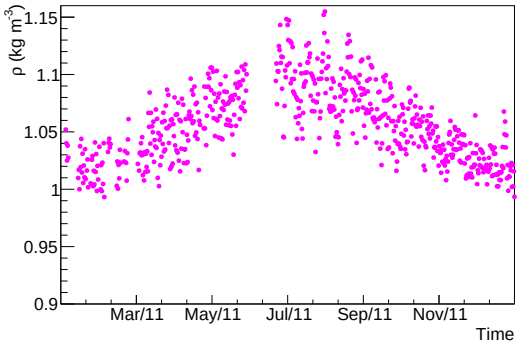
(b) Air density variation in 2006



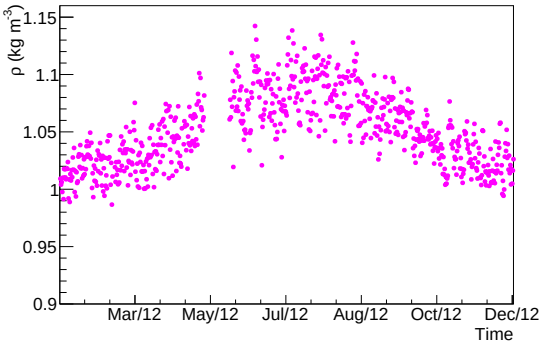
(c) Air density variation in 2009



(d) Air density variation in 2010

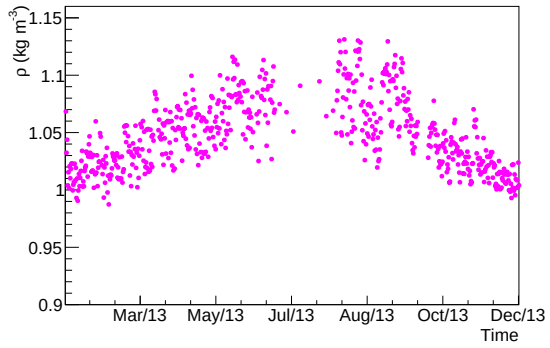


(e) Air density variation in 2011

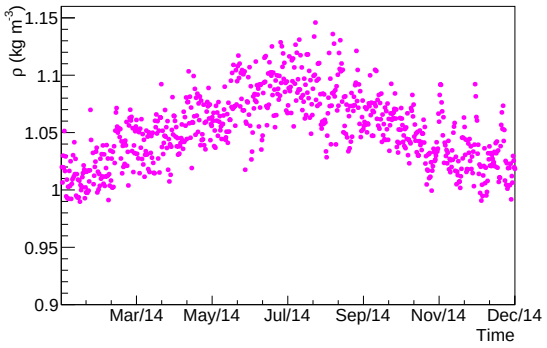


(f) Air density variation in 2012

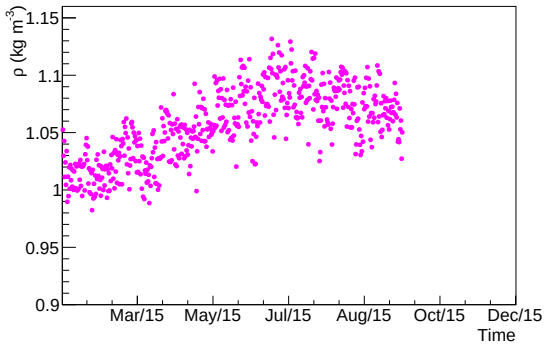
Figure B.14: Air density as a function of time each year from 2005 to 2012 at Los Leones.



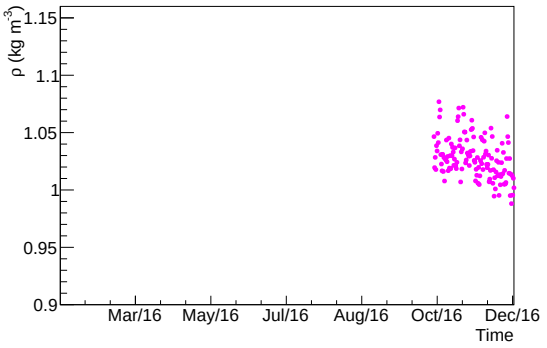
(a) Air density variation in 2013



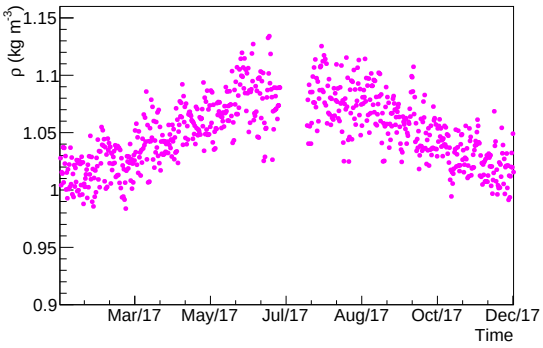
(b) Air density variation in 2014



(c) Air density variation in 2015

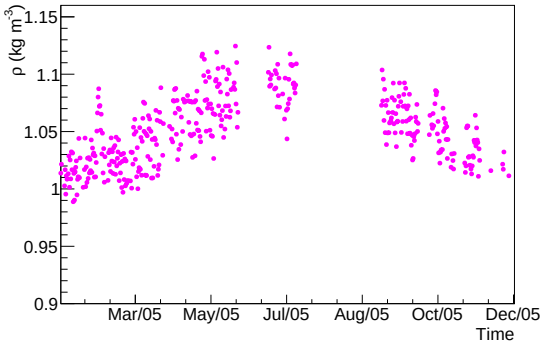


(d) Air density variation in 2016

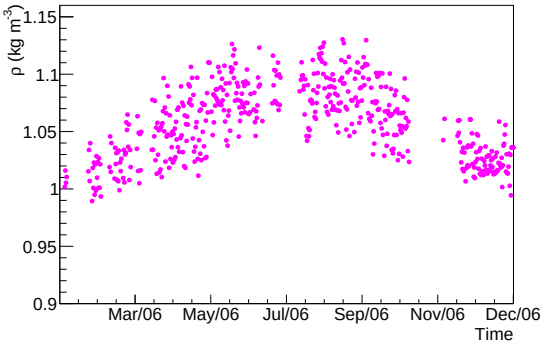


(e) Air density variation in 2017

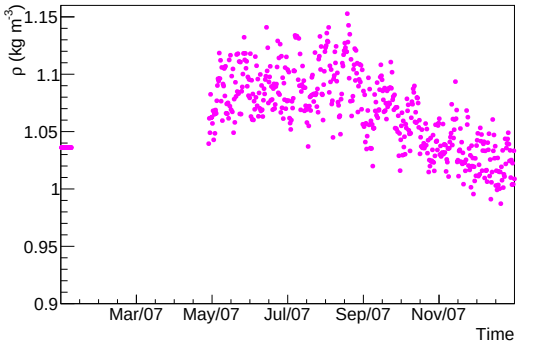
Figure B.15: Air density as a function of time each year from 2013 to 2017 at Los Leones.



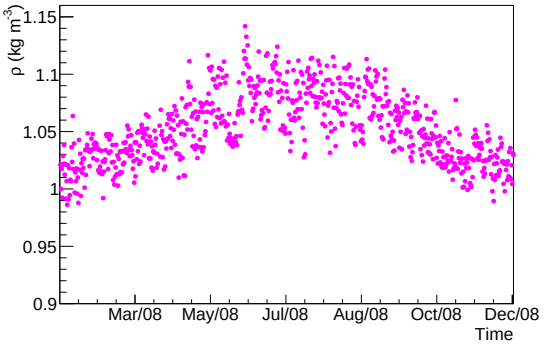
(a) Air density variation in 2005



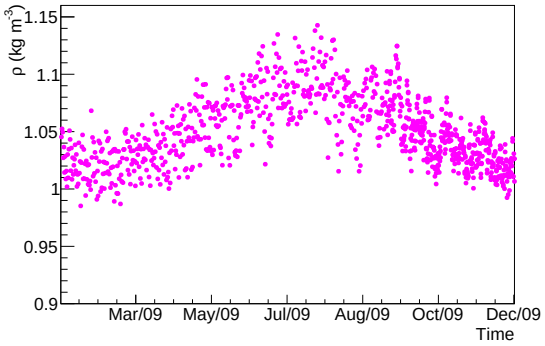
(b) Air density variation in 2006



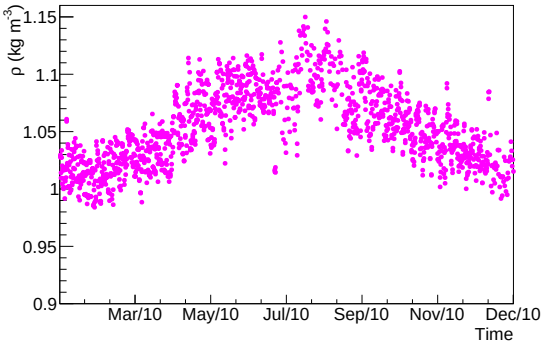
(c) Air density variation in 2007



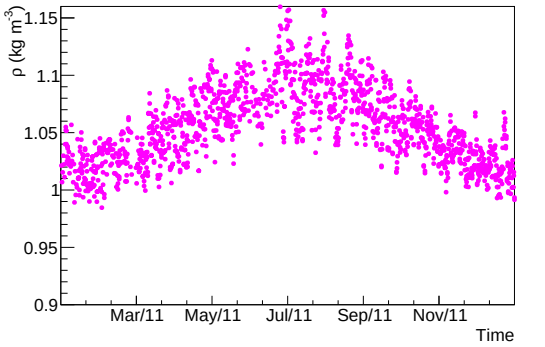
(d) Air density variation in 2008



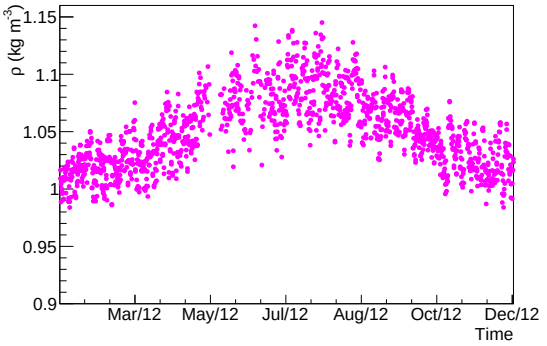
(e) Air density variation in 2009



(f) Air density variation in 2010

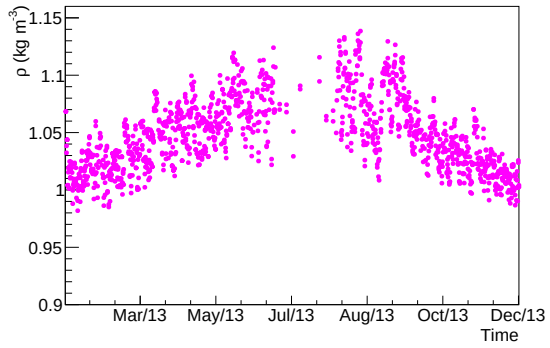


(g) Air density variation in 2011

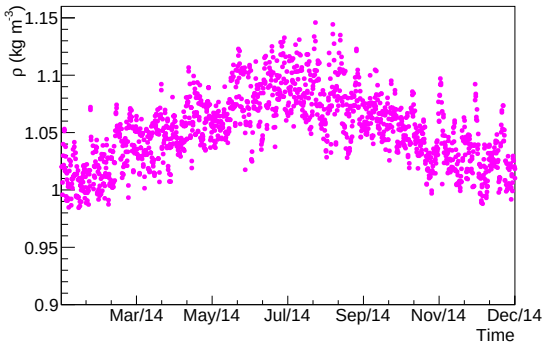


(h) Air density variation in 2012

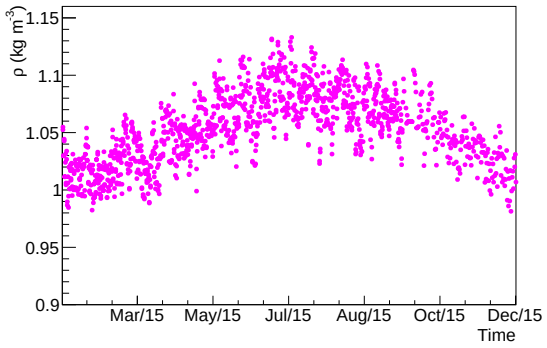
Figure B.16: Air density as a function of time each year from 2005 to 2012 at Los Morados.



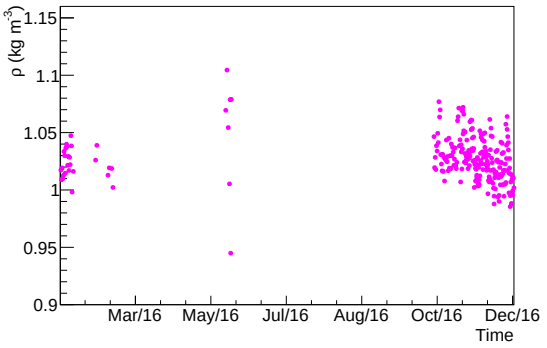
(a) Air density variation in 2013



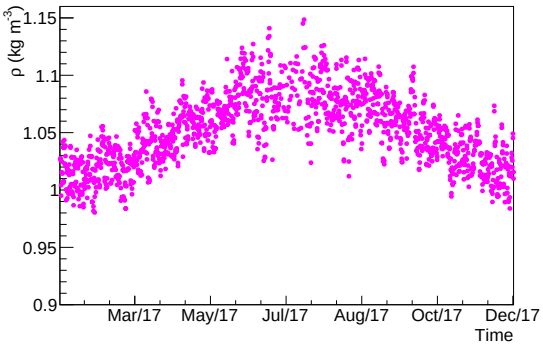
(b) Air density variation in 2014



(c) Air density variation in 2015

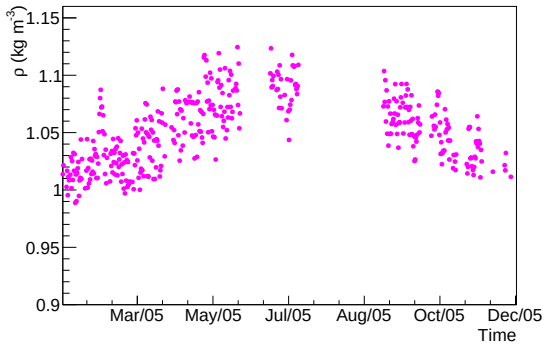


(d) Air density variation in 2016

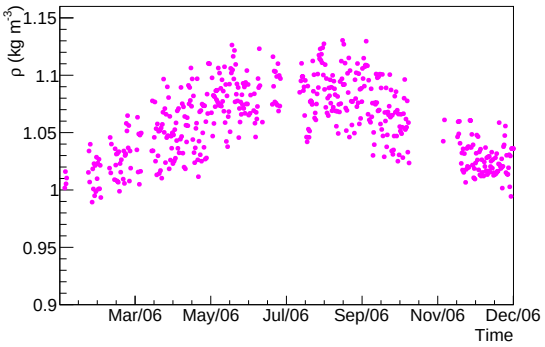


(e) Air density variation in 2017

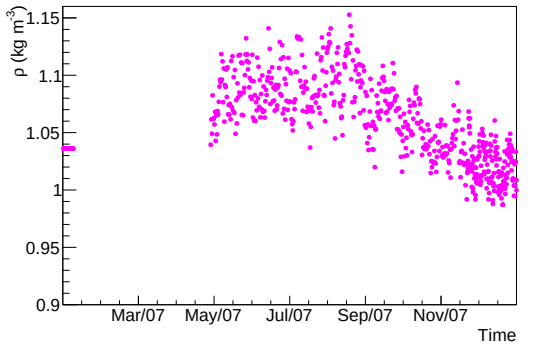
Figure B.17: Air density as a function of time each year from 2013 to 2017 at Los Morados.



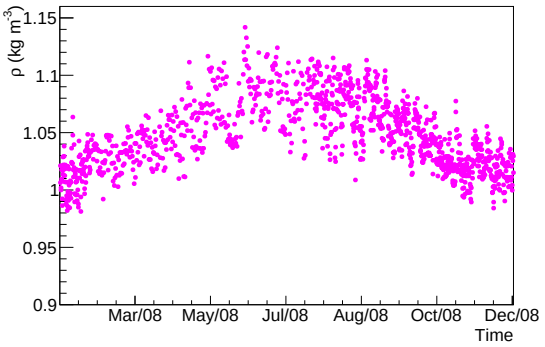
(a) Air density variation in 2005



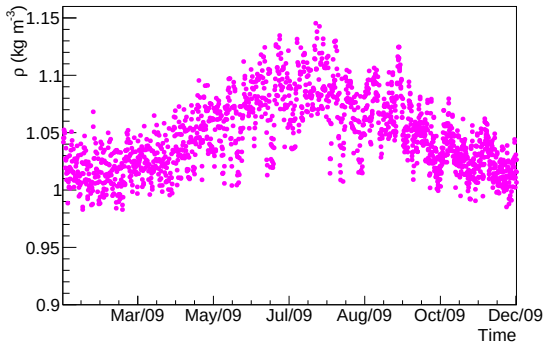
(b) Air density variation in 2006



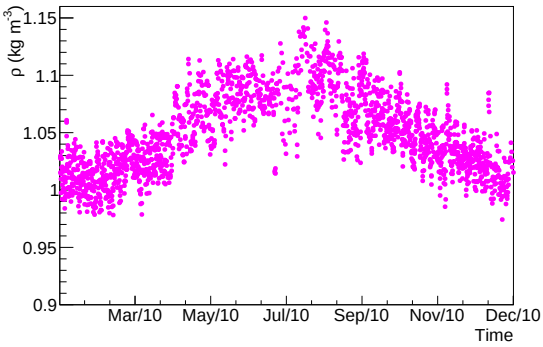
(c) Air density variation in 2007



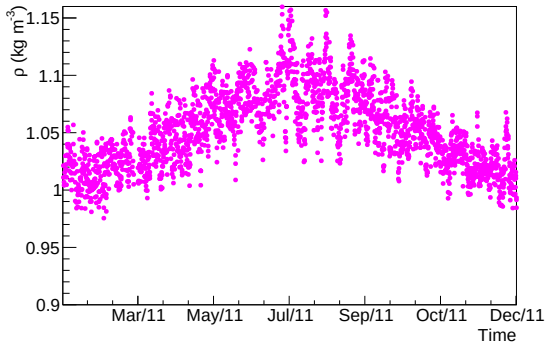
(d) Air density variation in 2008



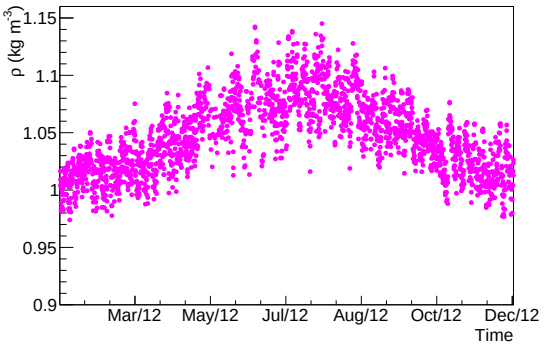
(e) Air density variation in 2009



(f) Air density variation in 2010

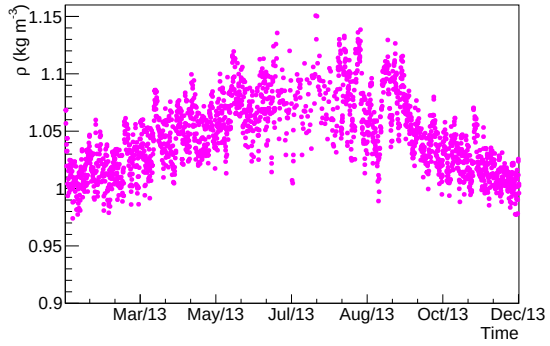


(g) Air density variation in 2011

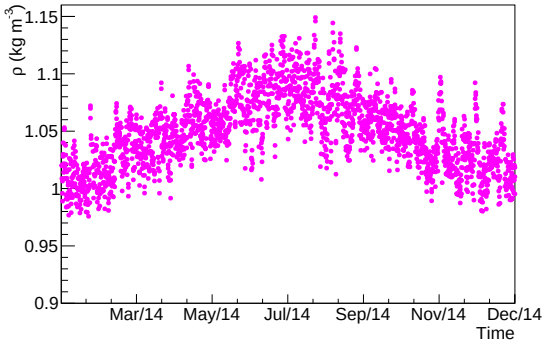


(h) Air density variation in 2012

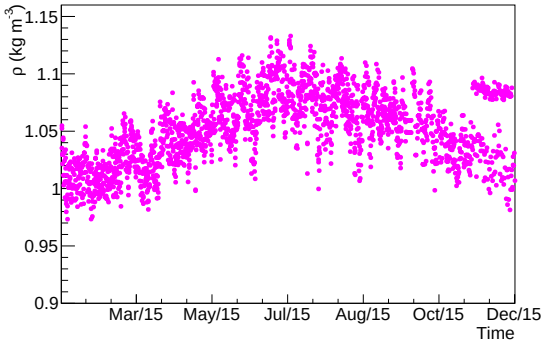
Figure B.18: Air density as a function of time each year from 2005 to 2012 at Loma Amarilla.



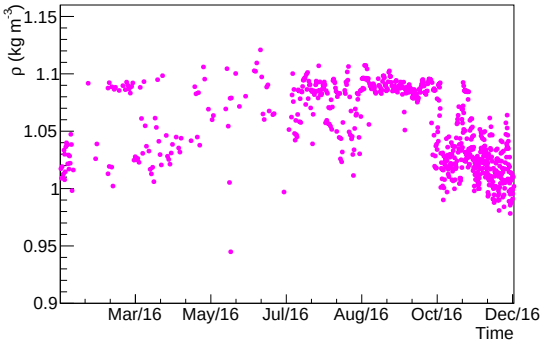
(a) Air density variation in 2013



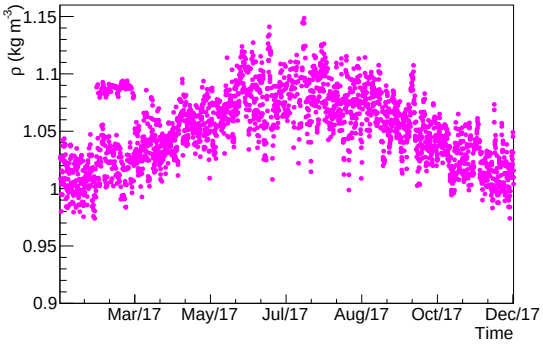
(b) Air density variation in 2014



(c) Air density variation in 2015

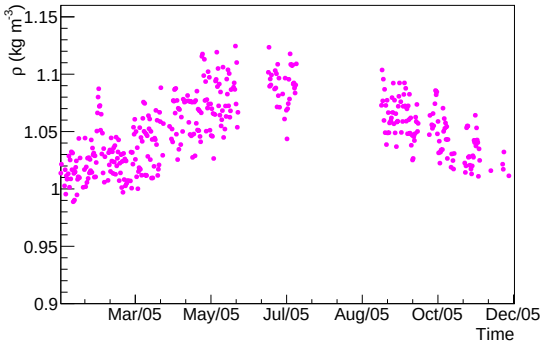


(d) Air density variation in 2016

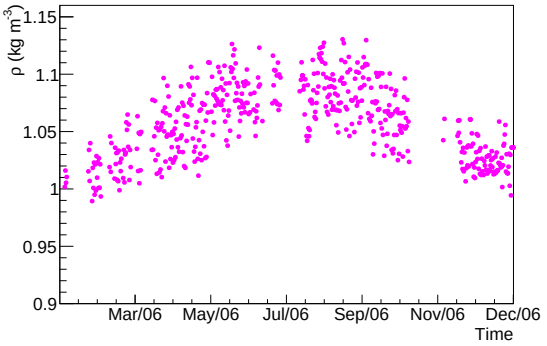


(e) Air density variation in 2017

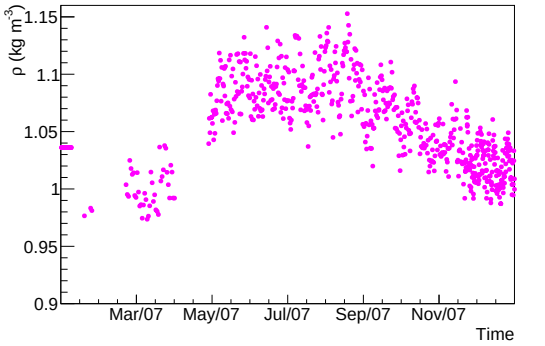
Figure B.19: Air density as a function of time each year from 2013 to 2017 at Loma Amarilla.



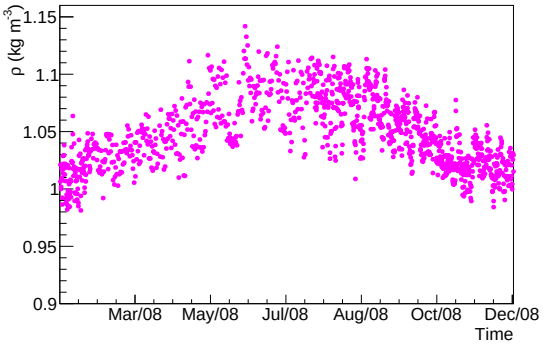
(a) Air density variation in 2005



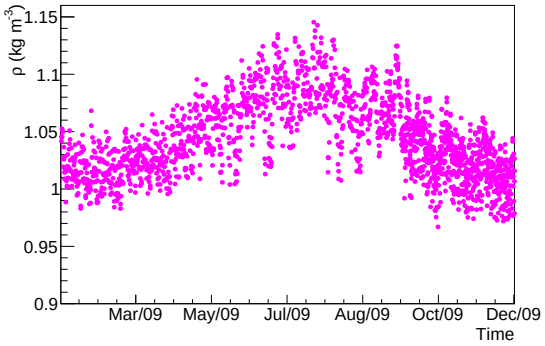
(b) Air density variation in 2006



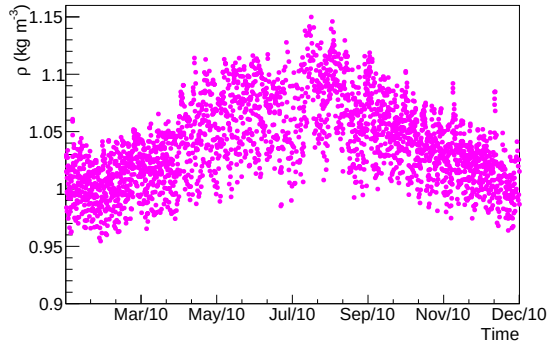
(c) Air density variation in 2007



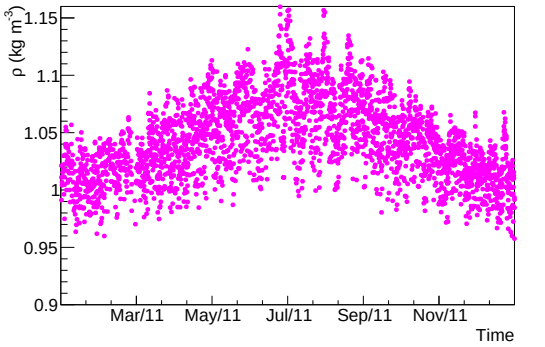
(d) Air density variation in 2008



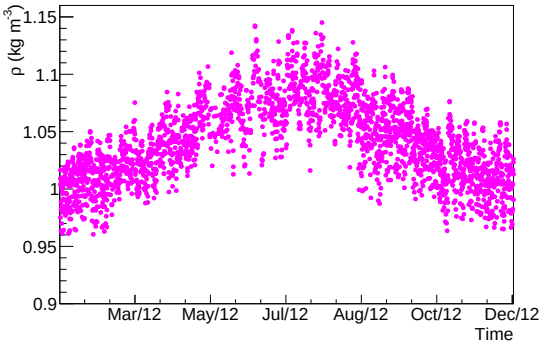
(e) Air density variation in 2009



(f) Air density variation in 2010

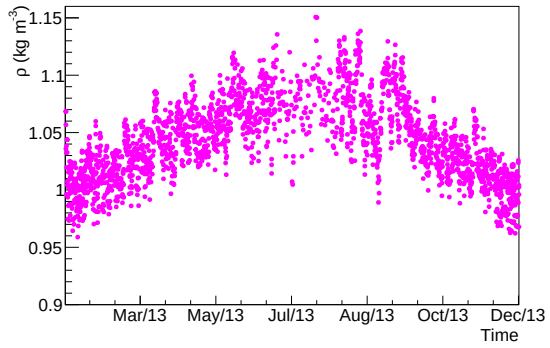


(g) Air density variation in 2011

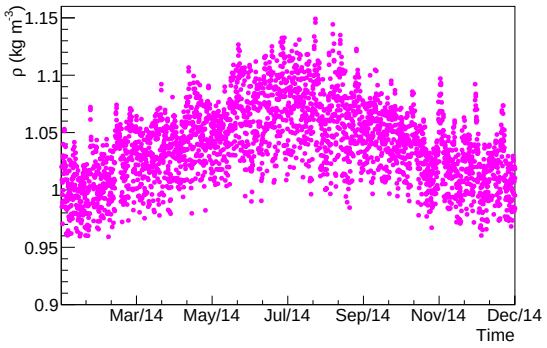


(h) Air density variation in 2012

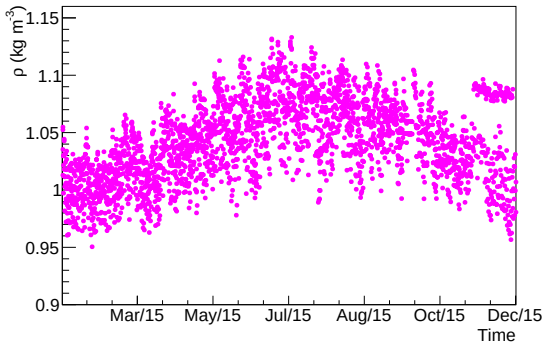
Figure B.20: Air density as a function of time each year from 2005 to 2012 at Coihueco.



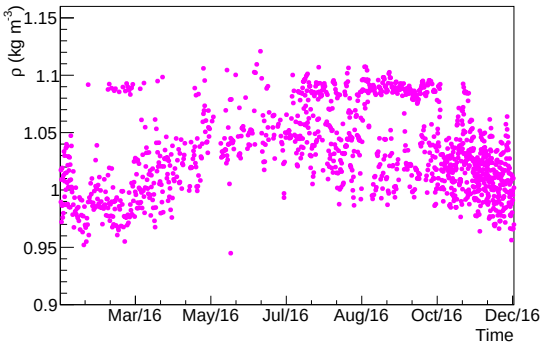
(a) Air density variation in 2013



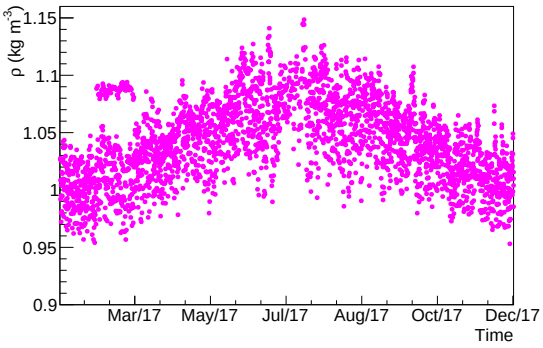
(b) Air density variation in 2014



(c) Air density variation in 2015

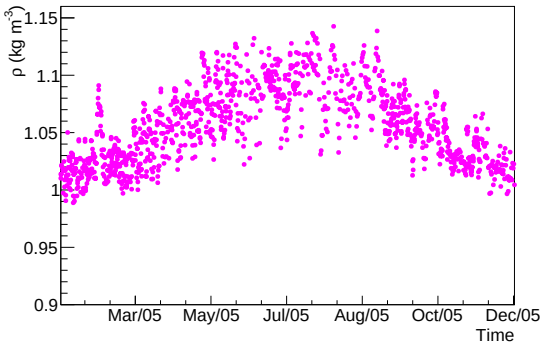


(d) Air density variation in 2016

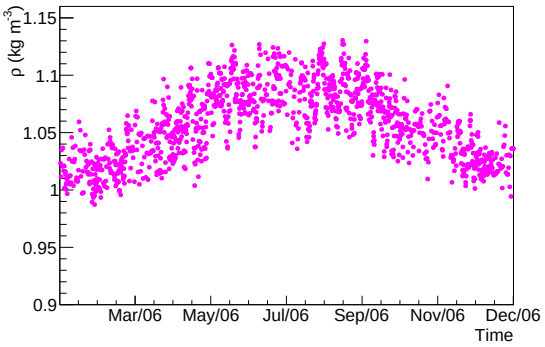


(e) Air density variation in 2017

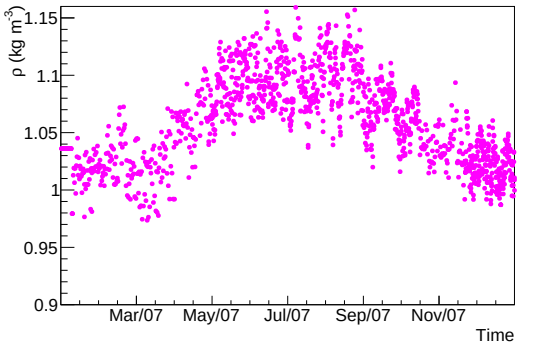
Figure B.21: Air density as a function of time each year from 2013 to 2017 at Coihueco.



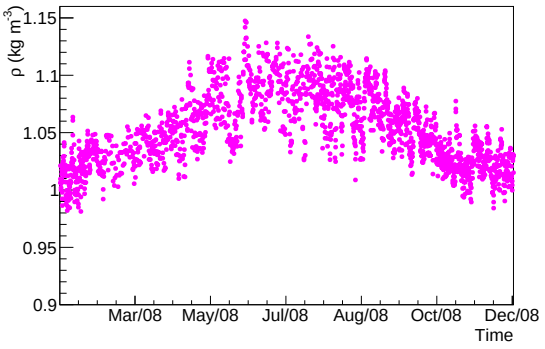
(a) Air density variation in 2005



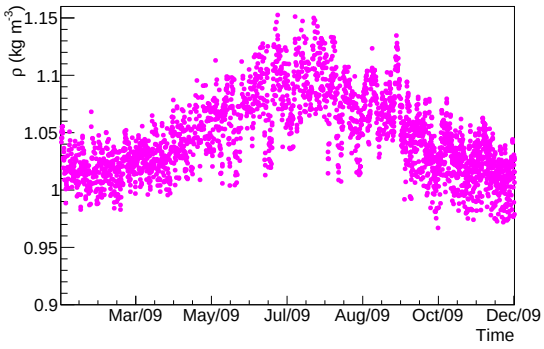
(b) Air density variation in 2006



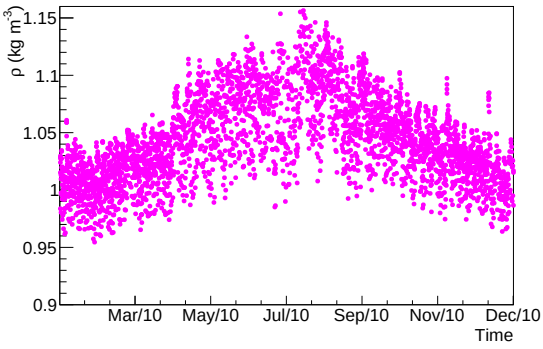
(c) Air density variation in 2007



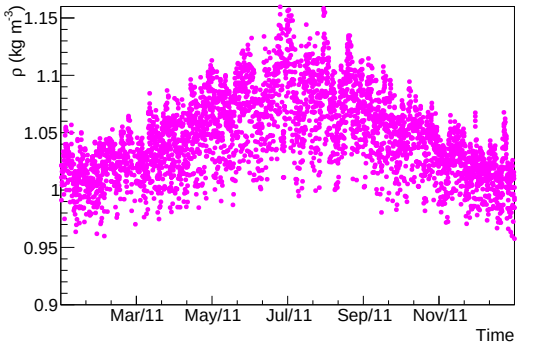
(d) Air density variation in 2008



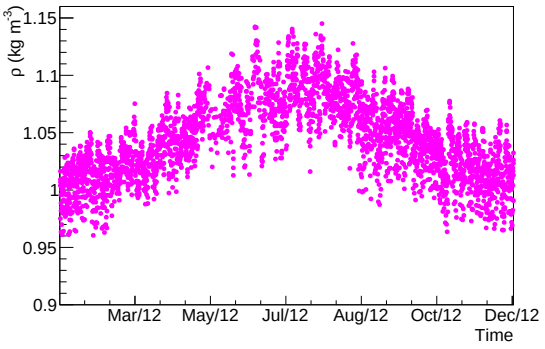
(e) Air density variation in 2009



(f) Air density variation in 2010

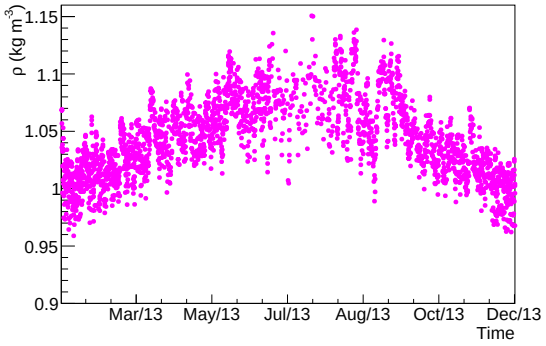


(g) Air density variation in 2011

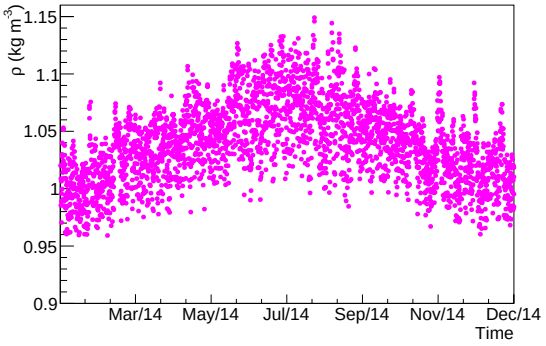


(h) Air density variation in 2012

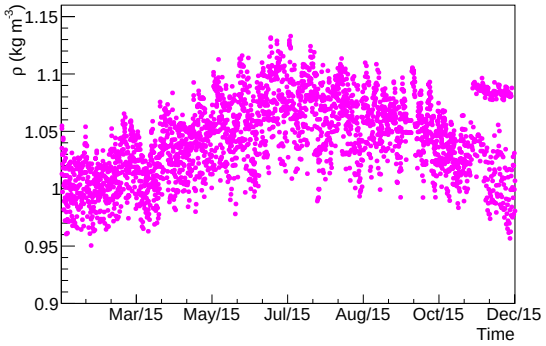
Figure B.22: Air density as a function of time each year from 2005 to 2012 at the Central Laser Facility (CLF).



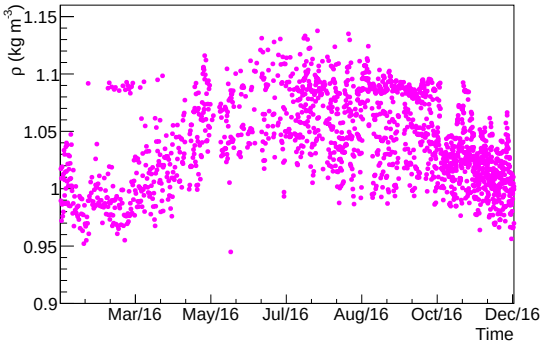
(a) Air density variation in 2013



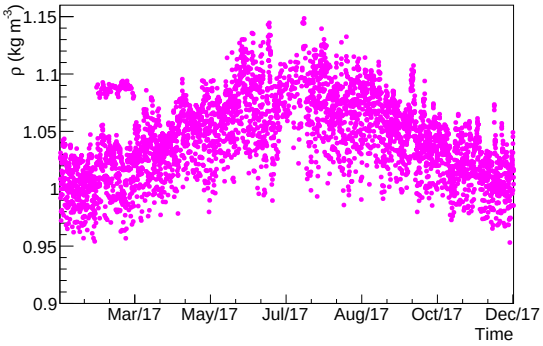
(b) Air density variation in 2014



(c) Air density variation in 2015



(d) Air density variation in 2016

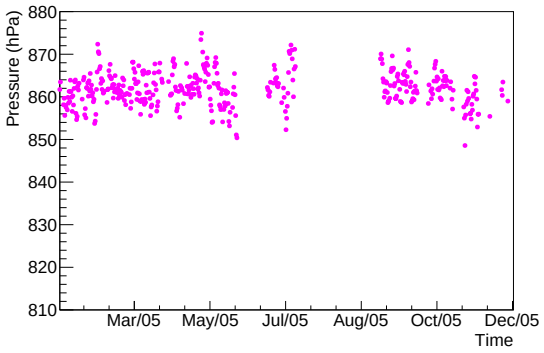


(e) Air density variation in 2017

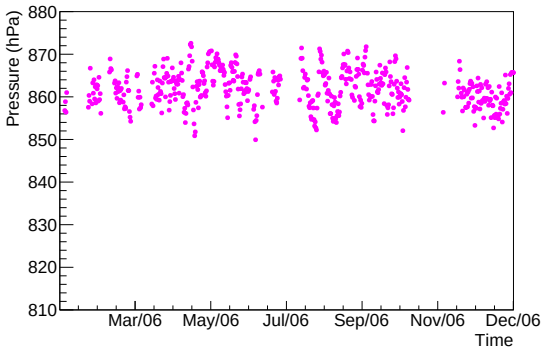
Figure B.23: Air density as a function of time each year from 2013 to 2017 at the Central Laser Facility (CLF).

B.6 Pressure as a Function of Time

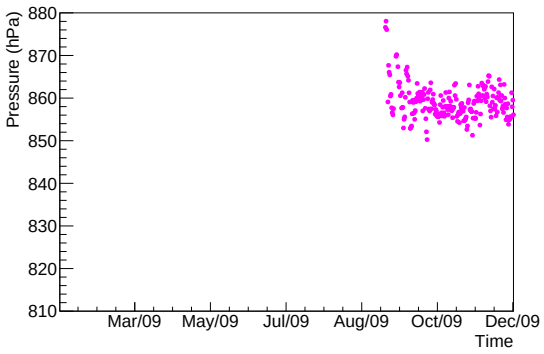
Figures B.14-B.23 show plots of the pressure as a function of time as recorded by the weather stations at Los Leones, Los Morados, Loma Amarilla, Coihueco and the Central Laser Facility. It was noted that no data was taken in 2007 and 2008 at Los Leones.



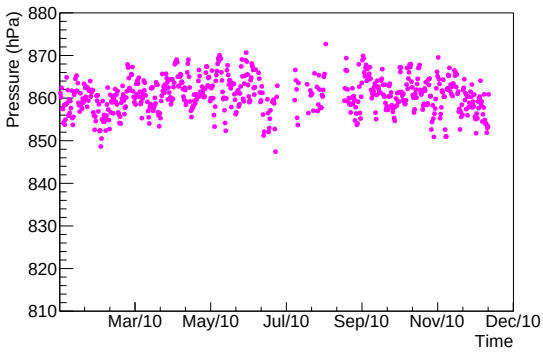
(a) Pressure variation in 2005



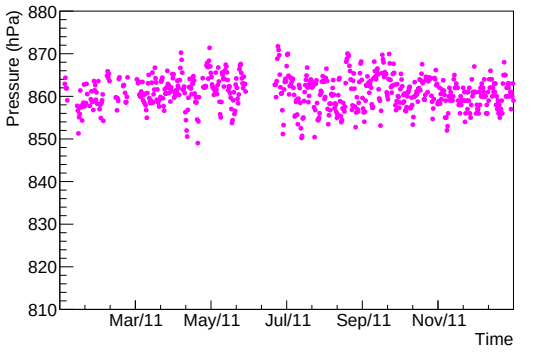
(b) Pressure variation in 2006



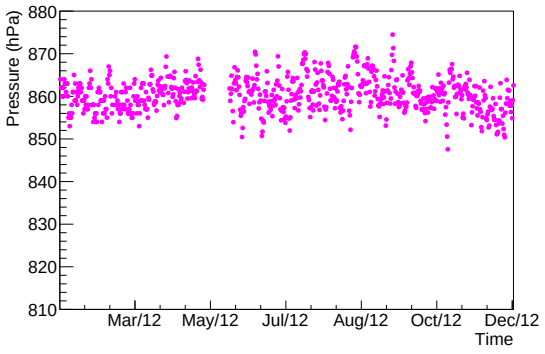
(c) Pressure variation in 2009



(d) Pressure variation in 2010

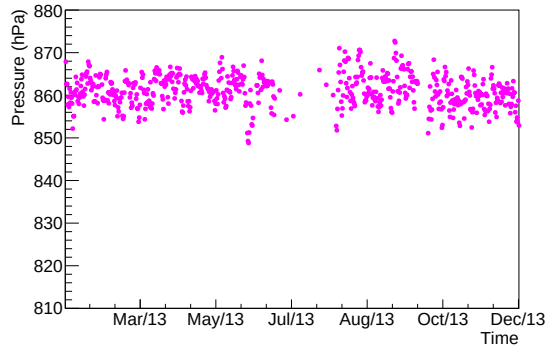


(e) Pressure variation in 2011

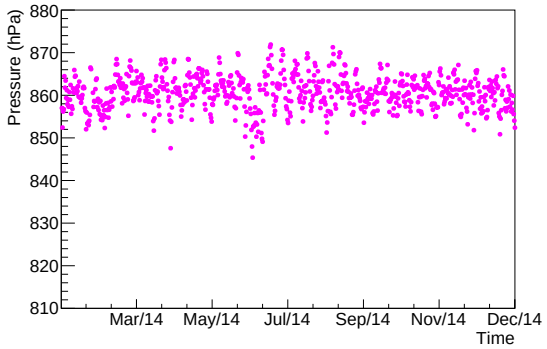


(f) Pressure variation in 2012

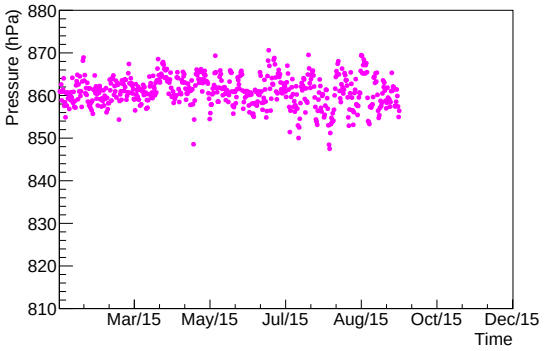
Figure B.24: Pressure as a function of time each year from 2005 to 2012 at Los Leones.



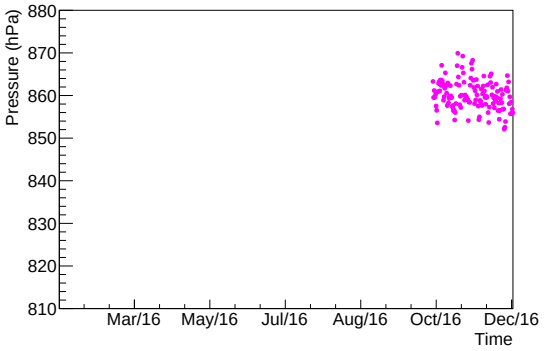
(a) Pressure variation in 2013



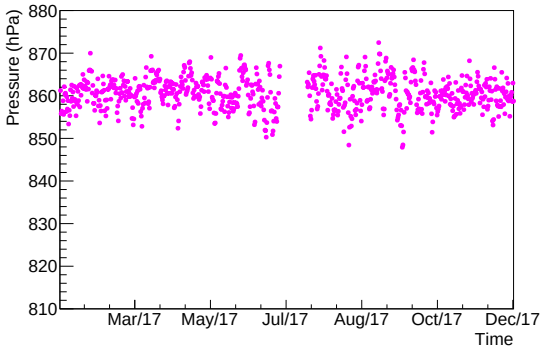
(b) Pressure variation in 2014



(c) Pressure variation in 2015

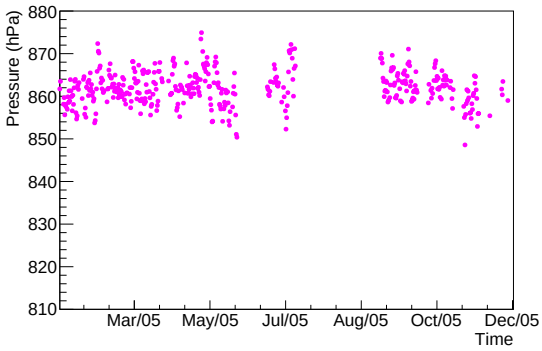


(d) Pressure variation in 2016

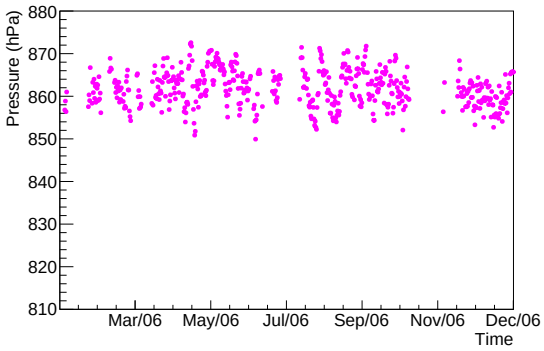


(e) Pressure variation in 2017

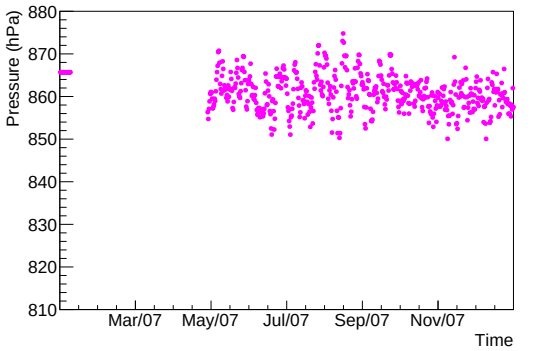
Figure B.25: Pressure as a function of time each year from 2013 to 2017 at Los Leones.



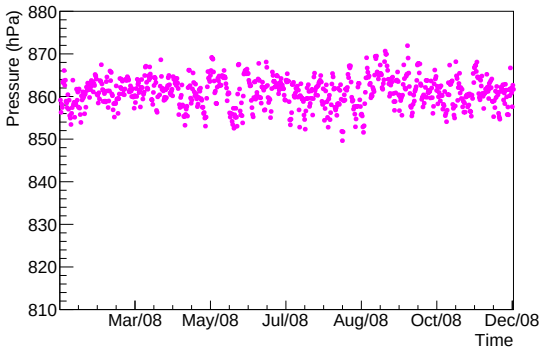
(a) Pressure variation in 2005



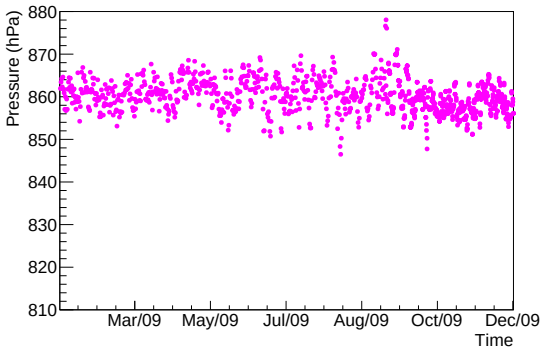
(b) Pressure variation in 2006



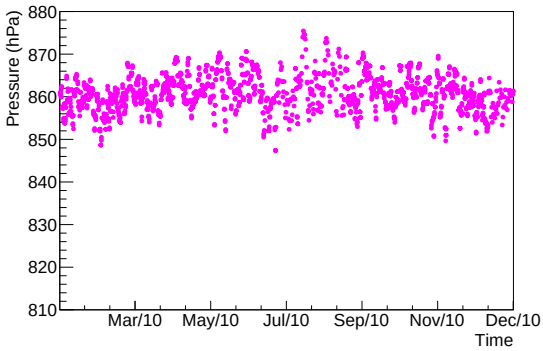
(c) Pressure variation in 2007



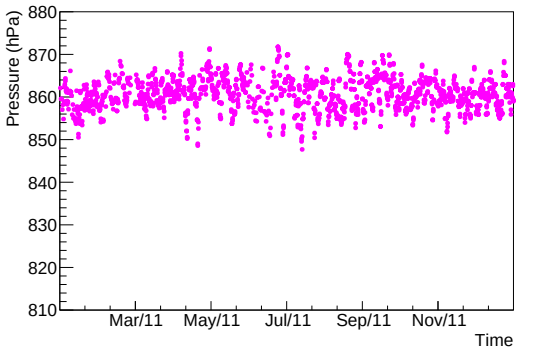
(d) Pressure variation in 2008



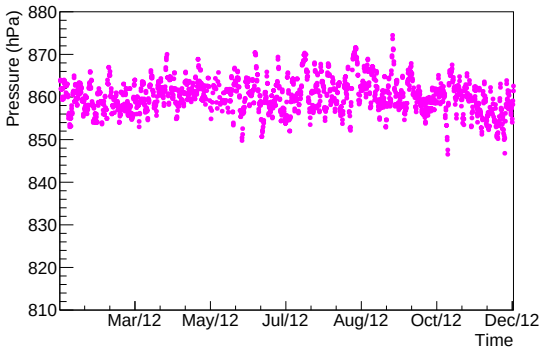
(e) Pressure variation in 2009



(f) Pressure variation in 2010

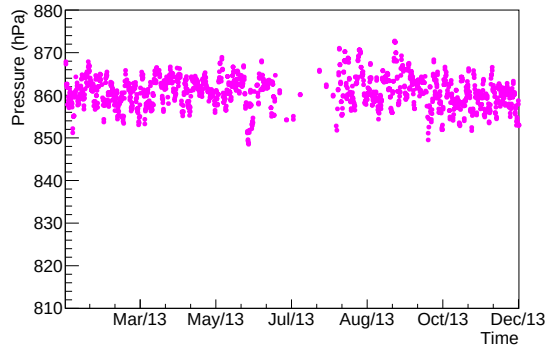


(g) Pressure variation in 2011

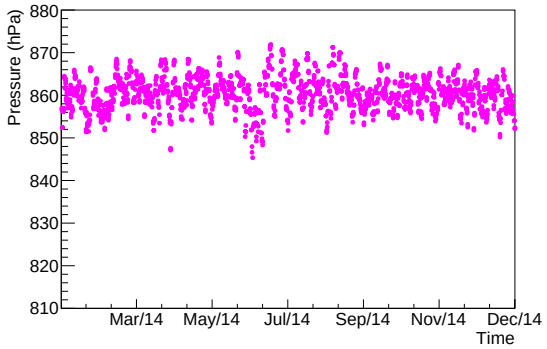


(h) Pressure variation in 2012

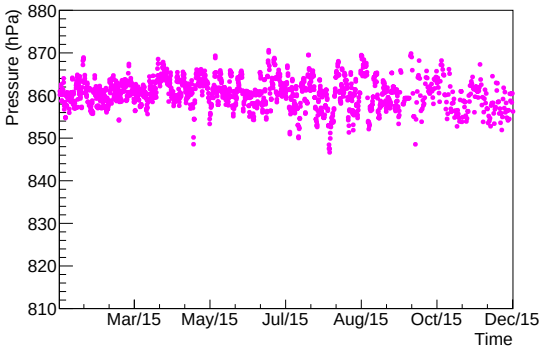
Figure B.26: Pressure as a function of time each year from 2005 to 2012 at Los Morados.



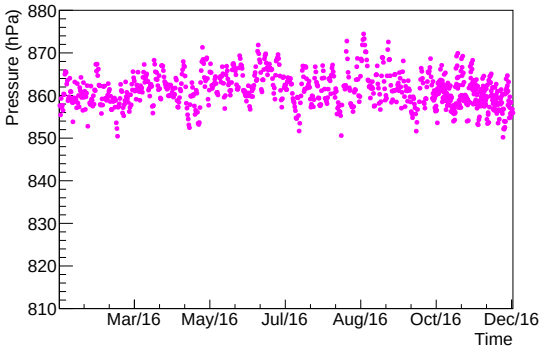
(a) Pressure variation in 2013



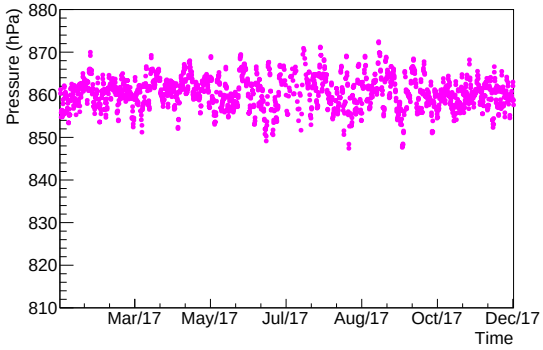
(b) Pressure variation in 2014



(c) Pressure variation in 2015

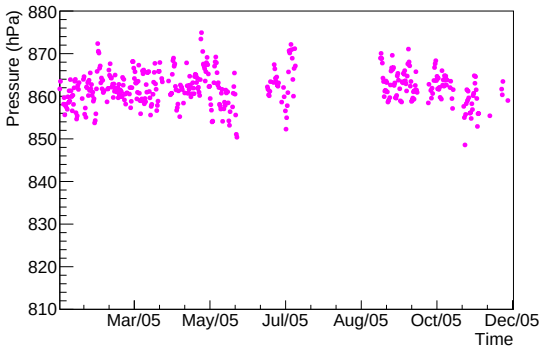


(d) Pressure variation in 2016

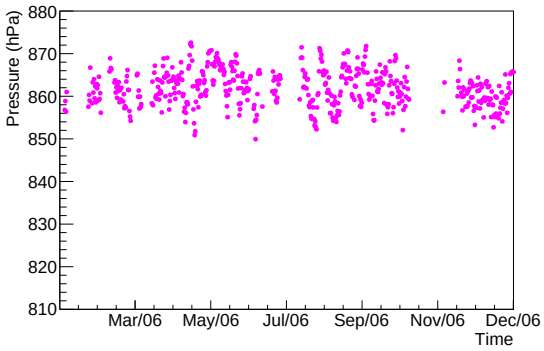


(e) Pressure variation in 2017

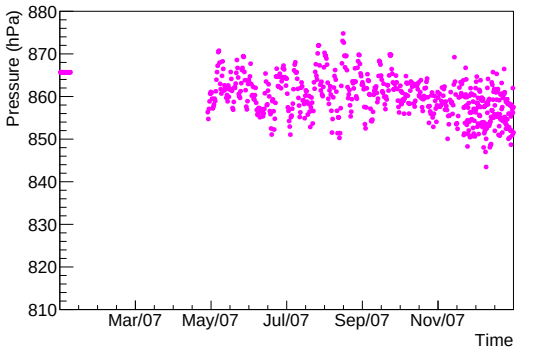
Figure B.27: Pressure as a function of time each year from 2013 to 2017 at Los Morados.



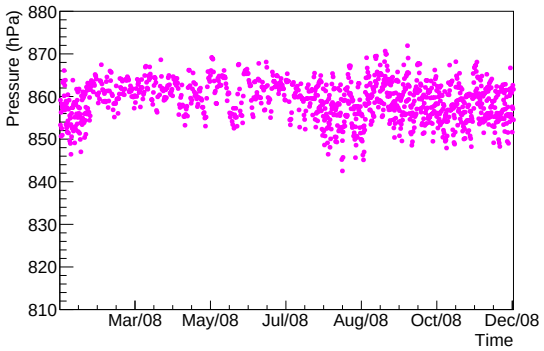
(a) Pressure variation in 2005



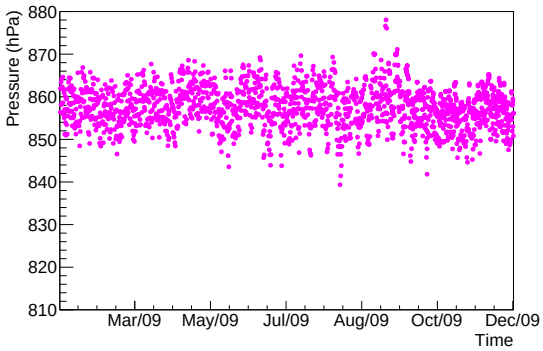
(b) Pressure variation in 2006



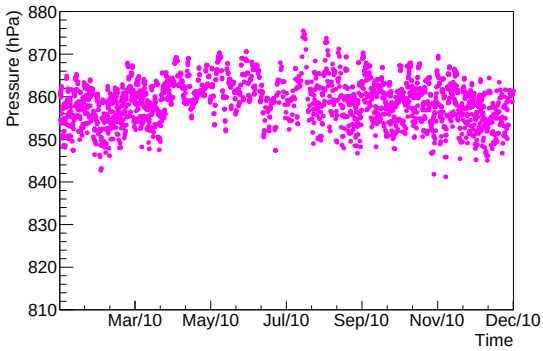
(c) Pressure variation in 2007



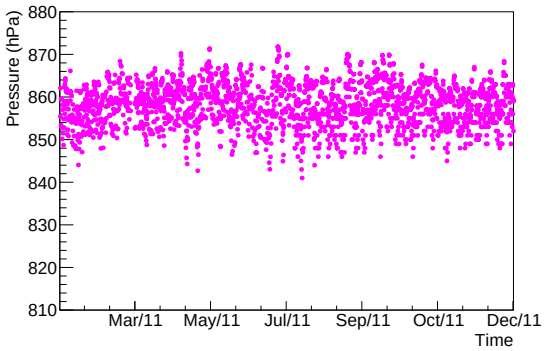
(d) Pressure variation in 2008



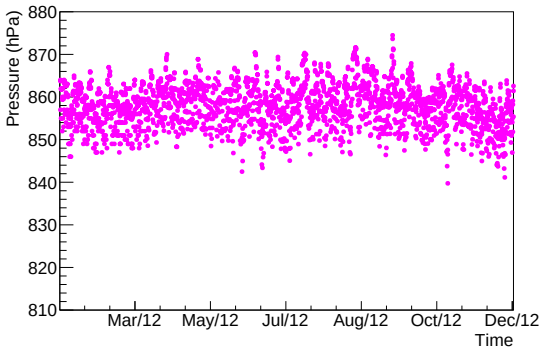
(e) Pressure variation in 2009



(f) Pressure variation in 2010

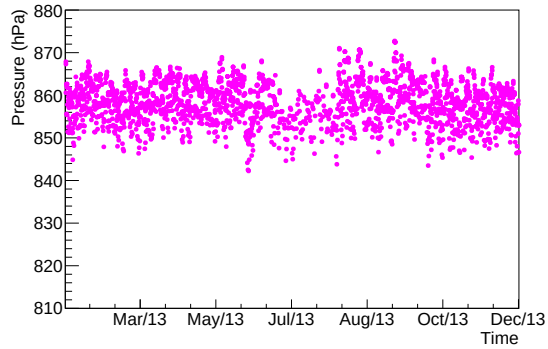


(g) Pressure variation in 2011

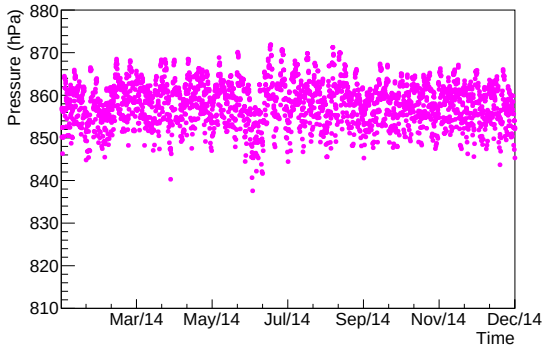


(h) Pressure variation in 2012

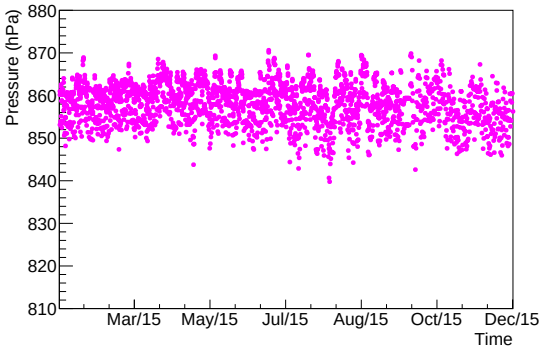
Figure B.28: Pressure as a function of time each year from 2005 to 2012 at Loma Amarilla.



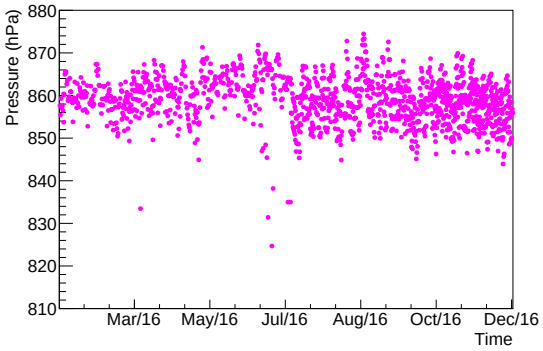
(a) Pressure variation in 2013



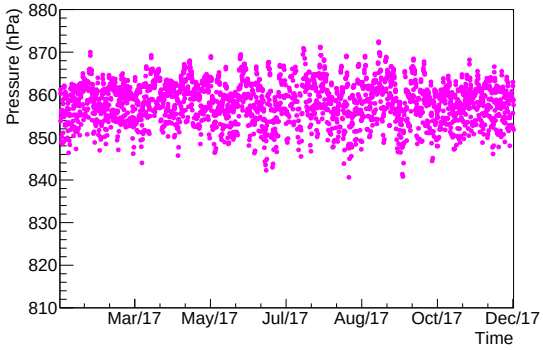
(b) Pressure variation in 2014



(c) Pressure variation in 2015

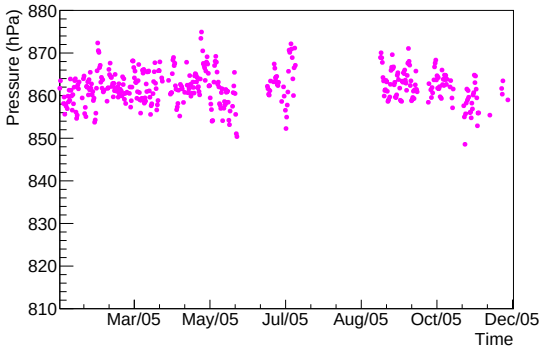


(d) Pressure variation in 2016

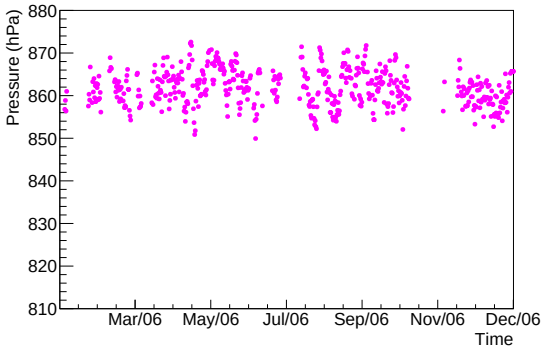


(e) Pressure variation in 2017

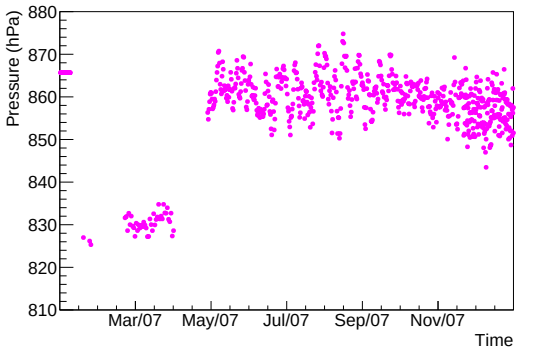
Figure B.29: Pressure as a function of time each year from 2013 to 2017 at Loma Amarilla.



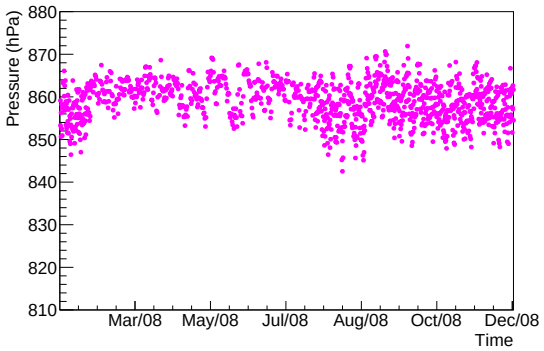
(a) Pressure variation in 2005



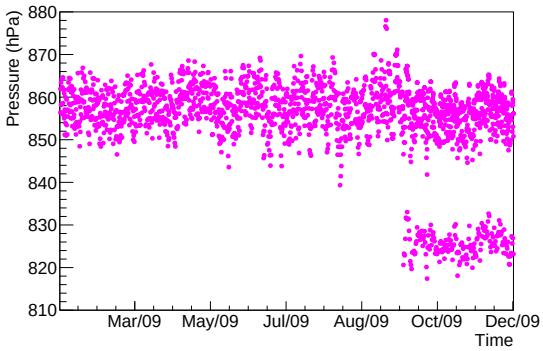
(b) Pressure variation in 2006



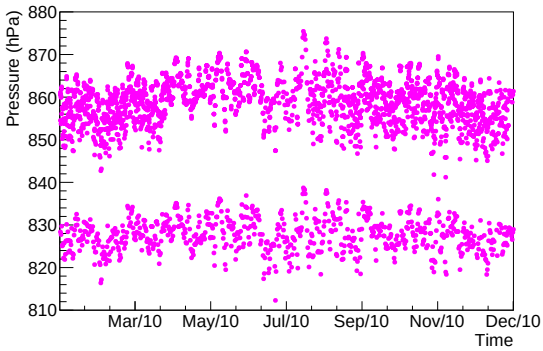
(c) Pressure variation in 2007



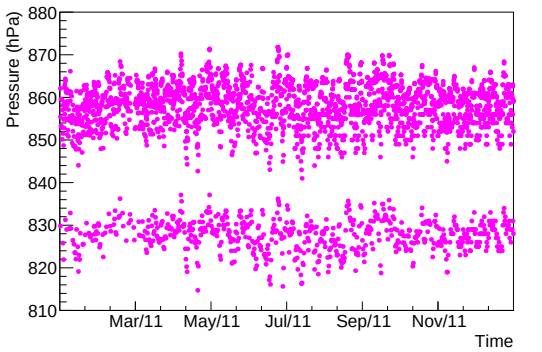
(d) Pressure variation in 2008



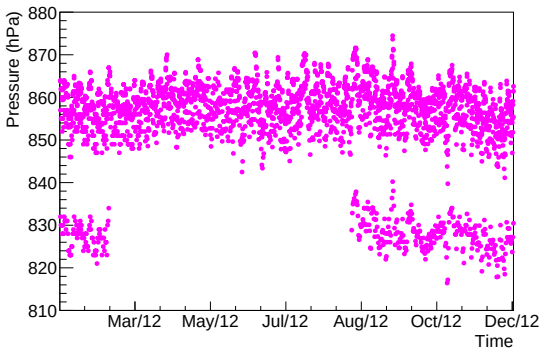
(e) Pressure variation in 2009



(f) Pressure variation in 2010

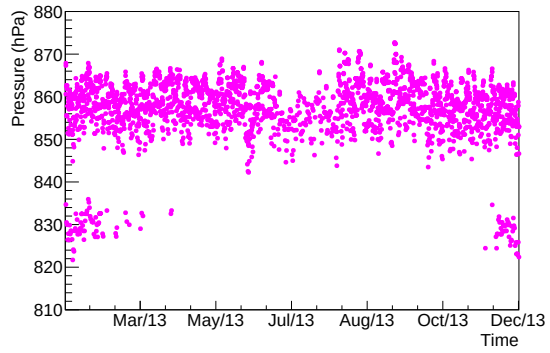


(g) Pressure variation in 2011

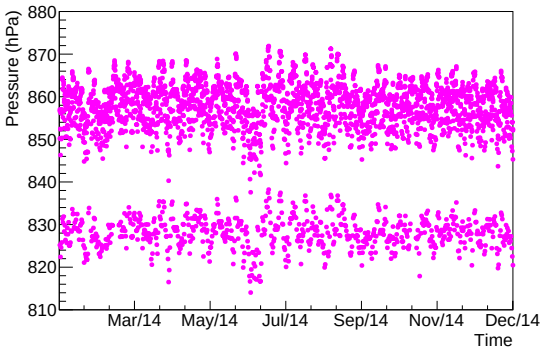


(h) Pressure variation in 2012

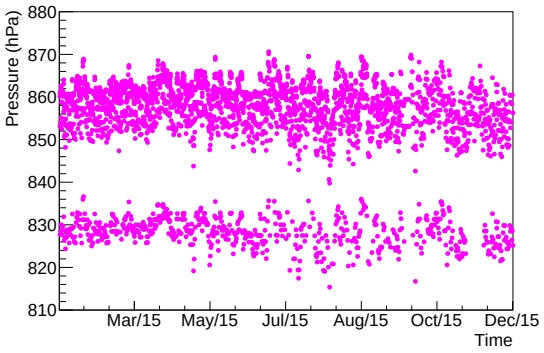
Figure B.30: Pressure as a function of time each year from 2005 to 2012 at Coihueco. For data in 2009 onwards, there is a defect in the pressure recorded at exactly 1159 and 2359 hours UTC time, which fluctuates at around 830 hPa. This data is ignored.



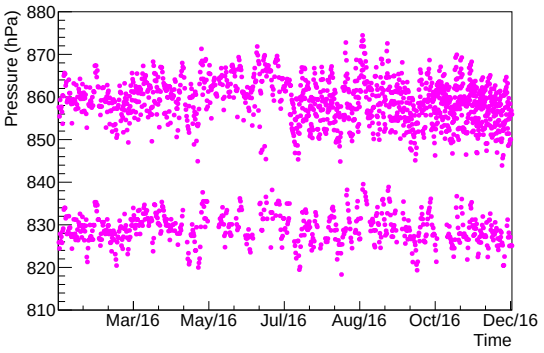
(a) Pressure variation in 2013



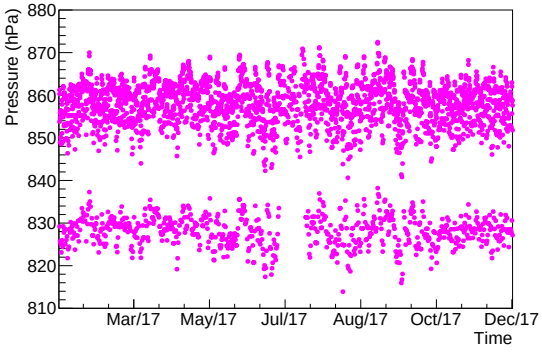
(b) Pressure variation in 2014



(c) Pressure variation in 2015

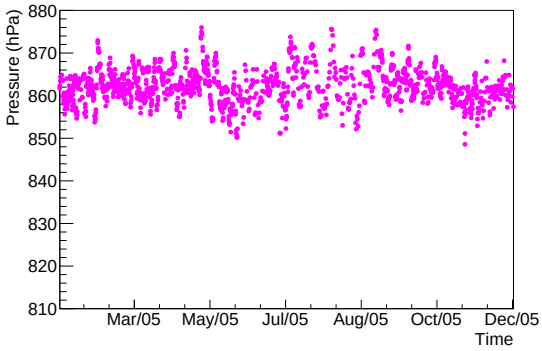


(d) Pressure variation in 2016

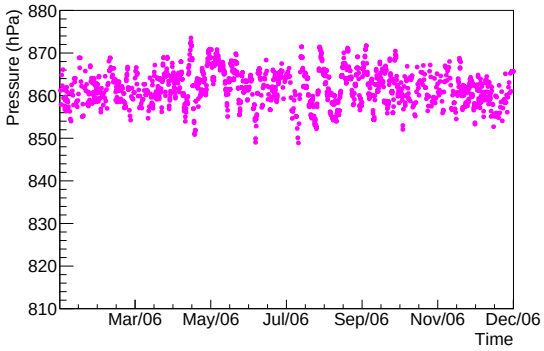


(e) Pressure variation in 2017

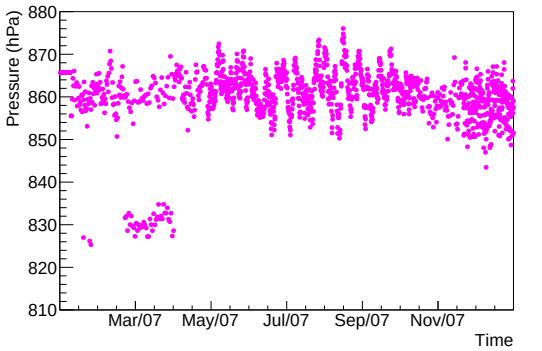
Figure B.31: Pressure as a function of time each year from 2013 to 2017 at Coihueco. Noted that there is a defect in the pressure recorded at exactly 1159 and 2359 hours UTC time, which fluctuates at around 830 hPa. This data is ignored.



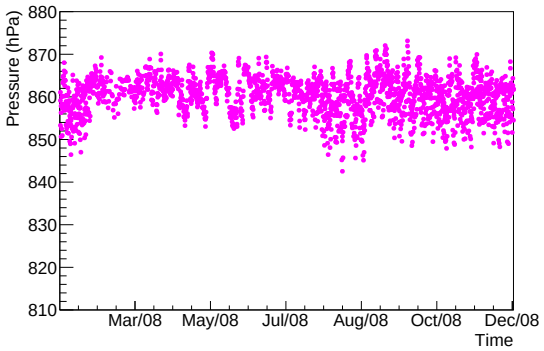
(a) Pressure variation in 2005



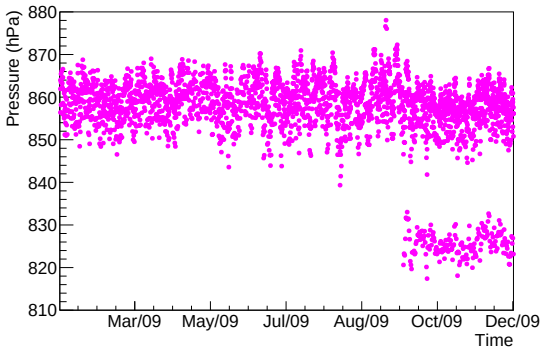
(b) Pressure variation in 2006



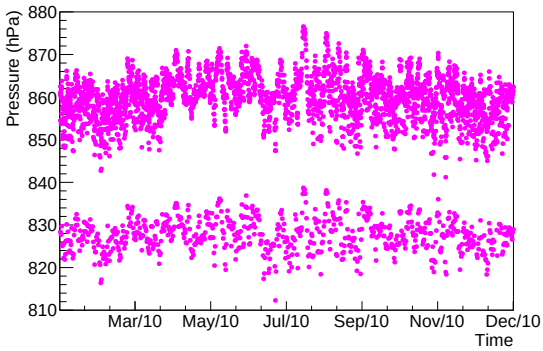
(c) Pressure variation in 2007



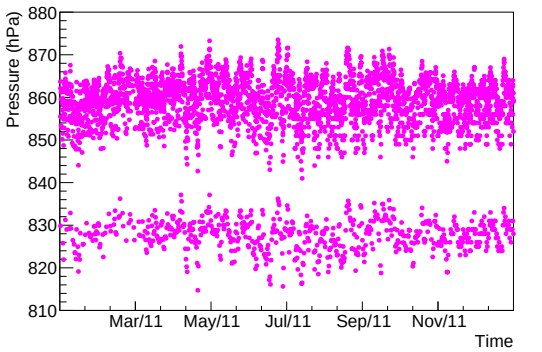
(d) Pressure variation in 2008



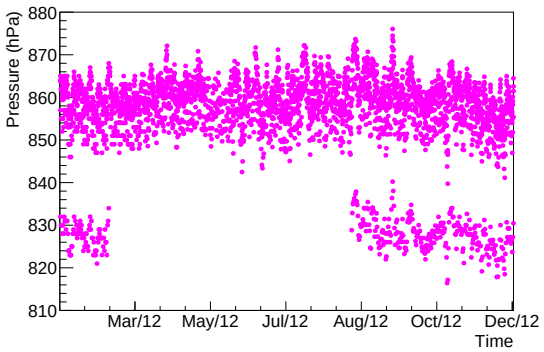
(e) Pressure variation in 2009



(f) Pressure variation in 2010

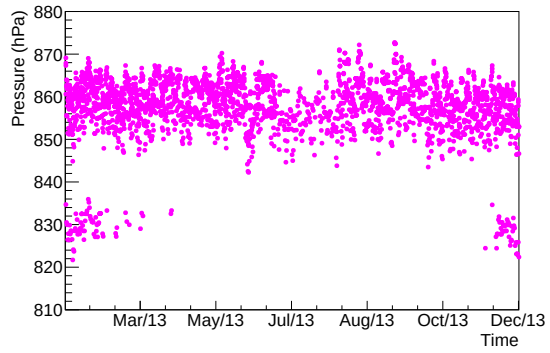


(g) Pressure variation in 2011

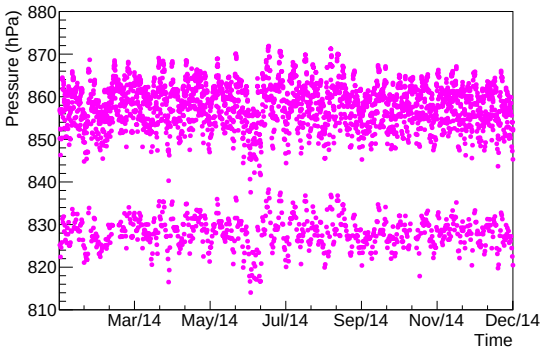


(h) Pressure variation in 2012

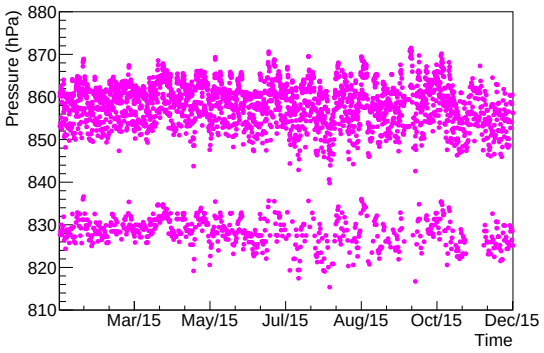
Figure B.32: Pressure as a function of time each year from 2005 to 2012 at the Central Laser Facility (CLF). It is noted that there is a defect in the pressure recorded at exactly 1159 and 2359 hours UTC time, which fluctuates at around 830 hPa. This data is ignored.



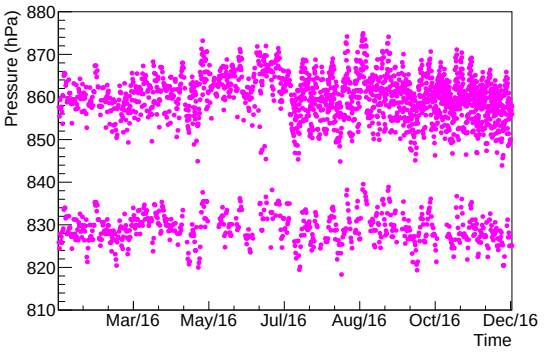
(a) Pressure variation in 2013



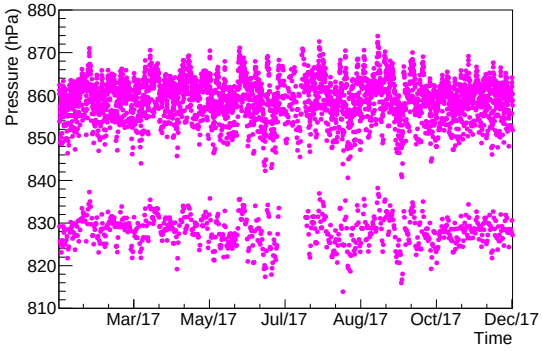
(b) Pressure variation in 2014



(c) Pressure variation in 2015



(d) Pressure variation in 2016

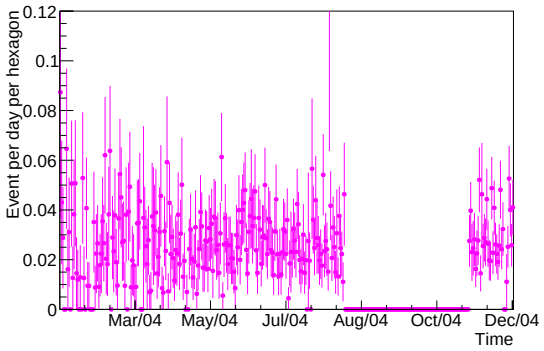


(e) Pressure variation in 2017

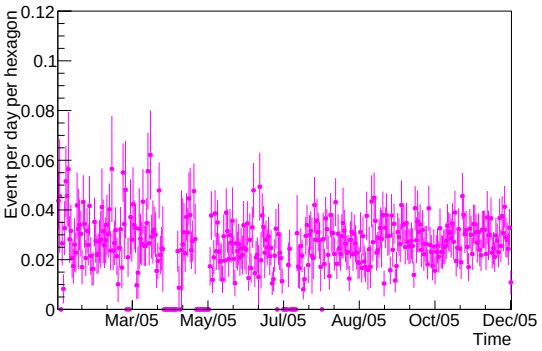
Figure B.33: Pressure as a function of time each year from 2013 to 2017 at the Central Laser Facility (CLF). It is noted that there is a defect in the pressure recorded at exactly 1159 and 2359 hours UTC time, which fluctuates at around 830 hPa. This data is ignored.

B.7 Daily rates from 2004 to 2017

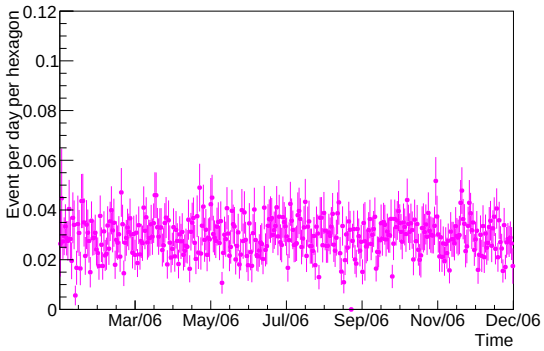
Figures B.34 and B.35 show daily rates observed every year from 2004 to 2017. Some data points show zero daily rate, which means that there is non-zero uptime for a particular station on that day, but no events detected. A notable feature is that 2010 is the only year which does not exhibit a zero daily rate.



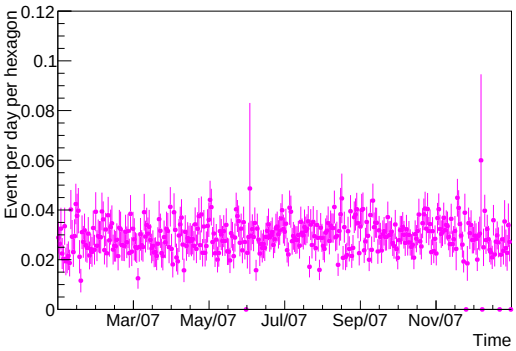
(a) Daily rates in 2004



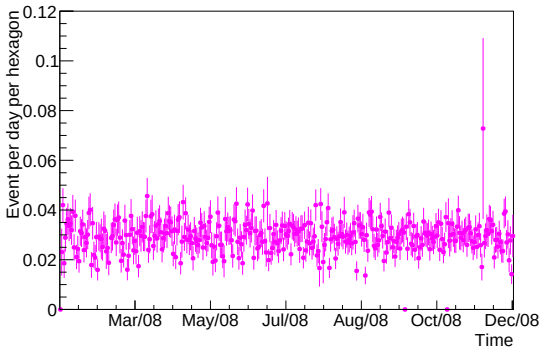
(b) Daily rates in 2005



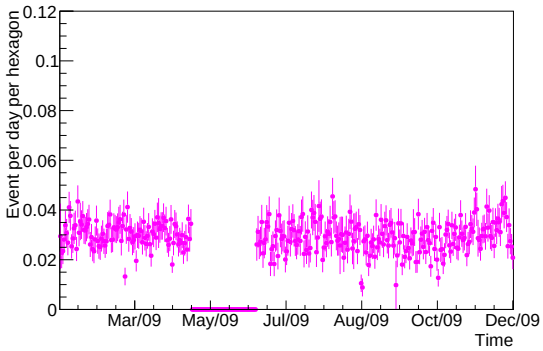
(c) Daily rates in 2006



(d) Daily rates in 2007

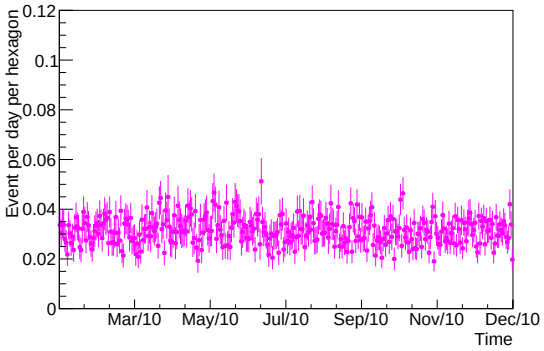


(e) Daily rates in 2008

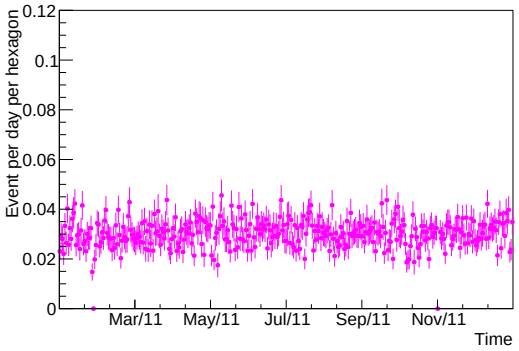


(f) Daily rates in 2009

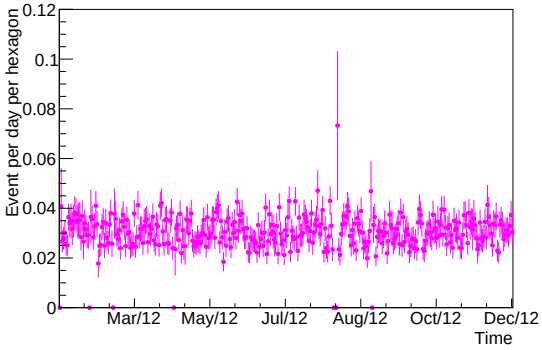
Figure B.34: Daily rates for each year, from 2004 to 2011.



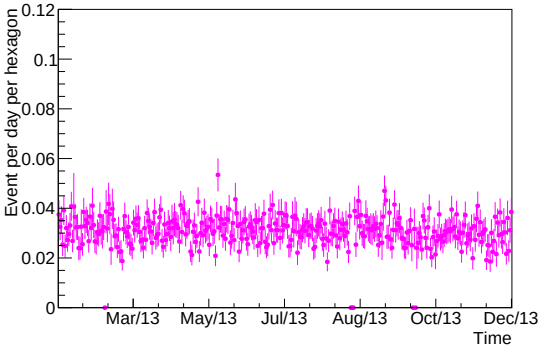
(a) Daily rates in 2010



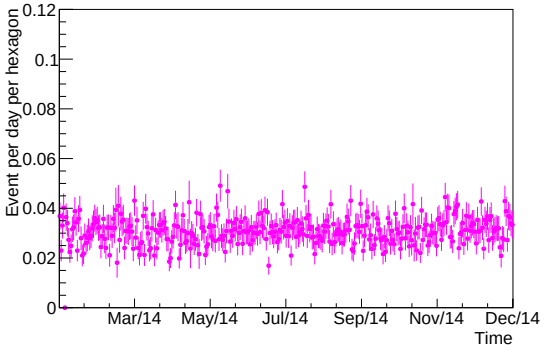
(b) Daily rates in 2011



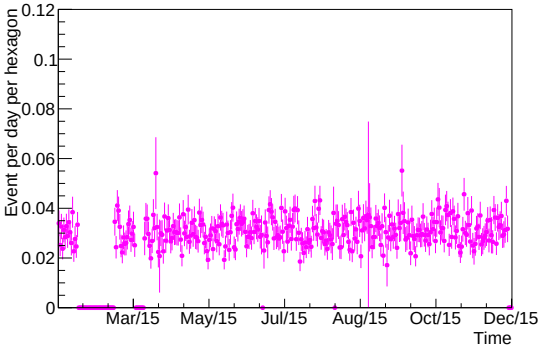
(c) Daily rates in 2012



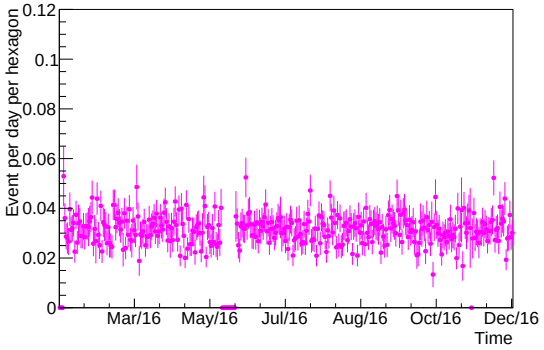
(d) Daily rates in 2013



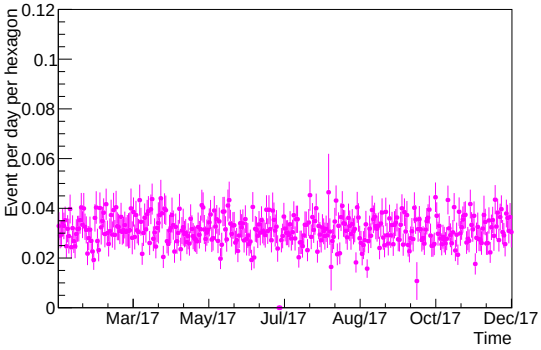
(e) Daily rates in 2014



(f) Daily rates in 2015



(g) Daily rates in 2016



(h) Daily rates in 2017

Figure B.35: Daily rates for each year, from 2012 to 2017.

References

[1] A. A. Watson. Cosmic Rays of the Highest Energy. *Royal Astronomical Society, Quarterly Journal*, 21:1–13, 1980.

[2] K. H. Kampert and A. A. Watson. Extensive air showers and ultra high-energy cosmic rays: A historical review. *European Physical Journal H*, 37(3):359–412, 2012.

[3] C. Coulomb. *Premier-[troisième] mémoire sur l'electricité et le magnétisme*. Nineteenth Century Collections Online (NCCO): Science, Technology, and Medicine: 1780-1925. Académie Royale des sciences, 1785.

[4] T. Stanev. *High Energy Cosmic Rays*. Springer, Chichester, UK, 2nd edition, 2010.

[5] K. Sakurai. High-energy phenomena on the sun: An introductory review. *Space Science Reviews*, 51(1-2):1–9, 1989.

[6] L. Miroshnichenko. *Solar cosmic rays: Fundamentals and Applications*, volume 405. Springer, 2015.

[7] E. Fermi. The origin of cosmic radiation. *Physical Review*, 75(8):1169–1174, 1949.

[8] R. J. Protheroe. Acceleration and Interaction of Ultra High Energy Cosmic Rays. *Topics in Cosmic Ray Astrophysics*, 230:247, 1998.

[9] R. D. Blandford and J. P. Ostriker. Particle acceleration by astrophysical shocks. *The Astrophysical Journal*, 221:L29, 1978.

[10] L. O'C. Drury. Origin of Cosmic Rays. *Astroparticle Physics*, 39-40:52–60, March 2012.

[11] A. M. Hillas. The Origin of Cosmic Rays. *Annual Review of Astronomy and Astrophysics*, 22:425–444, 1984.

[12] J. M. Thomas-Albury. *Extending the Energy Range of Ultra-High Energy Cosmic Ray Fluorescence Detectors*. PhD thesis, University of Adelaide, 2020.

[13] S. G. Lucek and A. R. Bell. Non-linear amplification of a magnetic field driven by cosmic ray streaming. *Monthly Notices of the Royal Astronomical Society*, 314(1):65–74, May 2000.

[14] J. Blümer, R. Engel, and J. R. Hörandel. Cosmic rays from the knee to the highest energies. *Progress in Particle and Nuclear Physics*, 63(2):293–338, 2009.

[15] J. Kolata. *Neutron stars, Black Holes and Gravitational Waves*. Morgan & Claypool Publishers, San Rafael, CA,USA, 2019.

[16] P. Blasi, R. Epstein, and A. Olinto. Ultra- High Energy Cosmic Rays from Young Neutron Stars. *The Astrophysical Journal*, 533:123–126, 2000.

[17] A. Abramowski et al. Acceleration of petaelectronvolt protons in the Galactic Centre. *Nature*, 531:476–479, March 2016.

[18] Y. Fujita, K. Murase, and S. S. Kimura. Sagittarius A* as an origin of the Galactic PeV cosmic rays? *Journal of Cosmology and Astroparticle Physics*, 2017(037), April 2017.

[19] R. Abbasi et al. Observation of anisotropy in the galactic cosmic-ray arrival directions at 400 TeV with IceCube. *Astrophysical Journal*, 746(1), 2012.

[20] A. Chiavassa et al. KASCADE-Grande experiment measurements of the cosmic ray spectrum and large scale anisotropy. *Nuclear and Particle Physics Proceedings*, 279-281:56–62, 2016.

[21] IceCube Collaboration. Anisotropy in Cosmic-Ray Arrival Directions in the Southern Hemisphere with Six Years of Data from the IceCube Detector. *The Astrophysical Journal*, 826(2):220, March 2016.

[22] Z. Cao et al. Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray Galactic sources. *Nature*, 594(7861):33–36, 2021.

[23] M. Amenomori et al. First Detection of sub-PeV Diffuse Gamma Rays from the Galactic Disk: Evidence for Ubiquitous Galactic Cosmic Rays beyond PeV Energies. *Physical Review Letters*, 126(14):1–7, 2021.

[24] A. U. Abeysekara et al. HAWC observations of the acceleration of very-high-energy cosmic rays in the Cygnus Cocoon. *Nature Astronomy*, 5(5):465–471, 2021.

[25] F. Halzen and E. Zas. Neutrino Fluxes from Active Galaxies: A Model-Independent Estimate. *The Astrophysical Journal*, 488(2):669–674, October 1997.

[26] P. S. Conti, P. A. Crowther, and C. Leithere. *From luminous hot stars to starburst galaxies*. Cambridge astrophysics series ; 45. Cambridge University Press, Cambridge, UK, 2008.

[27] E. Waxman. Cosmological Gamma-Ray Bursts and the Highest Energy Cosmic Rays. *Physical Review Letters*, 75(3):386–389, 1995.

[28] M. Vietri. The Acceleration of Ultra-High Energy Cosmic Rays in Gamma-ray Bursts. *The Astrophysical Journal*, 453:883–889, 1995.

[29] E. Roulet. Large-scale anisotropies above 0.03 EeV measured by the Pierre Auger Observatory. In *Proceedings of Science ICRC2019*, Madison, Wisconsin, USA, 2019.

[30] A. Aab et al. Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8 x 10^{18} eV. *Science*, 357:1266–1270, 2017.

[31] A. Aab et al. Large-scale cosmic-ray anisotropies above 4 EeV measured by the Pierre Auger Observatory. *The Astrophysical Journal*, 868(4):1–12, 2018.

[32] K. Kawata et al. Updated results on the UHECR hotspot observed by the Telescope Array experiment. In *Proceedings of Science ICRC2019*, Madison, Wisconsin, USA, 2019.

[33] L. Caccianiga. Anisotropies of the highest energy cosmic-ray events recorded by the Pierre Auger Observatory in 15 years of operation. In *Proceedings of Science*, 2019.

[34] P. Abreu et al. Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter. *Astroparticle Physics*, 34:314–326, 2010.

[35] R. U. Abbasi et al. Indications of intermediate-scale anisotropy of cosmic rays with energy greater than 57 EeV in the northern sky measured with the surface detector of the Telescope Array experiment. *Astrophysical Journal Letters*, 790(2), 2014.

[36] A. Haungs, H. Rebel, and M. Roth. Energy spectrum and mass composition of high-energy cosmic rays. *Reports on Progress in Physics*, 66:1145–1207, 2003.

[37] W. Heitler. *The quantum theory of radiation*. International series on monographs on physics. Clarendon Press, Oxford, 3rd ed. edition, 1954.

[38] A. Saftoiu. Overview of the Pierre Auger observatory dedicated to the study of cosmic rays. *AIP Conference Proceedings*, February 2019.

[39] J. Matthews. A Heitler model of extensive air showers. *Astroparticle Physics*, 22(5-6):387–397, 2005.

[40] H. W. Atherton et al. Direct measurement of the lifetime of the neutral pion. *Physics Letters B*, 158(1):81–84, 1985.

[41] A. J. Greenberg et al. Charged-pion lifetime and a limit on a fundamental length. 23(21):1267–1270, 1969.

[42] A. Yushkov. Mass composition of cosmic rays with energies above $10^{17.2}$ eV from the hybrid data of the Pierre Auger Observatory. In *Proceedings of Science ICRC2019*, Madison, Wisconsin, USA, 2019.

[43] W. Hanlon. Telescope Array 10 Year Composition. In *Proceedings of Science ICRC2019*, Madison, Wisconsin, USA, 2019.

[44] B. Peters. Primary cosmic radiation and extensive air showers. *Il Nuovo Cimento*, 22(4):800–819, 1961.

[45] D. Allard, N. G. Busca, G. Decerprit, A. V. Olinto, and E. Parizot. Implications of the cosmic ray spectrum for the mass composition at the highest energies. *Journal of Cosmology and Astroparticle Physics*, 2008(10), 2008.

[46] R. Aloisio, V. Berezinsky, and A. Gazizov. Ultra high energy cosmic rays: The disappointing model. *Astroparticle Physics*, 34(8):620–626, 2011.

[47] D. Hooper and A. M. Taylor. On the heavy chemical composition of the ultra-high energy cosmic rays. *Astroparticle Physics*, 33(3):151–159, 2010.

[48] K. Fang, K. Kotera, and A. V. Olinto. Newly born pulsars as sources of ultrahigh energy cosmic rays. *Astrophysical Journal*, 750(2), 2012.

[49] A. Aab et al. Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. Measurements at energies above $10^{17.8}$ eV. *Physical Review D*, 90(122005):1–25, 2014.

[50] A. Aab et al. Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II: Composition implications. *Physical Review D*, 90(122006):1–12, 2014.

[51] K. Greisen. End to the cosmic-ray spectrum? *Physical Review Letters*, 16(17):748–750, 1966.

[52] G. T. Zatsepin and V. A. Kuz'min. Upper Limit of the Spectrum of Cosmic Rays. *Soviet Journal of Experimental and Theoretical Physics Letters*, 4:78, August 1966.

[53] P. Sommers and S. Westerhoff. Cosmic ray astronomy. *New Journal of Physics*, 11, 2009.

[54] F. G. Schröder. Radio detection of cosmic-ray air showers and high-energy neutrinos. *Progress in Particle and Nuclear Physics*, 93:1–68, 2017.

[55] A. Putze. Direct cosmic-ray detection. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 662(SUPPL. 1):S157–S163, 2012.

[56] R. Poggiani. *High Energy Astrophysical Techniques*. Springer, 2017.

[57] T. Francke. *Position-sensitive gaseous photomultipliers : research and applications*. Advances in chemical and materials engineering (ACME) book series. Engineering Science Reference, Hershey, Pennsylvania, 2016.

[58] M Nagano and A A Watson. Observations and implications of the ultrahigh-energy cosmic rays. *Reviews of Modern Physics*, 72:689–732, 2000.

[59] P. Billoir. The Cherenkov Surface Detector of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 766:78–82, 2014.

[60] G. Rowell. Gamma-ray astronomy, August 2019. Honours Astrophysics 2019.

[61] T. C. Weekes. The atmospheric Cherenkov technique in Very High Energy Gamma-Ray Astronomy. *Space Science Reviews*, 75(1-2):1–15, January 1996.

[62] M. Ave et al. Measurement of the pressure dependence of air fluorescence emission induced by electrons. *Astroparticle Physics*, 28(1):41–57, 2007.

[63] M. Nagano, K. Kobayakawa, N. Sakaki, and K. Ando. New measurement on photon yields from air and the application to the energy estimation of primary cosmic rays. *Astroparticle Physics*, 22:235–248, June 2004.

[64] H. Falcke et al. Detection and imaging of atmospheric radio flashes from cosmic ray air showers. *Nature*, 435(7040):313–316, 2005.

[65] G. A. Askaryan. Excess negative charge of an electron-photon shower and the coherent radio emission from it. *Zhur. Eksptl'. i Teoret. Fiz.*, 41, 1961. Available at `https://www.osti.gov/biblio/4833087`.

[66] T. Huege and D. Besson. Radio-wave detection of ultra-high-energy neutrinos and cosmic rays. *Progress of Theoretical and Experimental Physics*, 2017(12):1–50, December 2017.

[67] C. Chen et al. Taiwan astroparticle radiowave observatory for geo-synchrotron emissions (TAROGE). In *International Cosmic Ray Conference 2015*, 2015.

[68] F. G. Schröder. *Instruments and Methods for the Radio Detection of High Energy Cosmic Rays*. Springer, 1 edition, 2019.

[69] J. Linsley, S. Livio, and R. Bruno. Extremely energetic cosmic-ray event. *Physical Review Letters*, 6:485–487, 1961.

[70] L. Bonolis. From cosmic ray physics to cosmic ray astronomy: Bruno Rossi and the opening of new windows on the universe. *Astroparticle Physics*, 53:67–85, 2014. Centenary of cosmic ray discovery.

[71] J. Linsley, L. Scarsi, and B. Rossi. Energy Spectrum and Structure of Large Air Showers. *Journal of the Physical Society of Japan Supplement*, 17:91, January 1962.

[72] J. Linsley. Evidence for a Primary Cosmic-Ray Particle with Energy 10^{20} eV. *Phys. Rev. Lett.*, 10:146–148, February 1963.

[73] J. Linsley. Primary cosmic rays of energy 10^{17} to 10^{20} eV, the energy spectrum and arrival directions. In *International Cosmic Ray Conference*, page 77, January 1963.

[74] H. E. Bergeson et al. Measurement of light emission from remote cosmic-ray air showers. *Physical review letters*, 39(13):847–849, 1977.

[75] R. M. Baltrusaitis et al. The Utah Fly's Eye detector. *Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment*, 240(2):410–428, 1985.

[76] D. J. Bird et al. Results from the fly's eye experiment. *AIP Conference Proceedings*, 338(1):839–854, 1995.

[77] T. Abu-Zayyad et al. The prototype high-resolution Fly's Eye cosmic ray detector. *Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment*, 450(2):253–269, 2000.

[78] R. U. Abbasi et al. First observation of the Greisen-Zatsepin-Kuzmin suppression. *Physical Review Letters*, 100, March 2008.

[79] M. A. Lawrence et al. The cosmic ray energy spectrum above 4 x 10^{17} eV as measured by the Haverah Park Array. *Journal of Physics G: Nuclear and Particle Physics*, 17:733–757, 1991.

[80] A. J. Bower, G. Brooke, D. Pearce, J. C. Perrett, and A. A. Watson. The lateral distribution of scintillator densities in air showers produced by cosmic rays with $E > 5 \times 10^{18}$ eV. *Journal of Physics G: Nuclear Physics*, 9(12):1569–1576, December 1983.

[81] A. A. Ivanov, S. P. Knurenko, and I. Y. Sleptsov. Measuring extensive air showers with Cherenkov light detectors of the Yakutsk array: The energy spectrum of cosmic rays. *New Journal of Physics*, 11, June 2009.

[82] A. Ivanov. Some results from the Yakutsk array experiment and the status of the modernization program. *EPJ Web of Conferences*, 208:08005, 2019.

[83] A. Ivanov. The Yakutsk array experiment: Main results and future directions. *EPJ Web of Conferences*, 53, 2013.

[84] P. Sokolsky, P. Sommers, and B. R. Dawson. Extremely high energy cosmic rays. *PHYSICS REPORTS (Review Section of Physics Letters)*, 217:225–277, August 1992.

[85] T. Stanev et al. Arrival directions of the most energetic cosmic rays. *Physical Review Letters*, 75:3056–3059, October 1995.

[86] L. J. Kewley, R. W. Clay, and B. R. Dawson. Arrival directions of the southern highest energy cosmic rays. *Astroparticle Physics*, 5:69–74, January 1996.

[87] M M Winn, J Ulrichs, L S Peak, B A Mccusker, and L Horton. The cosmic-ray energy spectrum above 10^{17} eV. *Journal of Physics G: Nuclear Physics J. Phys. G: Nucl. Phys*, 12:653–674, 1986.

[88] N. Chiba et al. Akeno Giant Air Shower Array (AGASA) covering 100 km^2 area. *Nuclear Instruments and Methods in Physics Research A*, 311:338–349, July 1992.

[89] M. Nagano. Results from AGASA experiment in the extremely high energy region. In *AIP Conference Proceedings*, 1998.

[90] M. Takeda et al. Small-Scale Anisotropy of Cosmic Rays above 10^{19} eV Observed with the Akeno Giant Air Shower Array . *The Astrophysical Journal*, 522(1):225–237, 1999.

[91] H. Tokuno et al. New air fluorescence detectors employed in the Telescope Array experiment. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 676:54–65, June 2012.

[92] S. Ogio. Telescope Array Low-energy Extension (TALE) hybrid. In *Proceedings of Science ICRC2019*, Madison, Wisconsin, USA, 2019.

[93] H. Kawai et al. Measurement of ultra-high energy cosmic rays by Telescope Array (TA). *Journal of the Physical Society of Japan*, 78(SUPPL. A):108–113, 2009.

[94] T. Abu-Zayyad et al. The cosmic-ray energy spectrum observed with the surface detector of the Telescope Array experiment. *Astrophysical Journal Letters*, 768:1–5, May 2013.

[95] S. Udo et al. The Telescope Array Low-energy Extension. In *Proceedings of Science ICRC2017*, Bexco, Busan, Korea, 2017.

[96] E. Kido. The TAx4 experiment. In *Proceedings of Science ICRC2017*, Bexco, Busan, Korea, 2017.

[97] E. Kido. Status and prospects of the TAx4 experiment. In *Proceedings of Science ICRC2019*, Madison, Wisconsin, USA, 2019.

[98] A. Aab et al. The Pierre Auger Cosmic Ray Observatory. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 798:172–213, 2015.

[99] P. Auger. *What Are Cosmic Rays?* University of Chicago Press, Chicago, 1953.

[100] M. Settimo. Review on extragalactic cosmic rays detection. In *25th European Cosmic Ray Symposium*, Paris, 2016.

[101] X. Bertou et al. Calibration of the surface array of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(2):839–846, 2006.

[102] J. Abraham et al. Trigger and aperture of the surface detector array of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 613(1):29–39, 2010.

[103] J. Abraham et al. The fluorescence detector of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 620(2-3):227–251, 2010.

[104] J. Linsley and L. Scarsi. Arrival Times of Air Shower Particles at Large Distances from the Axis. *Physical Review*, 128(5):2384–2392, 1962.

[105] A. A. Watson and J. G. Wilson. Fluctuation studies of large air showers: The composition of primary cosmic ray particles of energy Ep∼1018 eV. *Journal of Physics A: General Physics*, 7(10):1199–1212, 1974.

[106] A. Coleman. *Measurement of the Cosmic Ray Flux above 100 PeV at the Pierre Auger Observatory*. PhD thesis, Pennsylvania State University, 2018.

[107] K. Kamata and J. Nishimura. The Lateral and the Angular Structure Functions of Electron Showers. *Progress of Theoretical Physics Supplement*, 6:93–155, 1958.

[108] J. Hersil, I. Escobar, D. Scott, G. Clark, and S. Olbert. Observations of extensive air showers near the maximum of their longitudinal development. *Physical Review Letters*, 6(1):22–23, 1961.

[109] A. Schulz. The measurement of the energy spectrum of cosmic rays above 3 x 10^{17} eV with the Pierre Auger Observatory. In *Proceedings of the 33rd International Cosmic Rays Conference, ICRC 2013*, 2013.

[110] P. Nguyen. *Energy Systematics and Long Term Performance of the Pierre Auger Observatory's Fluorescence Telescopes*. PhD thesis, University of Adelaide, 2018.

[111] P. Abreu et al. The effect of the geomagnetic field on cosmic ray energy estimates and large scale anisotropy searches on data from the Pierre Auger Observatory. *Journal of Cosmology and Astroparticle Physics*, 2011(11), 2011.

[112] P. Bauleo et al. Absolute calibration of the Auger fluorescence detectors. In *29th International Cosmic Ray Conference, ICRC 2005*, 2005.

[113] J. Abraham et al. A study of the effect of molecular and aerosol conditions in the atmosphere on air fluorescence measurements at the Pierre Auger Observatory. *Astroparticle Physics*, 33(2):108–129, 2010.

[114] A. Aab et al. Impact of atmospheric effects on the energy reconstruction of air showers observed by the surface detectors of the Pierre Auger Observatory. *Journal of Instrumentation*, 12(2), 2017.

[115] C. Meurer and N. Scharf. HEAT - A low energy enhancement of the Pierre Auger Observatory. *Astrophysics and Space Sciences Transactions*, 7(2):183–186, 2011.

[116] T. K. Gaisser and A. M. Hillas. Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers. In *International Cosmic Ray Conference*, January 1977.

[117] V. Verzi. The Energy Scale of the Pierre Auger Observatory. In *Proceedings of the 33rd International Cosmic Rays Conference, ICRC 2013*, 2013.

[118] B. Fick et al. The central laser facility at the Pierre Auger Observatory. *Journal of Instrumentation*, 1(11):P11003–P11003, November 2006.

[119] S. Y. BenZvi et al. The Lidar system of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 574(1):171 – 184, 2007.

[120] V. Rizi, C. Medina Hernandez, M. Iarlori, and E. Pietropaolo. The Auger Raman Lidar: several years of continuous observations. *EPJ Web of Conferences*, 197(02003), 2019.

[121] V. M. Harvey. Real-time measurements with atmospheric instruments at the Pierre Auger Observatory. In *Proceedings of Science ICRC19*, 2019.

[122] A. Aab et al. Prototype muon detectors for the AMIGA component of the Pierre Auger Observatory. *Journal of Instrumentation*, 11(02):P02012–P02012, February 2016.

[123] A. Castellina. AugerPrime: the Pierre Auger Observatory Upgrade. *EPJ Web of Conferences*, 210:06002, May 2019.

[124] E. M. Holt. Recent Results of the Auger Engineering Radio Array (AERA). In *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*, Trieste, Italy, August 2017. Sissa Medialab.

[125] B. Pont. A large radio detector at the Pierre Auger Observatory - Measuring the properties of cosmic rays up to the highest energies. In *Proceedings of Science ICRC19*, Madison, Wisconsin, USA, 2019.

[126] J. Abraham et al. Atmospheric effects on extensive air showers observed with the surface detector of the Pierre Auger Observatory. *Astroparticle Physics*, 32(2):89–99, 2009.

[127] The Pierre Auger Collaboration. The Pierre Auger Observatory I : The Cosmic Ray Energy Spectrum and Related Measurements. In *32nd International Cosmic Ray Conference*, Beijing, 2011.

[128] B. R. Dawson. The Energy Scale of the Pierre Auger Observatory. In *Proceedings of Science ICRC19*, Madison, Wisconsin, USA, 2019.

[129] K. Daumiller, M. D. Rio, H. Mathes, R. Smida, and L. Tomankova. Towards a continuous calibration A during FD data taking. Technical report, Karlsruhe Institute of Technology, 2016.

[130] The Pierre Auger Collaboration. Merged list of cleanings. Available at `https://www.auger.unam.mx/AugerWiki/MergedListOfCleanings`, Accessed: December 2019.

[131] V. Verzi. Measurement of the energy spectrum of cosmic rays using the Pierre Auger Observatory. In *Proceedings of Science ICRC19*, Madison, Wisconsin, USA, 2019.

[132] P. R. Bevington and D. K. Robinson. *Data reduction and error analysis for the physical sciences*. McGraw-Hill, New York, 2nd ed. edition, 1992.

[133] I. Lhenry-Yvon. The "communications crisis" in 2009: proposal for handling data between June and November. Technical report, 2010. GAP2010-020.

[134] I. Lhenry-Yvon. T2 files in different formats. Available at `http://ipnwww.in2p3.fr/~augers/AugerProtected/AcceptT2.html`, Accessed: December 2019.

[135] I. Lhenry-Yvon. SD Performance, Online Auger Collaboration Meeting, April 2020.

[136] K. Choi. Long term performance of the surface detectors of the Pierre Auger observatory. In *Proceedings of Science ICRC2019*, 2019.

[137] R. Sato. Long term performance of the surface detectors of the Pierre Auger observatory. *Proceedings of the 32nd International Cosmic Ray Conference, ICRC 2011*, 3:204–207, 2011.

[138] O. Zapparrata and I. Mariş. Modelling the ageing of WCD using simulations, Online Auger Collaboration Meeting, March 2021.

[139] I. Lhenry-Yvon. Area over Peak ratio and event rate in SD Auger detectors, 2016. GAP2016-033.

[140] J. Roberts G. R. Farrar. SD rate as a diagnostic of energy reconstruction drift, 2013. GAP2013-043.

[141] A. Aab et al. Measurement of the cosmic-ray energy spectrum above 2.5 × 10^{18} eV using the Pierre Auger Observatory. *Physical Review D*, 102:62005, 2020.

[142] I. Mariş. Rejecting Bad-Silent Stations in the SD Reconstruction, 2019. GAP2019-049.

[143] O. Zapparrata. Describing the ageing effects on S1000 and risetime with simulations: preliminary results, September 2021.

[144] GEANT4: A Simulation Toolkit. Available at `https://geant4.web.cern.ch/node/1`, Accessed: October 2021.