Investigation of Cloud characteristics above the Pierre Auger Observatory

THE UNIVERSITY *of* ADELAIDE

Jarryd Day

School of Physical Sciences
The University of Adelaide

This dissertation is submitted for the degree of Doctor of Philosophy

August 2021

Contents

Abstract

The monitoring of the atmosphere above the Pierre Auger Observatory has been done extensively over its 14 year existence, and is important due to the effects of the atmosphere on measurements of cosmic ray air showers with the fluorescence technique.

Four infrared cloud cameras have been installed at the observatory site, and are used for remote sensing of night time cloud above the array. These cameras are radiometric, and sensitive to thermal radiation in the 7-13 μm wavelength band. There are difficulties in identifying clouds in this way since water vapour in the atmosphere absorbs much of the infrared radiation in this wavelength band.

In this dissertation, I present several methods used to estimate the height of clouds in the atmosphere using atmospheric monitoring instruments installed at the Pierre Auger Observatory. I also focus a large portion of my studies on modelling the response of the cloud cameras to this infrared radiation, with particular emphasis on the effects of water vapour and various types of cloud on this response. Seasonal variations of cloud cover and cloud base height are also investigated.

Declaration

I 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.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed: .. .. Date: 19/08/21

Acknowledgements

Firstly, I would like to thank my principal supervisor, Professor Bruce Dawson, as well as my co-supervisor, Emeritus Professor Roger Clay, for their support and encouragement during my candidature. Their passion and drive for research was invaluable, and I carry with me some important lessons learnt from their wisdom and guidance, which I will always be grateful for.

My thanks and gratitude must also go out to the wider Pierre Auger collaboration, in particular to those in the Auger Atmospheric monitoring group. Their input at collaboration meetings has been invaluable, as well as their expertise in being able to solve any issues I encountered and also to allow me the platform to present my research and take on valuable feedback. The collaboration has allowed me to grow my research and presentation skills, without which this thesis would not be possible.

I'd also like to mention that I was given the opportunity to travel to the Pierre Auger Observatory in Argentina in May/June 2017 as part of my research. In those 3 weeks I learnt a lot about how the Observatory functions, as well as all the pieces that go into place for the collaboration to find such exciting astrophysical discoveries. This opportunity is something that I will never forget, and I hope that I can find similar opportunities in my future career.

Next, I would like to thank everyone else in the High Energy Astrophysics Group at the University of Adelaide that I worked alongside during my candidature, their suggestions in group meetings and research presentations proved useful in many of my research endeavours, which I am thankful for. It was amazing to meet so many wonderful people with the same passion for Physics as I have, there are friendships I've made which I will cherish forever.

I also wish to extend my thanks to Trent Grubb and Patrick van Bodegom, for their maintenance of, and input into, all cloud camera related research, without which most of this thesis would not be possible. The many years of expertise and understanding of all software and hardware related issues with the cloud cameras has made interpreting and analysing the data much easier.

I would also like to thank all of my family; in particular, my father Gavin, mother Katrina, and sister Meagan. Their support and encouragement are forever appreciated, and I would especially like to thank them for keeping me grounded, and sane. There are many wonderful friends I have made, through University, high school and other social/community activities. Those friends have supported and encouraged me to pursue this goal, and I cannot thank them enough.

Lastly, I want to thank my wonderful girlfriend Rebecca. Without her love and support, much of this would not have been possible. I cannot put into words how thankful I am to have her in my life, to encourage me every step of the way, through success and failure, to achieve this goal.

Chapter 1

Introduction

This brief introduction will outline the details of every chapter of this dissertation.

Chapters 2 and 3 go through the background information required for this dissertation, namely what cosmic rays are and why we study them, the instrumentation employed at the Pierre Auger Observatory that we use to study these cosmic rays, as well as the atmosphere through which the cosmic rays traverse.

Chapter 4 will outline the basic characteristics of Earth's atmosphere, as well as the various types of clouds we see. On top of this, it will cover the radiative properties of clouds (important for later studies), before diving into my first analysis study. This was a model I developed to estimate the height of clouds in the atmosphere when our other monitoring instruments were unavailable, by using GDAS (Global Data Assimilation System), a software modelling tool that uses observed data predominantly for weather forecasting.

Chapter 5 outlines the cloud camera analysis done at Auger, specifically when it comes to corrections that must be made to the produced images. It will then go into detail about the cloud mask production and why it is important, with particular emphasis on improvements that have been made to the process (the extension to include HEAT [High Elevation Auger Telescopes] in the analysis, and the improvements made to the speed of production by taking advantage of a Graphical User Interface).

Chapter 6 will go through a study on the properties of clouds and how they vary over the different seasons of the year. Two relevant properties of cloud are the cloud base height and cloud coverage. Both of these properties will be measured and compared between the LIDARs and cloud cameras in operation at the Pierre Auger Observatory. These instruments are outlined in Chapters 3 and 5 respectively.

Chapter 7 contains information on a study of various cloud models I performed using an atmospheric transmission and propagation software tool called MODTRAN (Moderate Resolution Atmospheric Transmission). Analysis was done specifically on the difference between cirrus cloud (made of ice particles) and all other types of cloud (made up of water vapour), to study what differences exist in the radiative properties of the various types of cloud. Measurements with the cloud cameras themselves were compared to the MODTRAN model simulations in order to accurately depict the behaviour of clouds on our image analysis.

Chapter 8 details another cloud height study, this time utilising a variety of instruments (cloud cameras, weather stations and GDAS). Details are provided on the potential errors that may arise from this technique, as well as a comprehensive discussion of each instrument separately.

Lastly, Chapter 9 will be a short concluding chapter, discussing all the results found in the thesis as well as some potential future projects that may arise from this.

The hope is that all of these studies will provide useful information on the properties, and movement of clouds in the atmosphere, while also developing a better understanding of the effect of clouds on our cosmic ray shower analysis.

Chapter 2

UHE Cosmic Ray Physics

2.1 Cosmic Rays

Cosmic rays are highly energetic charged particles (namely protons and other heavier nuclei) that travel through space and interact with molecules in the upper atmosphere of Earth. Cosmic ray particles have been studied by physicists for over 100 years, starting in 1912. Victor Hess travelled in a high altitude balloon, reaching a height of 5000 m where he first discovered "penetrating radiation" coming from space. Hess is known as the father of cosmic ray research. [1] [2].

In 1927, Soviet physicist Dmitri Skobeltsyn was able to take the first photographs of the track left by a cosmic ray particle in a cloud chamber. A cloud chamber is simply a particle detector that can visibly track ions that are produced by collisions with other molecules in a moist gas of water vapour [3]. Carl Anderson used one such cloud chamber in order to discover the positron in 1932, and later the muon (in collaboration with Seth Neddermeyer) in 1937. The discovery of antimatter led to a debate over the nature of cosmic rays. Many suggested cosmic rays were in fact gamma-rays incident from space (which creates electrons and positrons through interactions with the atmosphere), however the muon discovery led to the birth of particle physics, which gave evidence that cosmic rays were in fact, energetic particles.

Just one year later, in 1938, physicists (led by Pierre Victor Auger) were able to discover the existence of air showers by noticing the coincident arrival time of particles to several ground detectors [1]. Being able to understand how these air showers develop in the atmosphere is crucial in obtaining an estimate of the energy of the primary incident particle [1] and the nature of this primary particle. The discovery of charged and neutral pions (and their decay products) in 1947 and 1950 respectively meant that scientists were able to recognise the primary particle as being either a proton or charged nucleus (before this, electrons and photons were thought of as the primary particle) which led to a much deeper understanding of the development of air showers in the atmosphere. More information on air showers will be discussed in a later part of this chapter.

At the same time, two teams of physicists led by Bruno Rossi and Georgi Zatsepin began building arrays of detectors in order to detect and study air showers. In 1949, Enrico Fermi proposed an explanation on how cosmic rays may be accelerated to such high energies, by the "shock acceleration" mechanism. Fermi's cosmic ray accelerator model proposes that particles such as protons are accelerated by colliding with moving magnetic clouds in space. Supernovae were proposed as cosmic accelerators, however

supernovae alone are not able to account for the very highest energy cosmic rays, which require much more powerful accelerators to reach these energies (above 10^{20}eV).

John Linsley and his team of scientists detected the first cosmic ray with energy $> 10^{20}$eV in the Volcano Ranch array in New Mexico, USA in 1962. This was significant since cosmic rays at this energy are extremely rare (about 1 per square kilometre per century arrive at Earth). Later in the 60's, Arno Penzias and Robert Wilson discovered that low-energy microwaves permeate across the universe. As a result, Kenneth Greisen, Vadem Kuzmin and Georgi Zatsepin pointed out that high-energy cosmic rays would interact with this microwave background, reducing their energy. This became known as the Greisen-Zatsepin-Kuzmin limit (or simply GZK limit), which is important for the energy spectrum discussed later (see section 2.3).

In 1991 the Fly's Eye cosmic ray research group in Utah observed a cosmic ray event with a primary energy of 3×10^{20}eV, clearly the highest energy event ever observed. There were also two other high energy events observed (about 2×10^{20}) by the AGASA group in Japan and the Yakutsk group in Russia. So where did such high energy cosmic rays come from? None of them seemed to point back to any astrophysical objects that could produce particles of such high energy. So, in 1995, an international group of researchers came together in order to propose a new cosmic ray observatory that could answer this (and other) questions. Thus, the Pierre Auger Observatory was born. For more information on Auger, see Chapter 3.

2.2 Extensive Air Showers

Cosmic rays travel from space until they interact high up in the Earth's atmosphere. The resulting interactions between the incident cosmic ray and the atmosphere create a cascade of particles that disperse through the atmosphere as an EAS (Extensive Air Shower). Most commonly, the initial interaction between the primary cosmic ray and the atmosphere produces both neutral and charged pions. The neutral pions decay quickly (mean lifetime of 8.5×10^{-17}s) into two gamma ray photons that kick-start the electromagnetic component of the shower, whilst the charged pions, which take longer to decay (mean lifetime of 26 ns), mainly decay into muons and neutrinos. The results of these reactions are shown by equations 2.1, 2.2 and 2.3

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{2.1}$$

$$\pi^- \rightarrow \mu^- + \nu_\mu \tag{2.2}$$

$$\pi^0 \rightarrow \gamma + \gamma \tag{2.3}$$

The result is three main components to such an air shower; an electromagnetic component, a hadronic component, and a muonic component. These will be discussed in the subsequent sections.

2.2.1 Electromagnetic (EM) Component

For the EM component, lets consider a neutral pion, produced from the interaction of the primary cosmic ray with the top of the atmosphere. These pions decay quickly, each forming two gamma ray photons. The photons then undergo pair production, which creates an electron-positron pair. Electrons produced this way can then undergo

bremsstrahlung interactions, losing energy and producing more photons. Each successive secondary photon has less and less energy as the shower penetrates deeper into the atmosphere. This is shown by the equations below;

$$\pi^0 \rightarrow \gamma + \gamma \tag{2.4}$$

$$\gamma \rightarrow e^+ + e^- \tag{2.5}$$

$$e^- \rightarrow \gamma \tag{2.6}$$

The EM shower can be modelled by a simple Heitler model, where for each "step" either an electron-positron pair or a bremsstrahlung photon is being produced, after travelling one interaction length d. After travelling this length (on average) in a medium, a photon will produce an electron/positron pair (with each particle retaining half the energy of the original photon), and an electron will undergo bremsstrahlung radiation, producing a photon [4]. Note that this interaction length corresponds to the distance where an electron loses half of its energy (on average) due to radiation losses [4]. After the "nth" step (each one corresponding to a distance λ - equal to a single radiation length) the number of particles present in the shower and the distance travelled by each one can be written as;

$$N = 2^n \tag{2.7}$$

$$d = \lambda ln(2) \tag{2.8}$$

This process continues until the particles reach a critical energy (given the symbol ζ), below which radiative energy losses become less likely than collisional energy losses [4]. In the Heitler model, at this critical energy, a simple relation can be used to find the energy E_0 of the initial pion, where the number defined as N_{max} is the maximum number of particles in the shower at this energy.

$$E_0 = \zeta N_{max} \tag{2.9}$$

Rearranging equation 2.9 yields

$$N_{max} = \frac{E_0}{\zeta} \tag{2.10}$$

Since the number of radiation lengths at the shower's maximum size n_{max} is reached at the critical energy ζ, one can write

$$\zeta = \frac{E_0}{2^{n_{max}}} \tag{2.11}$$

As a result, the number of radiation lengths before the shower reaches maximum development is given by;

$$n_{max} = \frac{ln(E_0/E_c)}{ln(2)} \tag{2.12}$$

Figure 2.1 is a schematic of such a shower as it travels through the atmosphere. When the particles in the shower all reach this critical energy, the cascade has reached an atmospheric depth labelled as X_{max}, which is the penetration depth into the atmosphere at this maximum shower size. The X_{max} value is given by

$$X_{max} = \lambda\, n_{max} = \lambda \frac{ln(E_0/E_c)}{ln(2)} \tag{2.13}$$

where λ is the radiation length in air. From equation 2.13 one can see that $X_{max} \propto ln(E_0)$, that is, the maximum depth is proportional to the logarithm of the initial pion energy. A typical value of this critical energy ζ is around 80 MeV in air [5].

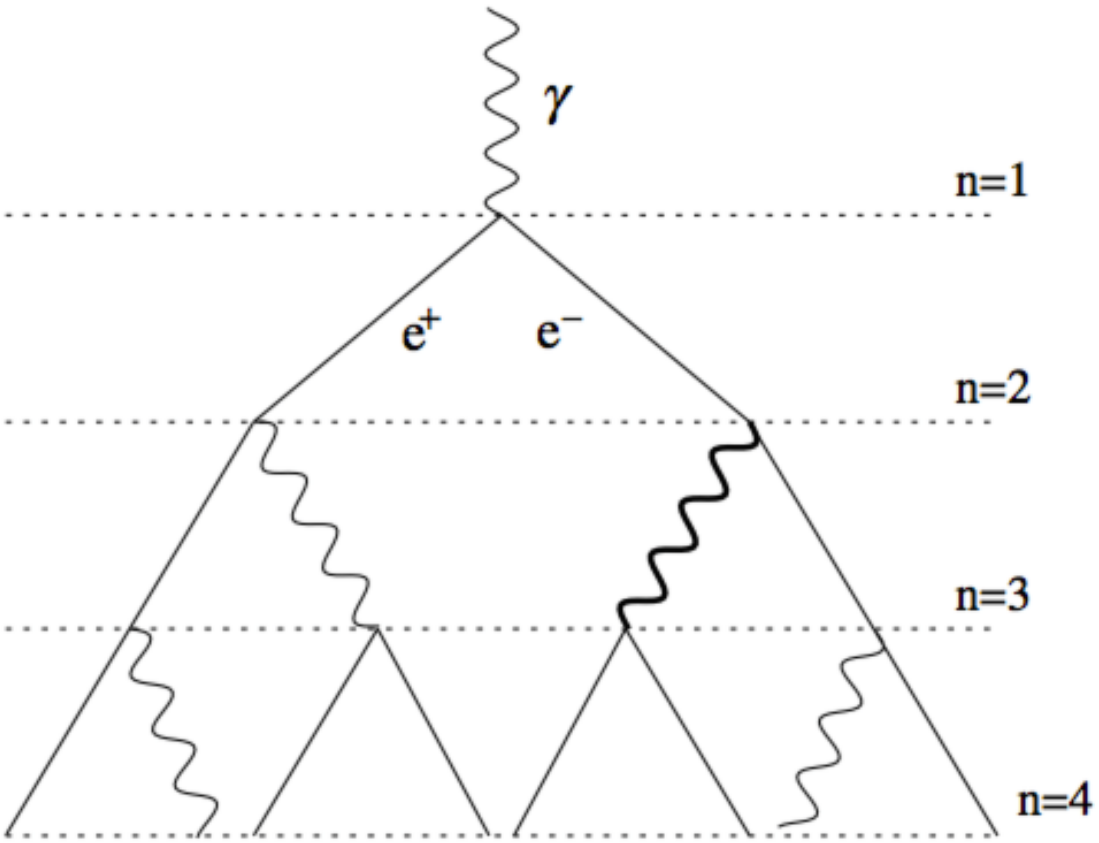

Figure 2.1: The EM Shower model proposed by Heitler. Each n level is a new interaction, producing either a photon or electron/positron pair.

2.2.2 Hadronic Component

Due to the extended lifetime of the positive and negative pions, some exist long enough to interact with molecules in the atmosphere to produce more pions. A schematic of the development of a full extensive air shower is shown in Figure 2.2. As was shown for the electromagentic component of an EAS (see equation 2.13), the depth of shower maximum X_{max} is proportional to $ln(E)$, where E is the energy of the primary cosmic ray particle. Furthermore, a higher initial energy E will mean more interactions can take place before reaching the critical energy ζ. As a result, the extensive air shower can penetrate deeper into the atmosphere, increasing the value of X_{max}. Additionally, the value of X_{max} hints at the composition of the primary incident cosmic ray (for showers initiated by protons or nuclei), which is one of the biggest mysteries in cosmic ray astrophysics at this time. For further discussion on mass composition, refer to section 2.4.

2.2.3 Muonic Component

Muons produced in the shower are formed as decay products of pions. Due to minimal multiple scattering and bremsstrahlung interactions by the muons, they travel in straight lines along the shower axis. As a result, information on their arrival direction can be retained as well as the distribution of production distances. The contribution of the muons that reach the ground depends on a variety of parameters (such as energy loss rate and decay probability) but simple assumptions can be made such that the arrival times give an idea on the distribution of muon production distances along the shower axis. Muons are produced by the decay of pions and kaons and, as such, the distribution of muon production distances provides information about the longitudinal development of the hadronic component of the EAS [6].

Muons produced in the shower decay with a mean lifetime of $\sim 2.2\mu s$, and their decays give electrons, positrons and neutrinos as products.

$$\mu^- \rightarrow e^- + \nu_\mu + \bar{v_e} \tag{2.14}$$

$$\mu^+ \rightarrow e^+ + \bar{\nu_\mu} + v_e \tag{2.15}$$

In this way, the muonic component of the EAS contributes a small portion to the electromagnetic component of an air shower (the electrons and positrons produced are able to undergo bremsstrahlung interactions and produce photons). However, despite their short lifetime, due to relativistic effects (time dilation), some muons are able to travel through the atmosphere and reach the ground. At the Pierre Auger Observatory, many of them pass into the volume of pure water in surface detectors (see section 3.2), resulting in a forward cone of Cherenkov light being emitted. For more information on Cherenkov light emission as a means of particle detection see the discussion in section 3.2.

2.3 Energy Spectrum

The flux of cosmic rays arriving at Earth as a function of energy follows a steep power law of $E^{-\gamma}$ with spectral index γ, a constant with a value of about 2.7 [7]. Figure

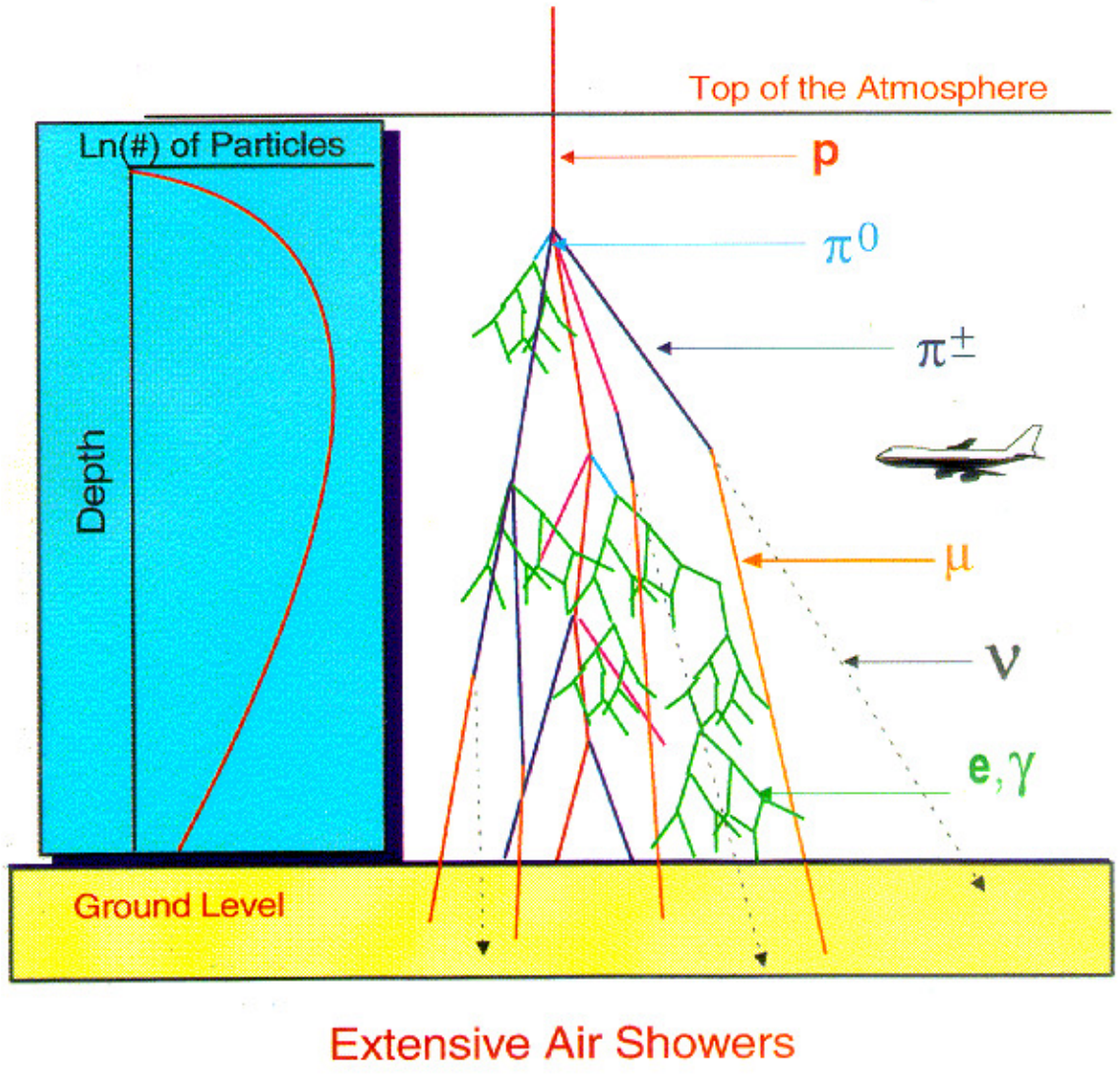

Figure 2.2: The development of an EAS through the atmosphere. The depth at which we have the highest number of particles in the red curve corresponds to X_{max} (furthest right point). The bottom 10 km of the atmosphere are the most relevant for extensive air shower development.

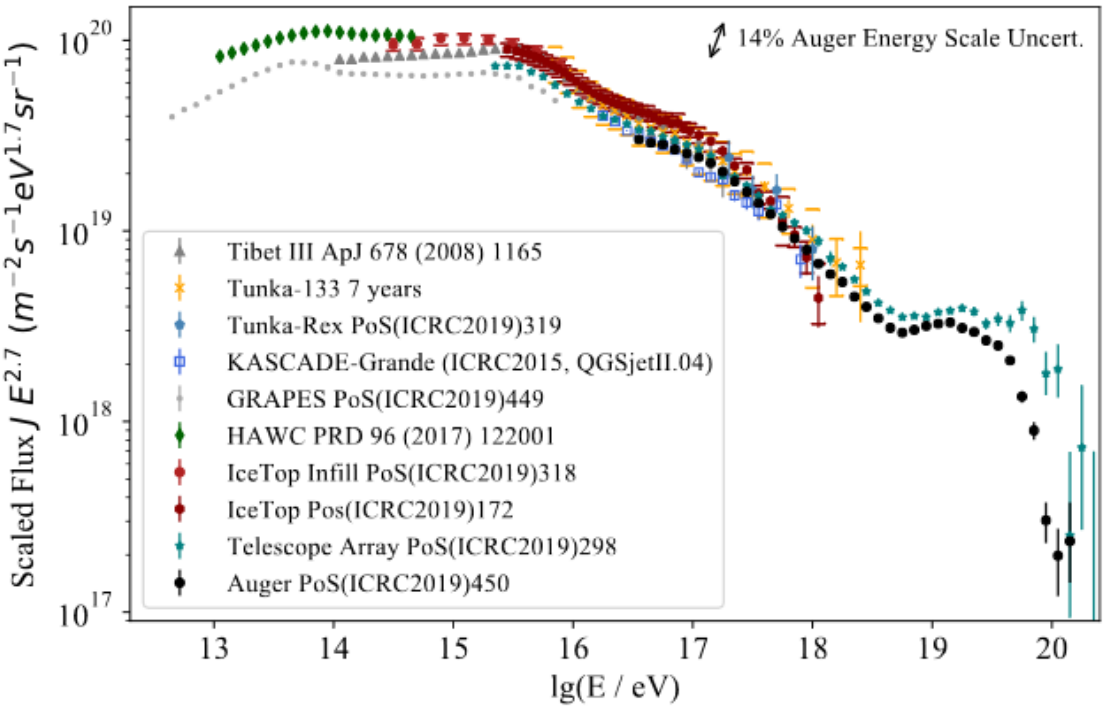

Figure 2.3: The energy spectrum of cosmic rays (for data up to 2017). The flux is multiplied by $E^{2.7}$ in order to emphasise the features around the knee, second knee, and ankle. Source: [9]

2.3 shows this flux as a function of the primary particle energy E [8]. In the spectral region above 10^{19}eV, the flux of cosmic rays is extremely low (approximately 1 particle per square km per year) so one needs a detector with extremely high collection area in order to study the very highest energies. For more on the Pierre Auger Observatory (used for the purpose of studying the highest energies), see chapter 3. The spectrum from Figure 2.3 has some interesting features, starting with the "knee" ($\sim 3 \times 10^{15}$eV), where the spectrum steepens (the value of γ here is roughly 3.1) [7]. The spectrum also steepens again at $\sim 4 \times 10^{17}$eV, which is referred to as the "second knee" of the spectrum. The "ankle" of the spectrum refers to another region of changing spectral index (at around $10^{18.5}$eV), where there is thought to be a transition from galactic sources of cosmic rays to extragalactic sources of cosmic rays. At these extremely high energies (above 10^{19} eV), a suppression in the energy spectrum is observed. This may be due to the GZK limit discussed earlier, since protons will interact with the cosmic microwave background (CMB) and lose energy due to pion photo-production [10]. For nuclei, the energy loss mechanism is photodisintegration, whereby a nucleus will absorb a high energy gamma-ray, which splits the nucleus by releasing either a proton, neutron or alpha particle. The value of the cut-off energy depends on the composition of the primary cosmic ray (see section 2.4 for more). Figure 2.4 illustrates the rate of decrease in energy for protons as they propagate through space. From Figure 2.4, above 100 Mpc the observed energy is always below 10^{20} eV, regardless of the initial energy of the proton [11]. This means that any cosmic ray particles detected with an energy $> 6 \times 10^{19}$ eV have come from a source that is < 100 Mpc away.

As it stands currently, the sources of these extra-galactic UHECRs are unknown, however there are several likely candidates that are theorised to generate such high

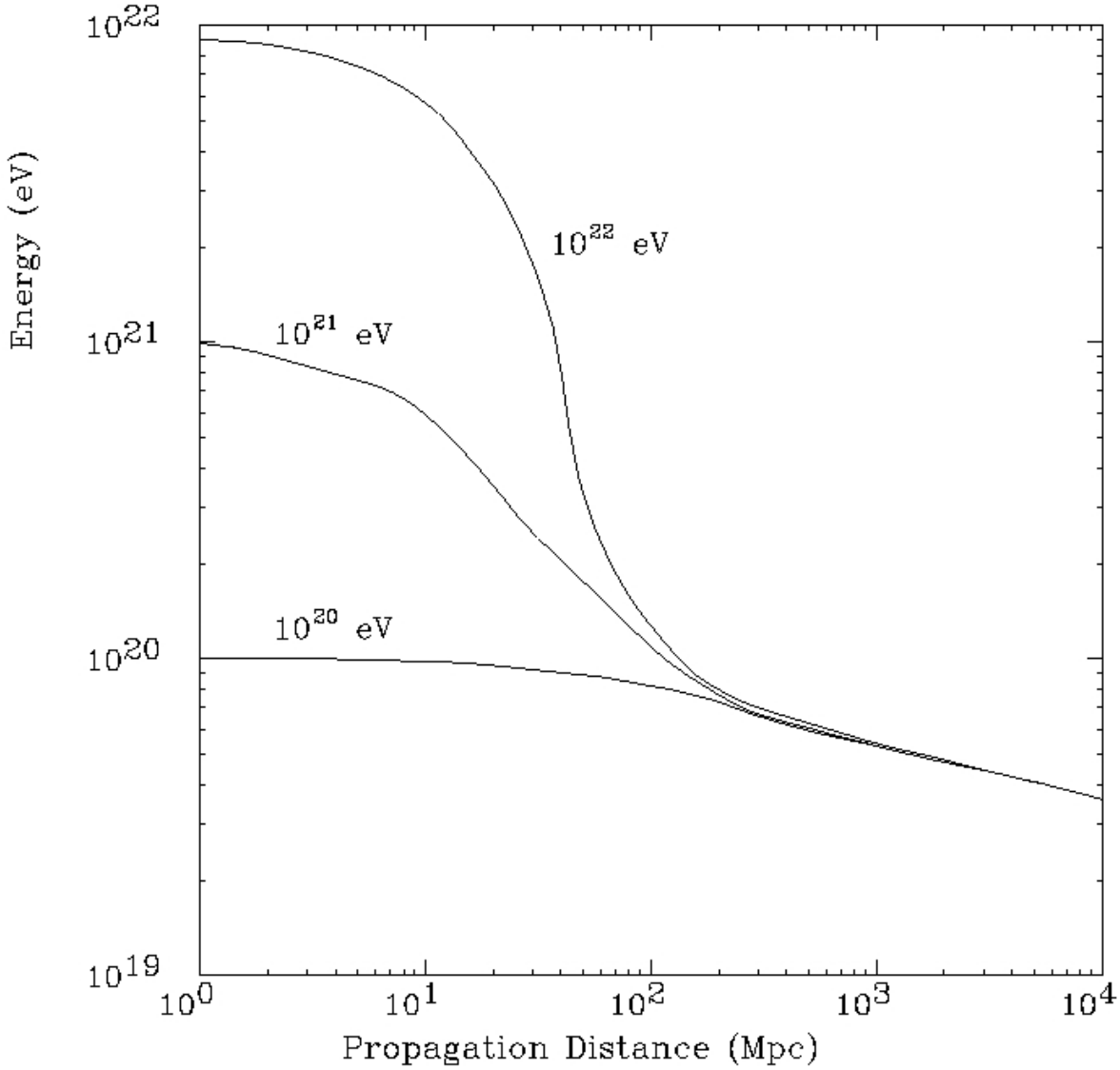

Figure 2.4: Energy of a proton as a function of the propagation distance through the CMB for different initial energies. Higher initial energies correspond to a faster decrease in energy however there is a common convergence for distances $>$ 1000 Mpc. Source: [11]

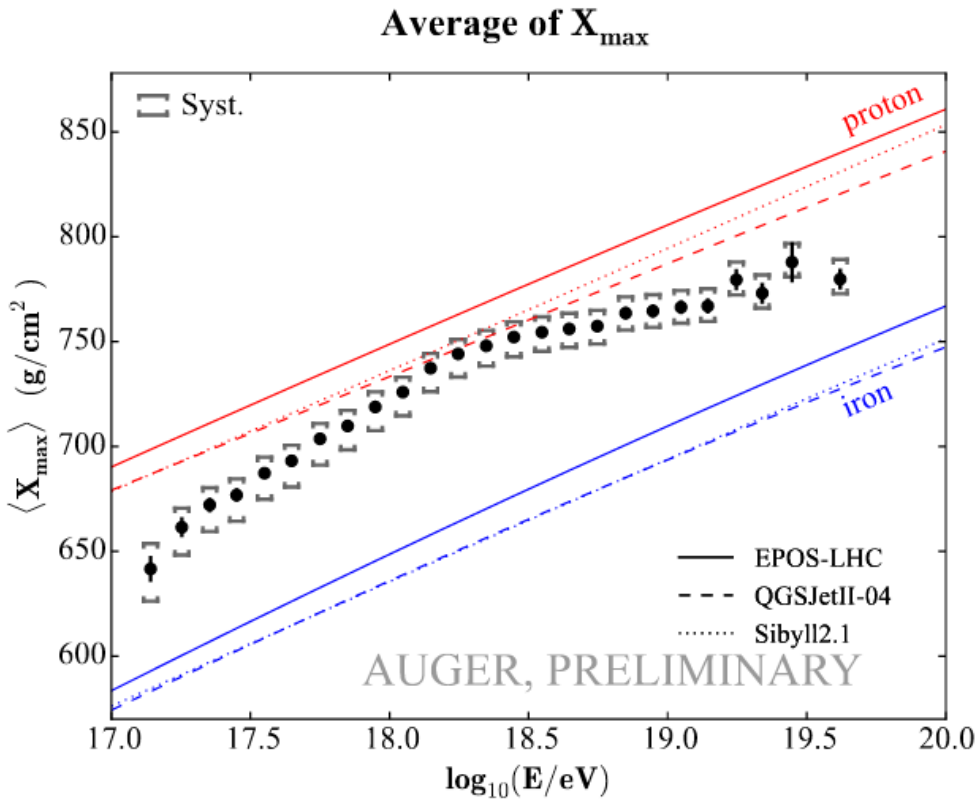

Figure 2.5: Variation of average X_{max} as a function of energy. The red/blue lines correspond to the simulated distributions for three hadronic interaction models for proton and iron primaries respectively. Source: [12]

energy particles. One potential candidate source for UHECR is the jets of Active Galactic Nuclei (AGN). Powered by supermassive black holes at their centre, AGN produce relativistic jets along which particles travel at speeds very close to the speed of light c, accelerated by the powerful magnetic fields present. These objects are intense emitters of X-rays, radio waves, infrared and optical light. Electrons accelerated by these fields produce synchrotron emission.

2.4 Mass Composition

Attempts to measure the mass composition of incident cosmic ray particles have proven to be an ongoing scientific challenge. In order to effectively determine the chemical composition of the primary cosmic ray particles at the highest energies, we need a strong knowledge of the hadronic interactions [12] and the development of the air shower in the atmosphere. Our most sensitive parameter to measure the mass composition of cosmic ray primaries is X_{max}, the atmospheric depth at which maximum energy of the extensive air shower is deposited. Figure 2.5 shows the distribution of measured X_{max} from the Fluorescence Detectors at the Pierre Auger Observatory (see section 3.1) as a function of energy. The proton (red) and iron (blue) expected distributions for three different hadronic interaction models are also shown. Figure 2.5 illustrates a tendency for heavier composition at the lower energy range (around 10^{17}eV, which is consistent with our knowledge of galactic sources [13]). This behaviour is also apparent at the

highest energies ($> \sim 10^{18.3}$eV) where the composition appears to get heavier with increasing energy past $10^{18.3}$eV [14]. The hope is that upgrades to the Observatory such as HEAT, AMIGA and AugerPrime (see Chapter 3) will give us a deeper understanding of the composition of cosmic rays and extensive air showers. In the intermediate energy range ($10^{17.5} - 10^{18.5}$) it appears as if proton primaries are the most likely composition (see Figure 2.5). This point is around the ankle of the energy spectrum, and is also the same energy where we see the anistropy in arrival direction of the primary cosmic ray. For more information on this anisotropy see section 2.5.3.

2.5 Origin of Cosmic Rays

There is a great mystery surrounding the origin of cosmic rays and how they are able to be accelerated to such extreme energies. There are two major issues encountered when trying to determine the origin of these particles. The first of these is due to the presence of magnetic fields in our own Milky Way galaxy and in intergalactic space. The first issue with these magnetic fields is that we do not know their strength or direction, making it extremely difficult to determine their effects on charged particles passing through them. Secondly, the composition of an individual primary cosmic ray particle is not known, as the gyroradius of a charged particle (i.e. how much it is deflected in the presence of a magnetic field) is dependent upon the particle's charge and mass (see equation 2.16), something that is unknown unless the primary particle composition is known.

$$r = \frac{mv_{\perp}}{|q|B} \tag{2.16}$$

In equation 2.16, B is the magnetic field strength and $v_{\perp}$ is the particle velocity component perpendicular to the direction of the field. Since we do not know either, it means that we are unable to determine how much a cosmic ray particle has been deflected, thus we are unable to trace back to its point of origin. For gamma rays this is not an issue since they carry no electrical charge, thus are not deflected when travelling through magnetic fields [15]. This makes tracing back to their origin point a much easier task [15]. These two reasons have resulted in scientists struggling for decades to determine the origin of cosmic rays, most specifically at the highest energies, since these high energy events do not occur often. As a result, there are insufficient statistics to properly correlate the highest energy cosmic rays to a known source in the Universe. There have been multiple recent studies [15], [16], [17] of data from the Pierre Auger Observatory looking for a directional correlation of galactic and extra-galactic sources of cosmic rays. These will be discussed in sections 2.5.1 and 2.5.2.

Figure 2.6 illustrates the difficulty in determining the origins of charged cosmic ray particles. As can be seen from Figure 2.6 [18], cosmic rays undergo many deflections before arriving at Earth. In comparison, due to the uncharged nature of gamma rays and neutrinos, they are able to travel in straight lines, undeflected. This is why these particles are effective for probing the high energy environments such as the supermassive black holes at the centres of AGN. Multi-Messenger astronomy is often used for point correlation studies since typically high energy cosmic rays can produce gamma rays and neutrinos through collisions and decays [15], [16]. Models of cosmic ray particles that propose galactic origin for the highest energies (above EeV) are shown by recent observations to not fit the data well [17]. This gives evidence of those highest

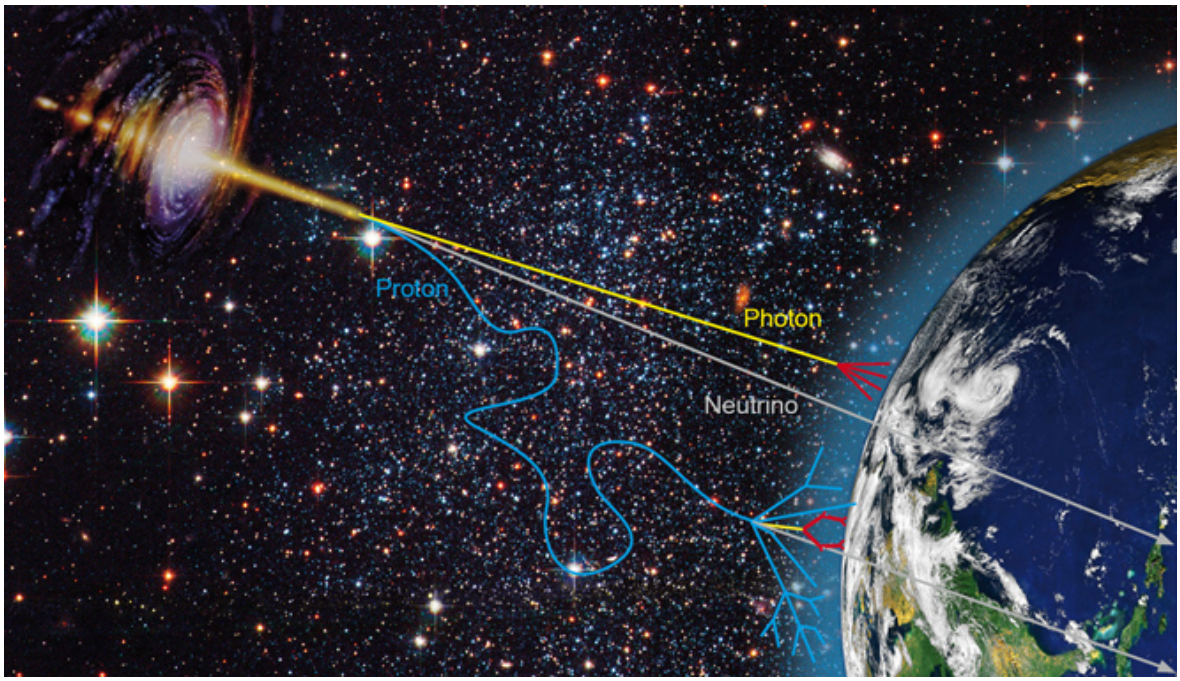

Figure 2.6: An artists impression of particles streaming from the centre of an Active Galactic Nucleus. Source: [18]

energy cosmic rays coming from outside of our own galaxy. Sections 2.5.1 and 2.5.2 will go into possible sources for cosmic rays produced both inside and outside the galaxy.

2.5.1 Galactic Cosmic Rays

It is quite well accepted that lower energy particles (below the knee of the spectrum - see Figure 2.3) come from sources within our own Milky Way galaxy. The most common progenitors of such particles are thought to be supernovae (possibly up to energies of around 10^{17}eV) [13]. Other sources have been discussed, such as evidence of a source in the Galactic Centre capable of accelerating particles up to 10^{15}eV [19].

The primary mechanism of acceleration for these particles up to PeV (10^{15}eV) energies is known as Diffusive Shock Acceleration (DSA) [13]. In 1949, Fermi first proposed that cosmic rays could be accelerated via diffusion between collisions of cosmic rays with interstellar clouds [20]. If these interstellar clouds have random directions of motion, then the frequency of "head-on" collisions between the cosmic rays and the clouds would exceed the rate of "tail" encounters, leading to a net acceleration. This has become known as Fermi second order acceleration [20]. Fermi's idea was founded on the fact that, when particles are confined to a specific region in space, such a diffusive process naturally produces power-law cosmic ray distributions, thereby modelling the observations well (see section 2.3 [20]. It was subsequently realized that shocks in plasma could also provide such diffusive acceleration in a more efficient manner. This works by transferring the shock's kinetic energy to particles both upstream and downstream of the shock, which accelerates the particles each time they cross over the shock [20]. Thus, the notion of diffusive shock acceleration, or first-order Fermi acceleration, was born [20]. This back and forth acceleration is repeated many times until the particles have enough energy to escape the environment and begin their journey towards Earth.

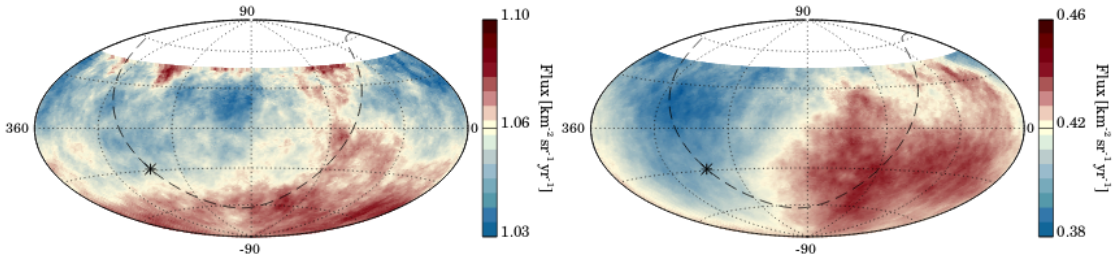

Figure 2.7: A map in equatorial co-ordinates of the cosmic ray flux for energies of 4-8 EeV (left) and $\geq$ 8EeV (right). The galactic plane is the dashed line, with the black star indicating the position of the galactic centre. Source: [17].

2.5.2 Extra-galactic UHE Cosmic Rays

Some cosmic rays (depending on their charge) above the ankle in the spectrum (see Figure 2.3) cannot come from within our own galaxy, since their Larmor radius within the galactic magnetic field is larger than the size of the Galaxy itself ($\sim$ 16 kpc) [13]. These cosmic rays, if produced in the galaxy at that energy, would have small particle deflections, hence their origin/source could be traced (which has not been observed). However, studies and observations have now shown the distribution of ultra-high energy cosmic rays are anisotropic in nature, indicating their extra-galactic origin (see section 2.5.3) [13]. One possible source of these UHECRs are starburst galaxies - dense regions with abnormally high star formation rates (much higher than a typical galaxy). The directions of starburst galaxies have been found to have a 4.0σ statistical significance above 39 EeV (relative to an isotropic background) as potential point-like sources of cosmic rays [16]. This significance was not as high either at lower energies or considering AGN as sources [16].

2.5.3 Anisotropy

We wish to understand what separates the galactic cosmic rays from the extra-galactic cosmic rays, so that we have a grasp on the directions in the sky where they arrive at Earth from. If we are able to know the sources of UHECRs, we can trace their arrival directions based on known positions of sources (from sky surveys and catalogues) such as AGN. A recent data analysis study by Auger [17] indicates that the highest energy cosmic rays (above 8×10^{18}eV) are arriving from outside of our own galaxy [17]. Figure 2.7 shows the results from that study.

An excess in the cosmic ray flux appears in a position away from the galactic centre above 8 EeV. If the direction is considered to have a dipolar distribution, the dipole has a direction of 233° galactic longitude and -13° galactic latitude, well away from the galactic centre at 0° galactic latitude and longitude. This means that cosmic ray particles above this energy are most likely extra-galactic in origin, with a 5.2σ significance. This means that there is a less than 1 in a million chance that this dipole is found by random chance [17].

It was realised early on in the study of cosmic rays that protons with energies higher than 10^{19}eV should come from extra-galactic sources since their gyroradius in the galaxy's magnetic field is of the same order as the size of the galaxy (radius of $\sim$ 16.2 kpc). What this means is that protons at such high energy are unable to stay confined to the galaxy and there's no acceleration mechanism possible to achieve this

energy [7]. The radius of gyration of a particle with charge Z can be written as in equation 2.17 for a typical galactic magnetic field strength of 1 μG [7].

$$R_g = 100kpc \times (E/10^{20}eV) \times (1\mu G/B)/Z \tag{2.17}$$

Plugging in the values for a proton at 10^{19}eV with charge = 1 unit, gives R_g = 10 kpc, close to the radius of the galaxy. As a result, any cosmic ray particles detected at Earth with an energy greater than 10^{19}eV are likely coming from sources outside of the Milky Way Galaxy. This allows scientists to hopefully find where these UHECRs are coming from.

Chapter 3

The Pierre Auger Observatory

The Pierre Auger Observatory, located in Malargue, Argentina, is the world's largest cosmic ray observatory, covering an area of 3000 km^2 at an altitude $\sim$ 1400 m above sea level. The purpose of the observatory is to study the origin and composition of cosmic rays coming from both galactic and extra-galactic sources, as these are two of the biggest mysteries in astrophysics. Such a large detecting area is required in order to study the highest energy cosmic rays, with energies $> 10^{19}$ eV. This is due to the very low flux of particles at the highest energies. The array itself consists of 1660 surface detector stations (water-Cherenkov detectors - see section 3.2), separated by 1.5 km, and 4 Fluorescence Detector stations (see section 3.1), each with 6 telescopes with 30° azimuthal field of views. An extra 3 telescopes (see section 3.1.3) were also installed in 2010 at one of the sites, giving a grand total of 27 telescopes. A map of the array is shown in Figure 3.1.

3.1 Fluorescence Detectors (FDs)

Four Fluorescence Detectors are situated around the array (Los Leones, Los Morados, Loma Amarilla and Coihueco), with 27 telescopes in total. Each site has 6 telescopes, however there was an addition made at Coihueco. Three HEAT (High Elevation Auger Telescopes) were installed to increase the field of view of the array (in elevation) by giving an extra 30° of coverage (see section 3.1.3). All telescopes have a $30° \times 30°$ field of view, such that each site covers 30° in zenith and 180° in azimuth. The telescope mirrors are segmented, spherical mirrors with a total diameter of 3.4m [21]. Each camera is a grid (22×20) of 440 hexagonal photomultiplier tubes (PMTs) which detect the U.V. light entering the aperture. The telescope system is based on Schmidt optics which reduces the aberration present in large spherical optical systems (e.g. spherical aberration which is present for the mirrors used at Auger). Fluorescence light from Nitrogen (N_2) molecules (see section 3.1.1) is emitted isotropically from an air shower, which passes through the circular aperture (1.1 m diameter) of the telescope. A glass filter is placed in front to filter the UV light in the $310 - 390$nm range [21]. The main benefit of the filter is that much of the background light is removed, improving the Signal-to-Noise ratio (SNR) of the measured shower. The filter is also able to protect the aperture and keep the system climate controlled. Shutters over the aperture are closed during daylight hours as well as when high wind speeds or rain are detected. A diagram of one FD telescope with all of its components is shown in Figure 3.2. The FD data is reconstructed using a software program developed by the Pierre Auger

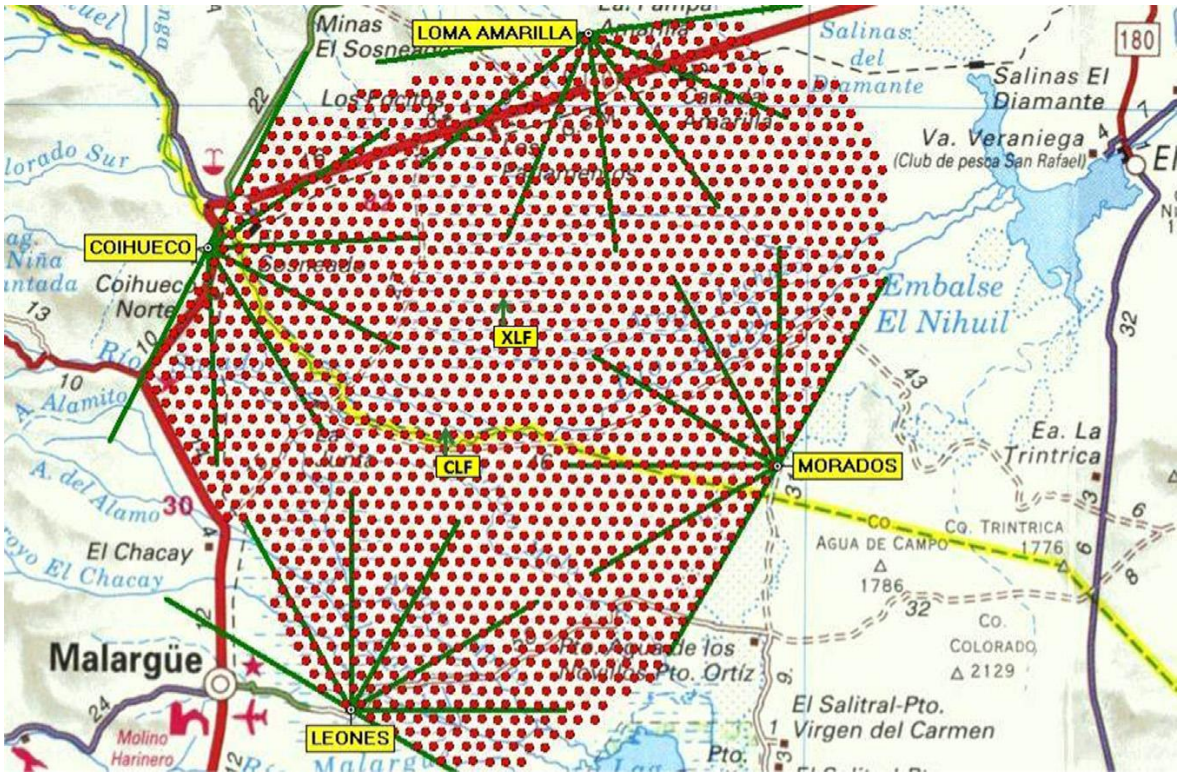

Figure 3.1: The Pierre Auger Observatory. Each red dot corresponds to a Surface Detector, whilst the green lines correspond to the field of view of each telescope. The Central Laser Facility (CLF) and Xtreme Laser Facility (XLF) are also labelled in the centre of the array [and are discussed in Section 3.5.3]. Source: [21]

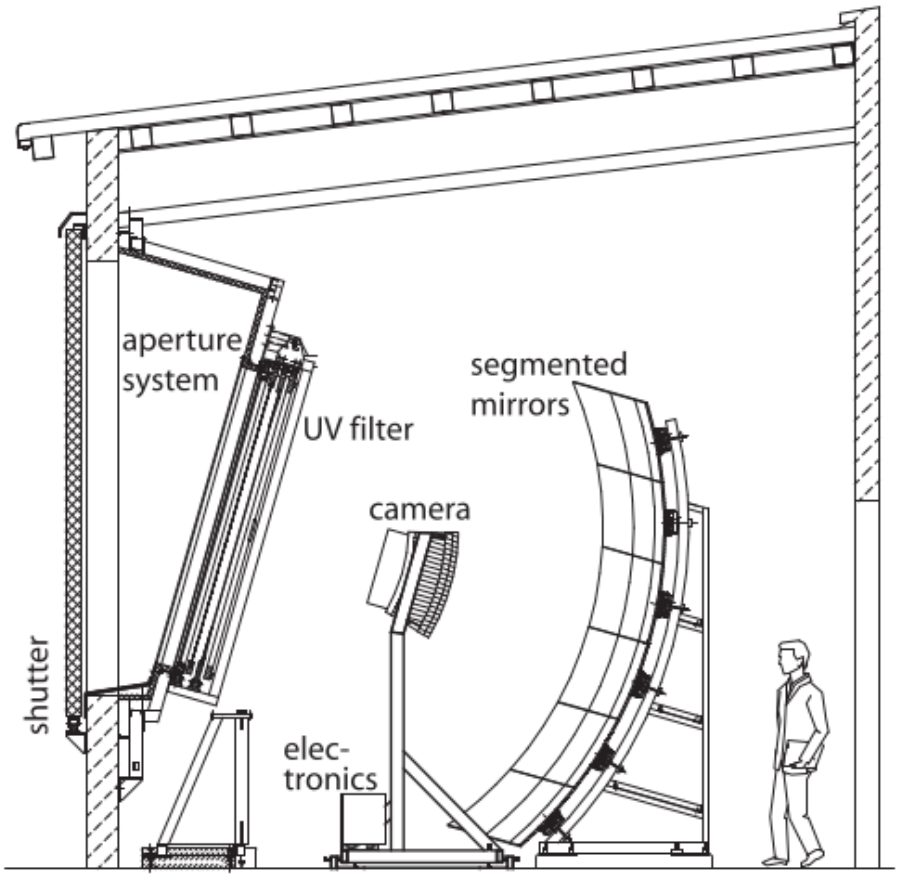

Figure 3.2: A diagram of a Fluorescence detector telescope with all of its components. [21]

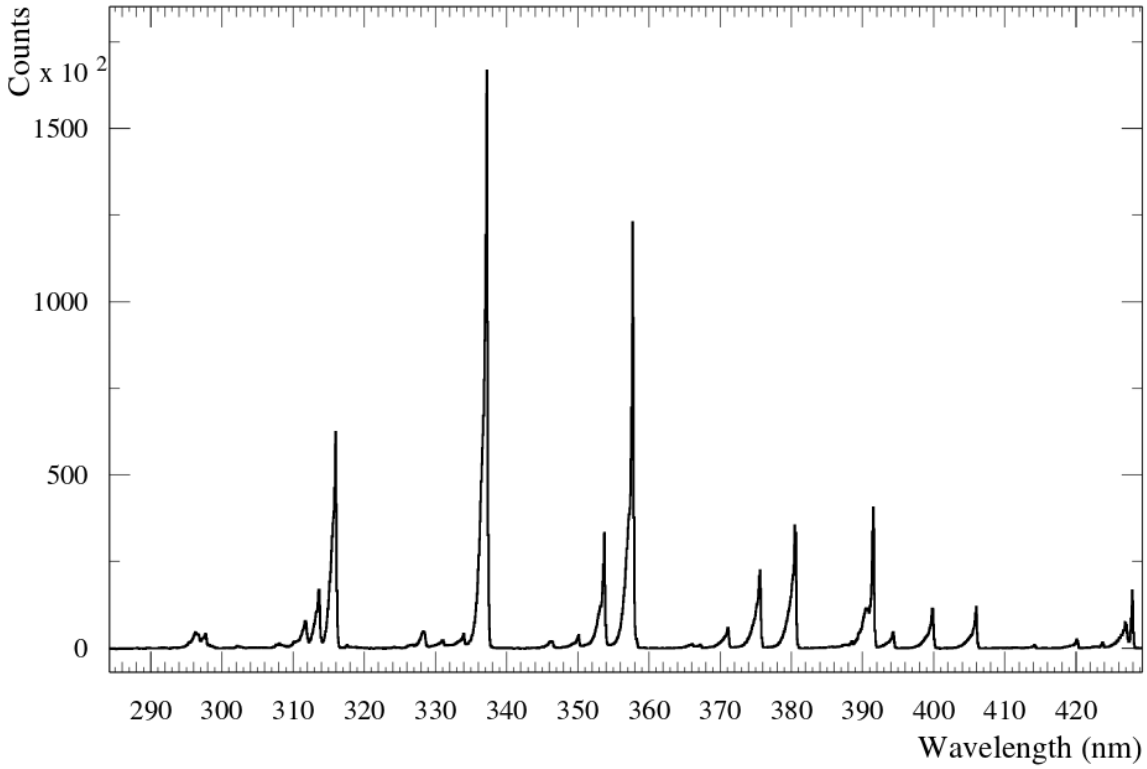

Figure 3.3: The nitrogen fluorescence spectrum in air at a given temperature (293 K) and pressure (800 hPa). Source: [22]

Observatory collaboration. Details of this are described in section 3.1.2.

3.1.1 Nitrogen Fluorescence

Due to the large number of electrons being produced in an air shower, there is an excitation of air molecules (particularly Nitrogen - N_2) caused by collisions between the electrons and molecules. This causes electrons within the molecules to jump up to a much higher energy level. After this excitation, they will de-excite - electrons return to their lowest energy level and release energy upon transitioning - which comes in the form of fluorescence light and is emitted isotropically. Transitions occur in several steps, with each step emitting a photon until the original energy level is reached. Each of these photons have wavelengths which range between 300 and 400 nanometres (thus they are UV photons). Figure 3.3 shows the nitrogen fluorescence spectrum in the UV band between 300 and 400 nm. The energy deposited in the atmosphere by charged particles is proportional to the number of fluorescence photons that are emitted by this process. It is this fluorescence that is detected - a means of indirectly measuring the presence of cosmic rays in the atmosphere [23].

3.1.2 Reconstruction of EAS and the effect of Clouds

Extensive Air Showers detected at the Pierre Auger Observatory are reconstructed via a large software program known as Auger $\overline{\text{Off}}\underline{\text{line}}$. Events present in the Fluorescence Detectors look like what is given by Figure 3.4.

The reconstruction is done by fitting a profile of the energy deposit per unit depth

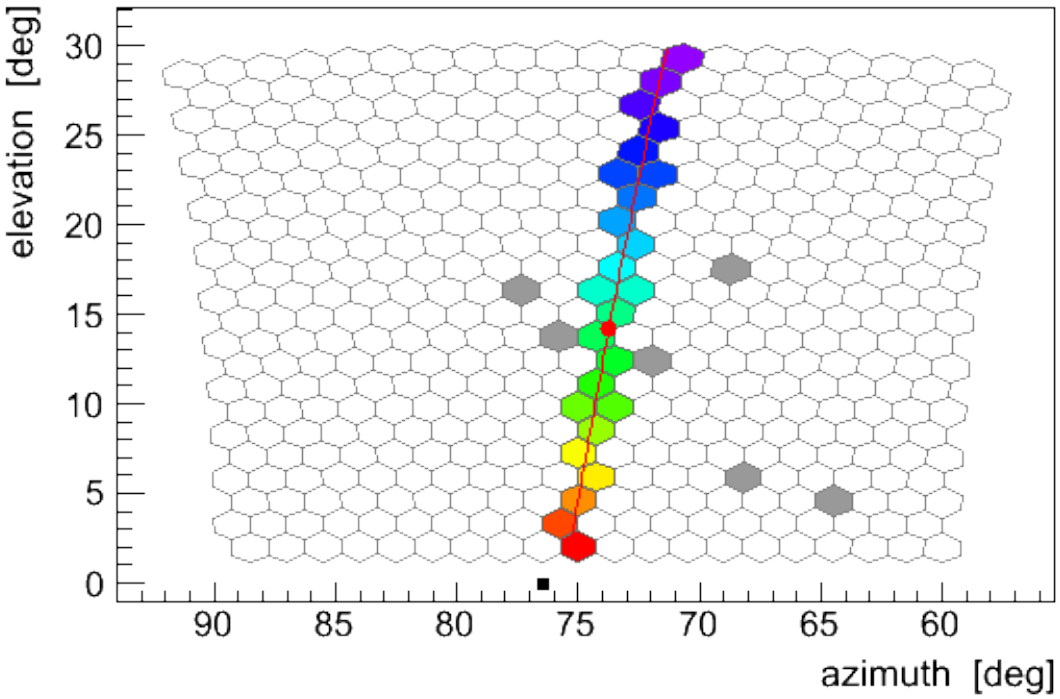

Figure 3.4: An example shower observed in the Field of View of an FD. The FD has collected fluorescence light emitted by nitrogen molecules in the atmosphere. The different colours represent timing information which is used for geometry reconstruction.

in the atmosphere, shown in Figure 3.5, known as a longitudinal profile. Energy deposit is extracted from the signal in each Photomultiplier tube, as the fluorescence yield is proportional to the amount of energy deposited into the atmosphere. The number of photoelectrons produced is proportional to the detected fluorescence light in the atmosphere; hence, the energy of the shower can be found from the number of photoelectrons emitted from the PMT. The fit to the energy deposit profile in Figure 3.5 (the red curve) is known as a Gaisser-Hillas (GH) fit [21]. The integration of this gives us the energy of the original cosmic ray particle that hit the top of the atmosphere. The different colours of the PMTs in Figure 3.4 give timing information of the air shower, meaning that we are able to retrace the footprints of the shower to determine where in the atmosphere it originated.

The reconstruction is based on fluorescence detector data with additional timing information provided by the surface detectors (see section 3.2) [21]. Each of the steps involved in this reconstruction will now be described.

To begin with, the ADC trace of each triggered pixel has a baseline and background noise subtracted from it. The background noise is estimated from the variance of the pixels that are free of air shower signal. Calibration constants from the FD drum and relative calibrations are then used to convert the ADC signal in each pixel into photon numbers at the telescope aperture [21]. Each FD pixel triggered by the air shower (by photons entering the camera) is subsequently searched for a shower signal by searching for pulse start and stop times that maximise the signal-to-noise ratio (SNR). Only those pixels with a high SNR are considered for the geometrical reconstruction [21]. The pulse time of the ith pixel is given by the signal weighted time (also known as "centroid time") of all bins in the pixel trace that belong to the pulse, given by:

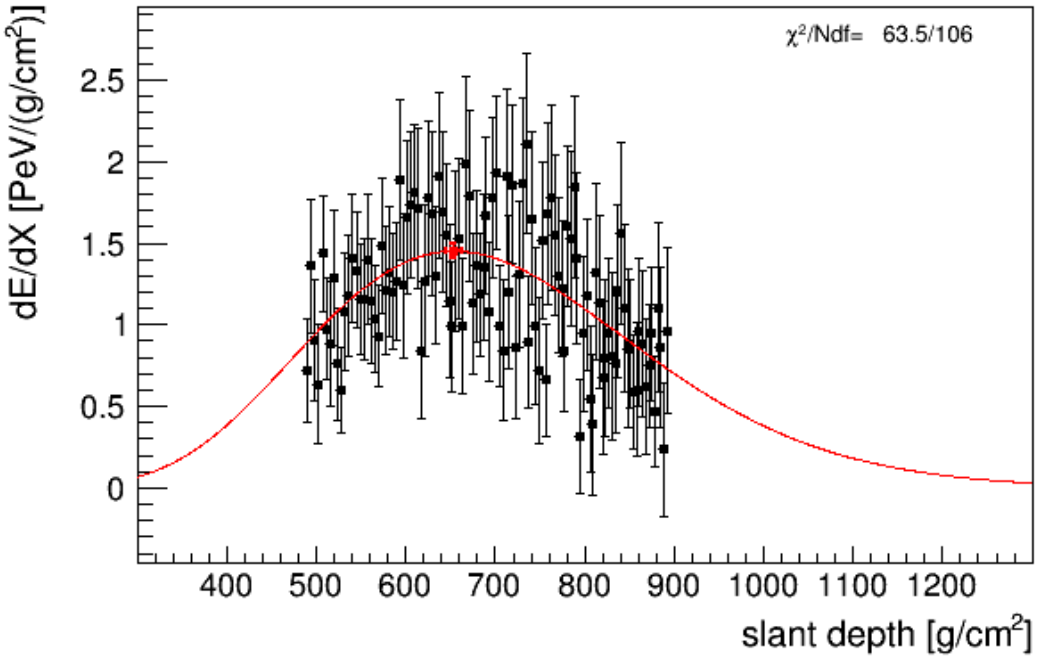

Figure 3.5: An example of an air shower energy deposit profile with a Gaisser-Hillas fit. The longitudinal profile is integrated to find the cosmic ray energy. This particular event was reconstructed with an energy of $(7.89 \pm 0.44) \times 10^{17}$eV, with an X_{max} of 654 ± 9 g/cm^2 at a zenith angle of $20.0 \pm 0.5°$. [24]

$$t_i = \frac{\sum \tau_k^i s_k^i}{q_i} \tag{3.1}$$

where the sum runs over the time bins that maximise the SNR. τ_k and s_k are the time and charge of the kth ADC bin respectively (each bin is 100 ns), and q_i is the pixel's integrated signal, given by:

$$q_i = \sum s_k^i \tag{3.2}$$

The next reconstruction step involves defining the Shower-Detector Plane (SDP) [see Figure 3.6] [21]. This is defined as the plane surface that contains the location of the triggered FD telescope and the line of the shower axis [25]. It is reconstructed from the data given by the telescope by minimising the S function, defined by equation 3.3 [21].

$$S = \frac{1}{\sum_i q_i} \sum_i q_i \left(\frac{\frac{\pi}{2} - arccos(\overrightarrow{p_i} \cdot \overrightarrow{n}_{\perp}^{SDP})}{\sigma_{SDP}} \right)^2 \tag{3.3}$$

This function is minimised over all pulses i with two free parameters θ_{SDP} and ϕ_{SDP} defining the vector $\overrightarrow{n}_{\perp}^{SDP}$ normal to the plane in spherical polar co-ordinates, and the pixel pointing direction $\overrightarrow{p_i}$. The pointing uncertainty for the SDP fit (σ_{SDP}) was determined to be 0.35° by studying SDP fits of the CLF (see section 3.5.3) laser shots that have a known geometry. This normalisation of the S function to σ_{SDP} allows one to interpret S as a χ^2 function where the SDP parameter uncertainties can be found [21].

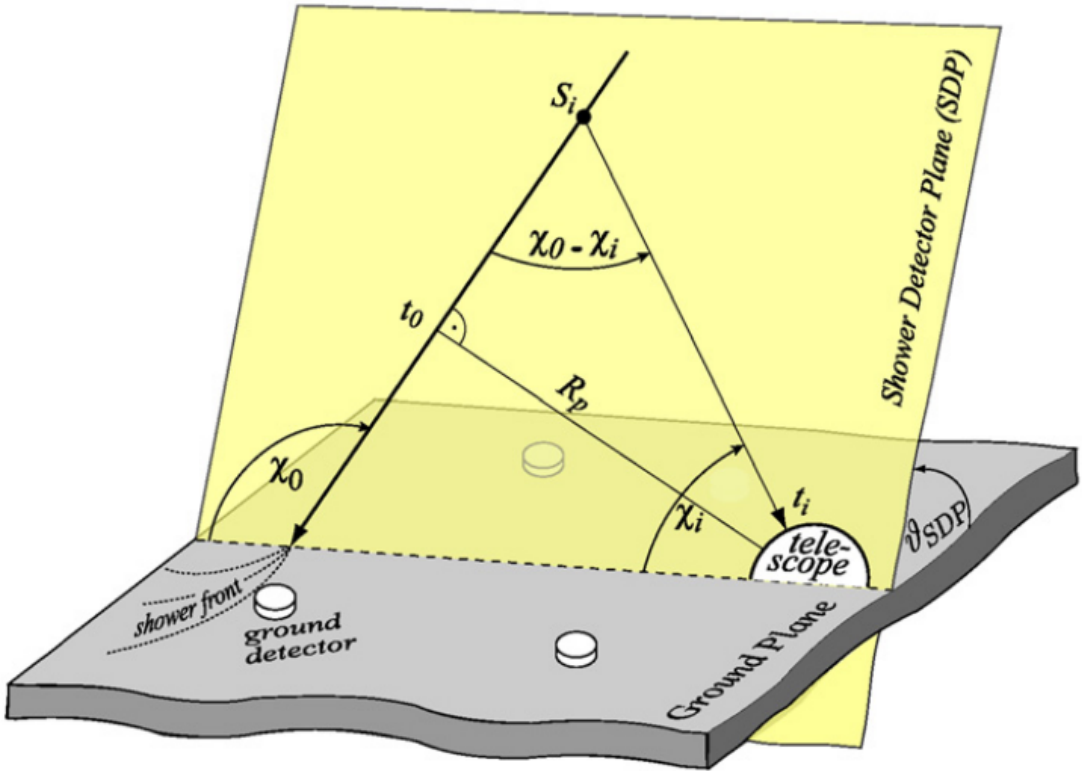

Figure 3.6: An illustration of the geometrical reconstruction from the Fluorescence Detector observables [25].

The next step in the reconstruction involves using a hybrid time fit, where each pixel that observes the shower track is associated with an angle χ_i along the SDP with respect to its horizontal axis at the telescope (see Figure 3.6). From this, each pixel that observes the track (the coloured pixels in Figure 3.4) points in a direction that makes an angle χ_i with this horizontal axis [25]. Light reaches the *ith* pixel at time t_i as given by equation 3.4

$$t_i = t_0 + \frac{R_p}{c} tan\left[\frac{\chi_0 - \chi_i}{2}\right] \tag{3.4}$$

where R_p is the perpendicular distance from the camera to the track, χ_0 is the angle that the track makes with the horizontal line in the SDP (shown in Figure 3.6), and t_0 is the time when the shower front on the axis passes the point of closest approach R_p to the camera.

To determine the three free parameters in this fit (t_0, χ_0 and R_p), the minimum of the function

$$\chi^2 = \sum_i \frac{(t_i - t(\chi_i))^2}{\sigma(t_i)^2} + \frac{(t_{SD} - t(\chi_{SD}))^2}{\sigma(t_{SD})^2} \tag{3.5}$$

has to be found. The sum in the χ^2 minimisation runs over all pixels i with a centroid time t_i (and its associated uncertainty $\sigma(t_i)$). An additional surface detector station time t_{SD} and its associated uncertainty $\sigma(t_{SD})$ is added as well. The shower front containing the surface detector station meets the shower axis at a point seen at an angle χ_{SD}, with $t(\chi_{SD})$ being the expected time when the shower centre would pass that point.

Once shower geometry is known, the total light at the aperture of the telescope as a function of time is obtained, by adding the signals s of the camera pixels j at each

time bin i. Only those pixels with a pointing direction within an angular distance ζ to the vector from the telescope to the shower centre at time i are included. ζ is chosen in order to maximise the signal-to-noise ratio. The light flux F_i arriving at the detector in time bin i is given by:

$$F_i = \frac{1}{A_{diam}} \sum_{j=1}^{N_{pix}} s_{ij} \tag{3.6}$$

where the sum here runs over all pixels N_{pix} within ζ at time bin i. Here, A_{diam} is the area of the diaphragm opening of the telescope.

The light collected at aperture as a function of time can then be converted into energy deposited by the shower as a function of slant depth (see Figure 3.5). The Cherenkov and fluorescence light produced by an air shower are connected to this energy deposit by a linear set of equations. As a result, the shower profile can be obtained by a linear least squares minimisation [21]. The full longitudinal energy deposit profile and its maximum $(\mathrm{dE/dX})_{max}$ at X_{max} is estimated by fitting a Gaisser-Hillas function, as described in equation 3.7.

$$f_{GH}(X) = \left(\frac{dE}{dX}\right)_{max} \left(\frac{X - X_0}{X_{max} - X_0}\right)^{(X_{max} - X_0/\lambda)} e^{(X_{max} - X_0/\lambda)} \tag{3.7}$$

The two parameters X_0 and λ are shape parameters that are constrained by the observed track length and the number of detected photons of the air shower event. The proportionality between the fluorescence intensity (i.e. the number of fluorescence photons produced in an air shower) and the energy deposit is given by the fluorescence yield. Knowledge of the absolute value of the fluroescence yield, as well as its dependence on wavelength, temperature, pressure and humidity in the atmosphere is essential in order to reconstruct the longitudinal profile. This yield is affected by aerosol attenuation in the atmosphere, which introduces a systematic uncertainty in the amount of light detected at the aperture (hence, the amount of energy deposited in the atmosphere).

The calorimetric energy of the shower is then obtained by integrating equation 3.7 before estimating the total energy by correcting for the "invisible energy" that is carried away by neutrinos and high energy muons produced in the shower [21].

Now that the reconstruction steps have been outlined, one can discuss quality cuts made to the data. There are three categories of quality cut made to the data; however, for the purposes of this dissertation, only the cloud cuts will be discussed. These cloud cuts are described in section 3.5.8.

Air shower events seen by the FD are given a description based on what instruments have seen the shower track. A description of each is given as follows [26]:

- *Mono*: At least one telescope of a single FD views the shower
- *Multi-Mirror*: Two or more adjacent telescopes in an FD view the shower
- *Stereo*: At least two telescopes in each of two different FDs saw the shower
- *Hybrid*: A mono event that is also viewed in coincidence with a signal from the Surface Detectors (SD)
- *Stereo-Hybrid*: At least two telescopes from two different FDs view the event in coincidence with an SD signal

For more information on the Surface Detectors and their role in the reconstruction see section 3.2. Hybrid events are also split into two subcategories: Brass Hybrids and Golden Hybrids. A golden hybrid event is an event that has been reconstructed independently by both the Fluorescence Detector and the Surface Detector, and is used in energy calibrations between the FD and SD. In contrast, a brass hybrid event is an event that has been triggered by the FD, however the signal in the SD is not large enough (i.e. only a single station triggered) to form its own event and as a result the SD is forced to trigger based on the geometry reconstructed by the FD.

An extension to the Observatory was done in 2010, adding in three extra telescopes next to the Coihueco site, known as HEAT.

3.1.3 High Elevation Auger Telescopes (HEAT)

HEAT (known as High Elevation Auger Telescopes) is another Fluorescence Detector comprised of 3 telescopes (rather than the usual 6) placed at the Coihueco site in 2010. It is used primarily for the reconstruction of low energy extensive air showers (together with the infill array, which is discussed in section 3.3) [21]. HEAT has two modes of operation, one being the low setting (where its FOV is 0-30° in elevation, as is the case with the other FDs) and the other being the high setting (telescopes are tilted in order to view the atmosphere from 30-60° elevation). The higher elevation view is useful for low energy reconstruction since these types of showers will develop closer to the detectors and higher up in the atmosphere, thus X_{max} is outside the standard FOV of the 4 FD stations. Combining these telescopes with the Infill Array allows us to study the lower energy showers (between 10^{17} and $10^{17.5}$ eV). Figure 3.7 is an image of one of the three HEAT telescopes, in the high setting (viewing elevation angles between 30 and 60 degrees). Figure 3.8 shows a schematic of an air shower that was tracked across the Field of View of both HEAT and Coihueco, especially important given that the shower would not have been reconstructed well without the HEAT telescopes. As outlined in section 3.1 and above, a quality cut rejects data when X_{max} is not in the field of view of the Fluorescence Detector. With HEAT added, lower energy showers will have X_{max} in the field of view of HEAT, rather than Coihueco. Figure 3.9 is an example of such an event detected by both HEAT and Coihueco in coincidence (known as a "HECo" event). As shown by Figure 3.9, the addition of HEAT has allowed us to extend the energy range of our air shower analysis as well as to increase the statistics of our existing analysis. An upgrade to the cloud cameras has been done to include the HEAT telescopes in the cloud database, which is outlined in Section 5.4.1.

Figure 3.7: One of the three HEAT telescopes. This image was taken when I attended the Pierre Auger Observatory for an FD shift in May/June 2017.

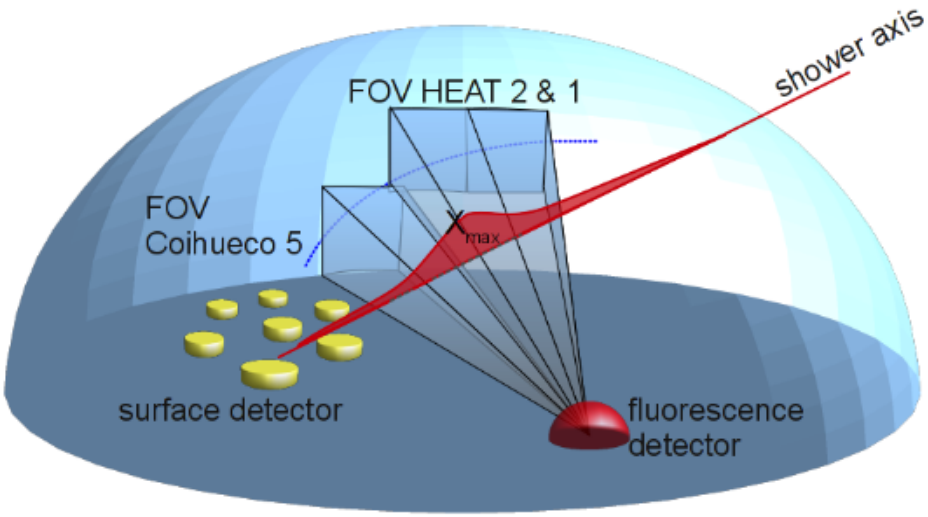

Figure 3.8: An example shower observed in the Field of View of both HEAT and Coihueco. Due to the geometry of such a shower, it would not be seen and reconstructed using the standard FD at Coihueco. The FOV of Coihueco telescope 5 and HEAT telescopes 2 and 1 are depicted [27].

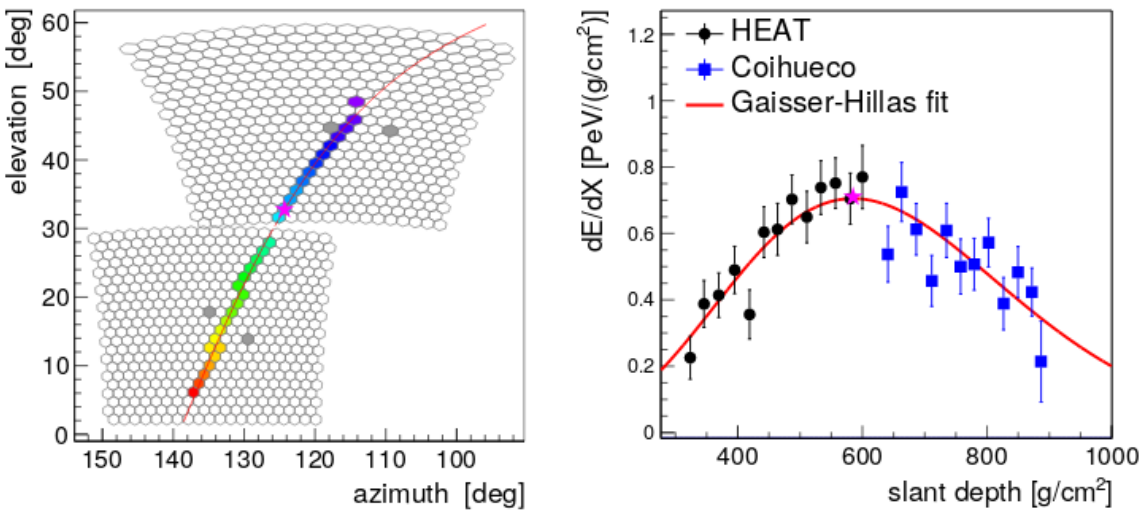

Figure 3.9: An example shower with a reconstructed energy of $(4.7 \pm 0.2) \times 10^{17}$ eV observed in the Fields of View of both HEAT and Coihueco. On the left is the shower track through the telescope in HEAT (top) and Coihueco (bottom). The right hand figure is the longitudinal profile (HE (black dots) and Coihueco (blue squares)) with the Gaisser-Hillas fit once again in red. The magenta star indicates the position of X_{max}, which as can be seen, lies outside of Coihueco's FOV [28].

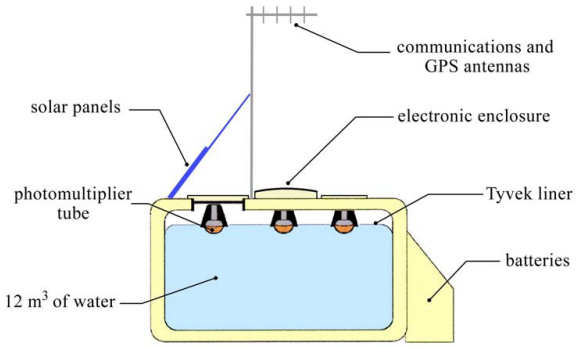

Figure 3.10: The interior of a surface detector tank. Source: [29]

3.2 Surface Detector (SD)

Each surface detector station is a 3.6 m diameter tank, filled to a depth of 1.2 m with highly purified water and lined with a reflective bag. Inside the tank, three photomultiplier tubes are positioned 1.2 m from the centre of the tank and pointed downwards in order to detect the induced Cherenkov light by particles travelling through the water.

Measurements of the amount of Cherenkov light emitted by particles within the tank are done using the VEM unit - "Vertical Equivalent Muon". This is the light a muon would emit in the tank given a direct vertical path through the water at the centre of the tank. Calibration of the PMTs is done using this unit, and signals from the tanks are expressed in VEM for the reconstruction process.

Reconstruction of events seen in the SD tanks is done by fitting a Lateral Distribution Function (or LDF) to the shower signal as a function of core distance. Figure 3.11 shows the SD signal fitted by an LDF for a shower seen by several SD stations. In

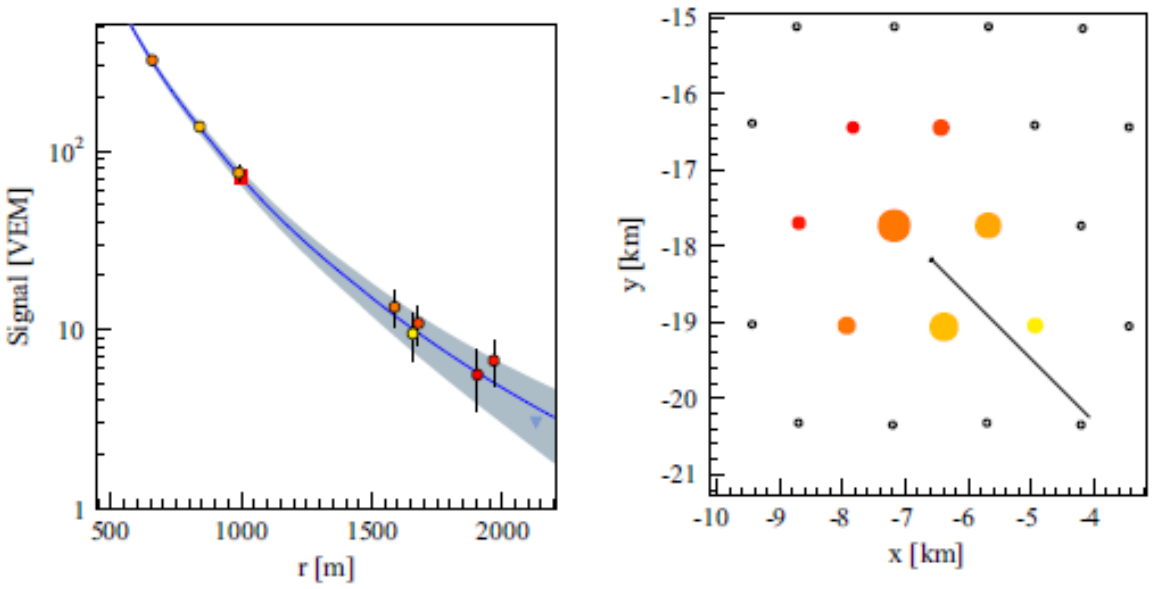

Figure 3.11: A Lateral Distribution Function fitted to a shower signal (in VEM) as a function of distance from the shower core (left). On the right is a schematic of where the shower hit the ground, showing the 8 SD stations that saw signal (red, yellow and orange dots), with the line illustrating the azimuthal direction of the SD reconstruction. The energy of this shower was 3×10^{19} eV.

Figure 3.11, the line pointing out to the bottom right is the azimuthal direction of the shower. The two lines of each separate reconstruction (FD and SD) intersect on the ground at the shower core, and the difference between the two lines in each case is an estimate of the uncertainty on the azimuthal angle of the shower and its core position on the ground.

Section 3.2.1 covers details on the recent upgrades of the Surface Detectors, known as AugerPrime.

3.2.1 AugerPrime

A recent upgrade to the surface detectors is rolling out, where a 3.8×1.3 m scintillator detector is placed on the roof of each surface detector - these are known as an SSD (Surface Scintillator Detector) [30]. Figure 3.12 shows an SSD on top of an existing water-Cherenkov tank. The goals and aims of the upgraded surface detector array are to answer some important questions, such as [30]:

- Understanding the origin of the flux suppression in the cosmic ray spectrum at the highest energies
- Evaluate the possibility that there is a proton fraction at the highest energies
- Find out new information about the hadronic interactions in this high energy range

The scintillator is able to sample particles from the shower with very different responses to both the electromagnetic and muonic components of an air shower than the water-Cherenkov detector [30]. There are also different responses of the SD to each of these

Figure 3.12: An SSD placed on top of a surface detector. Source: [30]

particle types (photons, electrons and muons) [30]. A small PMT inside the water-Cherenkov detector (WCD) has been added (replacing one of the orignal PMTs) to increase the dynamic range of the WCD. A new electronics board, known as a UUB (Upgraded Unified Board) was also installed as part of the upgrade [30]. The UUB is able to read the original 3 PMTs of the surface detector tank as well as the PMT for the SSD simultaneously, allowing for more information to be extracted about a single event.

3.3 AMIGA

The Auger Muon and Infill Ground Array (AMIGA) is an upgrade to the Observatory that directly measures the muon component of air showers at ground level. This detection enhancement is a combined installation of extra water-Cherenkov tanks and several scintillation detectors buried in the ground over a hexagonal area of 23.5 km^2. This array, located 6 km from the Coihueco FD site, has the Surface Detectors spaced 750 m apart (instead of the conventional 1500 m) [21]. The combination of this array with the HEAT telescopes (see section 3.1.3) enables the study of lower energy cosmic ray showers, since they develop higher up in the atmosphere and do not produce a strong enough signal to be detected by the regular 1500 m array.

Despite the small size of this infill, there is sufficient area to study a high number of events at energies above 3×10^{17} eV at zenith angles $\leq 55°$ (where it is fully efficient) given that the cosmic ray flux increases rapidly with decreasing energy. As such, the main purpose of AMIGA is to study the region between the second knee and the ankle of the cosmic ray spectrum (see Figure 2.3). The 750 m array was completed in September 2011 while the first prototype hexagon of buried scintillators, the "Unitary Cell", has been operational since March 2015. This array consists of 7 water-Cherenkov detectors paired with scintillators that cover 30 m^2 on the ground. In addition, two positions on the ground are equipped with twin scintillator detectors (that cover an extra 30 m^2 on the ground) to allow the accuracy of the muon counting technique to be assessed. To effectively shield out the electromagnetic component, the Muon

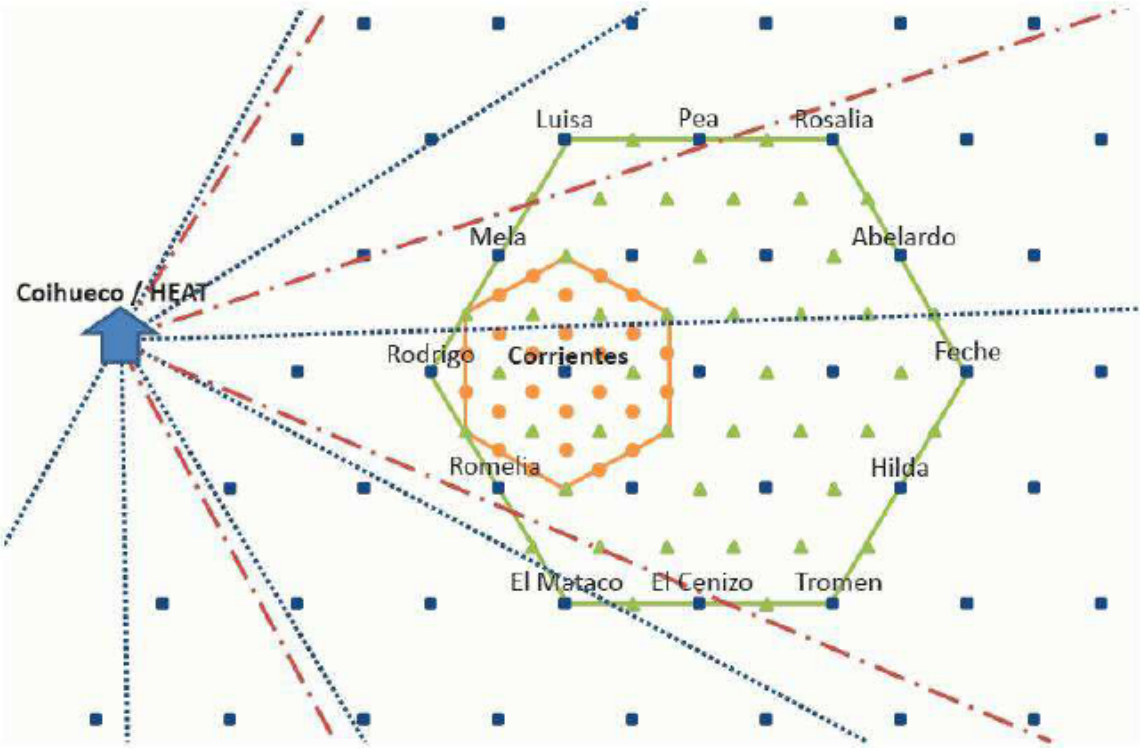

Figure 3.13: The schematic of the surface infill array. The blue dots correspond to the regular surface detector array, where each dot is separated by 1.5 km, while the green hexagon corresponds to the position of the infill array (where surface detectors are separated by 750 m). The blue arrow indicates the North direction. Source: [31]

Detector is placed under $\approx 540 \mathrm{g/cm^2}$ of vertical mass corresponding to a depth of 2.3 m in the ground (which corresponds to ≈ 20 radiation lengths). In addition to these muon detectors, there are extra scintillators buried at even shallower depths, which allows analysis of the shielding features. These shallower scintillators are at a depth of $\approx 310 \mathrm{g/cm^2}$ (or 1.3 m below ground). In total, $290\,\mathrm{m^2}$ of fully equipped plastic scintillators are operative in the Unitary Cell. The proven tools and methods used for the analysis of the regularly spaced 1500 m array have been extended to lower energy events in the 750 m array reconstruction. The energy reconstruction is based on the lateral density of particles at a distance 450 m from the shower core.

The objective of the AMIGA array is simple: to be able to probe air showers with an incident primary energy much lower than can be accurately reconstructed with the 1500 m array, as well as to better understand the role played by muons as they develop in showers and propagate towards the ground. The position and schematic of the 750 m array is shown in Figure 3.13. In combination with the HEAT and Coihueco Fluorescence Detectors (see section 3.1.3), we are able to probe the lower energy cosmic rays and better understand the various components of air showers, and also study the mass composition of incident cosmic ray particles.

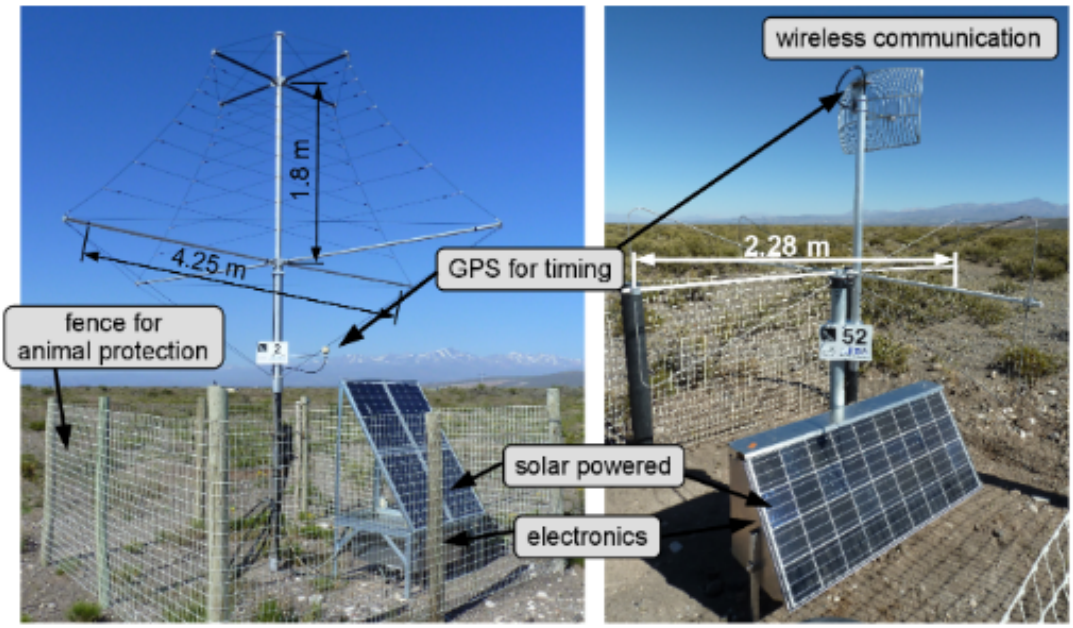

Figure 3.14: The designs of the AERA radio detectors. Each component is labelled accordingly. On the left is a Log Periodic Dipole Antenna (or LPDA) and on the right is a butterfly antenna.

3.4 AERA

The Auger Engineering Radio Array (AERA) is another recent extension of the Observatory, which consists of 153 radio detector stations that cover an area of $\sim 17\,\mathrm{km}^2$ ([32]). It is located in the same area as the AMIGA detectors (see section 3.3). Figure 3.14 shows the 2 different designs of the detectors. Each station runs off solar power and battery power storage allows for 24 hour usage. The first 24 LPDA stations were installed in 2011, spaced in a triangular grid that covers an area of $\sim 0.6\,\mathrm{km}^2$. In 2013, 100 butterfly antennas were installed, on both 250 m and 375 m grids, and a final 29 stations were installed in 2015 on a 750 m grid. Each station has a wireless communication system that operates at a frequency of 5 GHz. All AERA stations operate in the frequency band of 30-80 MHz with a digitised radio signal. The FDs and SDs can accurately measure the energy of the primary cosmic ray and can also be used as a cross-check for the energy calculated by AERA. Similarly, the position of the shower maximum (X_{max}) can be calculated and reconstructed using radio measurements [33]. Several methods are being developed to derive the mass composition of cosmic rays using measurements by AERA to reconstruct X_{max}. By pairing the AERA radio array with the buried scintillators, a measurement of the electromagnetic component (from radio) and muonic component (from underground) of the air shower can be done simultaneously, as a measure of the sensitivity of the mass composition of cosmic rays [33]. A diagram of the position of AERA within the array is shown by Figure 3.15. The goal of AERA is for cross-calibration of measurements taken by the FDs and SDs [32]. Specifically, it is for the calibration of our absolute energy scale measurements (i.e. the ratio of reconstructed SD energy to FD energy).

As part of the AugerPrime upgrade, radio antenna are planned to be installed

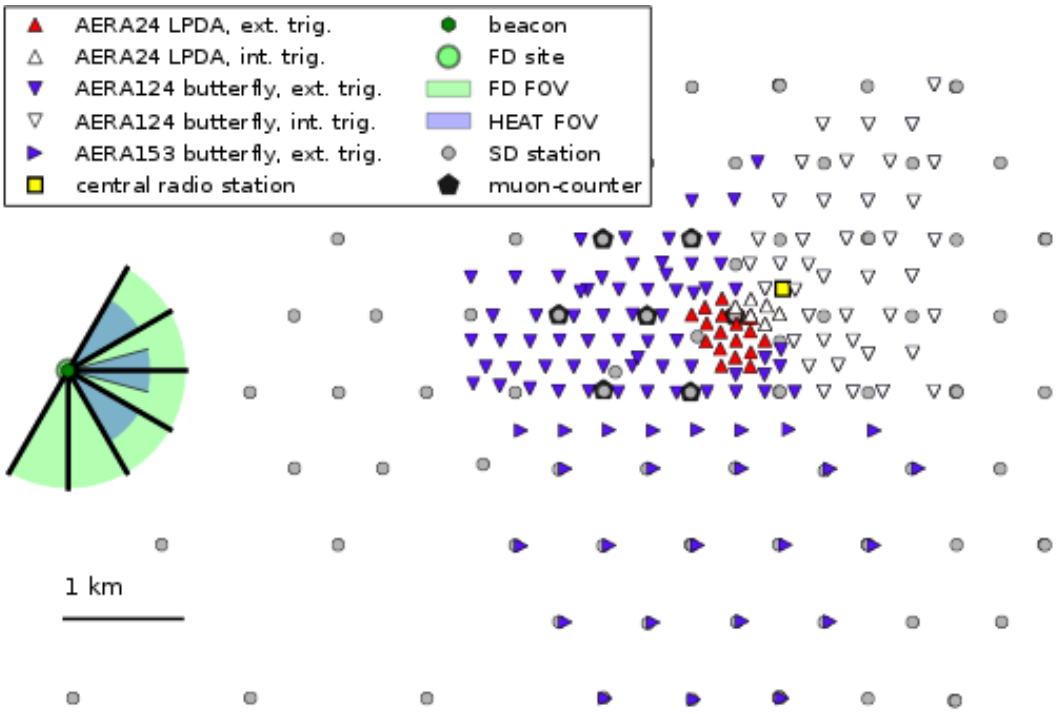

Figure 3.15: AERA's position within the SD array. Each antenna (coloured triangles) is surrounded by buried scintillators (black pentagons) and SD stations (gray circles). The Coihueco and HEAT telescopes are also shown as they are in close proximity to this array.

onto each of the 1660 surface detectors in order to improve the mass sensitivity of cosmic ray measurements to zenith angles $> 60°$ [34]. These antenna are sensitive in two polarisation directions and will have an operational bandwidth of between 30 and 80 MHz, which when fully installed will form the largest radio array in the world for the detection of cosmic ray air showers [34]. Since November 2018, an engineering array comprised of ten of these stations has been installed at Auger. This engineering array is also fully integrated into the data acquisition system at Auger, and records air shower measurements regularly [34].

The following sections go in depth on the instruments at Auger that monitor the state of the atmosphere, which is very important for the results presented in this dissertation.

3.5 Atmospheric Monitoring

The Pierre Auger Observatory treats the atmosphere as a giant calorimeter. Detailed knowledge of the troposphere above the array is required in order to accurately reconstruct the air showers observed in the FDs. Parameters such as temperature, pressure and relative humidity are needed in order to understand the longitudinal development of extensive air showers. They're also needed to better understand the propagation of isotropically emitted fluorescence light from particles in the shower - specifically, the molecular and aerosol scattering of the fluorescence and Cherenkov light.

Clouds will either attenuate or block light originating from an air shower, which produces a dip in the longitudinal profiles observed in the FD (the intensity of light coming from the shower has decreased). In terms of air shower reconstruction, this means that the reconstructed energy of cosmic rays will be underestimated. Conversely, when a shower passes through a cloud, increased scattering of the Cherenkov light can cause a bump in the observed profile. This serves to increase the reconstructed energy of cosmic rays. In a similar vein, aerosols present in the atmosphere can also scatter fluorescence and Cherenkov light emitted from an air shower, which creates uncertainty in the reconstructed air shower profiles.

This is why atmospheric monitoring of the state of the atmosphere is required - in order to better understand and allow for the adverse effects caused by cloud and aerosols in cosmic ray air shower analysis. Figure 3.16 shows all of the atmospheric monitoring equipment in use at the observatory, which will be discussed in the following sections. The main instruments of interest in this dissertation are the CLF and XLF (Central/eXtreme Laser Facility respectively), the LIDARs, the Weather Stations, and the Infrared Cloud Cameras. For details on each of these instruments, see sections 3.5.3, 3.5.6, 3.5.1 and 3.5.8 respectively.

3.5.1 Weather Stations

Weather Stations installed alongside the FD buildings, CLF, and XLF, are used to monitor changes in the local atmospheric environment, which is useful for a number of purposes. Every 5 minutes, each station is able to read out the local temperature, relative humidity, pressure and wind speed, as well as tracking other phenomena such as rainfall [21]. The last two are especially important since the FD cannot operate under high wind speeds (typically a speed exceeding 50 km/h) or wet conditions [21]. The data from each station is stored into a database that can be read out, to check if FD data acquisition was interrupted due to weather. Figure 3.17 shows a picture of one such Weather Station located at the Observatory. It is also useful to do some cross-calibration of the Weather Stations with themselves. Figures 3.18 and 3.19 show the differences in both temperature and relative humidity between two weather stations located at different Fluorescence Detector sites. From Figures 3.18 and 3.19, it is clear that there is good agreement between the two weather stations. However, it is good to check this consistency across the entire array. Figure 3.20 illustrates the temperature and relative humidity differences when comparing Los Leones to two other Fluorescence Detectors (Los Morados and Loma Amarilla). The average values of the relative humidity are 44% for Los Leones, 46.9% for Los Morados, 47.8% for Loma Amarilla, 41.6% for Coihueco and 50.6% for the CLF.

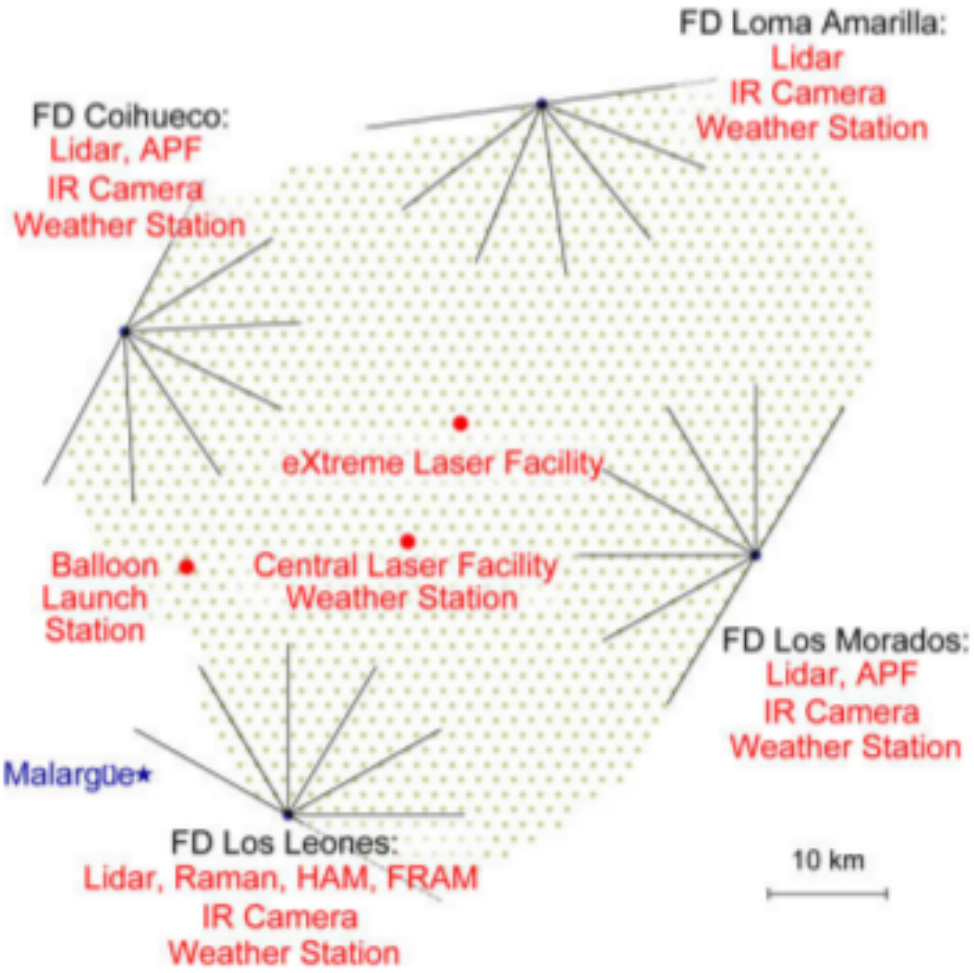

Figure 3.16: A summary of all of the atmospheric monitoring facilities that are available at the Auger observatory. A summary of each individual instrument is given in the subsequent sections. [35]

Figure 3.17: Weather Station installed at Loma Amarilla. This image was taken when I visited the Pierre Auger Observatory in May/June 2017.

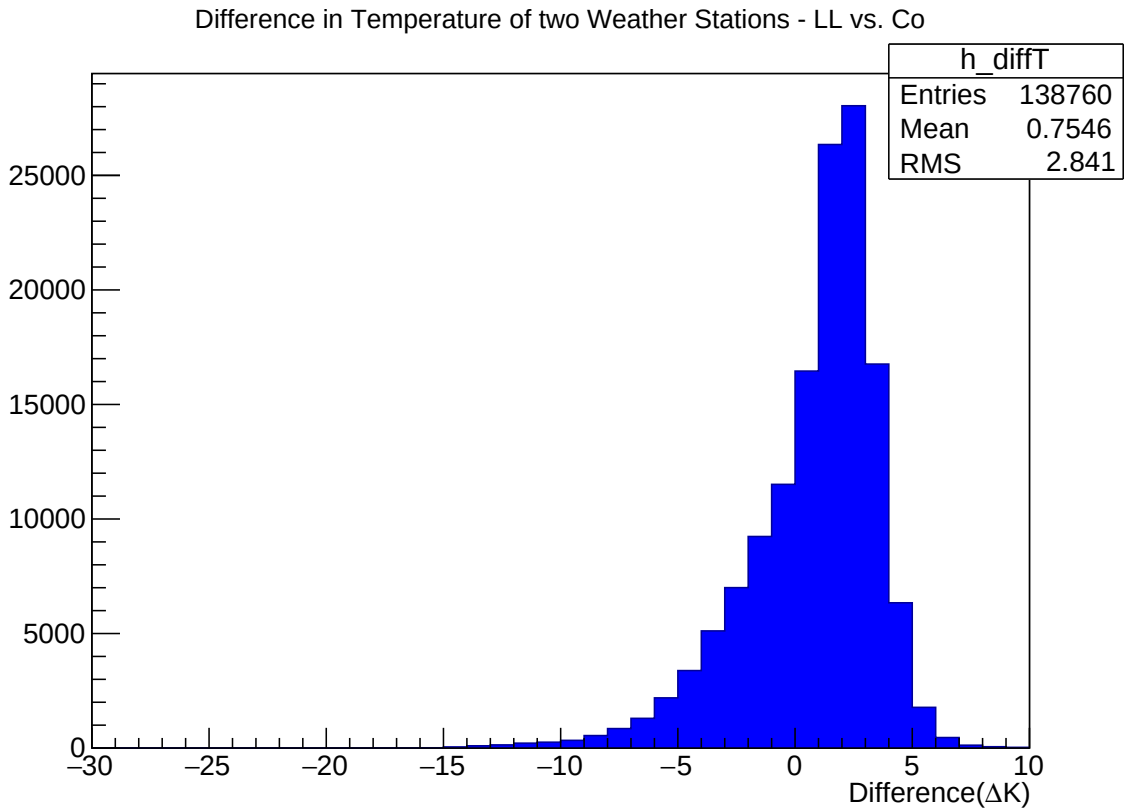

Figure 3.18: Difference in Ground Temperature between the Los Leones (LL) and Coihueco (Co) Weather Stations for 2017/2018 (Calculated as LL - Co). This shows good consistency across the array, with a slight offset due to the 300 m altitude difference.

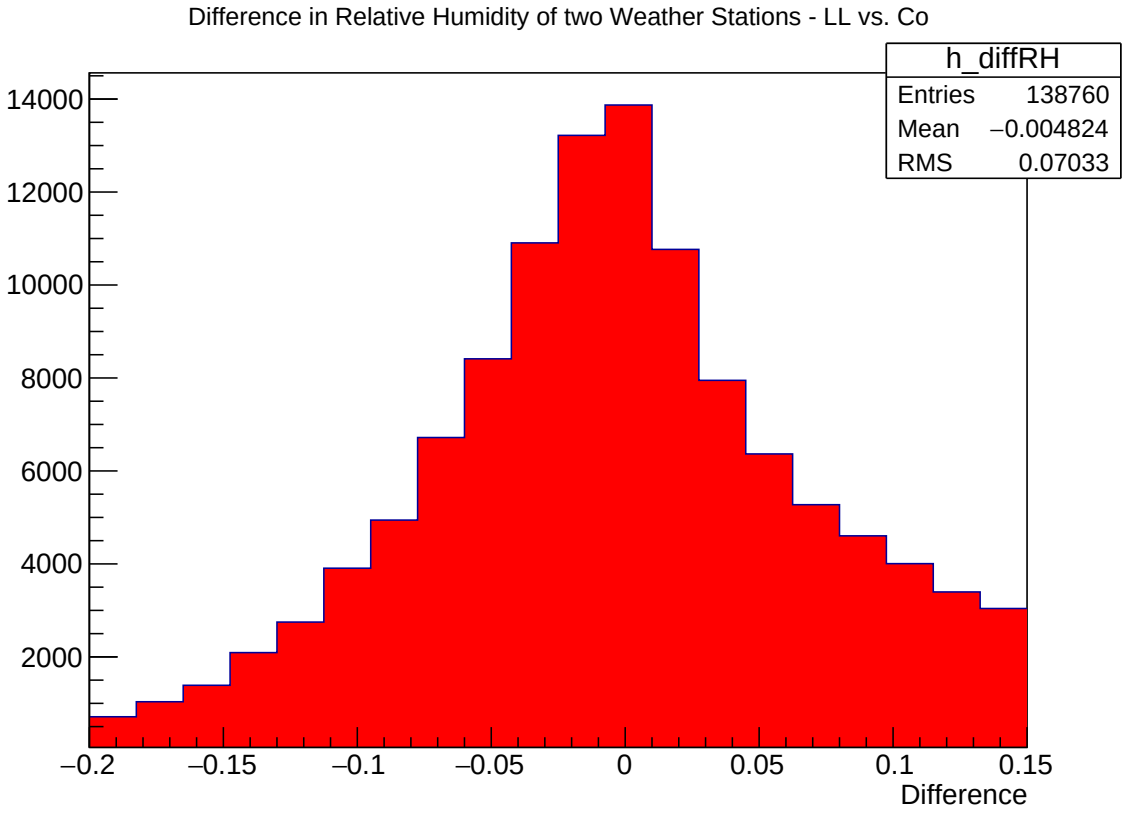

Figure 3.19: Difference in Relative Humidity at ground level between the Los Leones and Coihueco Weather Stations for 2017/2018. Here, relative humidity is represented as a fraction. Good consistency is once again achieved across the array.

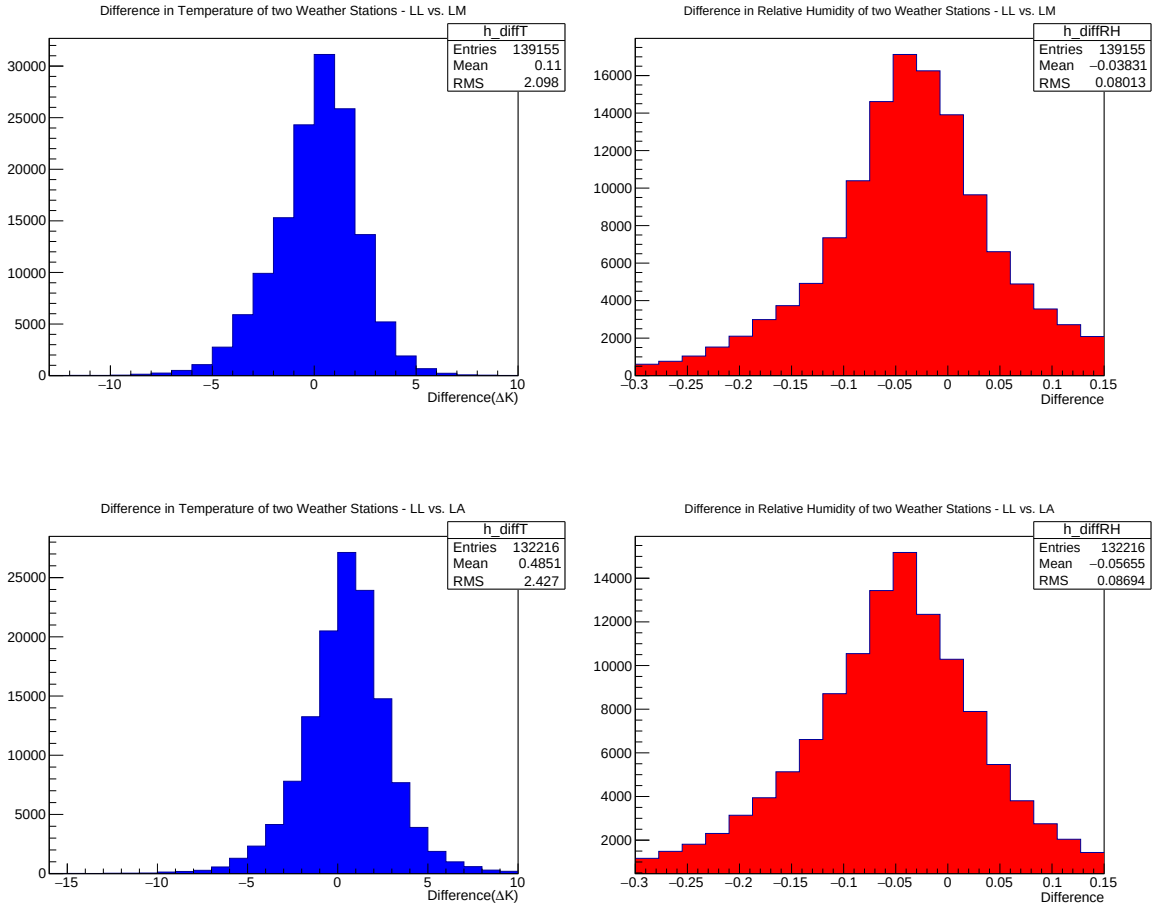

Figure 3.20: Difference between Temperatures (left) and relative humidities (right) between Los Leones and Los Morados (LM -top)/Loma Amarilla (LA - bottom). Relative humidities are represented as fractional differences. Temperature and Relative Humidity are both calculated as LL - LM/LA.

As can be seen by Figure 3.20, the weather stations at the Observatory agree quite well with each other. The uniformity of the temperature and humidity at Auger is encouraging for analyses done in later chapters. This consistency in temperature and relative humidity is important for the cloud height analysis done in Chapter 8.

3.5.2 GDAS

GDAS stands for "Global Data Assimilation System". Developed by the National Oceanographic and Atmospheric Administration (NOAA) in 2003, GDAS is a numerical model of atmospheric conditions calculated over the entire globe in a degree-by-degree latitude-longitude grid. Data from many instruments are fed into it and interpolated; radiosondes, satellites and weather station instruments are a few examples of such instruments. A radiosonde is a telemetry instrument carried into the atmosphere (usually by a weather balloon) that measures atmospheric parameters such as temperature, pressure and relative humidity. Each radiosonde carries a radio transmitter in order to transmit the values of these atmospheric parameters by radio to a ground receiver every few seconds as it rises in the atmosphere.

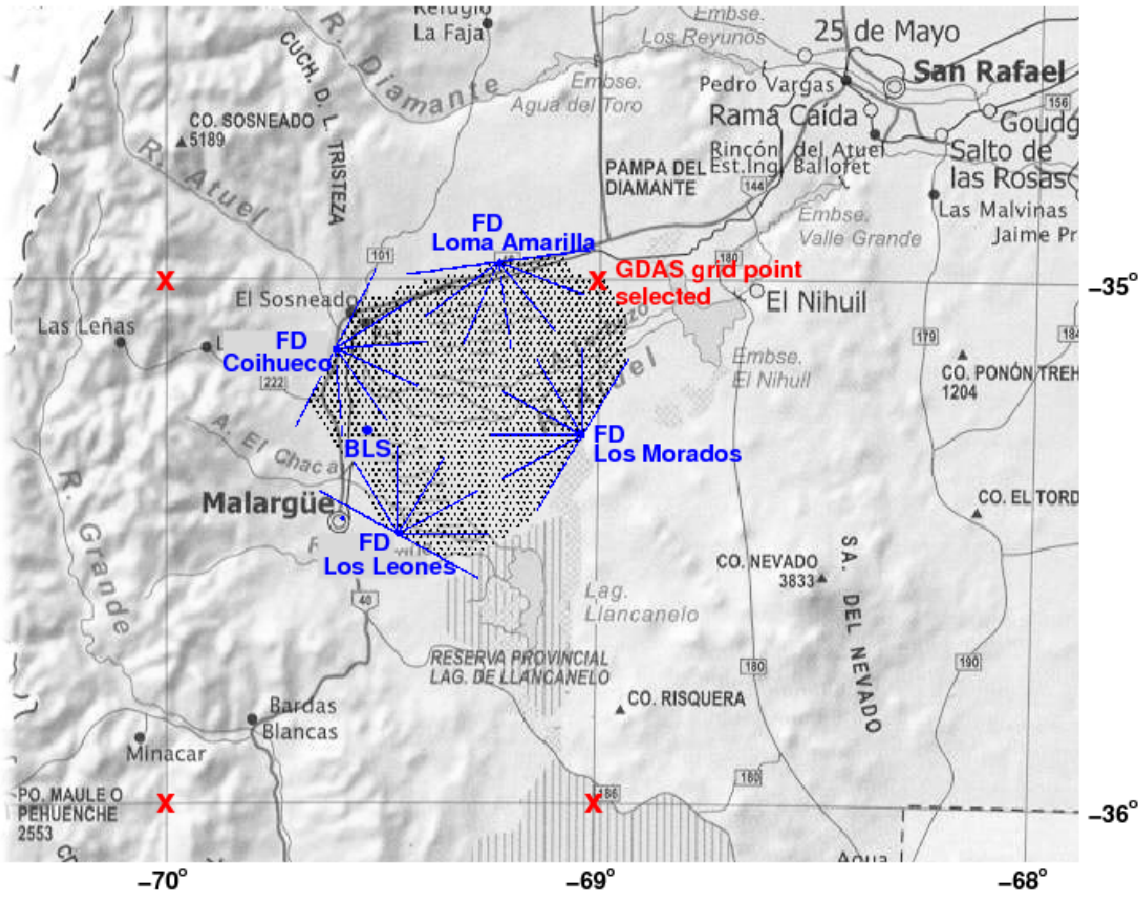

Figure 3.21: An overhead view of the array with the selected GDAS grid point. The grid point is close to the Loma Amarilla FD site.

The output of GDAS is produced every 3 hours starting from 00:00 UTC. As it happens, one such latitude/longitude grid point lies just inside of the observatory's surface array. This is shown in Figure 3.21 at 35°S, 69°W. The databases for GDAS are publicly available on the NOAA website [36]. The Pierre Auger Offline Framework also has the data for that specific grid point, which is what is used for all analysis subsequently.

Some cross-checks of the GDAS data need to be done in order to confirm consistency of the model over time. Firstly, we look at the height difference between two consecutive GDAS scans at the same temperature. This distribution is shown in Figure 3.22, where we have found the height in the atmosphere where the temperature is 285K and plotted the height difference. This is done to check the consistency and stability of the model over time for the purposes of cloud height estimation. Figure 3.22 shows that GDAS is stable over short time scales which is encouraging for the cloud height estimation technique outlined in Chapter 8. We can also check the difference between state variables (temperature, pressure, humidity) for the GDAS model at the selected grid point, with local radiosonde flights done at the Balloon Launching Station at Auger [37]. Figure 3.23 shows these differences for a number of key variables, comparing GDAS interpolations [black dots] to monthly mean profiles from radiosonde measurements taken over several years at Auger [red squares]. Figure 3.23 shows that the GDAS model is able to effectively estimate the conditions and state variables of

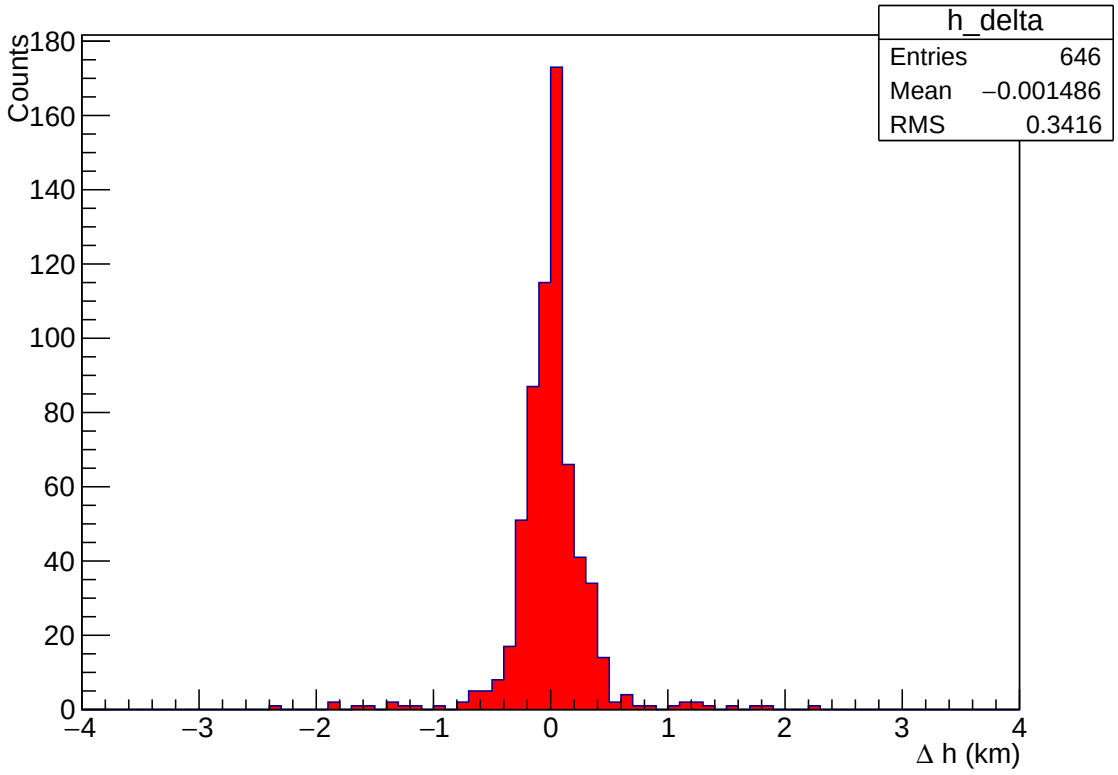

Figure 3.22: Distribution of the height difference between consecutive GDAS models for a temperature of $\sim$ 285 K. 285 K was chosen as it is an average temperature on a typical night of observation. What this means is we calculate the height in the atmosphere where GDAS interpolates a temperature of 285 K, and repeat this 3 hours later and calculate the difference in heights between these two times, shown here. The difference is consistent with 0, however there is a 342 m spread.

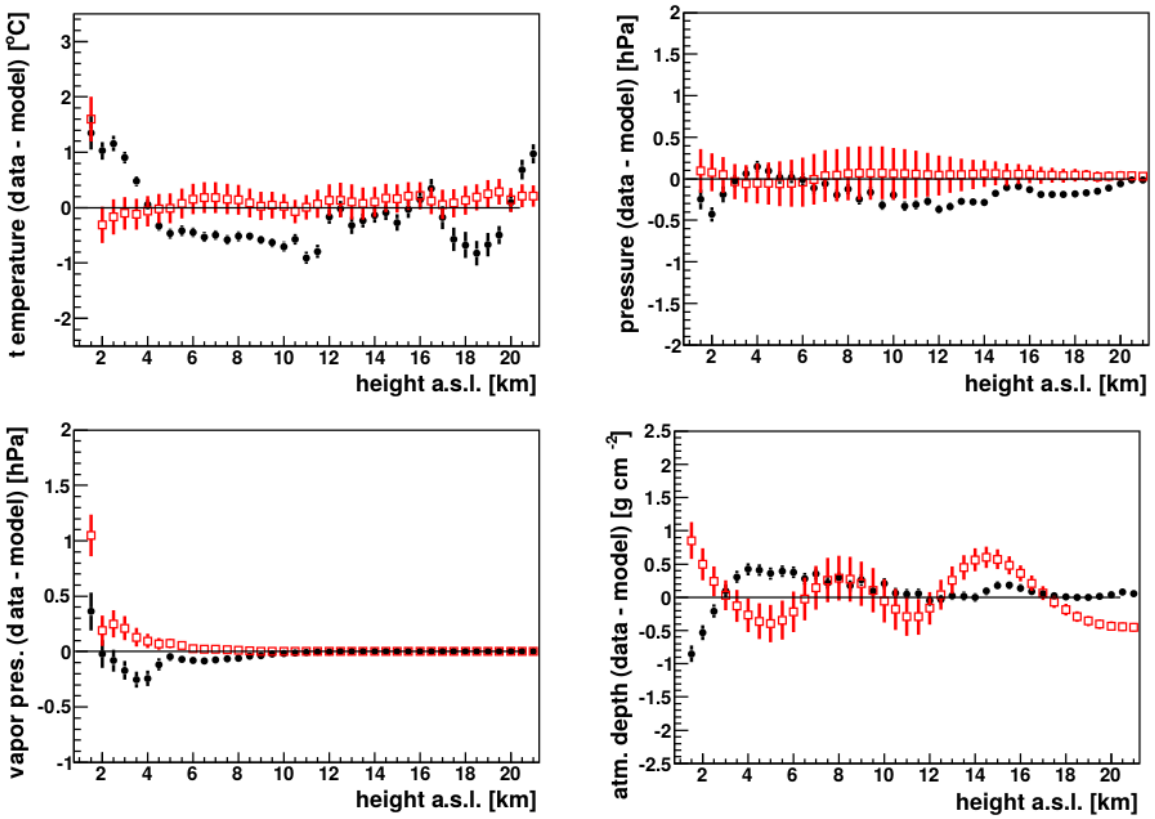

Figure 3.23: Parameter differences between the state variables as measured by local balloon flights and GDAS. The black dots correspond to the difference between GDAS interpolations and the radiosonde data for each parameter, whilst the red squares are the differences between the monthly mean profiles and the sounding data at Auger between 2005 and 2008. Top left is temperature, top right is pressure, bottom left is vapour pressure, and finally bottom right is atmospheric depth, all as a function of height above sea level [37].

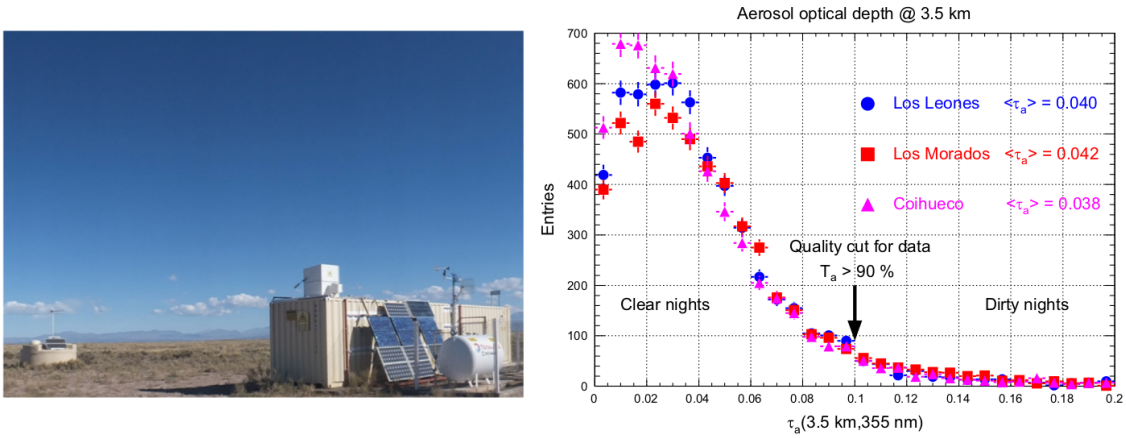

Figure 3.24: A picture of the CLF on the left, with measurements of the VAOD at the reference height (3.5 km above ground level). The legend indicates the average VAOD value for each of the three FD sites [21].

the local environment. GDAS is the main tool used to complement the Cloud Cameras and is used in much of the subsequent analysis, particularly in Chapters 6 and 8.

3.5.3 Central Laser Facility (CLF) and eXtreme Laser Facility (XLF)

The Central Laser Facility (CLF) and eXtreme Laser Facility (XLF) are two facilities placed near the centre of the SD array (see Figure 3.1), sitting at an altitude of 1401 m. These facilities are used to monitor the atmosphere for the presence of aerosols and clouds. The two sites contain a laser which fires vertically up into the atmosphere to track the properties of the atmosphere over the course of the night. CLF and XLF laser scans are performed by doing a series of vertical shots every hour to measure aerosol optical depth profiles. These are 7 mJ pulses fired at a wavelength of 355 nm [38]. Sets of 50 vertical shots are measured by the 4 FDs every 15 minute block within the hour (they appear as tracks in the telescopes at each site), and average values of the measured parameters (the Vertical Aerosol Optical Depth - known as VAOD, and the base height of clouds above the array - known as CBH) are fed into an atmospheric monitoring database used by the observatory. Vertical Aerosol Optical Depth is the optical depth of aerosols (particles such as dust, smoke, and haze) in the atmosphere, measured from ground level to all heights in the atmosphere [39]. The 50 shot averages are compared to clear sky averages (known as Data Normalised Analysis [39]) for the aerosol optical depth profiles. These CLF laser shots can be used to measure minimum cloud heights directly overhead, and are needed in the hybrid reconstruction. Figure 3.24 shows the CLF at Auger as well as some measurements of the aerosols up to a reference height. Figure 3.24 shows the consistent nature of the average VAOD values for 3 FD sites.

3.5.4 Aerosol Phase Function Monitor (APF)

Scattered Cherenkov and fluorescence light covers a large variety of angles, thus the scattering angular distribution (phase function) is of interest for studies of extensive air showers [40]. This function can be analytically estimated for the atmospheric molecular

component, however it is dependent on the size and shape of aerosols in the atmosphere for the aerosol component. The monitors are a collimated xenon lamp that fires light horizontally across the array, at a wavelength of 350-390 nm. The FDs measure the amount of light for scattering angles between 30° and 150°. A parametrization is required for the scattering behaviour of the shower reconstruction, which is given by;

$$P_a(\theta) = \frac{1-g^2}{4\pi}\left(\frac{1}{(1+g^2-2gcos(\theta))^{\frac{3}{2}}} + f\frac{3cos^2\theta - 1}{2(1+g^2)^{\frac{3}{2}}}\right) \tag{3.8}$$

This equation describes the total amount of aerosol scattering above the array. The quantity g describes the asymmetric behaviour of the aerosol scattering - it has been estimated as 0.56 ± 0.10 for the Auger analysis [21]. The parameter g is set to zero for clear nights (since only molecular scattering occurs). A very small aerosol content can be present during almost clear nights and as such small asymmetries occur in the phase function. To counter this, an uncertainty of 0.2 is added to the value of g. The first term of the parameterisation accounts for the forward scattering, whilst the second term accounts for the peak in scattered aerosol light at large θ. Lastly, the quantity f is related to the strength of the forward and backward scattered peaks of the aerosol scattering.

3.5.5 FRAM

The Photometric ("F"otometric) Robotic Atmospheric Monitor (FRAM) is an optical telescope with a 0.3 m diameter mirror that measures the integral light extinction in a given direction using stellar photometry [41]. The instrument measures the amount of light lost in the atmosphere due to absorption and scattering of light from molecules and aerosols in the atmosphere by comparing the measured brightness of a background star to its known brightness in the sky. The difference between measured and known brightnesses of a star gives us an indication of the amount of extinction caused by the constituents of the atmosphere (for example aerosols). The passive nature of the instrument means that it can be run in the field of view of the FDs during data acquisition. FRAM was installed at the Los Leones FD site (see section 3.1) in 2005 and has been doing passive optical measurements of the atmosphere ever since. The primary purpose of this instrument is the rapid observation of atmospheric conditions soon after a fluorescence detector records a shower with an anomalous longitudinal profile (much like the Shoot-the-Shower scan from the LIDAR - see section 3.5.6), so as to eliminate aerosols and clouds as the cause of this anomalous profile [41]. Figure 3.25 is a diagram of the FRAM telescope installed at Auger and the typical sequence of images that are captured by FRAM during data acquisition. With a field of view approximately 7° × 7° and a 30 second exposure, the FRAM telescope records images along the reconstructed shower axis (right hand figure in Figure 3.25), which takes several minutes to complete. For events where four or fewer images are captured, the FRAM data is discarded as it has not covered a large enough portion of the air shower in the FD FOV [41]. Once all of the images have been collected, background stars are extracted from the images and matched with sources in the Tycho2 star catalog [42]. Light fluxes obtained using star photometry from FRAM are compared to the values listed in Tycho2 to extract a light extinction for each source, before a global fit is performed based on the predicted dependence of the extinction on altitude [41].

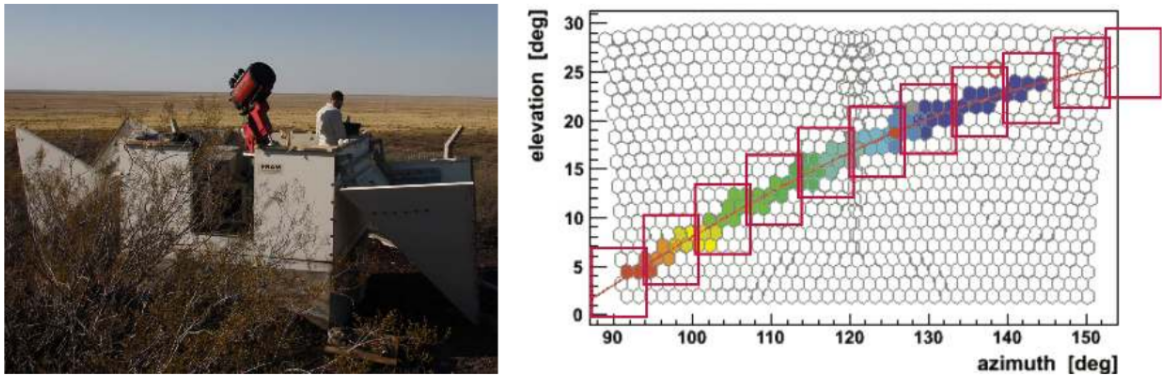

Figure 3.25: *Left*: The FRAM telescope installed at Auger. *Right*: The positions along the shower axis as viewed by FRAM (the FOV of FRAM is the red boxes) [41].

Figure 3.26: Building at the Observatory housing a LIDAR (Loma Amarilla). The dome opens up at night allowing vertical laser shots into the atmosphere above.

Any significant variations from the fit of this extinction can point to the presence of inhomogeneities in the longitudinal profile [41].

3.5.6 LIDAR

At the Pierre Auger Observatory, there are several LIDARs performing an azimuthal scan of the sky, as well as a 45° sky sweep (in zenith angle) every 15 minutes. The sky sweep is an inverted 45° cone in the sky over which the laser is scanning. This is done outside of the FOV of the FDs, to avoid the possibility of data contamination. From these scans, light profiles indicate cloud presence. Averaging is performed on the generated light profiles. Finally, various cloud parameters and read out and stored in a database - namely Cloud Base Height, Cloud Thickness/Vertical extent, and Cloud Coverage [for the 45° cone] (in %). Combined with the cloud cameras, better estimates of the position of the clouds in the atmosphere can be achieved. Figure 3.26 is an image of one such LIDAR building installed at Auger.

Now, a more detailed description of each scan type and the general operation of

Figure 3.27: Three parabolic mirrors of the Los Leones LIDAR. The box in the picture under the left hand mirror houses the laser. [44]

the LIDAR is required [43]. Each LIDAR is an Nd:YLF laser with a pulse energy of 100 μJ (compared to 7 mJ for the CLF), firing at 351nm with a repetition rate of 333 Hz [43]. Optical filters are also applied to reduce the effects of the NSB (Night Sky Background) light. Backscattered light from the atmosphere is collected by 3 parabolic mirrors with a diameter of 80 cm and focal length of 41cm [43]. The mirrors are steered on a frame driven by 2 DC motors with a pointing direction accuracy of 0.2°. A photograph of these mirrors is shown in Figure 3.27 [44]. The LIDARs are designed to be operated remotely from a central location at the Observatory, known as CDAS (Central Data Acquisition System). Operation of the LIDARs starts at the same time as the Fluorescence Detectors, at astronomical twilight. At this time the LIDAR goes into autoscan mode, which is a sequence of scans over the course of an hour that goes over the entire night. The main purpose of this scan sequence is to extract all the information needed for monitoring data quality and measure atmospheric properties such as aerosol attenuation [43]. Each of the other types of scans illustrated in Figure 3.28 are detailed below.

Horizontal shots: The laser fires horizontally in the direction of the Central Laser Facility which measure both the aerosol attenuation at ground level and the horizontal homogeneity of aerosols across the array (this is a key assumption in the event reconstruction).

Continuous scans: The laser is fired by scanning across the two rotation axes of the mirrors. The purpose of this scan is to find cloud layers and measure the cloud coverage over a 45° cone. The output of such a scan is shown in Figure 3.29.

Discrete scans: The laser is positioned at a few fixed angular points of particular interest, in order to gain higher statistics and calculate light attenuation from aerosols as a function of height.

Once these scans are done, the data are sent to CDAS in order to process the signals given from the laser shots. Figure 3.29 shows the output signal from a continuous scan performed at Auger. Clouds in the view of the laser are seen quite clearly as the region of high backscatter at $\sim$ 2000 m. From Figure 3.29 we are able to extract various properties about a cloud, specifically the cloud base height, cloud coverage and cloud thickness. All of these parameters are important for much of the analysis done in subsequent chapters.

There is one final scan that is performed, specifically during FD data acquisition. An online reconstruction program analyses cosmic ray events detected by both the

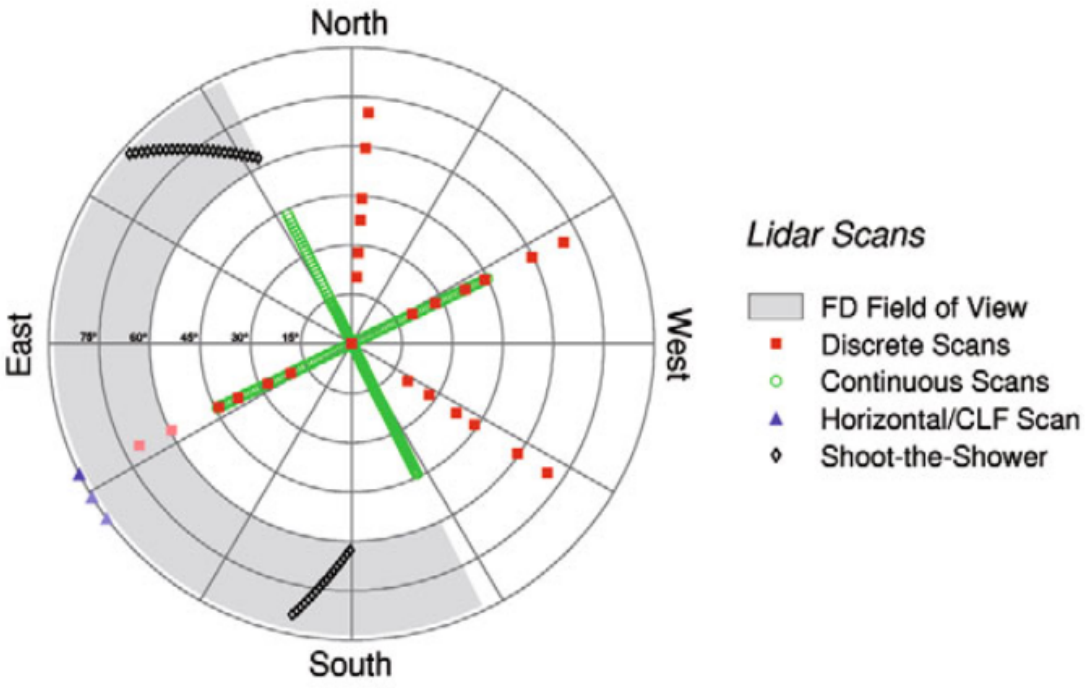

Figure 3.28: Diagram showing the various types of scans performed by the LIDARs at the Observatory. Continuous and Discrete scans are done outside of the FD field of view, whilst the horizontal shots (done once per hour) fire in the FOV in the direction of the CLF [43].

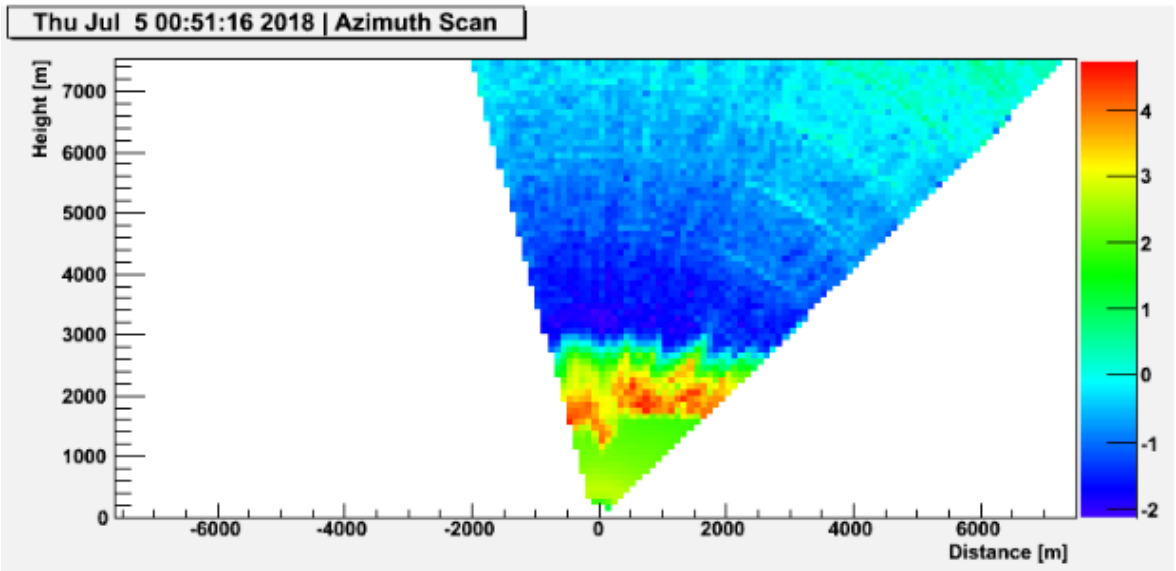

Figure 3.29: Raw Output from the LIDAR Continuous Scan. A cloud layer is present at around 2000m.

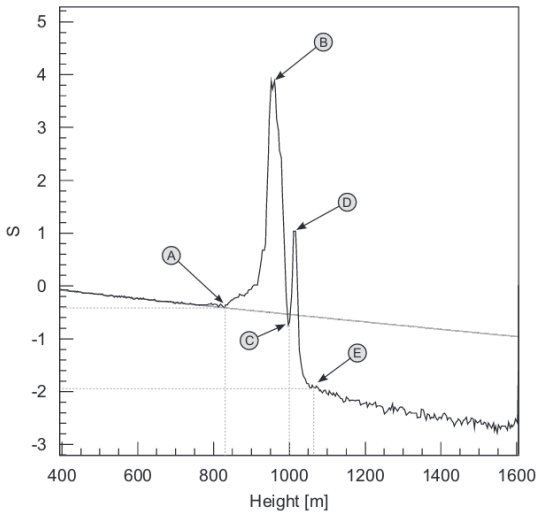

Figure 3.30: An example output signal $S(h)$ from the LIDAR. Clouds appear as strong backscatter regions such as the peak in the signal at point B. [44]

fluorescence and surface detectors in real time. If a high energy event or event with an anomalous longitudinal profile is found, the LIDAR auto-scan is stopped and the lasers are fired in the field of view of the FDs [43]. The LIDAR scan is done along the path of the shower track (see Figure 3.4) and is known as a Shoot the Shower (StS) scan [43]. A shoot the shower scan is done in order to check for the presence of clouds along the shower track in the FD FOV since this can distort the signal received at the FD, causing an under/over estimation of the shower energy during reconstruction. This is a form of calibration/cross-check to make sure that the high energy reconstructed is not caused by a cloud in the FOV of the FD. The backscattered signal from the laser shots as a function of height $[S(h)]$ can be extracted in order to determine the presence of clouds. Clouds in the atmosphere will appear as a spike in this backscatter signal, which is illustrated in Figure 3.30. From this backscatter signal $S(h)$, the cloud base height in the atmosphere can be extracted based on the peak position of the light profile.

3.5.7 GOES

GOES (Geostationary Operational Environmental Satellite) is series of satellites owned and operated by the NOAA (National Oceanographic and Atmospheric Administration), placed in a geostationary orbit to monitor atmospheric parameters over North and South America. Images are taken every 30 minutes in both the visible and infrared bands of the electromagnetic spectrum. Satellites are launched with a lifetime of a few years, by which time they are replaced with an upgraded version. The most recent of these (GOES-16) was launched on November 19, 2016 [45]. On the ground, the satellite data from GOES-13 is projected onto a pixelated grid with dimensions of $2.4 \times 5.5\,\mathrm{km}$. This is in the infrared band, as a means of detecting night time cloud. GOES is used at Auger in order to image clouds in the infrared, and produce cloud probability maps over each pixel. The pixel structure of the satellite over Auger is shown by Figure 3.31.

With the GOES system viewing the clouds from above, as well as other instruments

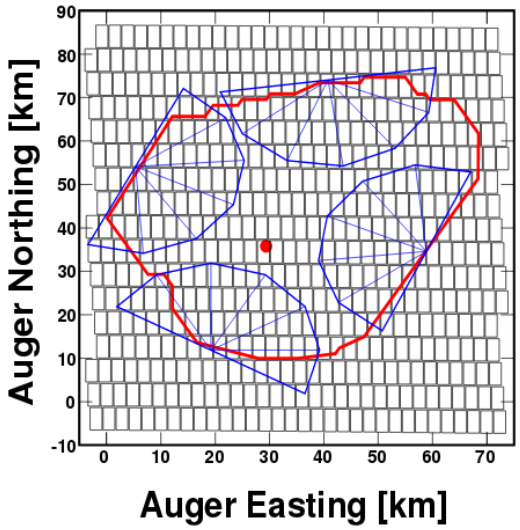

Figure 3.31: Pixelation of the satellite image data from GOES 13. The blue semi-circles represent the FD FOV, the red outlines the SD array with the CLF shown in the middle with the red dot. Source: [46]

viewing the clouds from below (at ground level), we are able to get a more complete picture of the cloud at Auger. All of these instruments play a role in the quality cuts done on event reconstruction discussed in section 3.5.9.

Figure 3.32: A Xenics Gobi 384 radiometric IR camera.

3.5.8 Infrared (IR) Cloud Cameras

In the atmosphere, clouds have an effect on the climate. Those that appear at lower altitudes (e.g. stratus cloud) have a high albedo, meaning that they reflect a large portion of the incoming radiation away from Earth's surface. This results in a net cooling effect on both the surface and the surrounding atmosphere. Higher altitude clouds, for example cirrus, are more transparent thus allowing more solar radiation to pass through. Additionally, infrared radiation reflected off Earth's surface is absorbed by clouds and re-directed towards the ground, which gives a net warming effect. The IR radiative signature is detectable by cloud cameras positioned at each of the various FD sites at the observatory. The cloud cameras are Xenics Gobi-384 model radiometric cameras, manufactured by Xenics Infrared Solutions [47]. Each one has 384×288 pixels, with each pixel being $\sim 25\,\mu$m in size [47]. The lens itself is 18 mm thick. A photograph of the camera is given in Figure 3.32.

Each camera scans the FD Field of View every 5 minutes, as well as an entire hemisphere image every 15 minutes (known as a Full Sky scan), and these images are processed and mapped onto FD pixels to look for the presence of cloud in the FOV. These cameras are sensitive to radiation in the 7-13 μm band - the blackbody spectrum peak of a thick cloud (the term "blackbody" refers to an ideal object that absorbs and re-emits 100% of incident radiation - there is no reflected radiation). However, there is an issue with this detection method. Namely, peaks in the emission lines of water vapour are also present in this band at $\sim 11\mu m$ - even for clear skies [21]. This can

Figure 3.33: One IR camera installed at an FD site, inside the gray box on a Pan and Tilt (Coihueco).

lead to false positives in the detection of cloud, compromising data obtained during that time.

An image of the camera hardware is shown in Figure 3.33, as well as the resulting images captured on a night of observation (Figure 3.34). Each image has an angular size of $\sim 40^\circ$ in azimuth and $\sim 30^\circ$ in zenith. The direction in which images are taken is controlled by two motors, which are used for the Pan (shift in azimuthal angle) and Tilt (zenith angle shift) of the camera. A FOV scan consists of a sequence of 5 images controlled by the Pan, which covers the FOV of the Fluorescence Detector. A full sky scan is a sequence of 19 images. The first 8 are taken in a 360° pan, with 5 corresponding to the FOV of the FDs and the rest covering the remaining azimuthal range. The tilt motor then moves the camera to view 30-60° in elevation. A sequence of 10 images is taken at this higher elevation, in the same vein as the first 8 images [48]. One final scan is taken with the camera pointed vertically (aptly named a vertical image), giving us 19 images in total. Figure 3.34 shows the full 19 images stitched together for each camera.

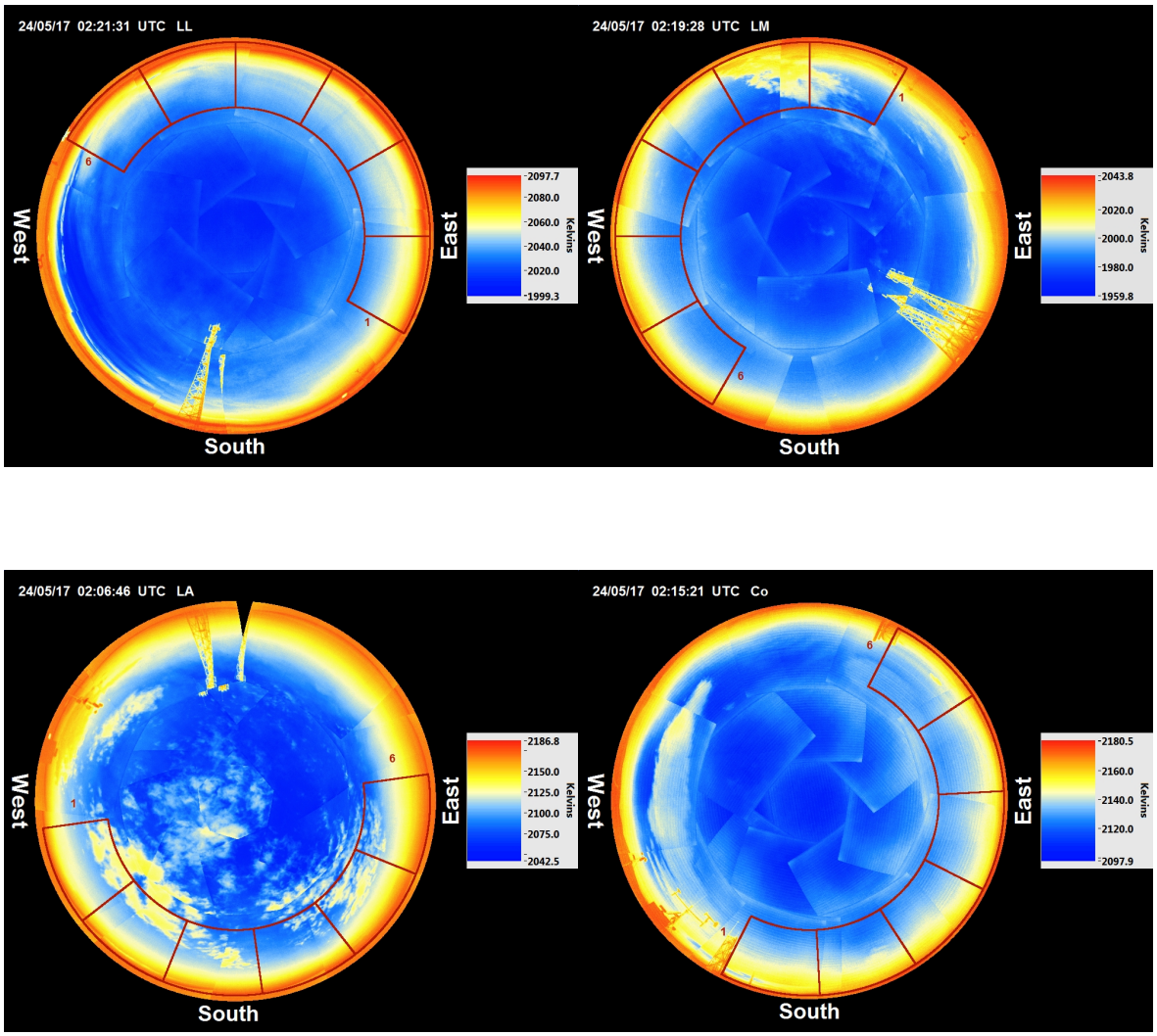

Figure 3.34: The resulting images of the Cloud Cameras running at each site on May 24th, 2017. The colour contours are output in ADC counts (of arbitrary units), which is converted into a brightness temperature later in the analysis (more on this in chapter 5). In order of site, there is: Top Left: Los Leones, Top Right: Los Morados, Bottom Left: Loma Amarilla and Bottom Right: Coihueco

In Figure 3.34, the red outlines correspond to the field of view of each telescope in the corresponding FD. For the colour scale, cold blue colouration represents a clear sky, whilst the whitish-yellow colours are the IR signature of the overhead cloud. The bright yellow-orange band forming a ring around the image appears warmer than the rest of the sky since this is air close to the horizon, thus more atmosphere is being imaged so that it appears warmer. As mentioned earlier, water vapour close to the ground can absorb and emit infrared radiation, which will warm the atmosphere close to ground. This is why cloud close to ground is difficult to detect, as it is often falsely identified.

Section 3.5.9 will go into detail about how each of these cloud monitoring instruments (Cloud Camera, LIDAR and GOES) are used to place cloud quality cuts on air shower reconstruction, as mentioned in section 3.1.2.

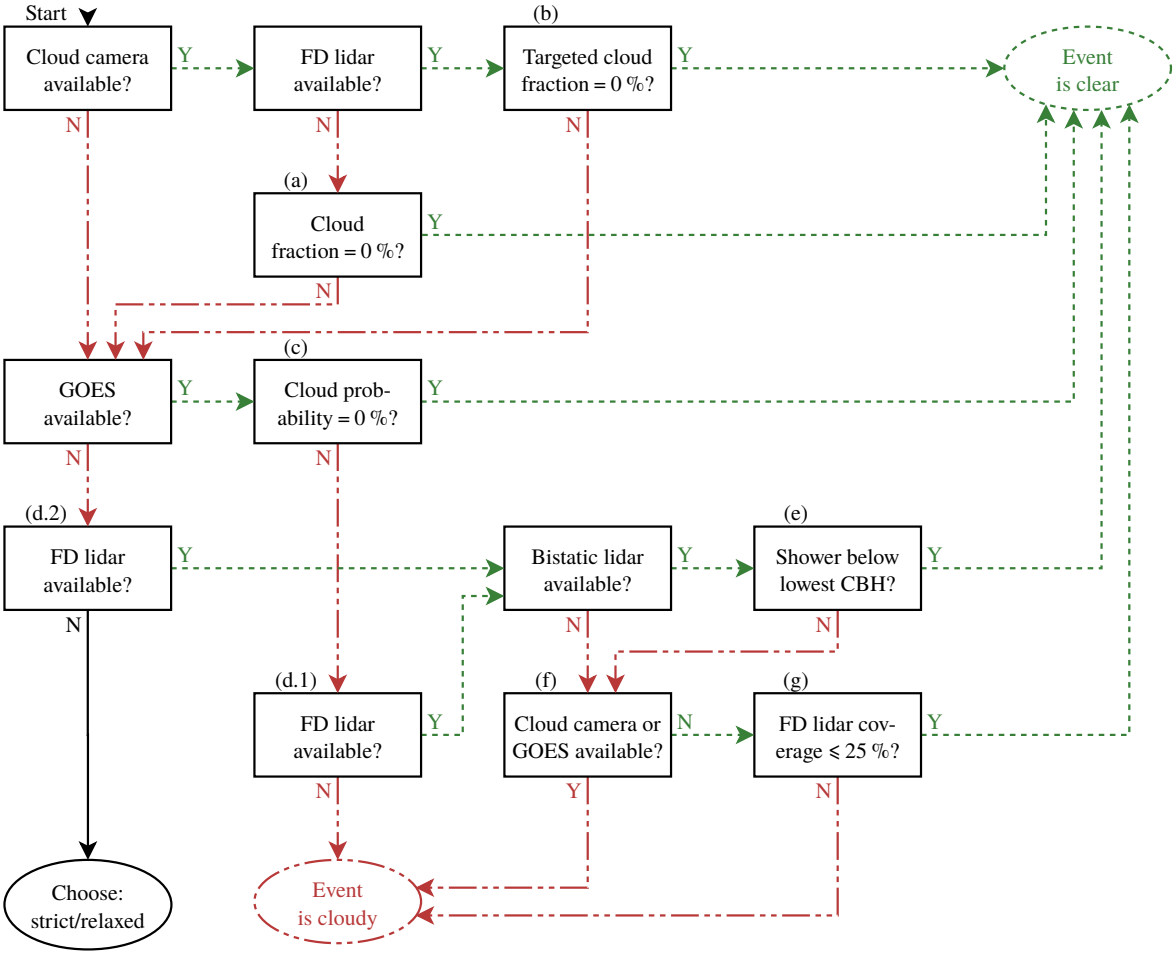

Figure 3.35: An overview of the cloud cuts employed at Auger. The boxes represent a step that has either a Yes or No (Y/N) based on the available instruments for a particular FD event [39].

3.5.9 Cloud Cuts

The cloud cameras at Auger are used to define cloud cuts on the air shower reconstruction. These cloud camera cuts are shown in Figure 3.35 (credit to Violet Harvey) [39]. Cloud cuts are done on each instrument at Auger, starting with the cloud cameras. In the rare case where the cameras are not available, the cloud cuts will move to the next available instruments, namely the LIDAR and GOES. These are used to decide whether or not a particular event in the reconstruction is cloudy or clear. Simply put, the event is not used in any other analysis unless it is flagged as clear. On those rare occasions where there are NO instruments available, then the choice of keeping/discarding the event is made independent of any quality cuts. The cloud cuts provided in Figure 3.35 are a series of simple Yes/No responses based on the equipment operating at Auger. These cuts are crucial in creating a dataset that contains high quality showers. Additionally, so called exotic events (with unusual longitudinal profiles) must be examined to determine if clouds are the cause of the strange behaviour seen [39].

More on the operation of the cloud cameras is outlined in Chapter 5.

Chapter 4

The Earth's Atmosphere

The Earth's atmosphere can be divided into 4 main regions; the troposphere ($\sim 1-20$ km), stratosphere ($\sim 20-50$ km), mesosphere ($\sim 50-65$ km) and thermosphere (65 km and beyond) [49]. Another known layer is the ionosphere which is incorporated into the thermosphere. The ionosphere is only a consideration when discussing the variability of electron density with height, which is outside of the scope of this thesis. The boundaries between the various layers are variable in height as well as in thickness. This variability is due to temperature fluctuations, and the flow of convective cells driving global climate through heat exchange and momentum transport. The height of these layers varies additionally depending on latitude. The four main regions of the atmosphere are shown in Figure 4.1. As Figure 4.1 shows, there are clearly defined boundaries where the overall vertical temperature gradient changes sign. What this means is that an increase or decrease of temperature as a function of height for each layer of the atmosphere is determined by different processes. Now to discuss the two lowest layers of the atmosphere in more detail.

Firstly, there is the troposphere, extending to $\sim$ 12 km (the exact height is dependent upon atmospheric temperature [50] and is a function of both latitude and season [51]). This is the layer housing the majority of the mass of atmospheric constituents (around 75-80 %) and almost all of the water vapour ($\sim$ 99%). Weather patterns (cloud formations, rain, thunderstorms among others) exist solely in this layer, as do the sources of convection cells which circulate heat and winds around the globe through the entire depth of the troposphere. Three convection cells exist in the atmosphere (known as Hadley, Ferrel and Polar cells) which drive atmospheric circulations. Air heats up as it approaches the Equator, causing it to rise. Air approaching the North and South poles cools down and subsequently sinks. A schematic of this convective circulation is given in Figure 4.2.

The overall steady decrease in temperature with height in the troposphere is due to a process known as adiabatic cooling [51]. Adiabatic cooling is the process of reducing heat through a change in air pressure caused by volume expansion. In this process, consider a parcel of air at a given temperature and pressure. As this parcel is lifted higher into the atmosphere and experiences a lower atmospheric pressure, the parcel will expand. The parcel continues to expand until it reaches a pressure equilibrium with the surrounding air (by decreasing the pressure). This in turn decreases the temperature of the parcel. Eventually, parcels of air may cool enough to reach the dew point temperature (the point at which it is cool enough for water vapour to condense into liquid droplets), which is when clouds can form [51]. This is why the troposphere

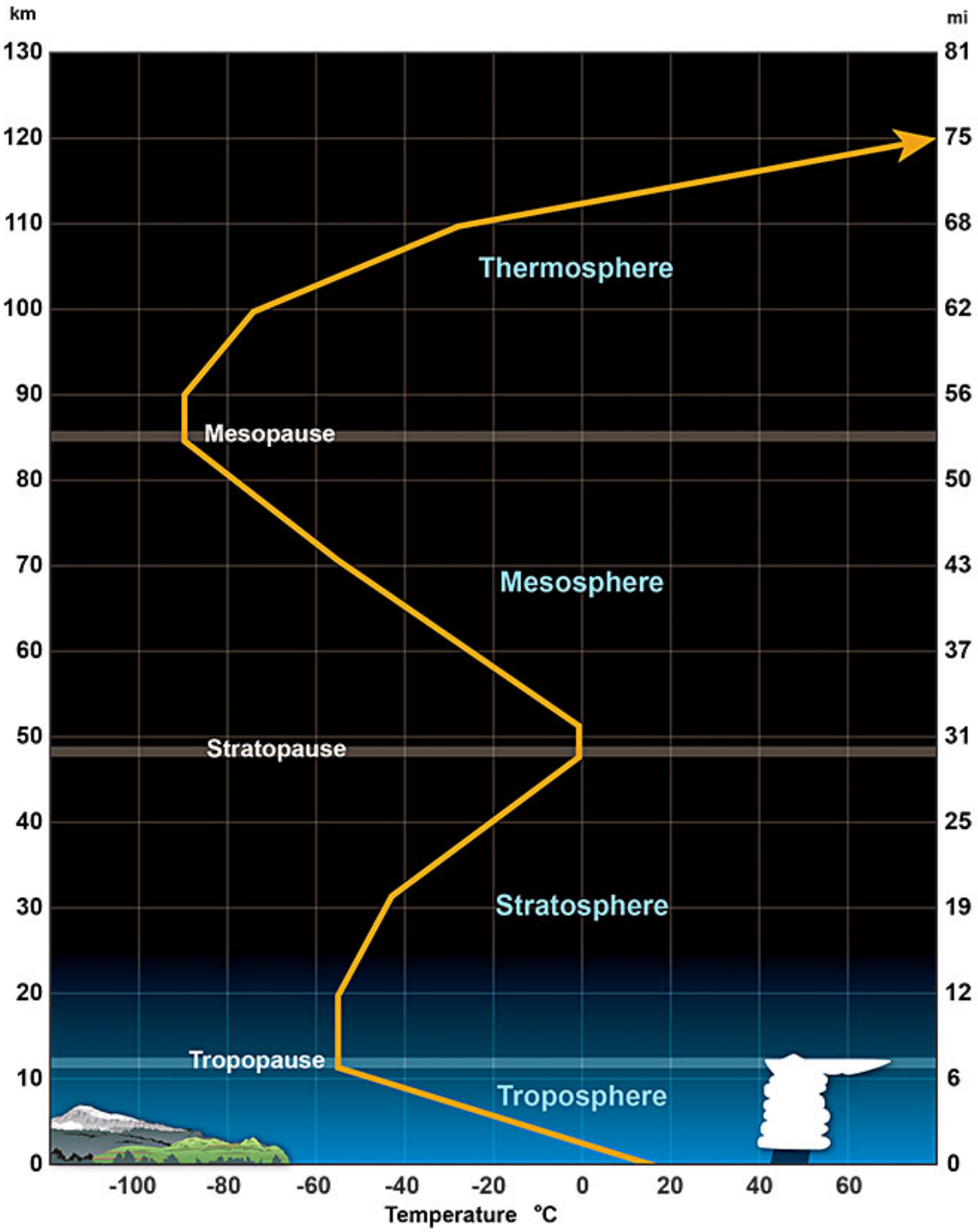

Figure 4.1: Temperature Profile of Earth's Atmosphere. There are clear changes in the overall vertical temperature gradient which define the boundaries between various layers. Source: [49]

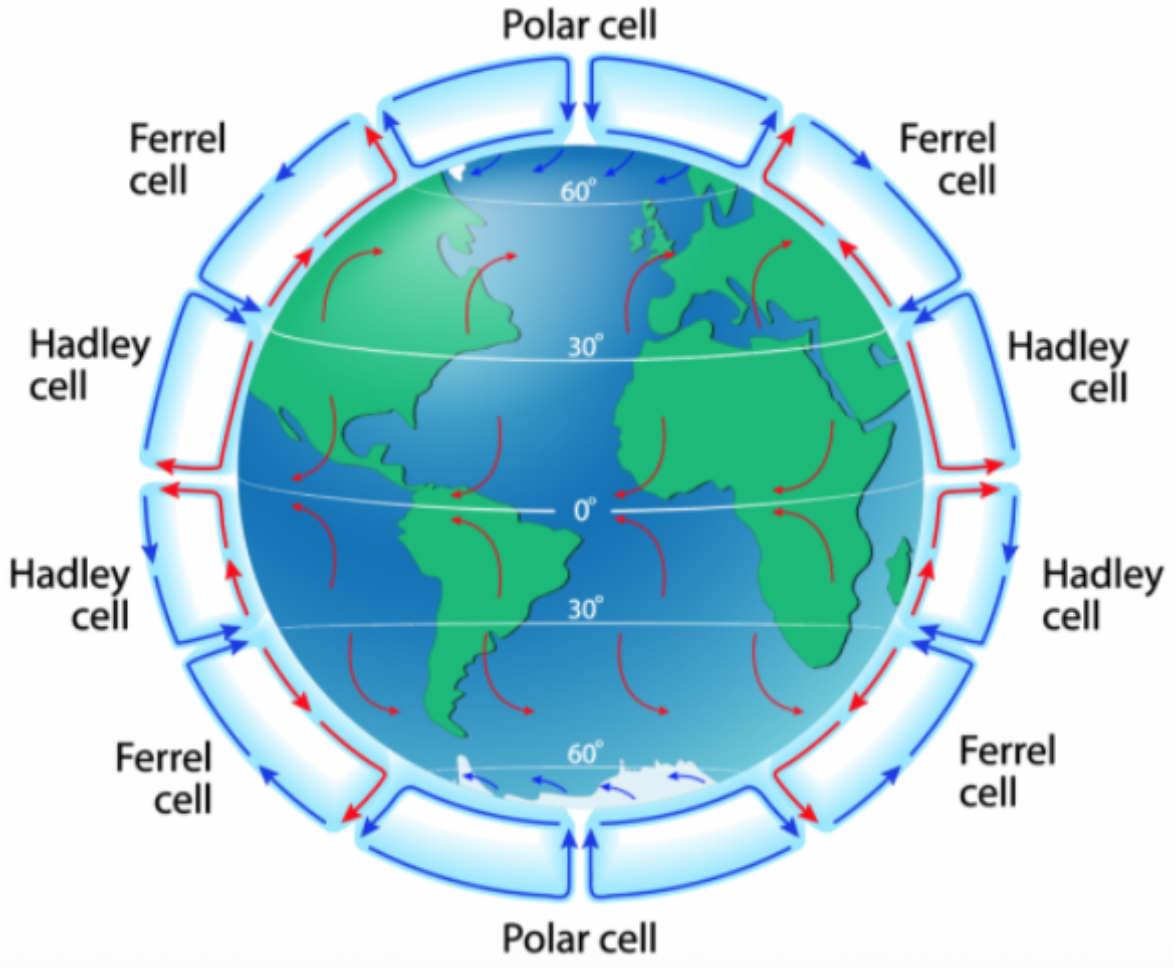

Figure 4.2: The presence of convection cells in the Earth's atmosphere which drive air circulations through heating imbalances. Source: [52]

is responsible for cloud formation in the atmosphere.

Above the troposphere lies the tropopause, a thin isothermal layer (i.e. a layer in which the temperature is constant at all heights inside the layer) which separates the troposphere from the stratosphere. The tropopause is located at the temperature minimum of the troposphere where the lapse rate (rate of change of atmospheric temperature as a function of height) is defined to be less than 2°C per km [53]. Atmospheric mixing and heat transport by convection only occurs when the temperature decreases with height. Due to the constant temperature of the tropopause, this lack of mixing and transport causes the density of this layer to be much lower than the troposphere below. It provides a cap for weather patterns such as clouds (this will be discussed further in section 4.1.2) as it is difficult for air from the troposphere to rise into the tropopause due to the lower density. This confines clouds to the troposhere. The tropopause height varies significantly through the seasons and as a function of latitude (i.e. it can be as low as 7 km at the poles, and as high as 20 km at the equator at various stages throughout the year) (on average it is about 12 km at the Pierre Auger Observatory) [54]. The tropopause can be classified as a "ceiling" or barrier for the global climate patterns observed in everyday life.

Moving up now to the stratosphere, a layer extending to around 50 km, some major differences from the atmosphere below are observed. The stratosphere is home to the Earth's ozone layer (between $\sim 15 - 50$ km), a major source of absorption of the Sun's UV (Ultraviolet) radiation, which protects organic life on the Earth's surface from this damaging UV radiation. Dissociation of Oxygen O_2 by incoming solar UV radiation breaks up the O_2 into atomic oxygen as described in equation 4.1 [55].

$$O_2 + energy \rightarrow O + O \tag{4.1}$$

After this, the atomic oxygen can recombine with molecular oxygen to form ozone as described in equation 4.2 [55].

$$O_2 + O \rightarrow O_3 \tag{4.2}$$

Due to the high concentration of oxygen in the atmosphere, some of the incoming UV radiation is absorbed in the stratosphere. During this process, the absorbed UV radiation is continuously converted into heat, which warms the surrounding stratosphere [56]. In a similar vein, the ozone can interact with the free atomic oxygen as shown in equation 4.3 [57] [58].

$$O_3 + O \rightarrow O_2 + O_2 \tag{4.3}$$

This reaction is much slower than those described by equations 4.1 and 4.2 which means the ozone layer is increasing in thickness over time. This leads to a larger net warming effect over time. This is why recent studies have noted an increase in the height of the stratosphere (where it begins) over the past 100 years, due to this enhanced warming [54].

For the purposes of this project, only the troposphere will be considered, since weather is entirely contained within this layer, and the development of extensive air showers in the atmosphere at Auger are only a consideration at low altitudes (below ~ 10 km, or atmospheric depths larger than $\sim 300\,\mathrm{g/cm^2}$) for Auger. For more information on Extensive Air Showers, see chapter 2. The following sections go into more detail about the formation and properties of clouds in the atmosphere, specifically how they are formed, the different types of cloud one can find in the atmosphere, and their radiative properties (which become important for the studies in Chapters 7 and 8).

4.1 Cloud Microphysics

The most common method of cloud formation in the atmosphere is the adiabatic cooling process, as described at the beginning of this chapter. If the air temperature matches the dew point temperature (hence relative humidity [RH] is 100%), condensation of water vapour is activated and, as a result, clouds can form. The process in which clouds form is known as nucleation, and it comes in two main forms: homogeneous and heterogeneous nucleation.

Homogeneous nucleation of water vapour is a random process with no nucleation sites, however it requires supersaturation in order for the water droplets to grow. Supersaturation is a regime where the relative humidity in the atmosphere is above 100%. This homogeneous nucleation is the result of the coalescence of water molecules [59], which is a probabilistic process (and an unlikely one [59]). If two water molecules are colliding, the collision tends to be elastic, making bonding less likely since kinetic energy is conserved (i.e. there is no energy released in the system, which would occur in a bonding situation) [59]. Three body collisions can overcome this bonding limitation, however they very rarely occur in nature. When analysing this nucleation process, the Gibbs Free Energy can be considered. Gibbs Free Energy (GFE) is a measure of the spontaneous nature of a thermodynamic process. When the change in GFE ΔG is negative, a process can occur spontaneously. In homogeneous nucleation, the spontaneous nature of it is described by the GFE parameter. For the homogeneous formation of water droplets (with radius R) the change in GFE ΔG is given by;

$$\Delta G = -V_d n_l \Delta\mu + 4\pi R^2 \gamma \tag{4.4}$$

Here, V_d is the volume of the droplet, with a given molecule concentration n_l and chemical potential $\Delta\mu$ (the difference between the chemical potential energy of water in its gas/vapour phase and its liquid phase). The second term is the work done against the surface tension (γ) in order to form the droplet with radius R [59]. Taking the derivative of equation 4.5 with respect to the Radius R gives equation 4.6.

$$\Delta G = -\frac{4}{3}\pi R^3 n_l \Delta\mu + 4\pi R^2 \gamma \tag{4.5}$$

$$\frac{d\Delta G}{dR} = -4\pi R^2 n_l \Delta\mu + 8\pi R\gamma \tag{4.6}$$

In order to find the critical radius R_c for droplet formation, the GFE change must be a maximum - this is shown in equation 4.7.

$$\frac{d\Delta G}{dR} = 0 \tag{4.7}$$

$$\therefore 4\pi R_c^2 n_l \Delta\mu = 8\pi R_c \gamma \tag{4.8}$$

$$\therefore R_c = \frac{2\gamma}{n_l \Delta\mu} \propto \frac{2\gamma}{n_l k T ln(\frac{e}{e_s})} \tag{4.9}$$

In equation 4.9, e is the partial pressure due to water vapour, and e_s is the saturated vapour pressure. Relative Humidity (RH) is then defined by equation 4.10 as

$$RH = \frac{e}{e_s} \tag{4.10}$$

As shown by equation 4.9, the critical radius is inversely proportional to the logarithm of the Relative Humidity (RH). When RH is 100%, the critical radius becomes equation 4.11.

$$\therefore R_c = \frac{2\gamma}{n_l k T ln(1)} \rightarrow \infty \tag{4.11}$$

Since the critical radius is approaching an extremely large value (tending to infinity), there is no physical radius that droplets can exist at in order to condense and form clouds. As a result, in the case of homogeneous nucleation, supersaturation is required in order for droplets to grow (with relative humidity $> 100\%$) and clouds to form. If $RH > 100\%$, it means that $ln(RH) > 0$; thus there is a maximum size R_c that droplets can reach (when supersaturated), at which point they will condense into clouds. Due to the highly unlikely nature of this nucleation to occur, it is typically heterogeneous nucleation that occurs in the atmosphere.

Heterogeneous nucleation is the process where water droplets will condense onto the surface of cloud condensing nuclei (CCN) in the atmosphere. There are several forms that these cloud condensing nuclei can take, such as chemical salts (Ammonium Sulfate and Sodium Chloride are common). Aerosols are one of the biggest types of CCN, and information about their effect on clouds and climate is a question constantly asked over the past 20-30 years [40]. Aerosols are known to affect cloud properties and can have a strong influence on climate [40]. Typically there are enough CCN present in the atmosphere such that the only requirement for cloud formation is 100% relative humidity (the requirement for condensation to occur). Heterogeneous nucleation occurs much more frequently than homogeneous nucleation, since the energy required for the gas-liquid phase transition (of water) to occur is greatly reduced. This process is responsible for the majority of cloud formation in the lower atmosphere.

Cloud can come in many forms, and the formation and behaviour of each class (and subclass) is dependent on many factors - where they form, what makes them up, and the atmospheric circulation and conditions that are present [60]. The following sections describe some of the more well known cloud types in more detail.

4.1.1 Stratus Cloud

The lowest cloud type in the sky (in altitude), stratus is the cloud that appears most commonly on cold, wet days. It is the low hanging grey cloud visible in the sky, with a typical height range of a few hundred metres up to $\sim$ 2 km. Stratus clouds may be associated with light precipitation. When they appear close to ground, there is another name for them - fog. Most commonly, stratus clouds are formed by thermal advection and radiative cooling. Air is displaced by mechanical means (e.g. being pushed up over a hill or mountain range) and as it rises, the decreased pressure causes an expansion of the air which cools it down. Eventually it will reach a point where the air cools enough to reach the dew point temperature. When this happens, condensation occurs via heterogeneous nucleation, as described in section 4.1, to form stratus cloud. Horizontal uniformity of stratus clouds is characteristic of this cloud type (i.e. the base altitude of the cloud remains unchanged, even over a large horizontal range).

4.1.2 Cumulus Cloud

Perhaps the most easily recognisable cloud form, cumulus clouds are the thick, puffy white clouds seen in patches of sky on clear days. They are primarily formed by thermal

Figure 4.3: The "anvil" at the top of a cumulonimbus cloud. Source: [61]

convection, where UV radiation from the Sun reaches Earth's surface, gets absorbed and is re-radiated back up into the atmosphere. This warm mass of air above the surface becomes less dense than the surrounding air. As a result, it will rise and move to a colder region where it proceeds to cool and water vapour begins to condense. If the updraft is strong enough (caused by strong vertical winds in the lower atmosphere) to push air rapidly upwards, with enough moisture present, large vertical development of the clouds can occur, forming a cumulonimbus cloud. These are associated with heavy rain, lightning and thunder. High wind speeds (up to 110 km/h) can rapidly push the warm air upwards and create long towers of this cloud which penetrate high up into the troposphere. At the top of the cloud, since the tropopause prevents air from the troposphere from rising any further (as discussed at the beginning of this chapter), clouds cannot move up into the tropopause. This is illustrated in Figure 4.3. As a result, the cloud top peaks and further spreads out horizontally. This creates the characteristic anvil shape seen in these clouds as shown in Figure 4.3.

4.1.3 Cirrus Cloud

The highest forming of all cloud types, cirrus appears as the thin wispy clouds. Since cirrus clouds form high up in the troposphere ($\sim$ 8 to 11 km), they are entirely composed of ice crystals of varying size and shape (as the temperature at these altitudes is below the freezing point of water). These high altitude clouds reflect a large portion of sunlight, and absorb long-wave radiation emitted by the warm Earth [62] [63]. The absorption and re-radiation of this long-wave radiation by cirrus clouds contributes a net warming effect on the Earth.

Ice in the atmosphere forms via two separate mechanisms, termed homogeneous and heterogeneous freezing [64]. Homogeneous freezing is better understood. It is the

dominant process forming these high altitude clouds. This involves a spontaneous sublimation of ice at sufficiently cold temperatures, however an extremely high Relative Humidity is required ($\sim 150 - 170\%$) in order to undergo this freezing. On the other hand, heterogeneous freezing can occur at $\sim 100\% RH_i$ (RH_i being relative humidity with respect to ice) at temperatures just below 0° [64]. This phenomenon is the depositional freezing of water droplets onto the surface of ice nuclei (in the same vein as cloud condensing nuclei discussed earlier) [64] [65]. Laboratory experiments have shown several candidates for these ice nuclei (INs), including mineral dust, organic compounds, and some metallic materials. However only ~ 1 in every 10^5 particles in the atmosphere can act as an ice nucleus (IN). It is not clear why the concentrations of IN are so low. As such, cirrus clouds are less common than their lower altitude cousins.

Now that we have established the different types of cloud that are formed in the atmosphere, we must discuss some other important behaviours of these clouds, which will become very important for our analysis later on in this dissertation. Section 4.2 discusses the radiative properties of clouds, particularly in the infrared (IR) portion of the electromagnetic spectrum, as this pertains to our cloud camera measurements at Auger (see section 3.5.8).

4.2 Radiative Properties of Clouds

Clouds in the atmosphere can absorb, reflect or re-emit both incoming solar radiation and outgoing terrestrial radiation into the surrounding atmosphere [66] [67]. Solar radiation is the incident radiation from the Sun, at different wavelengths on the electromagnetic (EM) spectrum. Typically, this incident radiation is at wavelength ranges in the visible portion of the EM spectrum. For a clear atmosphere with no cloud present, this short-wave radiation will pass through the gases present in the atmosphere, heating the Earth's land and oceans before being radiated from the Earth's surface at a longer infrared (IR) wavelength. Some of this IR radiation is absorbed by gases in the atmosphere (such as water vapour and carbon dioxide) since they have energy levels that correspond to these infrared wavelengths [68]. These gases re-radiate photons at the same infrared wavelengths into the atmosphere, distributing heat and energy into the atmosphere and onto the ground (this is known as the greenhouse effect) [68].

The effectiveness of absorption/emission is determined by the emissivity ϵ, which takes a value between 0 and 1. The emissivity is defined as the amount of incoming radiation that is absorbed by an object. An emissivity of 0 means no incoming radiation is absorbed (known as a "white-body") and a value of 1 means that all incoming radiation is absorbed (known as a "black-body").

As described in section 4.1.2, most clouds are primarily comprised of water vapour in the form of droplets. The one exception is high altitude cirrus clouds, which are comprised of ice crystals (see 4.1.3). Cirrus clouds have different radiative and reflective properties to both short-wave and long-wave radiation.

In this way, clouds (as well as aerosols) in the atmosphere affect the radiative balance of the atmosphere. This radiative balance is the difference between energy gained in the atmosphere (from incident radiation) and the energy lost back into space by reflection (and re-radiation) of radiation from the Earth's surface, clouds, and other molecules present in the atmosphere. Figure 4.4 summarises all of the net radiative effects caused in Earth's atmosphere, including those effects caused by the presence of aerosols and cloud. A positive radiative forcing value means the component creates a

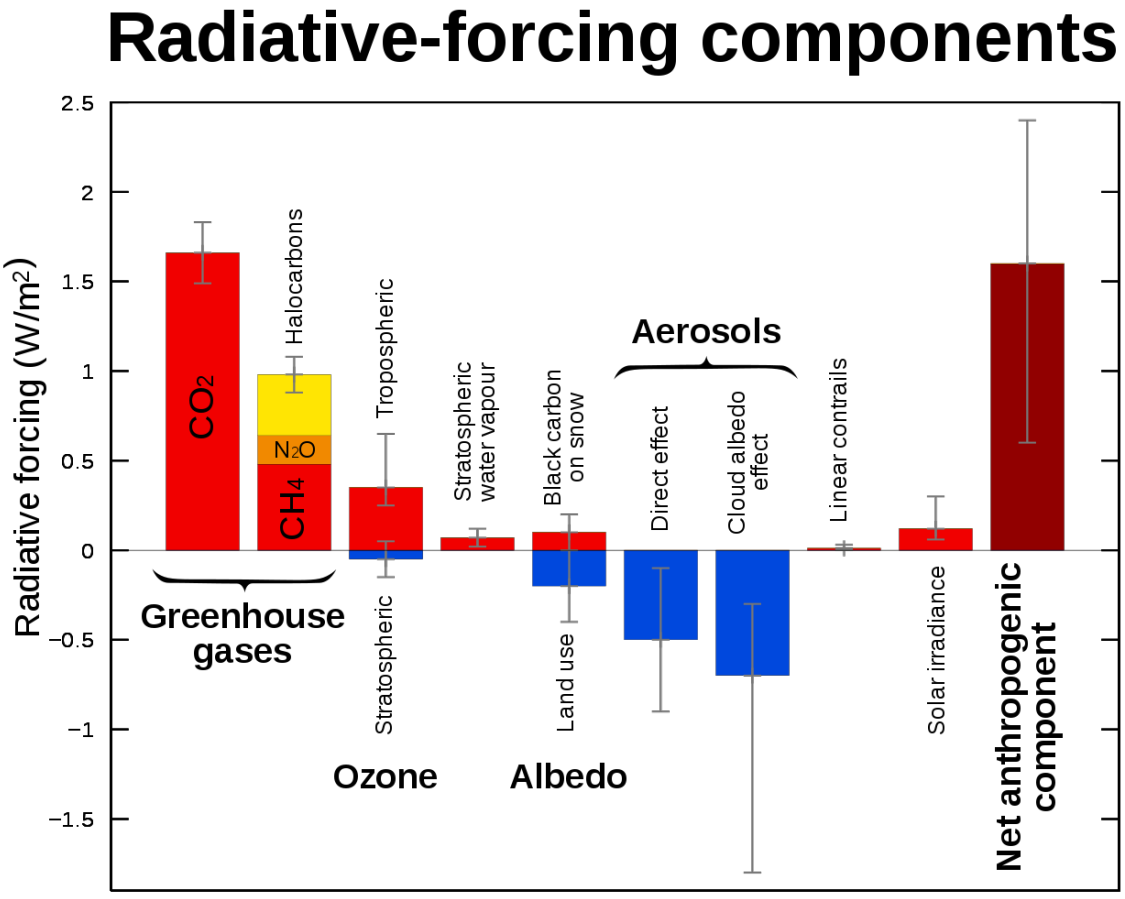

Figure 4.4: Net radiative forcing of the atmosphere due to several factors. Clouds and aerosols will be the focus of this dissertation. Source: [69]

net warming effect on the atmosphere (for example clouds) and a negative radiative forcing value means the component produces a net cooling effect on the atmosphere (for example aerosols).

Firstly, a discussion on clouds, since they overall produce a net warming effect on the atmosphere (as shown by Figure 4.4). This is achieved since low altitude thick clouds (below $\sim$ 4000 m) act as blackbody radiators in the infrared (wavelengths between $\sim$ 2 and 15 μm). The emissivity of the cloud approaches that of a perfect blackbody - studies have shown an emissivity of cumulus cloud (see section 4.1.2) to be in the range $0.97 - 1$ [70]. This means that thick, low altitude clouds are acting as almost perfect blackbodies. As a result, long-wave terrestrial radiation from the surface of Earth is mostly absorbed by clouds, before being re-emitted back towards the surface which acts to heat up the lower atmosphere. In this way, clouds can be thought of as a "blanket" which traps heat and raises the surface temperature of the Earth [66] [67]. It is estimated that around 60% of Earth's outgoing long-wave radiation is re-emitted by clouds back towards the ground [71]. In total, clouds and atmospheric molecules (water vapour and carbon dioxide) re-radiate about 80% of the terrestrial radiation back towards the ground [71].

Conversely, short-wave solar radiation can be either transmitted or scattered off the top of a cloud, the amount of which is determined by the cloud's albedo and optical thickness [66] [67] [72]. Optical thickness is a measure of the ratio of incident light

that is absorbed by an object, to the amount of light transmitted through that object. Albedo is a measure of the fractional amount of radiation incident on a surface that is reflected. For Earth's surface, the albedo factor is ~ 0.35 for short-wave radiation [73]. For most thick clouds in the atmosphere, the albedo to short-wave radiation is very high (typically around 0.9), meaning that a very small amount of incident solar radiation is allowed to pass through the cloud [66] [67]. On its own, this reflection would cause a net cooling of the Earth's surface over time [72] [68].

High altitude cirrus clouds (which can also persist for long periods of time) also have a profound effect on this short-wave radiation [72]. Due to the high reflectivity of a cirrus cloud (the albedo of ice is estimated to be between 0.5 and 0.7 [74]), they produce a large amount of reflection and scattering of solar short-wave radiation, which creates a net cooling effect on the atmosphere.

It is the long-wave radiation which is of most interest at the Pierre Auger Observatory, since this wavelength band lies within the operational range of our IR cloud cameras (see section 3.5.8). How terrestrial long-wave radiation relates to cloud height and temperature measurements is discussed in section 4.2.1.

It is also worth briefly discussing the effect of aerosols. Aerosols present in the atmosphere produce a net cooling effect due to their high degree of scattering of incident radiation (as shown in Figure 4.4). The effect of aerosols can be split into two separate effects; a direct and an indirect effect. The direct effect of aerosols is a combination of scattering and absorption of radiation by the aerosol particles themselves [69].

The indirect effect is due to the interaction of aerosols with clouds and their formation. The presence, abundance and lifetime of a cloud in the atmosphere is dependent upon the amount of aerosols present in the atmosphere [69]. As mentioned earlier in this section, aerosol particles act as nucleation sites for droplet formation in clouds. In a cloud, a larger concentration of aerosol particles creates more nucleation sites for the condensation of water vapour [69]. As a result, a larger number of small cloud droplets are present (rather than a collection of large droplets that have coalesced) which increases the reflection of solar radiation by the cloud, leading to a higher cloud albedo [69].

The second indirect effect due to aerosols relates to the lifetime of a cloud in the atmosphere. An increase in aerosol particle concentration leads to the increased number of smaller droplets within the cloud. Smaller water droplets take longer to grow to the size necessary to precipitate out of the cloud as rain. This leads to the cloud having a much longer lifetime in the atmosphere, thus increasing the total amount of radiation reflected [69].

These indirect aerosol effects are difficult to accurately quantify, however they do contribute a net cooling of the atmosphere (as shown in Figure 4.4) since these effects lead to more solar radiation being reflected back into space [69].

Now that the basics of cloud formation are known, as well as the radiative effects of cloud, one can use this information to determine cloud temperatures in the infrared.

4.2.1 Measurement of Cloud Temperature in the Infrared

An important characteristic of clouds for this dissertation is an estimate of the base height at which they exist in the atmosphere. This is important for the cloud quality cuts on air shower reconstruction as discussed in section 3.5.9. One method used in determining the base height of a cloud is to first measure its temperature, which can be

compared to the temperature of the atmosphere in order to give the cloud height. In the past, measurements of cloud temperature were done by flying an aircraft inside the cloud, utilising an immersion thermometer attached to the outside of the aircraft [75]. The main issue with measurements in this way is a phenomenon known as "sensor wetting" (cloud droplets condense onto the surface of the thermometer), which causes a warmer temperature of the cloud to be read compared to the surrounding atmosphere [75]. This is due to the release of latent heat when water vapour condenses onto the surface of the temperature sensor [66] [67]. Latent heat is the energy absorbed or released by a substance when it changes state. If the state changes from a lower density to a higher density (for example, from a gas to a liquid), energy is released. In the case of the atmosphere, the water vapour condenses into clouds and this condensation process releases latent heat into the atmosphere, which heats the cloud [66] [67]. This heat that is released will be detected by the temperature sensor as it heats up. The amount of energy released via the condensation process depends upon the total amount of water vapour in the atmosphere (locally) [67]. The release of latent heat is the more important effect to be considered in measuring a cloud's temperature.

Radiosondes (small radio transmitters attached to weather balloons and flown into the atmosphere) measure temperature in the same fashion, with the same sensor wetting issue. This effect is estimated to cause a temperature increase of $\sim 2\,\mathrm{K}$ [75]. In addition to the sensor wetting issue, immersion thermometers typically have a slow response, which affects the accuracy of temperature measurements within the cloud.

Figure 4.5 shows an example radiosonde temperature profile with a cumulus cloud present. Due to the condensation of water vapour via convection, latent heat is released inside the cloud, which causes the warmer air to rise. Sensor wetting of the radiosonde instrument also contributes to this effect. As more condensation occurs, this effect causes a temperature inversion as shown in Figure 4.5 [67].

The temperature and dew point match for the bottom 1 km of the atmosphere due to the condensation of water vapour (the atmosphere has reached a relative humidity of 100%). As can be seen, the temperature of the atmosphere above this point increases due to the release of latent heat by the cloud, an important factor for the analysis done in Chapter 8.

Figure 4.6 shows an example of a cloud slowing down the rate of temperature decrease due to the latent heat process. This differs from Figure 4.5 since it is a different type of cloud (stratus compared to cumulus for the previous example), formed from a slightly different process (advection as opposed to convection) [67]. The advection process causes the cloud to spread horizontally (see section 4.1.1) and as a result there is no temperature inversion. However, it does serve to decrease the rate of temperature change in the atmosphere (i.e. a lower lapse rate).

Cloud temperatures can also be measured by flying an aircraft below the base of a cloud, and pointing a temperature sensor vertically in order to measure the temperature at the cloud base [75]. There is an issue with this method also, due to the presence of intervening greenhouse gases between the sensor and the cloud, most notably water vapour and carbon dioxide, which both have emission bands in the infrared [75]. This also serves to raise the temperature that the sensor interprets, making the cloud appear warmer than it truly is. Correcting for this emission effect has been an ongoing study in meteorology [75]. A correction based on the emission from various greenhouse gases in the atmosphere using a ground based detector (the infrared cloud cameras at Auger) is shown in section 8.6.

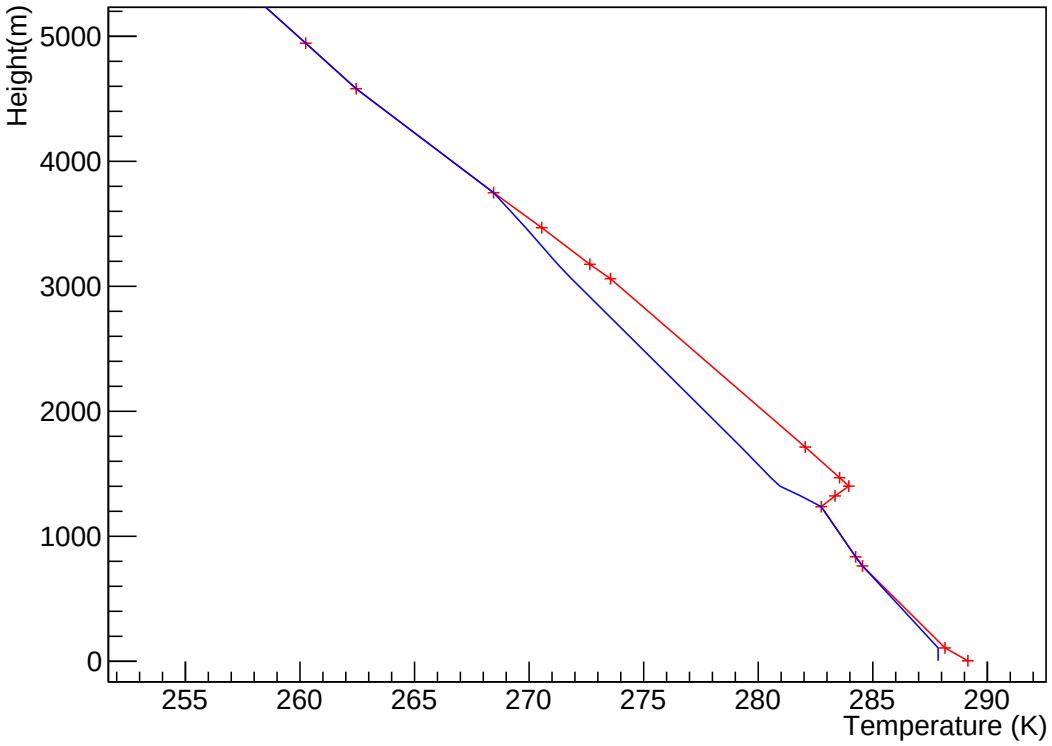

Figure 4.5: Temperature (Red) and Dew Point (Blue) Temperatures as a function of height from a radiosonde flight in Adelaide on October 16th 2020 at 0 UTC. A cloud present at $\sim 1200\,\mathrm{m}$ creates a temperature inversion (a region in the troposphere where temperature increases with height) of about $1\,\mathrm{K}$ in the atmosphere due to the release of latent heat.

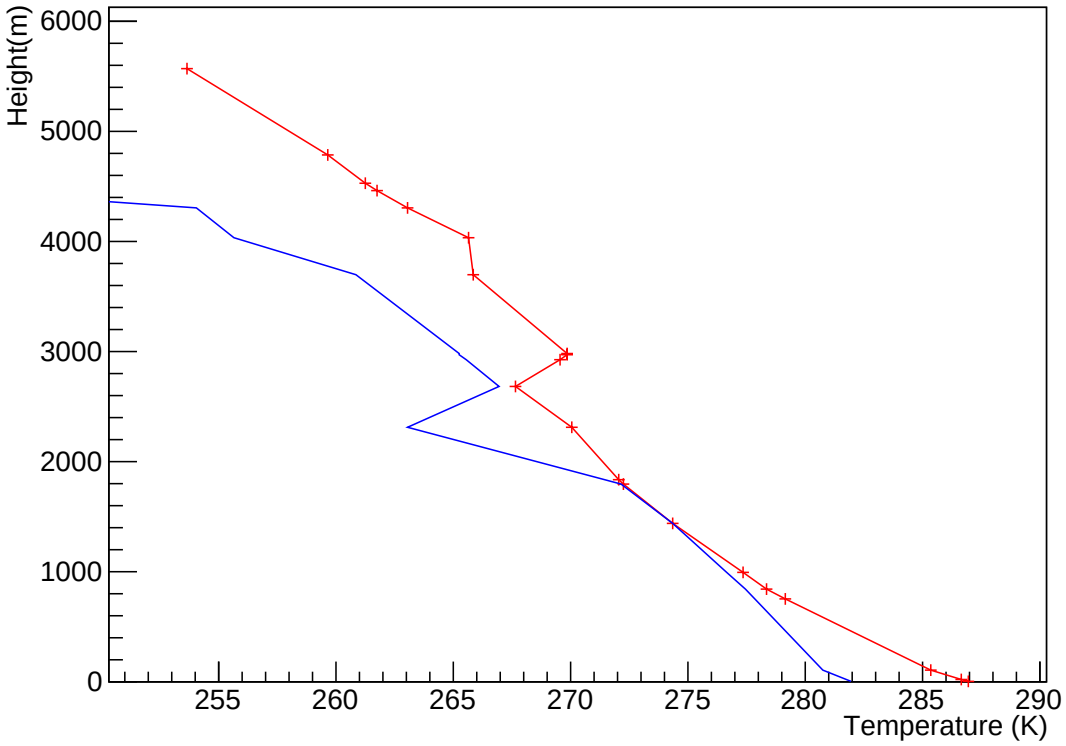

Figure 4.6: Temperature (Red) and Dew Point (Blue) Temperatures as a function of height from a radiosonde flight in Adelaide on October 8th 2020 at 0 UTC. A cloud present below $\sim$ 1500 m slows down the decrease of temperature with height due to the release of latent heat. Another cloud present at $\sim$ 2300 m causes a drop in the dew point temperature due to convection causing the saturated air to rise.

Another method of measuring cloud temperatures via a ground based detection system has been to use Raman LIDAR similar to what is in use at Auger (see section 3.5.6) [76]. Measurements of a 354.7 nm LIDAR laser pulse are split into two channels (which peak at 353.4 nm and 354.2 nm respectively) and the relative backscatter signal return between the two channels is used to determine the cloud temperature as in equation 4.12

$$T(z) = \frac{a}{\ln\left[\left(\frac{X_{r1}(z,\lambda_{r1})}{X_{r2}(z,\lambda_{r2})}\right)\right] + b} \tag{4.12}$$

where $T(z)$ is the cloud temperature, X_{r1} and X_{r2} are the amplitudes of the LIDAR signals of each channel, and a and b are derived calibration constants [76]. This temperature is compared to radiosonde measurements and is shown to agree quite well as shown by Figure 4.7.

Figure 4.7 shows good agreement with the derived temperature given by the LIDAR with the measured temperature from the radiosonde [76]. This method, utilising a LIDAR, is the most accurate method of measuring cloud temperatures in the atmosphere. The infrared cloud cameras at the Pierre Auger Observatory (see section 3.5.8) also measure sky temperature based on the long-wave radiation emitted by clouds, thus there is a way of estimating cloud temperatures in the cloud camera field of view. From the measured cloud temperature, a cloud height can be estimated, which is a main goal of this dissertation. The method used for this purpose is described in Chapter 8.

Of course, in order to accurately estimate this cloud base height, the cloud temperature must be matched to the temperature of the surrounding atmosphere. To do this, corrections have to be made to the measured cloud temperature to take into account the effects described previously that cause a cloud to be warmer than its surrounding environment. These effects include the release of latent heat by the cloud when water vapour condenses, as well as the absorption of Earth's long-wave radiation by both clouds and intervening atmospheric gases. The corrections for these effects are described in section 8.6.

As well as using the cloud cameras to estimate cloud base heights in the atmosphere, experimental studies utilising known atmospheric parameters can be used to determine the base heights of cloud. Section 4.3 describes an experimental study used to determine the base height of a cloud at any given time during the observing period of the Pierre Auger Observatory, by calculating a quantity known as the lifting condensation level (or LCL).

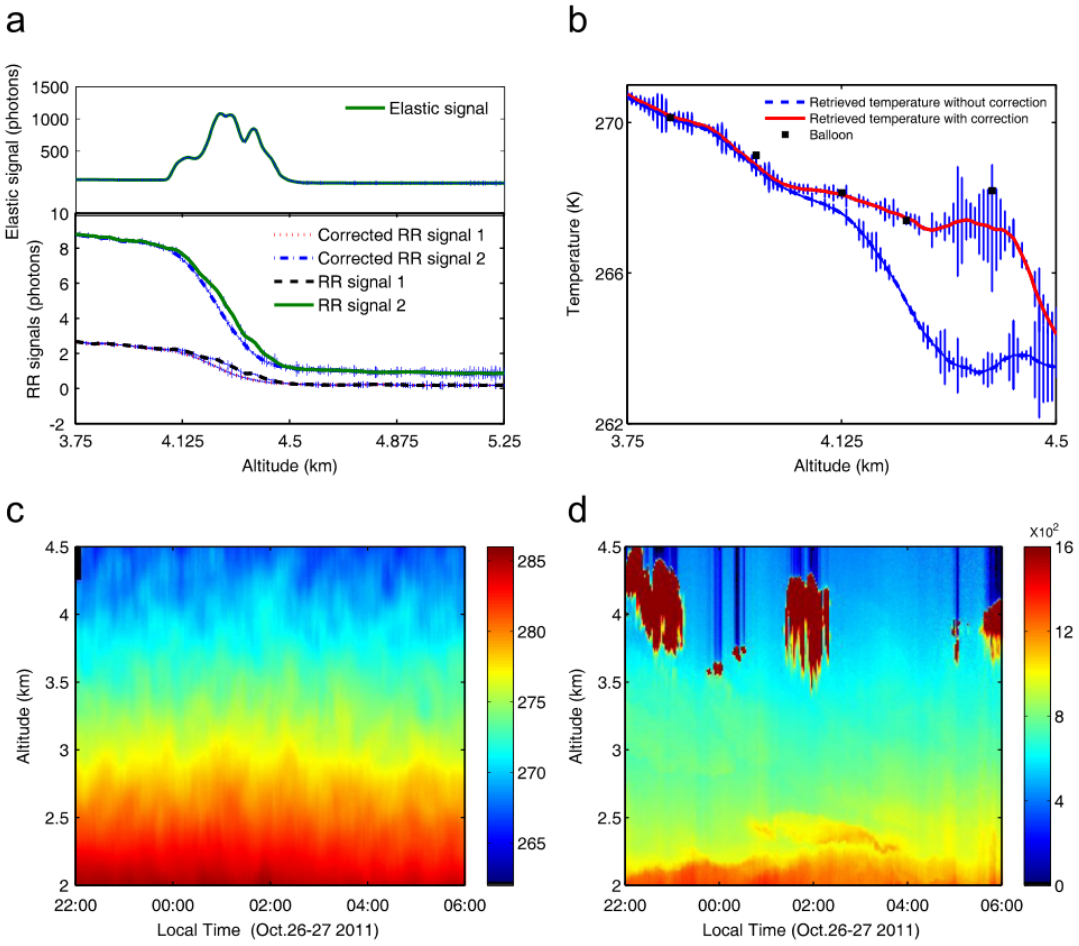

Figure 4.7: a) Elastic backscatter signal of channel 1 (353.4 nm - red dots) and channel 2 (354.2 nm - blue dashes) of the Rotational Raman (RR) LIDAR at 22:10 local time. b) Temperature $T(z)$ from equation 4.12 within a cloud at the same 22:10 local time [red line] compared to radiosonde measurements at the same time [black squares]. Error bars are derived from statistical uncertainties. c) Atmospheric temperature from equation 4.12 over the course of a whole night, as a function of height. d) Image of clouds as viewed by the LIDAR over the course of the night. The cloud of interest is in the top left corner at 22:10 local time. Colour bar represents the backscatter intensity from the LIDAR. Source: [76]

4.3 Estimating the Cloud Base Height

There are a number of instruments utilised in order to measure the heights and other properties of clouds in the lower atmosphere above Auger. However, there is unreliability in the estimations of cloud base height. The cloud detection instruments (the LIDAR, and the CLF as discussed in Chapter 3) have not always been running during the detection of extensive air showers. This can be due to any number of factors (such as unresponsive electronics or weather conditions including rain or high wind speed). Rain at the Pierre Auger site causes issues with both the FDs and LIDARs, meaning the equipment cannot be operated during these times. The two detectors (FD/LIDAR) have different thresholds of wind speed (35 km/h for the LIDAR and 50 km/h for the FD) before they cannot be operated, thus there are times where the FD is undergoing data acquisition but it is too windy to safely operate the LIDAR. This effects the fractional on-time of the LIDARs. Figure 4.8 shows the fractional on-time of each LIDAR station from 2004-2015. The vertical axis refers to the fraction of events detected by each FD which had LIDAR data corresponding to that time (i.e. the percentage of time that cloud base heights were calculated during a run, compared to how often one would expect these data to be calculated).

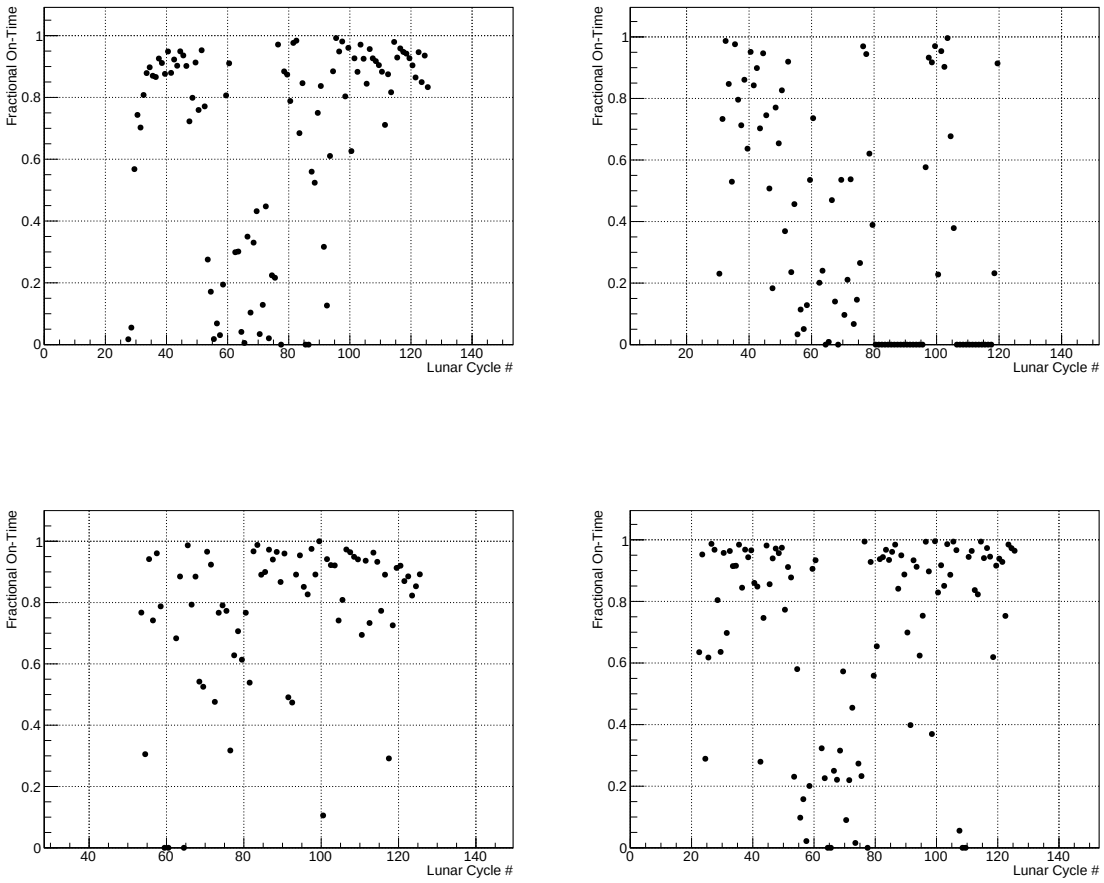

Figure 4.8: Fractional on time of the FD LIDARs over a 12 year span from 2004 (Lunar Cycle of 0) to 2015 (Lunar Cycle of 140). Each graph corresponds to a different FD site. Top row (left to right): Los Leones and Los Morados. Bottom row (left to right): Loma Amarilla and Coihueco. Each dot is the fractional on time for a given lunar cycle (which is roughly 29 days long).

The same analysis can be done for the Raman LIDAR located at the CLF (see section 3.5.3). The results for the CLF are shown in Figure 4.9. Ideally, the fractional on time of the CLF, Raman LIDAR, and the LIDAR at each FD site would be 1 (i.e. running 100% of the time). However, as can be seen from figures 4.8 and 4.9, it is not possible to rely solely on LIDARs or the CLF/XLF as a means of measuring cloud heights, due to the factors mentioned earlier. In fact, the largest factor in preventing the running of the CLF and LIDAR is the wind, as noted previously.

The infrared cloud cameras (which do not have their data acquisition influenced by wind and rain) are practical instruments to use in order to measure cloud base heights. However, whilst clouds are able to be detected with these cameras, for a given image it is unknown how far away the clouds are from the camera site. As a result, there is currently no way of determining the height of clouds from cloud camera data. An experimental technique utilising the temperatures of clouds in the cameras field of view to estimate cloud heights is discussed in Chapter 8.

In the meantime, one option for theoretically calculating cloud base heights is to use the lifting condensation level (or "LCL" for short). The following sections will go

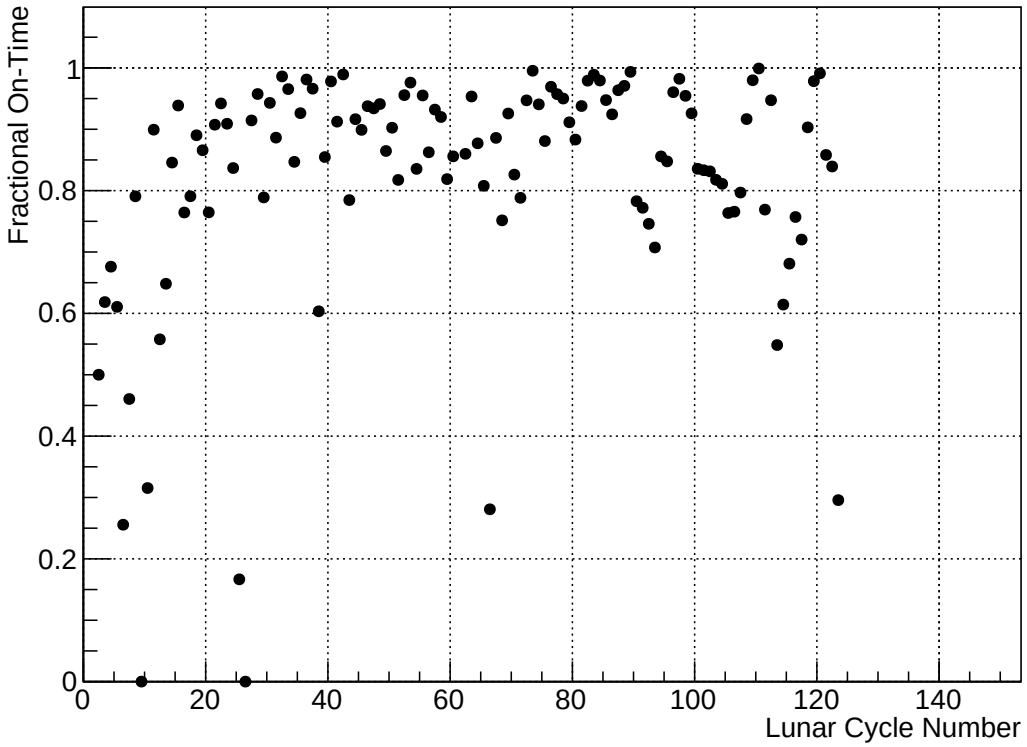

Figure 4.9: Total on time fraction for each lunar cycle ($\sim$ 29 days) of the CLF over a 10 year span.

into detail about the lifting condensation level and how it is used to estimate the height of clouds in the atmosphere above Auger.

4.4 Theory of the Lifting Condensation Level (LCL)

The LCL is the height in the atmosphere at which relative humidity reaches 100% and the air saturates, when a parcel of air is lifted by force. When this occurs, and enough water vapour is present, clouds will form via the condensation of vapour in the air. The calculation of the LCL is given by equation 4.13, utilising either radiosonde or GDAS data to calculate/measure each parameter.

$$LCL = \frac{1}{\Gamma - \Gamma_w}(T - T_d) \tag{4.13}$$

A derivation of the LCL is given in Appendix C. In equation 4.13, T is the atmospheric temperature at a given height, T_d is the dew point temperature at that same height, Γ is the environmental lapse rate of the atmosphere, and Γ_w is the wet adiabatic lapse rate of the atmosphere. In the troposphere it is known that the temperature decreases as a function of height due to adiabatic cooling. This rate of decrease is known as the "lapse rate", which is assumed to be linear in this lowest atmospheric layer (measured in units of K/km). The environmental lapse rate is the amount by which the temperature decreases in the true atmosphere, as opposed to the dry adiabatic lapse rate (a fixed value of 9.8 K/km, assuming the atmosphere is dry i.e. no water vapour). The environmental lapse rate is the value Γ in equation 4.13, which must be calculated based on the temperature profile from GDAS. The value of Γ ranges between 4 and 7 K/km. The wet adiabatic lapse rate (Γ_w above) is the temperature decrease which would be found for completely saturated air. This is also assumed to be a linear decrease, with a value of ~ 3.3 K/km [51]. The temperature of the atmosphere (at a given height) T is given by GDAS, and the final parameter that must be calculated in order to compute the LCL is T_d - the dew point temperature. This is the temperature to which air must be lowered, at a constant pressure, in order to become saturated - thus condensation can occur and clouds can form.

The calculation of dew point (units of °C) is given by equation 4.14 for air temperatures $T > 0$°C [37].

$$T_d = \frac{234.9 \ln(\frac{e}{6.108})}{17.15 - \ln(\frac{e}{6.108})} \tag{4.14}$$

This form of the Magnus formula is chosen to be consistent with the formula used in the processing of GDAS data [48]. The dew point calculation is empirical due to its highly variable nature with respect to humidity, temperature and air pressure in the atmopshere. In equation 4.14, the quantity e is the partial vapour pressure - the partial pressure exerted by water vapour present in the mixed gas atmosphere (in units of hPa). From the relative humidity values given by GDAS (see section 3.5.2), the water vapour pressure e can be calculated and fed into equation 4.14 (see Appendix C) to calculate the dew point temperature. The saturated vapour pressure is given by Tetens' formula as in equation 4.15.

$$e_s = 6.107 \times \exp\left(\frac{17.15T}{234.9 + T}\right) \tag{4.15}$$

T is the temperature in the atmosphere at a given height (in °C). Equation 4.15 only applies for $T > 0$°C. Usually, the relative humidity is quoted, rather than the vapour pressures. The relation of these pressures to humidity is given by equation 4.16.

$$RH = \frac{e}{e_s} \times 100(\%) \tag{4.16}$$

The calculation of the LCL requires knowledge of the temperature, lapse rate and relative humidity/vapour pressure. The LCL can be calculated at any pressure level of the atmosphere, and the value it gives is an estimate of the height of cloud, above the reference point at which it was calculated. For example, taking the temperature/humidity 2 km above sea level (a.s.l), with a calculated LCL of 1 km, means that the cloud layer is estimated to be 3 km above sea level. For the purposes of this study, the LCL was calculated at a height of 1.42 km a.s.l (as the Pierre Auger Observatory is located at an altitude of $\sim$ 1420 m a.s.l [21]). There are many ways to know these quantities - in this study the values of temperature, humidity and pressure are taken from GDAS (see section 3.5.2).

4.4.1 GDAS Data Analysis

The databases for GDAS are publicly available on the NOAA website, for any latitude/longitude point. A large number of parameters can be extracted from this website, a list of which can be found here: <https://www.ready.noaa.gov/gdas1.php>. The Auger Offline Framework (the tool used to extract atmospheric information) contains the data for the grid point shown in Figure 4.10, which is what is subsequently used here.

As can be seen by Figure 4.10, the grid point in question lies just inside of the surface detector array, closely positioned to the Loma Amarilla FD site. Previous studies of the atmosphere by J. Abraham et al. [2010] have shown consistency across the entire array in terms of aerosol content and parameters such as temperature, pressure and relative humidity [77]. Due to this fact, one can assume uniformity across the array making the data from that grid point a reasonable approximation to the behaviour at any point within the array.

From this GDAS database, the goal is to extract temperature and vapour pressure values at ground level. As it happens, GDAS itself calculates relative humidity, which is then converted to vapour pressure by the empirical formulae described by equation 4.16. In this case, relative humidity (RH) is given as a fractional value between 0 and 1 since relative humidity is simply the ratio of partial pressure of water vapour to the saturated water vapour pressure. Since one knows relative humidity and temperature from GDAS, and the saturated vapour pressure is calculated from equation 4.15, the partial water vapour pressure e can then be calculated as in equation 4.17.

$$e = RH \times e_s \tag{4.17}$$

All of these values are fed into the equations to calculate dew point temperature at ground level (given by equation 4.14). However, there is one value that is not immediately available from this GDAS database - the environmental lapse rate Γ. In order to do this, one use the assumption described earlier, that temperature decreases linearly with height in the troposphere. Hence, one must perform a linear fit to the GDAS temperature profile in order to calculate Γ. An example of such a profile is shown in Figure 4.11.

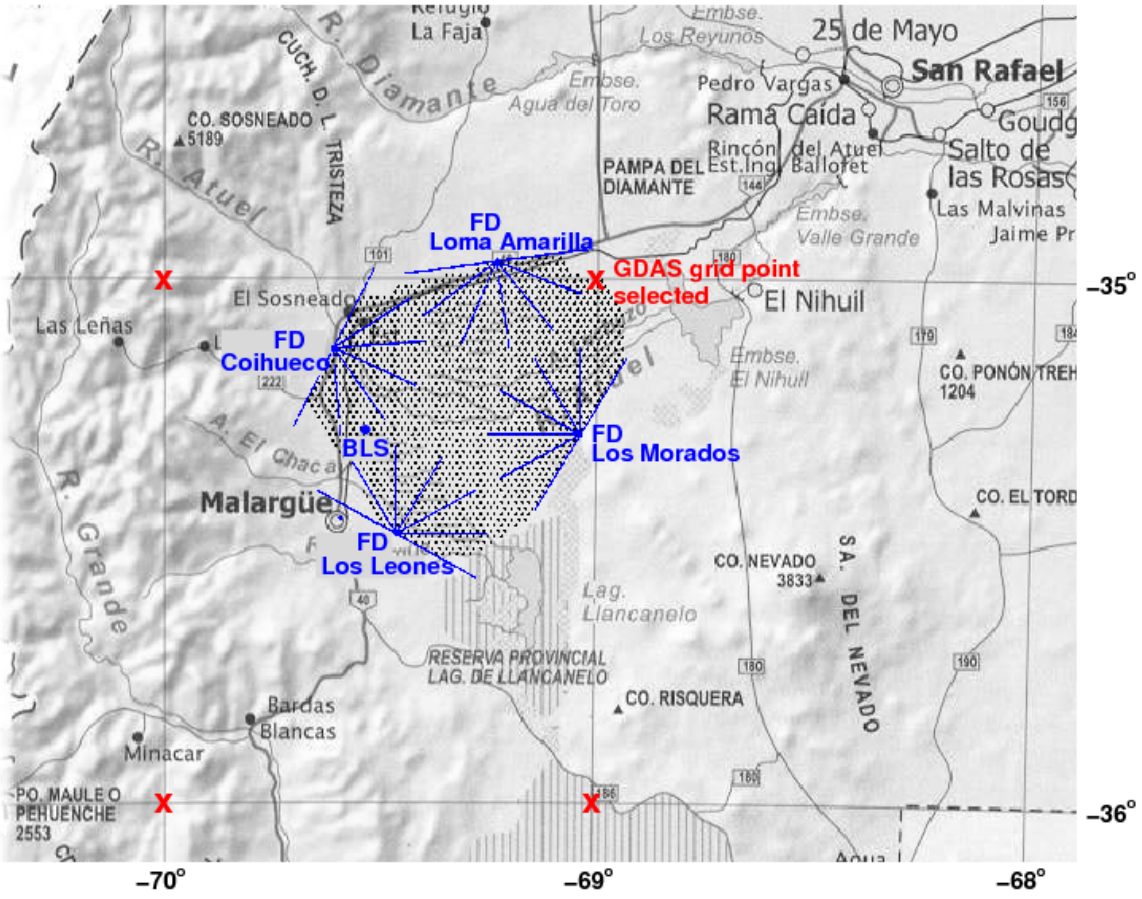

Figure 4.10: GDAS grid points (red crosses) with respect to the array. The selected grid point is indicated.

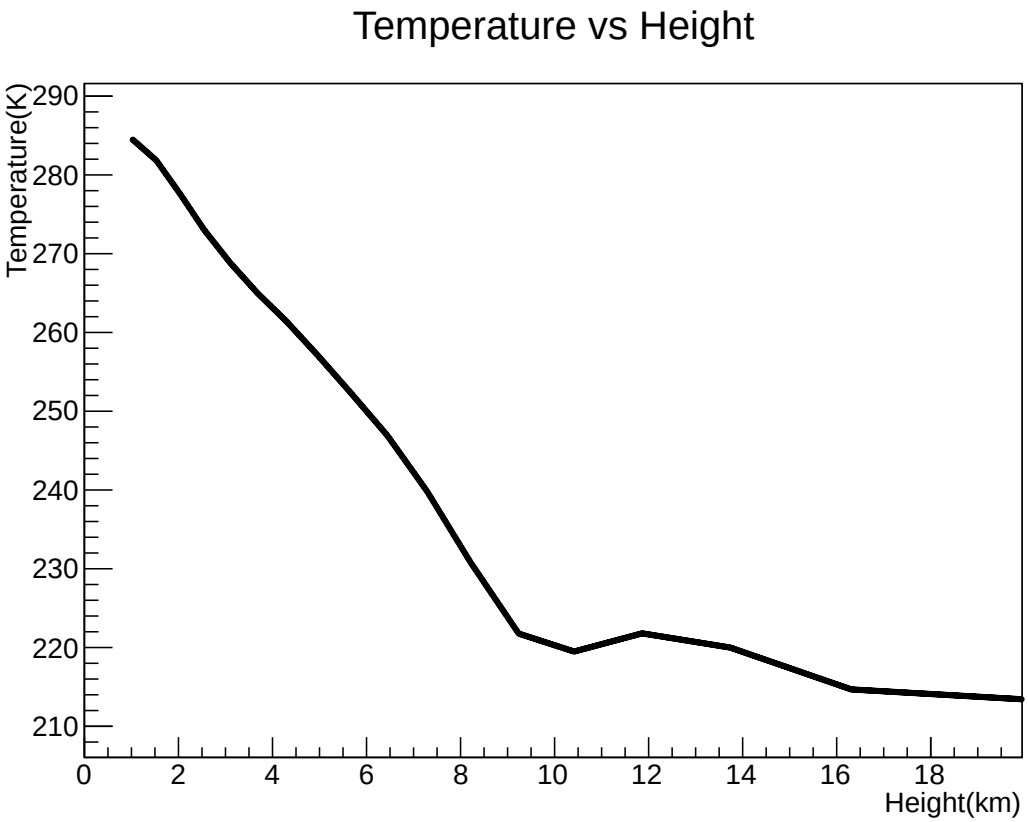

Figure 4.11: Temperature Profile of Earth's Atmosphere at Auger. Date: 17th August, 2012 at 03:00 UTC. Our approximation of linear temperature decrease with height is sound for this example. Height is taken to be above sea level (a.s.l).

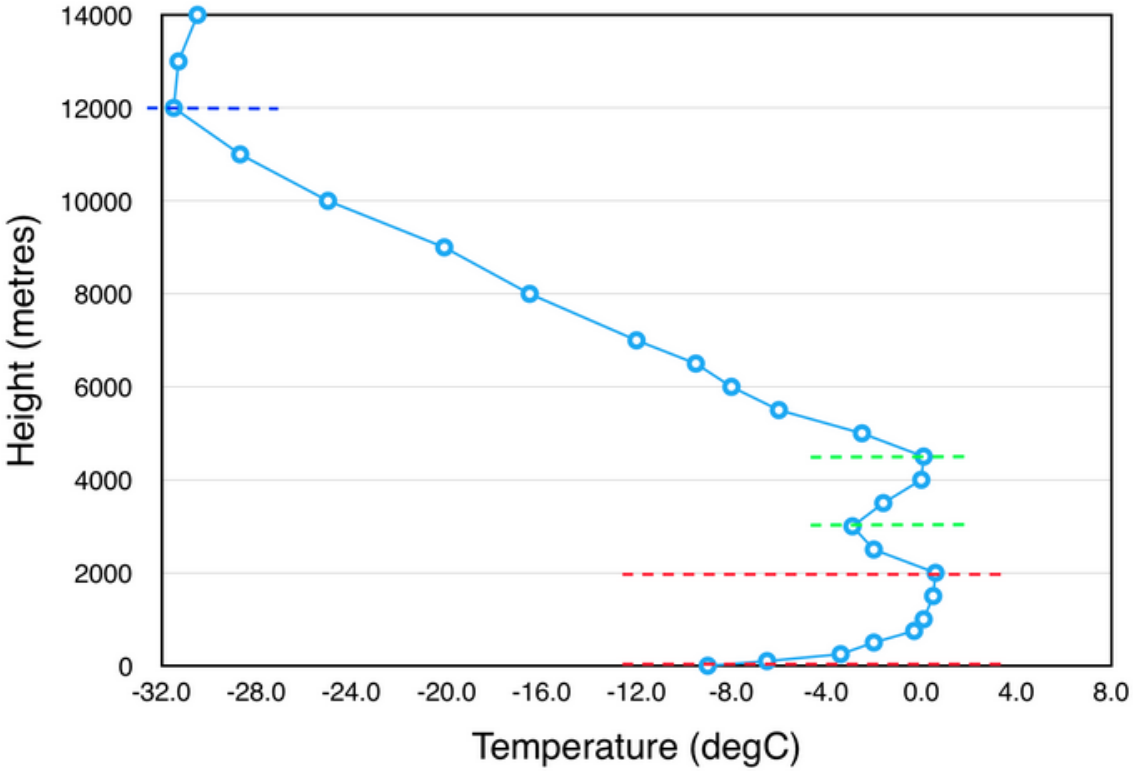

Figure 4.12: Temperature Profile of Earth's Atmosphere at Auger, with two clear temperature inversions present, one below 2000 m and the other between $\sim$ 3000 – 4500 m. In this example atmosphere, a linear fit would do poorly, complicating the analysis. Source: [78]

As can be seen by Figure 4.11, there is an approximately linear trend in the temperature change with height below $\sim$ 10 km. So, a linear fit to this region of the atmosphere is valid. For the fitting routine, a few considerations have to be made in order to get the best possible fit, and hence to best calculate the lapse rate. Since the GDAS grid point is $\sim$ 1.4 km a.s.l (above sea level), any values of temperature and pressure below this height are extrapolations by the GDAS model and may not be accurate representations of the true nature of the atmosphere (since these heights are below the ground level of the Pierre Auger Observatory). Thus, these points cannot be used to determine an accurate LCL. Also, since temperature inversions are prevalent below the boundary layer (especially in winter months), it is best to fit the data between $\sim$ 4 and 10 km - since above 4 km the atmosphere is known as the "free" troposphere. This layer is so called as it is "free" of temperature inversions which are small regions where temperature increases with height due to atmospheric circulation. Fitting below this region where temperature inversions are possible will invalidate the assumption of a linear decrease. The release of latent heat by clouds (as described in section 4.2.1) also changes the rate of decrease of temperature with height (see Figures 4.5 and 4.6), which means a linear lapse rate is no longer accurate. In addition, a tropopause height of 10 km or less (as shown by Figure 4.11) can also affect the lapse rate fit, however this is only a consideration in the winter months (on average the tropopause height at Auger is 12 km as discussed earlier), thus it only affects a small fraction of the results. Figure 4.12 shows an example temperature profile with such an inversion present.

There are three reasons for cutting off the fit at 10 km: firstly, due to the fact

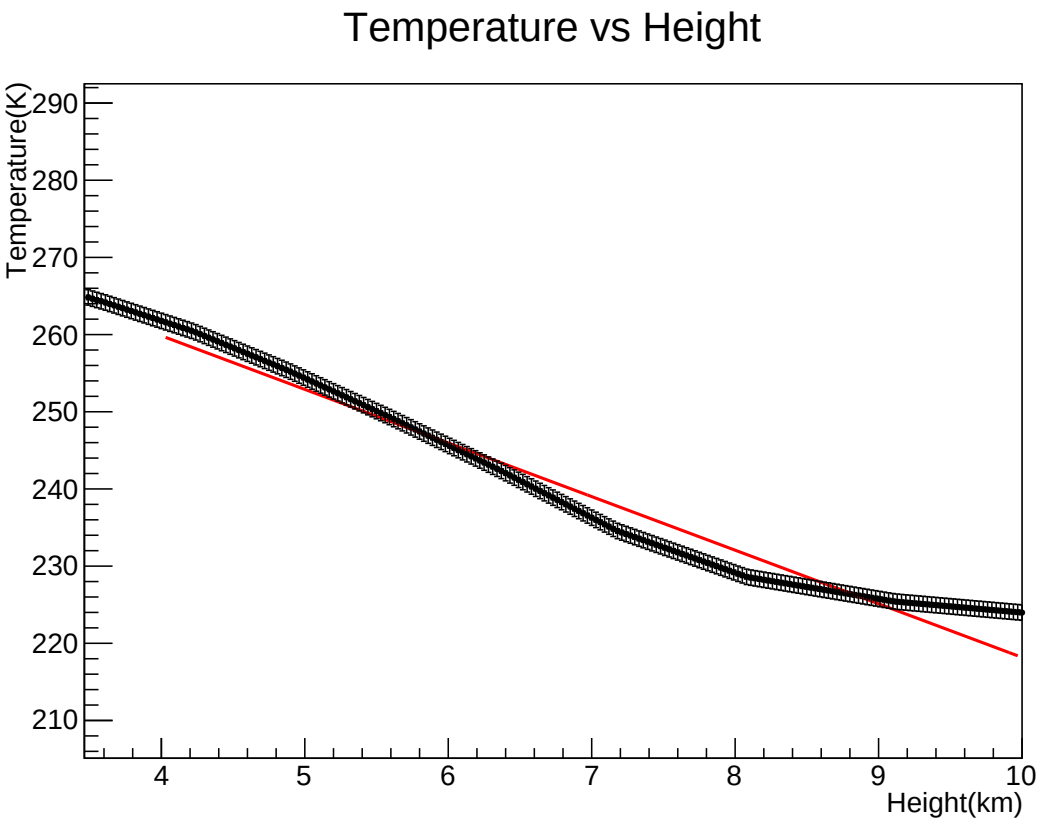

Figure 4.13: Temperature profile of Earth's Atmosphere at Auger, with a straight line fit between 4 and 10 km. The slope of the red fit line is 6.94 (in units of K/km). Height is given as above sea level (a.s.l). Date: 24th August, 2014 at 03:00 UTC

that extensive air showers mainly develop below this height, the atmosphere above 10 km can be ignored in the analysis/reconstruction of cosmic ray events and their development down to the ground (for more information on extensive air showers see section 2.2) for Auger. This is due to 10 km corresponding to an atmospheric depth of $\sim 300\ \mathrm{g/cm^2}$, and extensive air showers with an X_{max} value of $300\mathrm{g/cm^2}$ or lower are too low in energy to be detected by the hybrid detectors of the Pierre Auger Observatory. Secondly, water vapour molecules above this height are ice crystals due to the extremely cold temperatures, and in this case the relative humidity is handled in an entirely separate way (i.e. the assumptions and definitions of the LCL are not applicable in this regime). Lastly, the environmental lapse rate Γ is only linear within the troposphere. Above ~ 10 km, the tropopause begins, before moving into the stratosphere (where the atmospheric temperature begins to increase with height due to the absorption of solar radiation by ozone). This is also shown in Figure 4.12, as the temperature steadily increases in the atmosphere above ~ 12000 m.

Using all of this knowledge, a linear fit is performed and the slope of the fitted line is extracted - this is the Γ value. Figure 4.13 shows an example of this linear fit, used in order to calculate a value of the LCL.

As can be seen from Figure 4.13, there is a flattening of the temperature profile at ~ 8 km due to the saturation of the air - this has the effect of slowing down the decrease in temperature with altitude. Now, by extracting this slope, and calculating all the values (taking height, temperature and vapour pressure - adjusting for sea level), the

LCL (a.s.l.) can be calculated. In the example above, taking ground level values of height (1.42 km as discussed previously), temperature (280.9 K or 7.7° C) and vapour pressure (6.14 hPa), and plugging these values into the equations described in section 4.4, the LCL is calculated to be 2.9 km a.s.l. Once again, ground level of the Pierre Auger Observatory was chosen for calculation of the LCL since the LIDAR and CLF calculate cloud base height from the same reference point. This is done in order to simplify the comparison of the two cloud base heights. In the example here, the LCL (above ground level - a.g.l) is 1.48 km. This LCL value is below the height range over which the lapse rate is fitted (a consideration that is further investigated in Chapter 8).

Now that the LCL can be calculated, a comparison has to be made to the cloud base heights that are measured by the FD LIDARs and the CLF. The hope is to find some evidence of correlation between the values, assuming the height given by the other instruments (LIDAR,CLF) is the true cloud base height. Section 4.4.2 investigates the correlation between the LCL calculated height and the height output from the LIDARs and CLF.

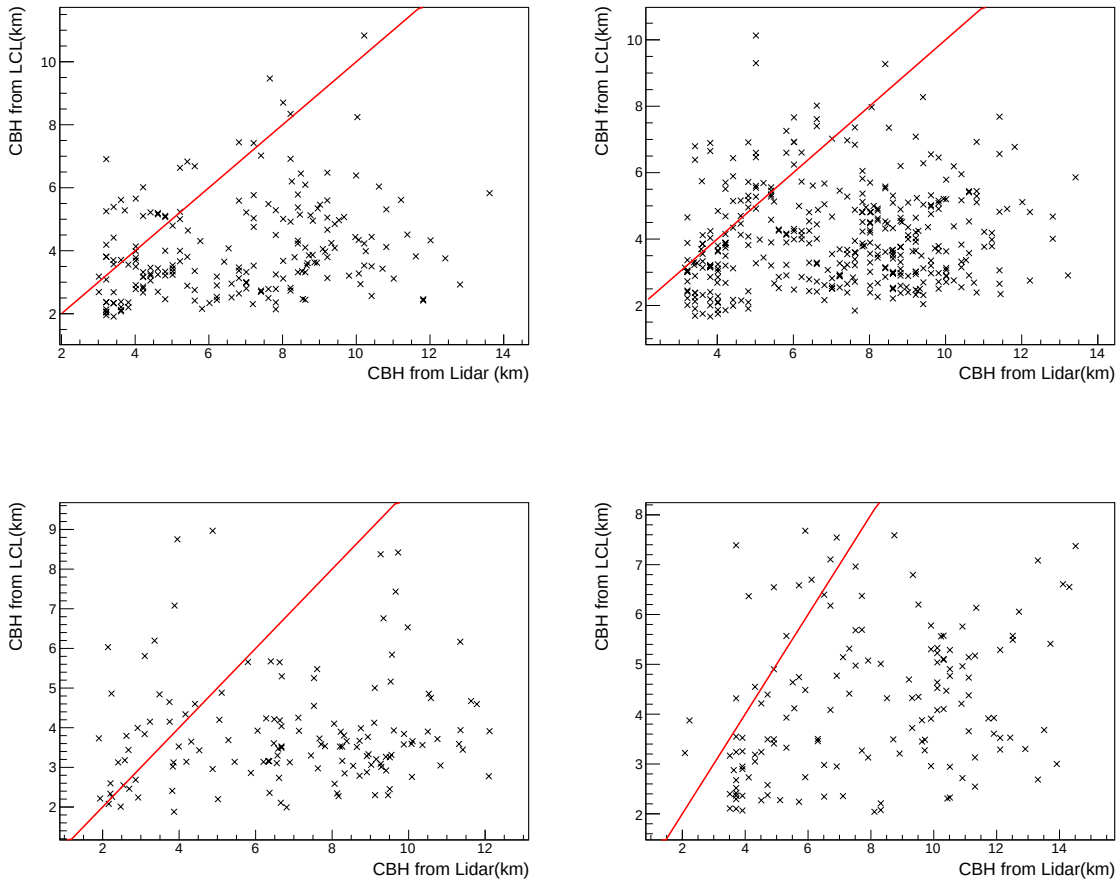

Figure 4.14: A direct comparison between the LCL calculated from GDAS (vertical axis) with the Cloud Base Heights read from the LIDARs (horizontal axis) between 2005 and 2014. The red line corresponds to the two values being equal. All values are given above sea level (in km). Top left shows LIDAR Heights recorded at Los Leones, top right is Los Morados, bottom left is Loma Amarilla, and bottom right is Coihueco.

4.4.2 Correlation between the LCL and the Cloud Base Height

Now that the basis of the LCL is covered, one can begin to find times where either the LIDAR or CLF measured the height of a cloud, and to compare this height to which is expected theoretically from the LCL. Since the LCL is only applicable to parcels of air that are forcibly lifted (before condensing), which is not very common, the expectation is that the LCL will not be able to estimate cloud base heights very well. Figure 4.14 shows the comparison of the LCL to the cloud base height given by the LIDAR at each FD site over a 10 year period.

What can be instantly recognised from Figure 4.14 is that the cloud base height (represented on the horizontal axis) and lifting condensation level (vertical axis) are not well correlated. As explained earlier, given that the lifting condensation level estimation of cloud base height only works for parcels of air that are lifted by force, the correlation between the two parameters (LCL and cloud base height) is not strong. However, there is a distinguishable pattern - the cloud base height is predominantly higher than the lifting condensation level. As a result, the lifting condensation level can be seen as a lower bound on the cloud base height in the atmosphere, as given by

the LIDAR. So why is this the case? Clearly, some cross-checking of the analysis must be done.

In order to calculate the LCL some assumptions needed to be made. Firstly, the temperature and the dew point were assumed to decrease linearly with height in the atmosphere. The temperature follows the environmental lapse rate, whilst the dew point follows the wet adiabatic lapse rate, both of which are assumed to be constant in the troposphere. Of course, the latent heat released by clouds when they form causes the rate of temperature decrease to slow down, an effect that, whilst known, cannot be corrected for in this analysis. This effect is taken into account in the techniques outlined in Chapter 5. In addition, the wet adiabatic lapse rate is not constant with height (as assumed by the LCL), rather it has its own intrinsic height dependence due to its depedence on the amount of water vapour present in the atmosphere [51]. Thus, an uncertainty is introduced into the LCL calculation, namely that a decrease in Γ_w will decrease the value of the condensation level (looking at equation 4.13). Since this linear assumption has failed, the calculation of the LCL changes from what is expected, which partially contributes to the differences in cloud base height shown by Figure 4.14. This is a small effect on the results in Figure 4.14, however the simplifying assumptions for the calculation have broken down.

In order to minimise the effect of the lapse rate assumptions, one can focus on the other two parameters in equation 4.13 - the temperature and the dew point. Using these, one can find the dew point depression (or DPD) which is the difference between the temperature and the dew point at a given pressure/altitude. The comparison of the temperature to the dew point for a typical GDAS profile is shown in Figure 4.15.

Figure 4.16 illustrates how the dew point depression (DPD) changes as a function of height. This is important for the analysis done in the next section.

From the earlier definition of the dew point (i.e. the point where all of the air saturates and water vapour condenses to form clouds), a dew point depression of 0 is defined to be the LCL. However, experimental studies have shown that the dew point depression does not always have to be 0 to define a cloud layer [79]. The idea is to repeat the study of Zhang et al. (see Figure 4.17) for the Auger data, which can be classified into two parts; estimating the base heights of clouds in the atmosphere, as well as using limits on the dew point depression (DPD) to estimate whether the conditions are clear or cloudy.

Using this empirical method of dew point depression decreases the uncertainty about the nature of the atmosphere. Section 4.5 goes into the detailed study done using these ideas, to hopefully estimate a more accurate cloud base height, and be able to correctly predict when clouds occur.

4.5 Thresholds on Dew Point Depression (DPD)

Studies into the correlation between atmospheric measurements of cloud base height (CBH) and cloud base height estimates from LIDAR have been done extensively, most notably by Zhang et al. [2010] [79]. By using empirical estimates of the DPD, there was agreement on the cloud base height to within 500 m between the radiosondes and micropulsed lasers (MPL) 77% of the time. A micropulsed laser is similar in design and function to the LIDAR in use at the Auger observatory. The results of this study are shown in Figure 4.17 [79].

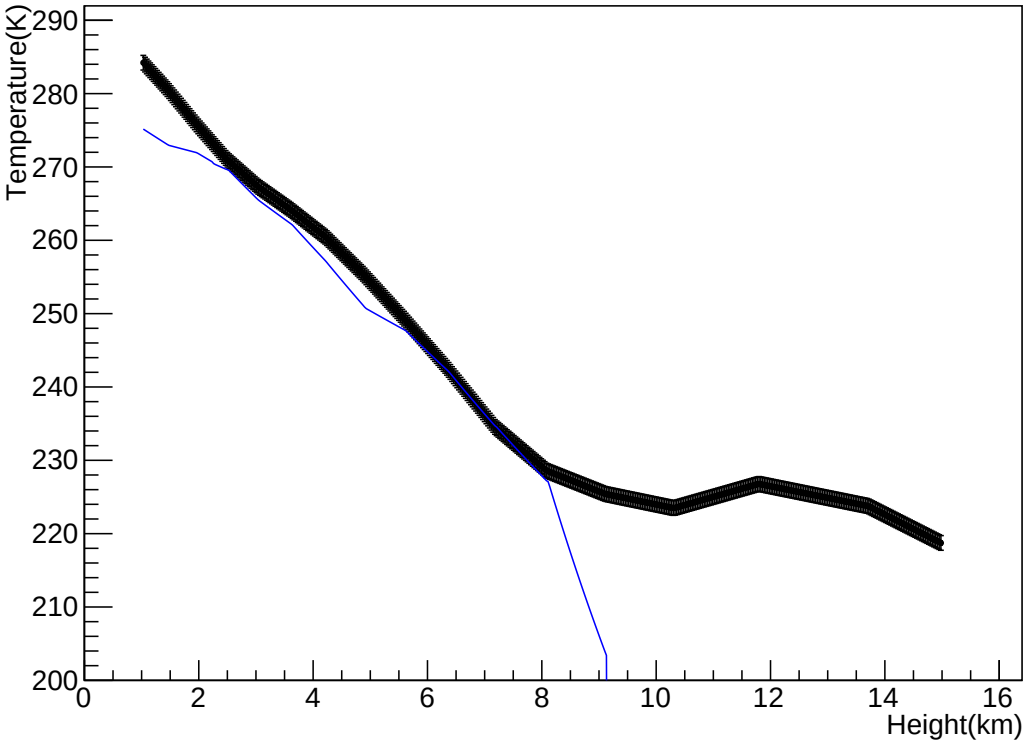

Figure 4.15: Temperature (black line) and dew point (blue line) at all altitudes for a given time from GDAS. Height is given above sea level. Date: 24th August, 2014 at 03:00 UTC

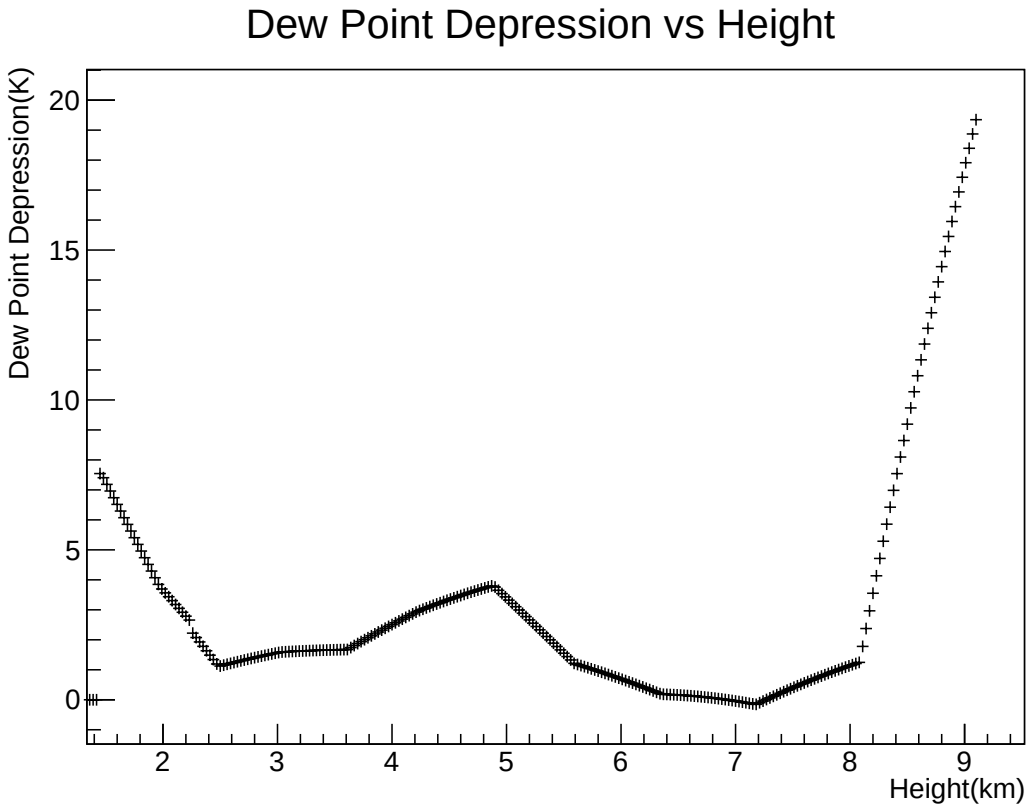

Figure 4.16: Dew Point Depression (DPD) at all altitudes for a given time from GDAS. This value is the difference between temperature and dew point at a given height, and is used to define cloud formation. Height is given above sea level. Date: 24th August, 2014 at 03:00 UTC

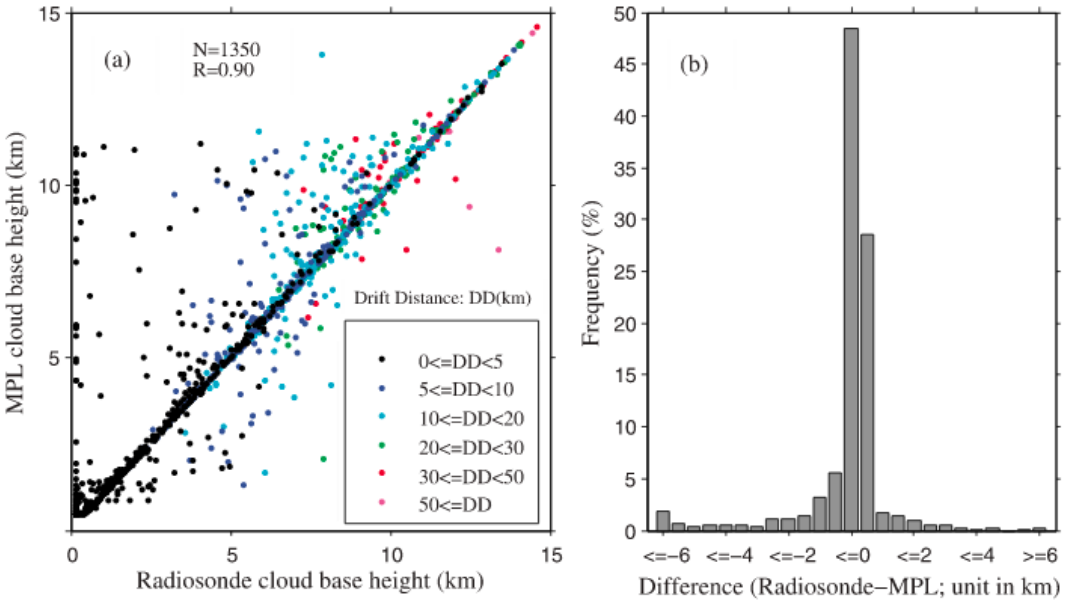

Figure 4.17: Results of the study performed by Zhang et al. in Shouxian, China in 2010 [79]. Left plot: The horizontal axis is the empirically derived Cloud Base Height from radiosonde soundings, and the vertical axis is the measured cloud base height from the Micropulsed Laser (MPL), an equivalent instrument to the LIDAR at Auger. Right Plot: A histogram of the difference between the two CBH measurements, which shows excellent agreement. All values here are taken above sea level.

The idea is to use the methods performed in the Zhang study and apply them to Auger data (in this case, comparing empirical CBH from GDAS profiles using the DPD, with CBH read at the FD LIDAR sites).

There are several conditions that must be considered in order to perform this study. Firstly, the estimate of a cloud base is derived from empirical estimates on the value of the Dew Point Depression (DPD) which is adjusted according to the air temperature. Dew point depression is the difference between the air temperature and the dew point temperature for a given pressure. The criteria for a cloud layer, using the DPD (from Zhang et. al.), are as follows [79]:

$$\begin{aligned} DPD &< 1.7^{\circ}C, T > 0^{\circ}C \\ DPD &< 3.4^{\circ}C, -20 < T < 0^{\circ}C \\ DPD &< 5.2^{\circ}C, T < -20^{\circ}C \end{aligned} \tag{4.18}$$

As shown in Figure 4.14, the LCL (i.e. the point where the DPD is 0) overestimated the cloud base height measured by the LIDAR. As a result, using the criteria defined in equation 4.18 will allow for a lower, more accurate cloud base height to be estimated.

As long as the above criteria are met, one is said to be within a cloud layer [79]. For each temperature/height in the atmosphere, the DPD is calculated. If equation 4.18 holds for a given temperature, then one is still within the same cloud layer. Breaking this threshold at a given temperature means that the cloud top has been reached, and one is no longer within a cloud. This information could be used to determine a cloud's thickness, however this is not of interest in this study. The reason that cloud thickness is not a consideration is due to the fact that base height is the only factor considered

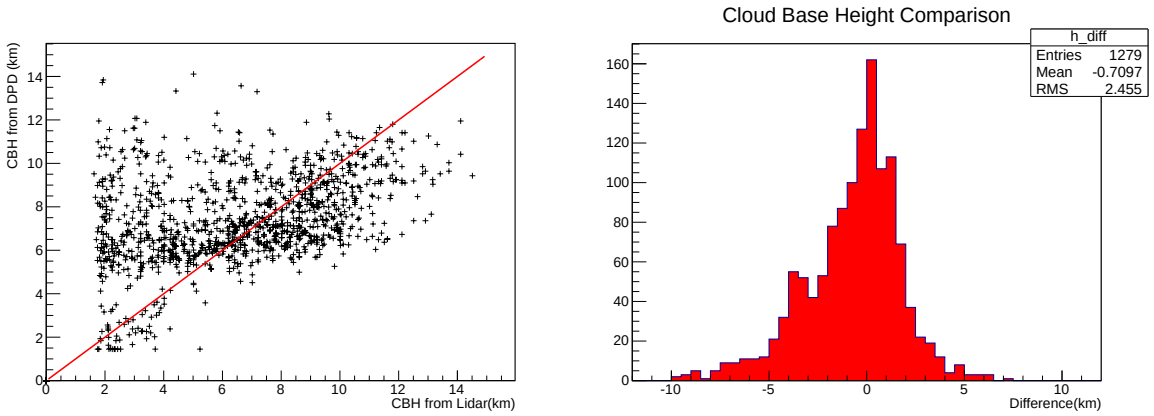

Figure 4.18: Left: Scatter plot of the CBH given by the LIDARs (horizontal axis) compared to the CBH given by the DPD method (vertical axis). A red 1:1 line is shown for comparison of the two cloud base heights. The DPD method has not improved the estimation of cloud base heights since there is a large scatter of the points, and this method still overestimates the cloud base height given by the LIDARs significantly. Right: Histogram showing the difference in cloud base height at Auger between the LIDAR and GDAS (using the DPD) [in km] for all data between 2010 and 2014. The difference is taken as Height (GDAS) - Height (LIDAR). As can be seen, a large spread is present (indicated by a standard deviation of 2.5 km). Similarly, there is a mean offset of 700 m, indicating that the DPD method estimates a lower height on average.

by the Auger cloud quality cuts mentioned in section 3.1. Specifically, there is a quality cut placed on data with regards to the cloud base. If an air shower developed below the cloud base, then the cloud will have no effect on the longitudinal profile reconstructed and that particular event is kept. Cloud thickness has no bearing on this cloud cut and as such is not required to be studied here.

The only consideration of a cloud's thickness is in what defines a cloud layer. The cloud thickness is the difference between Cloud Top Height (CTH) and Cloud Base Height (CBH) based on the criteria in equation 4.18. In order for a cloud layer to be defined, the thickness must be $>$ 30m for low cloud, and $>$ 61.5m for middle/high cloud. Otherwise, it is defined as a moist clear layer [79]. Similarly, any cloud bases that begin below 280 m above ground level (AGL) are flagged as false detections of cloud. The resolution of the laser instruments inhibits this measurement at very low altitudes due to the high degree of scattering of the laser beam [79].

By applying these thresholds to the Auger GDAS data, cloud base heights, top heights, and thicknesses can be estimated across the array. In this instance, the radiosonde data come from GDAS, which is compared to the CBH data from the LIDARs located at the four Fluorescence Detector sites. Applying these thresholds and plotting the difference in the cloud base height (LIDAR - GDAS) is shown in Figure 4.18.

As shown in Figure 4.18, whilst the DPD method has reasonable accuracy in guessing cloudy/clear conditions, it has still not improved the estimation of cloud base heights in comparison to the LIDARS. As a result, there is room to improve the accuracy of this empirical dew point depression method in measuring cloud base heights by utilising statistical optimisation techniques. The application of one such technique is outlined in section 4.6.

4.6 Optimisation of the Thresholds

The current estimated thresholds for DPD (1.7, 3.4 and 5.2 °C respectively) very accurately estimate cloud base heights in China with two separate instruments (up to 92% accuracy) [79]. However, those thresholds were empirically derived for the atmosphere above Shouxian, China. The optimum thresholds for the Pierre Auger Observatory could potentially be different, due to the differing climate/atmosphere present (the atmosphere at the observatory is much drier in comparison). In order to test this, a confusion matrix was created, which measures the accuracy of cloud detection from a predicted method (the DPD thresholds) compared to the "truth" (the LIDAR in this case). For the original thresholds as shown previously, the resulting confusion matrix (for Auger data) is shown in table 4.1.

n=2242	Lidar Cloudy	Lidar Clear	Total
DPD Cloudy	1085 (TP)	500 (FP)	1585
DPD Clear	211 (FN)	446 (TN)	657
Total	1296	946	2242

Table 4.1: Confusion Matrix on the dew point depression thresholds as defined by Zhang [79].

In table 4.1, the "Truth" is that a particular method detected cloud. For the LIDAR, it is whether or not a cloud base height was recorded into the database, and for the DPD, whether any of the 3 thresholds from equation 4.18 were met. False indicates that the LIDAR stated clear conditions, or that none of the DPD thresholds were satisfied. In this way, there are 4 subcategories into which the matrix is split; True Positives (TP for short, which is when both methods detect a cloud), True Negatives (TN; when both methods say clear conditions), False Negatives (FN; when the LIDAR says a cloud is there but the DPD thresholds are not met) and False Positives (FP; when the LIDAR says clear conditions but the DPD thresholds are met). Additional parameters can be calculated from this table, namely accuracy, precision, True Positive Rate (TPR) (also known as sensitivity) and False Positive Rate (FPR). These four parameters are summarised in Equations 4.19, 4.20, 4.21 and 4.22 respectively. Acuuracy is a measure of how often the two methods agree on the conditions (either clear or cloudy).

$$Accuracy = \frac{TP + TN}{Total} = \frac{1085 + 446}{2242} = 68.3\% \tag{4.19}$$

The precision is how often the threshold method correctly predicts a cloud layer.

$$Precision = \frac{TP}{TP + FP} = \frac{1085}{1085 + 500} = 68.5\% \tag{4.20}$$

The false positive rate is how often cloud layers are predicted by the DPD method, when the LIDAR returns clear conditions.

$$FPR = \frac{FP}{FP + TN} = \frac{500}{500 + 446} = 52.9\% \tag{4.21}$$

The true positive rate is how often cloud layers are predicted from the DPD thresholds when the LIDAR returns cloudy.

$$TPR = \frac{TP}{TP + FN} = \frac{1085}{1085 + 211} = 83.7\% \tag{4.22}$$

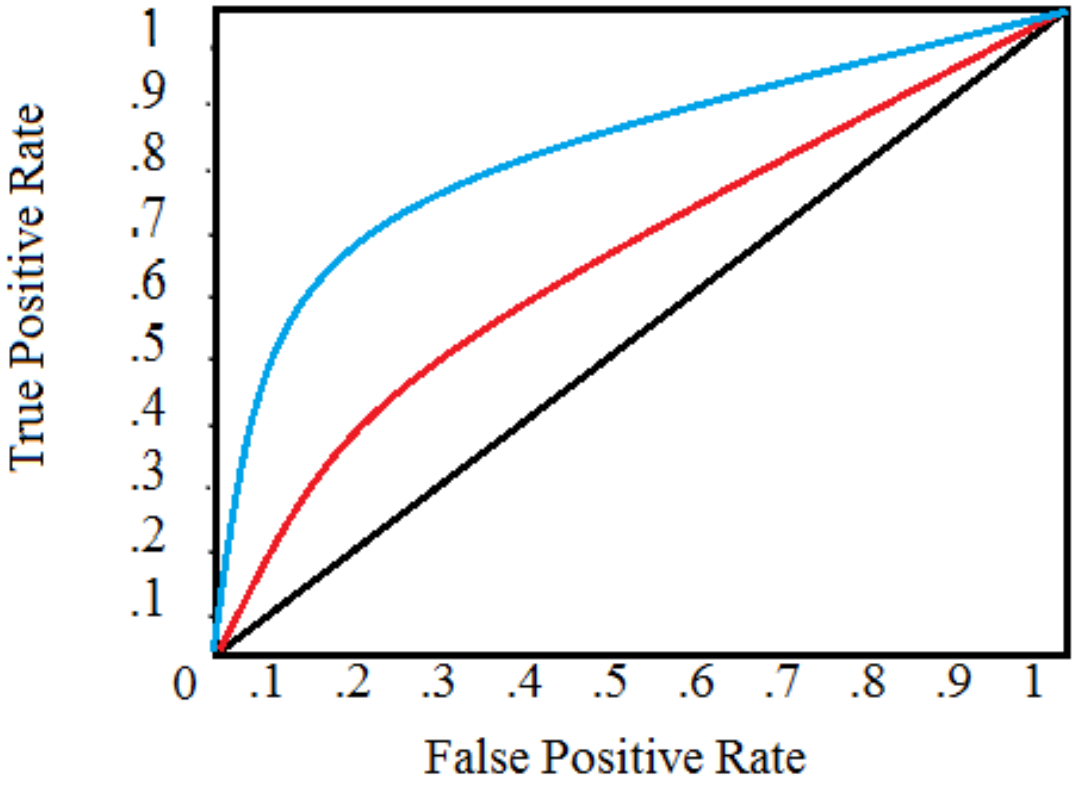

Figure 4.19: Typical receiver operating characteristic (ROC) schematic curve for two different model tests (shown by the red and blue lines). The 1:1 line in black is what would follow for a random guess of the conditions - being above this line means the model has improved upon a random guess method [80].

The accuracy value of 68.3% indicates that the thresholds applied here (from equation 4.18) can be improved. The goal, by now optimising these thresholds, is to improve the overall accuracy of the classification of conditions by finding the values that produce the best overall accuracy of cloud prediction. This will hopefully also enable this analysis to be able to better resolve cloud base heights and make the predictions of those much more accurate. To do so, one wishes to have as high a TPR as possible, while at the same time, having an FPR kept as low as possible. Due to the LIDAR only viewing a smaller patch of sky, there is the possibility of high rates of clear conditions being reported by the LIDAR when there is actually cloud present. This could potentially overestimate the false positive rate. However, there is a balance between the two parameters (TPR and FPR) that can be achieved for the best choice of DPD thresholds. To do this practically, an ROC (Receiver Operating Characteristic) curve is used, which is simply true positive rate against false positive rate for a given set of thresholds. This ROC curve is shown schematically by Figure 4.19 [80].

From Figure 4.19 the best measure of accuracy is to be as far above the random guess line as possible. The accuracy is measured by calculating the area under the curve in the chosen model. In order to achieve 100% accuracy, the area under the curve would be 1. There is a turning point at which the maximum distance from the 1:1 line is achieved. This is the point at which there is maximum accuracy of the model being tested. The aim is to find this maximum accuracy and the dew point depression

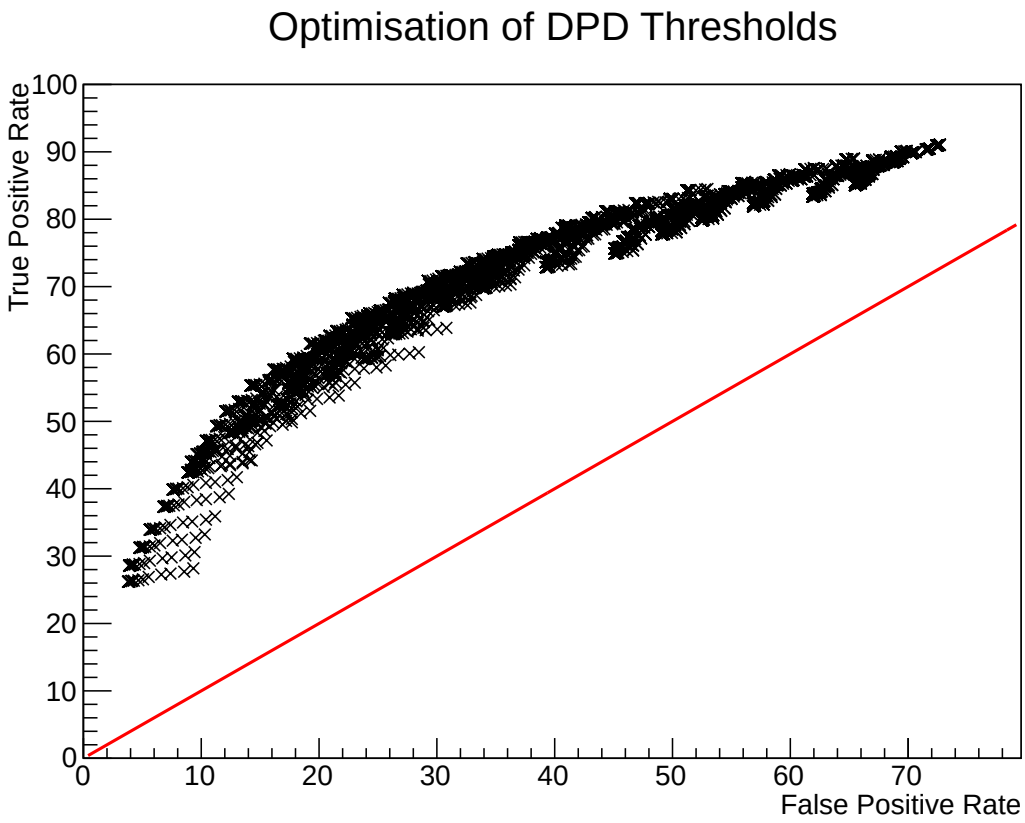

Figure 4.20: Test for several sets of DPD threshold values comparing the DPD and LIDAR methods. The general trend of each data point (black crosses) shows that the DPD method increases the accuracy of estimating cloudy conditions. This is due to all points (marked as crosses) lying above the 1:1 random guess line (the red line indicated here). Each black cross corresponds to a particular set of threshold values being tested.

thresholds that correspond to this.

Firstly, one can test the empirical DPD method on a wide range of threshold values, by taking all three thresholds from 0-7 °C in steps of 0.5°C, in order to narrow the range over which this accuracy can be achieved. The testing is done by splitting the data into two halves - the GDAS data used for training the thresholds for optimisation is taken at 00,06,12 and 18 UTC, whilst the test data (checking that the optimisation gave a better result) is taken at 03, 09, 15 and 21 UTC. These times are chosen since the GDAS model is calculated every 3 hours at regular intervals of these times, thus this made the testing easier (for more on the GDAS model see section 3.5.2). The test data is what was run on the previous confusion matrix (shown in table 4.1), giving the accuracy as calculated in equation 4.19. The results of the optimisation test on the training data is shown by the ROC curve in Figure 4.20.

As can be seen from Figure 4.20, the DPD method has validity in the estimation of cloud layers, as all points lie above the 1:1 line. Thus, this optimisation technique is performing better than a random guess of the conditions would. Next, the threshold ranges corresponding to the maximum difference between the black and red curves (from Figure 4.20) will be found, before the threshold values are made more precise (steps of 0.2 in DPD rather than 0.5) to narrow down the optimum threshold values. The maximum distance l between the black line and red line (see Figure 4.20) can be

found by the manipulation of vectors. The distance l is the length of a perpendicular line extending from the black line, meeting the red. The derivation of this distance is given in Appendix B, as given by equation 4.23.

$$l = \sqrt{2}(\frac{TPR - FPR}{2}) \tag{4.23}$$

The goal is to use this distance to find the best balance between reducing the overall false positive rate and increasing the true positive rate. Hopefully, this yields thresholds that can improve the overall accuracy of the DPD model. Performing this over a range of threshold combinations - in larger steps at first before narrowing down the range and taking smaller steps - hopefully leads to some new thresholds that can predict atmospheric conditions (clear or cloudy) at the Pierre Auger Observatory more accurately.

Starting from Figure 4.20, the distance l was maximised with values of 1, 4 and 1.5°C for the three thresholds. Now, by looking around those values and shrinking the step size to 0.1 (rather than 0.5), and repeating the steps above, the thresholds that give the largest distance l are found to be 1.7, 3.8 and 1.3°C respectively for the three temperature regimes (shown in equation 4.18). Creating the confusion matrix using these thresholds is shown in table 4.2.

n=2242	LIDAR Cloudy	LIDAR Clear	Total
DPD Cloudy	815 (TP)	202 (FP)	1017
DPD Clear	481 (FN)	744 (TN)	1225
Total	1296	946	2242

Table 4.2: Confusion Matrix on the optimised thresholds.

When looking at the overall accuracy on the optimised thresholds, these new thresholds achieve a value given by equation 4.24.

$$Accuracy = \frac{TP + TN}{N} = \frac{815 + 744}{2242} = 69.5\% \tag{4.24}$$

In equation 4.24 N is the total number of cases (i.e. N = TP + TN + FP + FN). It is shown that the accuracy has increased slightly, showing that the method has improved overall in determining if cloud is present (by $\sim 1\%$).

The next step in the analysis is to compare the cloud base heights recorded by the DPD method and by the LIDAR (originally shown as a comparison between the LCL and LIDAR as in Figure 4.18) for the optimised thresholds. These new differences are shown in Figure 4.21.

Running the optimised thresholds on the testing data (shown by equation 4.24 and Figure 4.21) produces a slightly higher overall accuracy, but also a larger spread in the difference between cloud base heights. The mean difference of $\sim 650\,\text{m}$ illustrates that the dew point depression method is still slightly overestimating the cloud base height on average, compared to the LIDAR. However, the left-hand plot in Figure 4.21 illustrates improvement in the estimation of cloud base height for the DPD method (compared to Figure 4.18), particularly when looking at low altitudes ($\sim 2\,\text{km}$) and high altitudes ($\sim 11\,\text{km}$). When comparing to the results from Zhang et al., it is clear that the technique has not worked as well as first hoped. The reasons for this need to be found by cross-checking all the steps along the way. These will be outlined in the following sections.

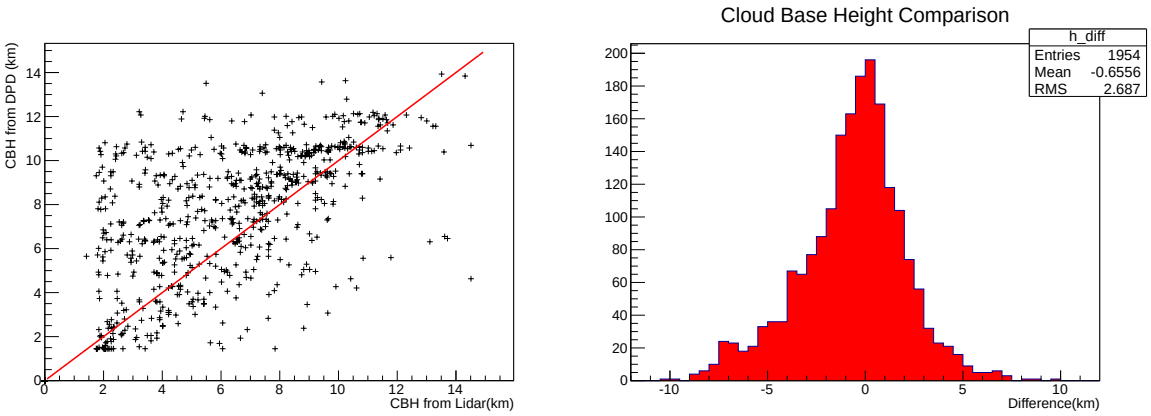

Figure 4.21: Left: Scatter plot showing the CBH given by the LIDARs (horizontal axis) compared to the CBH given by the DPD method (vertical axis) for the optimised thresholds. A red 1:1 line is shown for comparison of the two cloud base heights. There is improvement to the estimation of cloud base height for the DPD method, particularly at low altitudes ($\sim 2\,\text{km}$) and high altitudes ($\sim 11\,\text{km}$). Right: Histogram showing the difference in Cloud Base Height between the LIDAR and DPD method (in km) for the optimised thresholds. A larger spread is present (2.7 km compared to 2.5 km from Figure 4.18).

4.6.1 Error Analysis

The accuracy of this empirical model has uncertainties associated with it, which can be calculated. Since the distribution of cloudy/clear conditions is a discrete count, a Poisson distribution of values can be assumed. For such a distribution, the error σ is given by $\sqrt{N}$, where N is the number of measurements taken. Given the formula for accuracy (equation 4.24), the error σ_A^2 can be calculated as given in equation 4.25.

$$\sigma_A^2 = (\frac{\partial A}{\partial(TP)}\sigma_{TP})^2 + (\frac{\partial A}{\partial(TN)}\sigma_{TN})^2 \tag{4.25}$$

where the partial derivatives are given by

$$\frac{\partial A}{\partial(TP)} = \frac{1}{N} = \frac{\partial A}{\partial(TN)} \tag{4.26}$$

and the errors σ_{TP} and σ_{TN} are given by equations 4.27 and 4.28 respectively.

$$\sigma_{TP} = \sqrt{TP} \tag{4.27}$$

$$\sigma_{TN} = \sqrt{TN} \tag{4.28}$$

Plugging these errors into equation 4.25, one finds the error on the accuracy to be;

$$\sigma_A^2 = (\frac{1}{N}\sqrt{TP})^2 + (\frac{1}{N}\sqrt{TN})^2 = \frac{1}{N^2}(TP + TN) \tag{4.29}$$

As a result the error in the accuracy σ_A is

$$\sigma_A = \pm\frac{\sqrt{TP + TN}}{N} \tag{4.30}$$

Applying equation 4.30 to both confusion matrices given previously yields the following error estimates (for the original and optimised thresholds respectively).

$$\sigma_A = \pm\frac{\sqrt{1085+446}}{2242} = \pm 1.7\% \tag{4.31}$$

$$\sigma_A = \pm\frac{\sqrt{815+744}}{2242} = \pm 1.8\% \tag{4.32}$$

It has been shown that the optimisation technique has improved the overall accuracy of discriminating between cloudy and clear conditions. However, the results are still not perfect; for example, the number of false negative cases (the LIDAR detects cloud but clear conditions are predicted by the DPD thresholds) is still rather high, which can cause problems in air shower reconstruction. This is due to a false detection of cloud leading to the cloud quality cuts (as described in section 3.1.2) removing many more air shower events than would be removed otherwise. As a result, reducing these false detections is extremely important. An investigation into GDAS reveals why such issues with false detection occur. This is outlined in section 4.7.

Lastly, the results also show that being able to better detect clouds does not reduce the errors present in attempting to estimate cloud base heights, since the mean difference in cloud base height between the DPD method and the LIDAR was not significantly improved upon optimisation of the DPD thresholds. Chapter 8 illustrates a technique used to better estimate cloud base heights in the atmosphere.

4.7 Comparison of GDAS and Radiosonde launches

Radiosondes (defined in section 3.5.2) are another means of measuring atmospheric parameters as a function of height. Since radiosonde launches at Auger were only done sparingly (and completely discontinued after 2009), comparisons between GDAS and radiosondes is best done by using data from periodic radiosonde launches done at Adelaide International Airport (Airport Code 94672). Adelaide is a similar latitude to Malargue, with a similar climate, thus making it a good candidate for comparison to Auger. An example of a comparison of the GDAS model to its corresponding radiosonde launch at Adelaide is shown by Figure 4.22.

Noticeably from Figure 4.22, while the two sets of profiles (red with black for temperature and blue with green for dew point) generally agree well, some structural details can be observed. Firstly, the black curve (GDAS temperature profile) is a smoothed version of the radiosonde temperature profile (red line). This smoothing effect is a major reason why small-scale structures (such as the threshold trigger for cloud) are harder to identify in GDAS profiles. This serves to reduce the effect caused by the release of latent heat, since the characteristic bump seen in radiosonde profiles is suppressed by this smoothing. In any case, the temperature difference of a cloud caused by the release of latent heat is not able to be corrected for here (see Chapter 5 for a correction). This can cause inaccuracies in cloud base heights measured via the LCL method.

The second major difference between GDAS profiles and radiosonde profiles occurs between the two dew point profiles (blue for the radiosonde and green for GDAS). This is especially true below the freezing point of water (273 K). Across an entire $1^\circ \times 1^\circ$ grid, the dew point can be quite variable due to the turbulent nature of the atmosphere. The

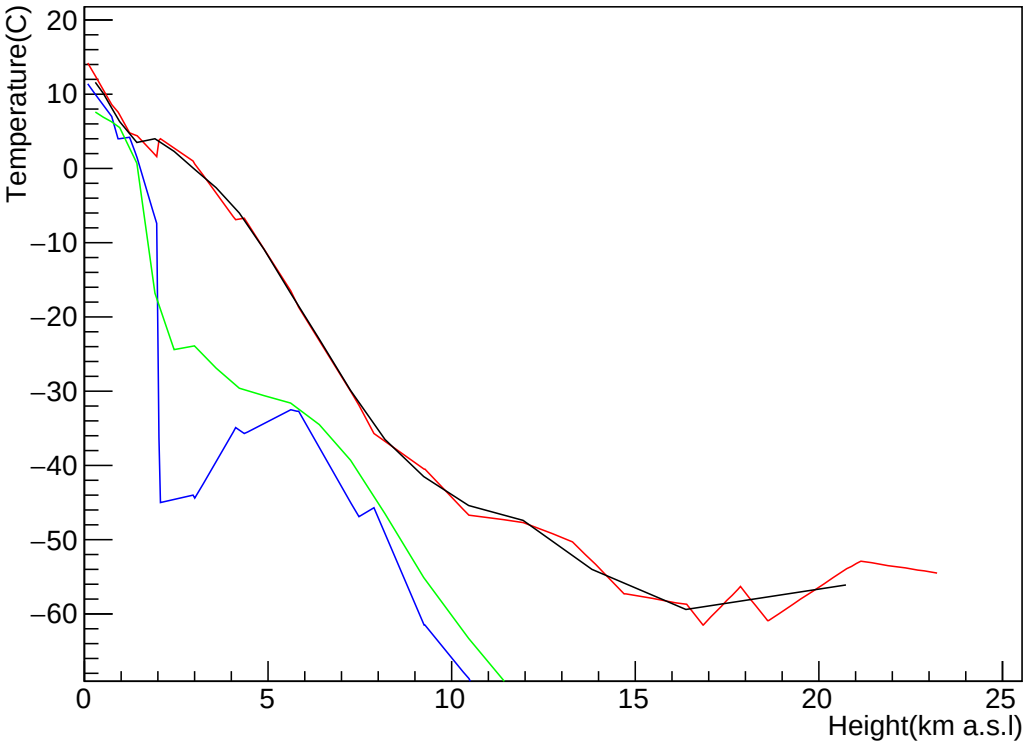

Figure 4.22: GDAS comparison to a radiosonde launch at Adelaide Airport. The various colours show the GDAS and radiosonde Air Temperature and Dew Point temperature profiles. Here, red is the Radiosonde air temperature profile, black is the GDAS air temperature profile, blue is the radiosonde dew point temperature profile, and green is the GDAS dew point temperature profile. The GDAS model is taken at the nearest latitude/longitude grid point to the Adelaide Airport, (located roughly 56 km from the airport). Date: 12/09/2017 12:00 UTC

grid point in question near Adelaide Airport is located $\sim 56\,\text{km}$ away, which can lead to large differences in the dew point temperature and the air temperature. Another possible explanation for the discrepancy is that the vapour pressure e can be more variable depending on the location on Earth or the time of year. The smoothing effect of GDAS temperatures is also seen in the dew point profiles, which may cause large differences in the measured dew point depression.

Some further investigation into GDAS is required to understand what other uncertainties are at play to influence the spread seen in the cloud base height measurements.

4.8 Further GDAS Analysis

The overarching assumption of this GDAS model is uniformity across each degree by degree square across the globe [37]. For the Pierre Auger observatory, the grid point used is located at 35° S latitude and 69° W longitude. However, the READY product website also provides 0.5×0.5 degree precision on GDAS data [81]. As a result, data can be taken at adjacent grid points in order to compare to the original point - the assumption is that they should produce the same results in terms of atmospheric parameters (temperature and dew point most importantly). Taking data from the point 35.5° S, 69.5° W and comparing to the original point for a given time, gives results as shown by Figure 4.23.

There is a general agreement between the two grid points, as is assumed in the Auger analysis. However, some small differences appear. For low temperatures ($T < -20°\text{C}$), the two dew point lines (green and blue curves in Figure 4.23) are separated by about 10°C- suggesting a difference between the two grid points. As such, can the assumption of uniformity still be valid? Differences in the two grid points may explain some of the discrepancy in cloud base height comparisons, since many of the cloud base height measurements take place far away from the located GDAS grid point (especially when comparing heights for the Los Leones, Los Morados and Coihueco FD LIDARs). The assumption of uniformity can be a good first guess at the conditions, however, to increase the accuracy of cloud base height measurements a more refined method is required. One such method to do this is outlined in Chapter 8, utilising the infrared cloud cameras (see section 3.5.8).

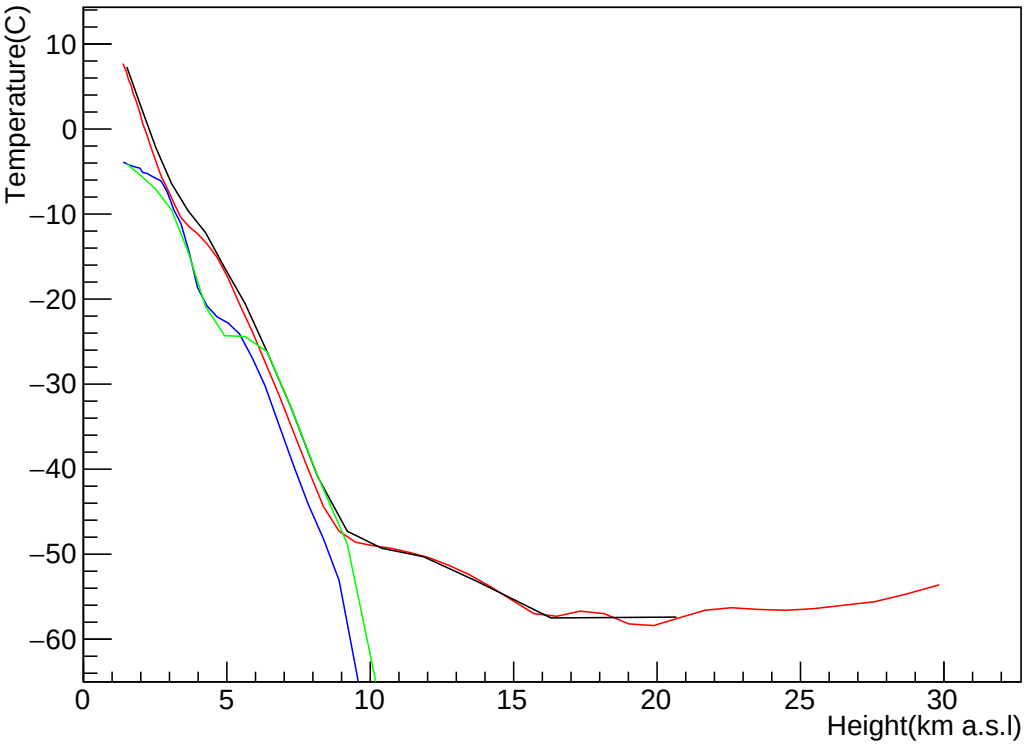

Figure 4.23: GDAS comparison of the full degree and half degree grid points at Auger. The various colours show the GDAS Air Temperature and Dew Point Temperature profiles (Black and Green for the full degree air temperature and dew point temperature, and Red and Blue for the half degree air temperature and dew point temperature respectively). Date: 12/09/2017 00:00 UTC.

4.9 Summary

To summarise this study, the resolution of the GDAS model is not sufficient in order to accurately determine the cloud base height (the spread in the results is too large - refer to Figure 4.18) by comparisons with cloud base heights given by the LIDAR using the two methods described previously (LCL and DPD). The results shown in this chapter display that solely relying on GDAS in order to measure cloud base heights will not be successful, however the atmospheric parameters that GDAS models (most notably temperature and relative humidity/vapour pressure), can be utilised in other cloud height estimation techniques. Chapter 8 builds upon this initial GDAS study, by utilising the temperature profile data in order to estimate cloud heights for clouds in the field of view of the infrared cloud cameras (see section 3.5.8). The effect caused by the release of latent heat by clouds when water vapour condenses can cause a discrepancy in the calculated lapse rate (hence cloud base height). This is further investigated and corrected for in Chapter 8.

There is one potential use of the dew point depression threshold method, and that is to estimate if the conditions above the array are cloudy or clear. Tests on the accuracy of this estimate reveal an overall success rate of $69 \pm 2\%$. Whilst there is merit to this method, more precise measurements are required to better validate the results, and apply them successfully to air shower reconstruction. In any case, this method can be used in the absence of other forms of cloud detection, to check atmospheric conditions and gauge whether extensive air shower events were likely to have been affected by the presence of cloud.

Chapter 5

Analysis of Cloud Camera Data and Implications

Infrared (IR) cloud cameras have been installed at the Pierre Auger Observatory to indicate the presence of cloud in the field of view of the Fluorescence Detectors.

Here, the focus will be on the operation and output of the IR cameras at the Auger site; namely the various corrections applied to the data and the overall processing steps required to prepare and store the cloud camera data for further analysis.

Previous work has been done to get the cloud cameras operational, and to calibrate and correct for the anomalies within cloud camera images [48]. Much of this work is summarised in the following sections.

The main focus of this chapter is to build upon the work that has already been done, improving the overall quality of cloud camera analysis for use in extensive air shower reconstruction, as well as to summarise all of the work that goes into creating and maintaining the cloud database. For more information on the use of these data in air shower reconstruction, see section 3.5.9.

5.1 Calibration of IR Cameras at Auger

Raw images obtained from the cameras at Auger are processed into an infrared colour band image. This does not always produce a sharp image as demonstrated by Figure 5.1

As can be seen by Figure 5.1, the image has not been properly rendered in order to view cloud information. To combat this, several calibration steps are applied to the image. This is done by applying two calibration corrections; namely a flat-fielding correction, and a temperature calibration [48]. The details of these corrections will be outlined in the following sections.

5.2 Flat Fielding Correction

One of the first steps in the image processing of the cloud cameras is to perform what is known as a flat field correction. A flat-field is the response of a camera to a uniform source of illumination [48] [82]. A vertical clear sky image is chosen as the reference since the infrared brightness of a clear sky remains fairly constant at small zenith

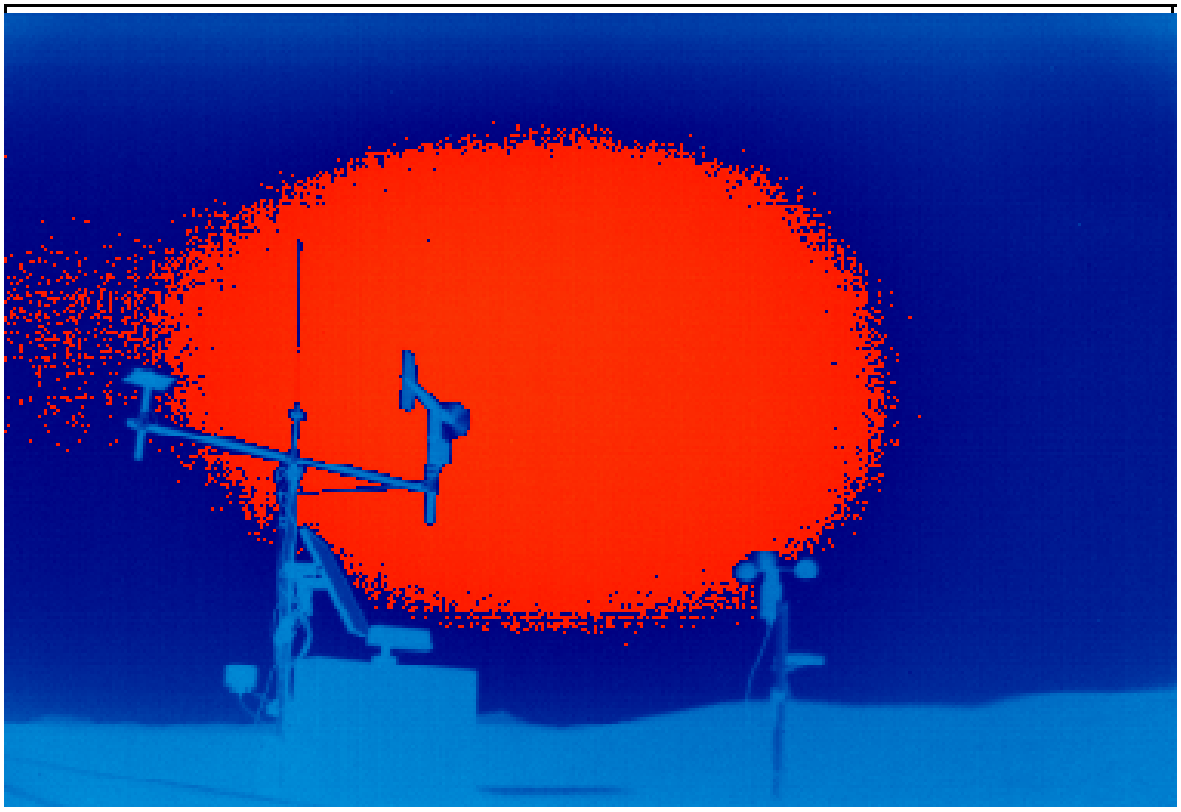

Figure 5.1: Processed image from the cloud camera at the Coihueco FD site. The image is not flat as evidenced by the various colours present in the image, even though the sky is clear. This must be corrected for. Date: 17/01/2017 01:02 UTC

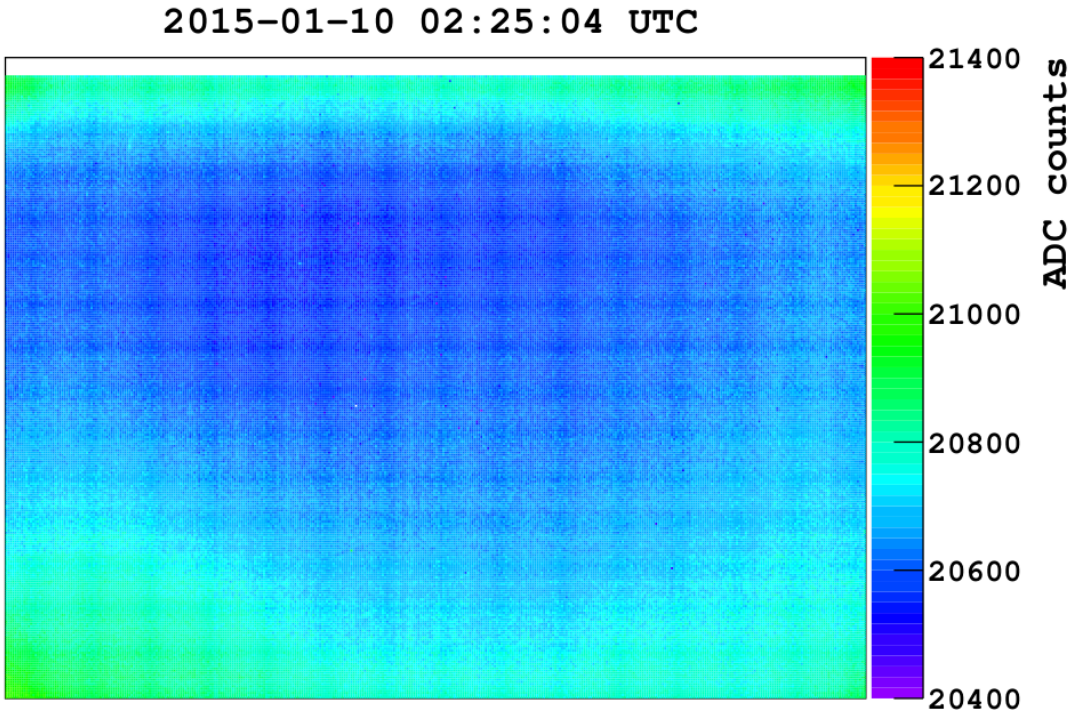

Figure 5.2: An example of the non-uniformity in the brightness of a cloud camera image away from the centre [48].

angles [48]. As a result, the thermal radiation emitted from a clear night sky in the vertical direction provides a uniform source of illumination [48].

The flat-field correction is then done by calibrating every pixel's digital signal (measured in ADC counts) for a vertical clear sky image to be equal to a central region of the camera. This region is chosen to be a circle centred on the middle of the image, with a 2° radius [48].

There are three effects to take into consideration when creating a flat image. Firstly, vignetting is a characteristic of many cameras, which is the reduction in brightness of an image further away from the image centre [83]. For the cloud cameras at Auger, the vignetting effect is known as pixel vignetting, characteristic of these types of cameras. Pixels in the center of the camera receive infrared radiation at an angle which is perpendicular to the plane of the camera, whilst pixels on the edge receive radiation at a slightly tilted angle. As a result, the pixels on the edges will receive slightly less radiation compared to the center, causing pixel vignetting.

Ideally, if the correction is applied, this would result in a flat image, where each pixel in the camera has the same intensity [48]. However, variations in pixel-to-pixel sensitivity produce image artefacts [48]. In addition, the cameras view the sky through a window. The window itself has non-zero emissivity, which means it emits thermal radiation. This increases the brightness of the infrared images [48]. Flat-field images are required to compensate for all these effects. Figure 5.2 shows an example of a non-flat image which must be corrected for.

Using Kirchhoff's Law for thermal radiation (wavelength specific emission and absorption of radiation as described in section 4.2), it can be shown that the average temperature difference between the centre and edge of a typical cloud camera image is $\sim 3\,\mathrm{K}$ [48]. Given the thresholds that are set on cloud in our camera images (more on

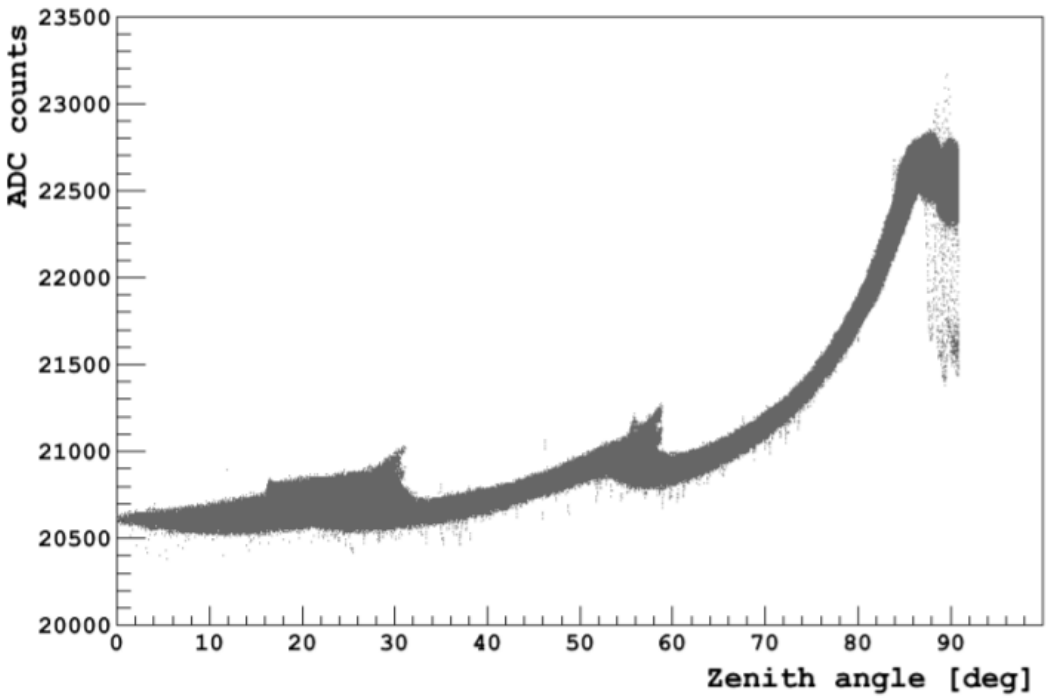

Figure 5.3: ADC counts (brightness) for all pixels of the camera as a function of zenith angle. This should form a smooth curve. However, due to these brighter edges of the images (seen at $\sim 30°$ and $60°$ zenith at the boundary of neighbouring images), bumps in the profile are created. The flat-field correction removes this effect, which will be shown later in this section [48].

this in Section 5.4) is of the same approximate order, this 3 K temperature difference is potentially an issue in the processing and analysis of cloud camera images. Not correcting for this effect can introduce false positive cloud detections, which causes the reconstructed dataset (see section 3.1) to be much smaller, with lower statistics. Thus, it is imperative that this difference is corrected for.

The final point is that the flat field correction becomes very important when stitching several images together (this process will be outlined in section 5.4). As an example, consider Figure 5.3. Due to the increased brightness away from the image centre, the sequence of images (used for a mosaic) causes adjacent images to not transition smoothly in brightness.

Each of these image artefacts can lead to false detection of clouds, as a result, we wish to remove these features for better quality cloud images and subsequent analysis. Section 5.2.1 will go into detail about how these effects are corrected for at Auger.

5.2.1 Correction Method

Removing these brightness features is desirable in order to obtain the best quality images from the cloud cameras for analysis. As noted above, clear sky brightness within the spectral range of the cloud cameras varies slowly for zenith angles less than 30°. Thus, vertical full sky images (more on this in section 5.4) of a clear sky are used for the flat-field correction as this is a good approximation of a uniform source of radiation. This is evident in looking at Figure 5.3, as all pixels in the 0-30° zenith

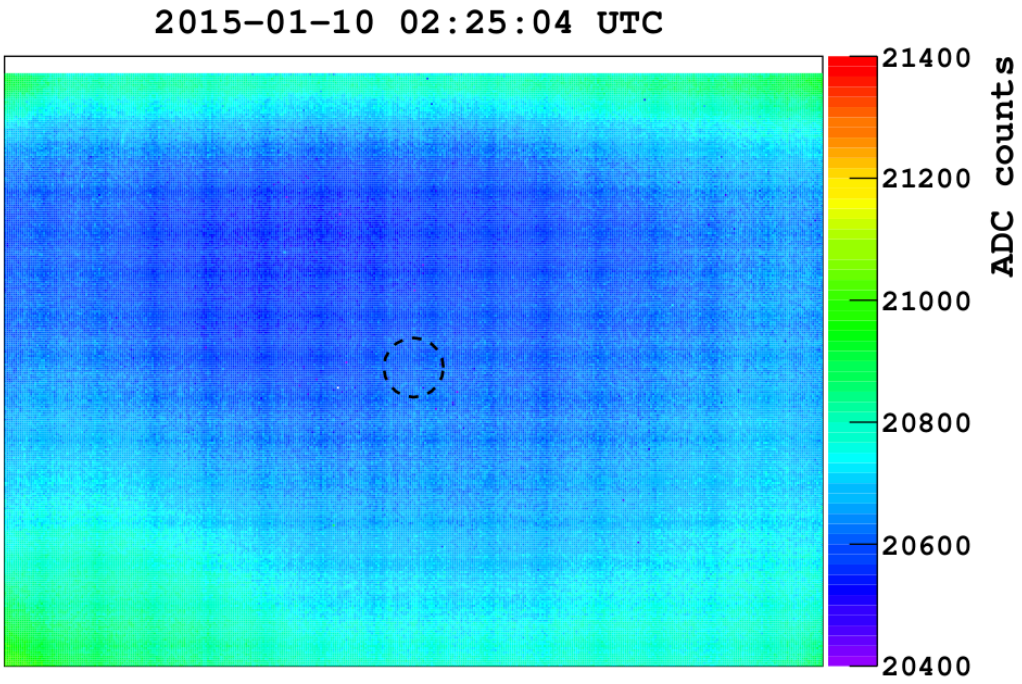

Figure 5.4: An example of a chosen reference region on the image. The dashed black circle corresponds to the 2 degree region around the image centre where the reference value is taken from.

angle range are at approximately the same brightness, thus one can use these pixels to calibrate the remainder of the image for larger zenith angles.

Flat-field corrections are determined for clear night skies, which have been previously catalogued [48]. Each camera is calibrated independently, since they all have a slightly different pixel response. The calibration of a single image is done by correcting the intensity of each separate pixel based on a reference value ADC_0. This reference value is chosen by averaging the intensity of each pixel in the centre of the image (which is chosen as a circle of radius 2 degrees - roughly 600 pixels). The image centre is chosen since this is the region of the image least affected by the window thermal effect described earlier in this chapter. More than a single pixel is also chosen in order to reduce the error in the reference value. Figure 5.4 illustrates how this reference value is chosen.

After the mean reference value ADC_0 is calculated, each pixel will be adjusted to have the same ADC brightness as this reference value ADC_0 [48]. Once this reference is calculated, the ratio R of the brightness of each pixel to this reference value is then evaluated. The relative difference is used since the effects that cause non-flatness in the images are characterised as a relative change in brightness away from the image centre. If this ratio R is known, then it can be deduced that the brightness of each pixel after the flat-field correction is applied (denoted ADC_{ff}) is given by equation 5.1 [48].

$$ADC_{ff} = ADC_0 = \frac{ADC}{R} \tag{5.1}$$

where ADC is the value of each pixel before the flat-field correction is applied. This is a circular calculation, however it is used in order to make sure that ratios can be determined and then can be applied to any cloud camera image [48].

One parameter that is found to be related to this ratio R is the temperature of the camera's sensor (denoted T_{cam}). This temperature is typically about 25 K warmer than the air temperature, thus it can be used as a proxy for the temperature of the window. The dependence of T_{cam} on the ratio R is parameterised as shown by equation 5.2 [48].

$$R = p_0 + p_1(T_{cam}) \tag{5.2}$$

where p_0 and p_1 are fit parameters that are determined for each pixel. Once this fit is done, the fit parameters are saved into a flat-field template text file which can be read in by a program to determine the flat-field correction for any image.

Now that the flat-field correction has been discussed, the next correction to discuss is the temperature calibration, outlined in section 5.3.

5.3 Temperature Calibration

The temperature of the sensor in each camera at the Auger FD sites is dependent upon the air temperature [48]. The radiation which is incident on the cloud cameras is converted into a digital signal (in ADC counts), before being converted into a brightness temperature. However, the settings and properties of the camera partially determine this digital output, meaning the final measurement has some associated uncertainty due to the camera itself [48]. As a result, one must perform temperature calibrations on our instruments to remove this dependence of the final temperature reading on the camera properties [48].

One wishes to remove the dependence of the temperature on the settings and properties of the camera, done by a conversion of ADC counts to temperature as shown in equation 5.3 [48].

$$\mathrm{ADC} = m(T_{\mathrm{cam}}) \times T_{\mathrm{obj}} + C(T_{\mathrm{cam}}), \tag{5.3}$$

In equation 5.3, $m(T_{\mathrm{cam}})$ and $C(T_{\mathrm{cam}})$ depend on the temperature of the camera's sensor (T_{cam}). T_{obj} is the temperature of an object (in kelvin) in the camera's field of view.

The work completed previously found the temperature calibrations given by equation 5.3 [48]. The method of this temperature calibration will be discussed in section 5.3.1.

5.3.1 Calibration method

Calibration of the cameras at Auger is not able to be achieved directly, however it can be done by taking advantage of other instruments installed at Auger, namely the single-pixel radiometers located at the CLF and XLF (see section 3.5.3) which have been calibrated to have a similar response to the cloud cameras in the operational bandwidth ($8 - 14\,\mu$m). Each radiometer is mounted to record an overhead sky temperature in this bandpass. In addition, a thermistor records the approximate local air temperature, which is a proxy for ground temperature since the thermistors are located close to ground level. Much like the cloud cameras, the radiometers collect data every 5 minutes, thus making temperature measurements perfect for calibration purposes [48].

The atmosphere is opaque when viewing close to the horizon, as such it resembles a black body radiator (due to the absorption and re-emission by various greenhouse gases such as water vapour). In directions close to the horizon, this infrared black

body temperature approximates the ground temperature given by the thermistor. If the radiometer was oriented to point at the horizon and measure sky temperature near the horizon, it would closely resemble the temperatures read by the thermistor [84].

In comparison, the atmosphere is quite optically thin overhead (mostly transparent), resulting in the measured sky temperature being much cooler. One finds on average that clear sky temperatures measured by the radiometers are $\sim 20\,\mathrm{K}$ lower than the thermistor temperatures [48]. Preliminary measurements of the camera response (in ADC counts) to object temperature is shown to be approximately linear as given by equation 5.3 [48]. As a result, a two point calibration can be performed on the radiometer measurements, with one point corresponding to the clear sky temperature measured vertically by the radiometer, and the other being the thermistor temperature (a proxy for sky temperature at the horizon). Clear sky is used to maximise the difference between these two points, since an overcast sky is optically thick at all zenith angles, meaning the temperature difference between the two points is small. This will reduce the effect of errors on the final calibration [48].

To compare the measured temperatures from the radiometer and thermistor, the signal from the cloud camera (in ADC) as a function of zenith angle is used. For the vertical sky temperature from the radiometer, the camera signal for all pixels with zenith angles $\theta < 20°$ is averaged. For the horizon measurement, the thermistor temperature is compared to the averaged camera signal for zenith angles between 89.5 and 90°. An example two point calibration is shown by Figure 5.5.

This must be repeated over several clear nights at different times during the year in order to get a wide range of operating temperatures $\mathrm{T_{cam}}$. This is shown in Figure 5.6 for the Los Leones cloud camera.

In Figure 5.6, the blue points can be compared to the overhead sky temperature measured by the radiometer, and the red points to the thermistor temperature [48]. One must now produce calibration curves that solve equation 5.3, which are unique to each cloud camera. The influence of $\mathrm{T_{cam}}$ on the temperature calibrations are tested using this two point calibration method. The next step is to plot the resultant $m(T_{\mathrm{cam}})$ and $C(T_{\mathrm{cam}})$ values as a function of $\mathrm{T_{cam}}$, shown in Figure 5.7 for the Los Leones cloud camera [48].

Figure 5.7 shows the relationship to T_{cam} for the slope m and offset C. These are given by equations 5.4 and 5.5.

$$m(T_{\mathrm{cam}}) = m_2 \times T_{\mathrm{cam}}^2 + m_1 \times T_{\mathrm{cam}} + m_0 \tag{5.4}$$

$$C(T_{\mathrm{cam}}) = C_2 \times T_{\mathrm{cam}}^2 + C_1 \times T_{\mathrm{cam}} + C_0 \tag{5.5}$$

The values of the fit parameters (for each FD site) have already been summarised [48]. By substituting the values from equations 5.4 and 5.5 one can then calculate the temperature of an object in the camera's field of view using equation 5.6.

$$T_{\mathrm{obj}} = \frac{ADC - C(T_{\mathrm{cam}})}{m(T_{\mathrm{cam}})} \tag{5.6}$$

Overall, this temperature calibration works very well, however there were some small discrepancies found for lower sensor temperatures [48]. This was compensated for by adding in an extra factor, $r(T_{\mathrm{cam}})$ which is given by equation 5.7.

$$r(T_{\mathrm{cam}}) = T_{\mathrm{obj}} - T_{\mathrm{true}} \tag{5.7}$$

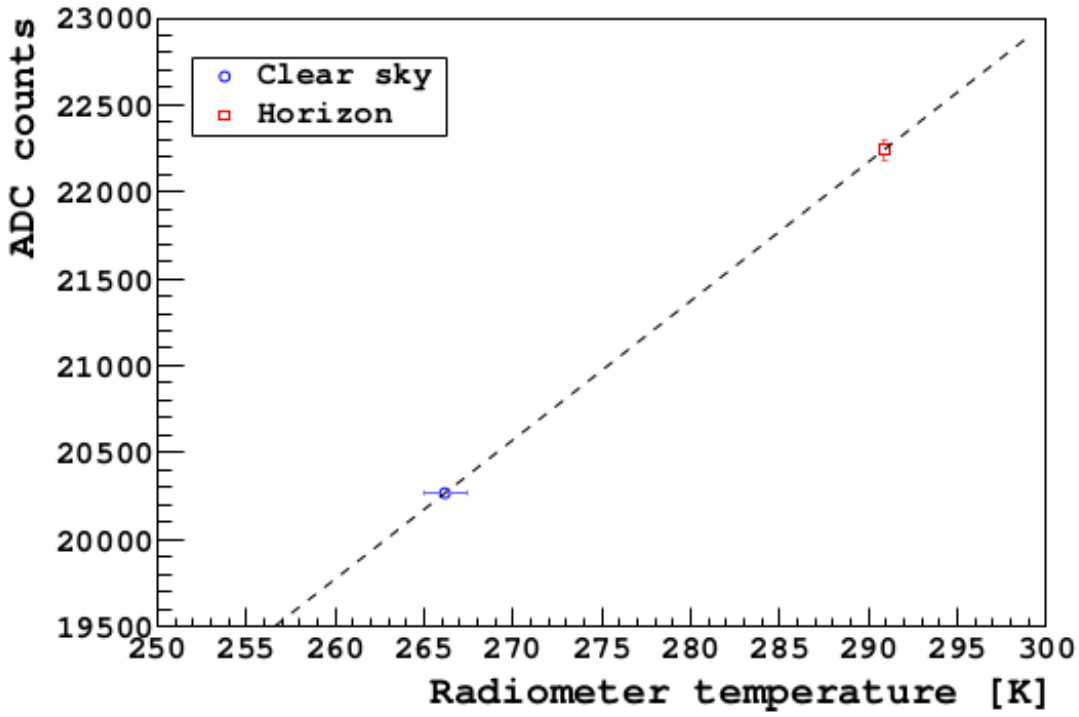

Figure 5.5: An example of a two-point temperature calibration comparing Los Leones cloud camera data (vertical axis) to the temperature given by the radiometer (horizontal axis). The blue point corresponds to the overhead sky temperature, while the red point corresponds to the temperature measured by the thermistor which is a proxy for sky temperature at the horizon [48]. The slope m and intercept c are taken for each of these two point calibration curves. Credit: Patrick van Bodegom [48].

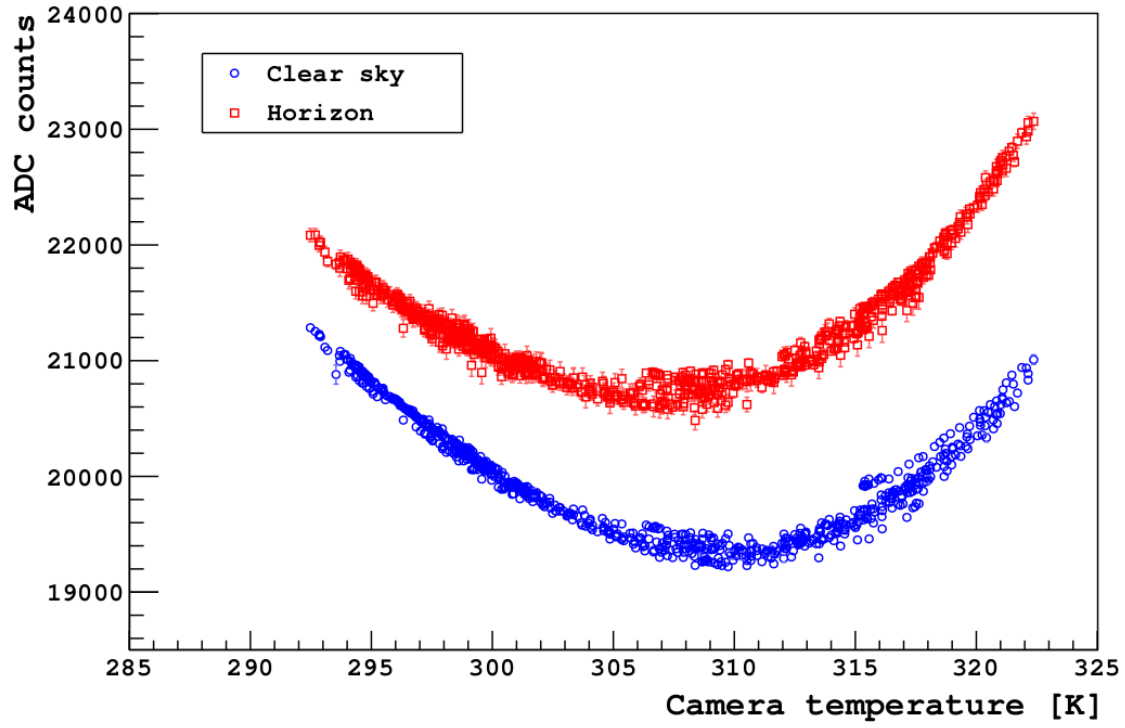

Figure 5.6: Measurements of infrared brightness as a function of camera sensor temperature (T_{cam}) for the Los Leones cloud camera, from full sky scans on clear nights. The blue points are the signals pointing the camera vertically, and the red points are field of view images close to the horizon. Credit: Patrick van Bodegom [48].

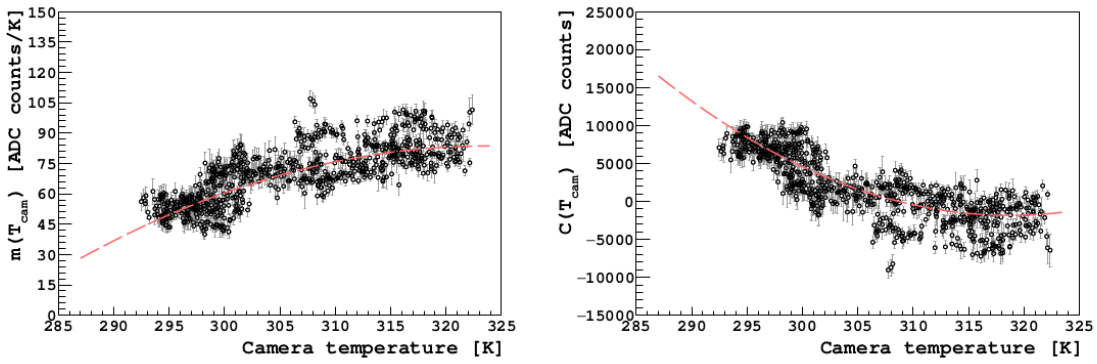

Figure 5.7: Various two point calibration slope values ($m(T_{cam})$ - left) and offset values ($C(T_{cam})$ - right) as a function of T_{cam} for the Los Leones camera. The red dashed lines correspond to the fit to the data. Credit: Patrick van Bodegom [48].

where T_{obj} is given by equation 5.6 and T_{true} is the true temperature measurement given by the radiometer [48]. A cubic function is used to parameterise $r(T_{\text{cam}})$ given by equation 5.8 [48].

$$r(T_{\text{cam}}) = r_3 \times T_{\text{cam}}^3 + r_2 \times T_{\text{cam}}^2 + r_1 \times T_{\text{cam}} + r_0 \tag{5.8}$$

Once again, the parameters of the fit (for each camera) are summarised in previous work [48]. This slightly adjusts equation 5.6 to become equation 5.9 [48].

$$T_{\text{obj}} = \frac{ADC - C(T_{\text{cam}})}{m(T_{\text{cam}})} - r(T_{\text{cam}}) \tag{5.9}$$

This added factor improves the accuracy of the temperature calibration, which is applied to all subsequent cloud camera images taken at Auger. The calibration is not quite as accurate for the Coihueco camera, due to the $\sim 300\,\text{m}$ altitude difference between Coihueco and the Central Laser Facility (where the radiometer is housed). At this higher altitude, the measured sky temperature at Coihueco will be lower [48]. Interestingly, the accuracy of these temperature calibrations is not affected by differences in humidity - which means that the radiometers share a similar spectral response to our cloud cameras [48].

This temperature calibration, combined with the flat-field correction discussed in section 5.2, are all of the corrections that need to be applied to cloud camera data in order to start our analysis of these images. The rest of this chapter will go into more detail on the analysis that has since been performed on these corrected images, in particular the updates to the cloud database and the extension of data analysis to include HEAT (see section 5.4.1).

5.4 Cloud Masks

As mentioned in Chapter 3, clouds in the field of view of the Fluorescence Detectors can block and scatter nitrogen fluorescence and Cherenkov light emitted from extensive air showers. This will change the total signal read by the detectors. As a result, the reconstructed energy of the incident cosmic ray particle will be affected. Cloud masks are the means to find where these clouds are in the field of view (FOV) of the FDs so that their effects on reconstruction are better understood (see section 3.1).

A cloud mask is the output of a series of images from the cloud camera, stitched together with the pixels of the FD camera as shown in Figure 5.8. In a cloud mask, each individual FD pixel is assigned a value known as the cloud fraction.

Cloud fraction is the amount of cloud that covers each pixel of the FD, represented as a fraction. The 440 pixel cameras of each FD telescope are mapped onto the sky in the IR camera field of view, as shown by Figure 5.8. As described in section 3.5.8, the cloud camera pixels are arranged in a 384×288 pattern, which means there are $\sim$ 100 cloud camera pixels for each FD pixel. Each FD covers a range of 30° in zenith (or elevation) angle and 180° in azimuthal angle. The cloud fraction for each FD pixel is calculated as a fractional value between 0 and 1, based on how many of the 100 cloud camera pixels viewed cloud.

Figure 5.8 is an example of the camera of the FD mapped onto the sky as viewed by the cloud camera.

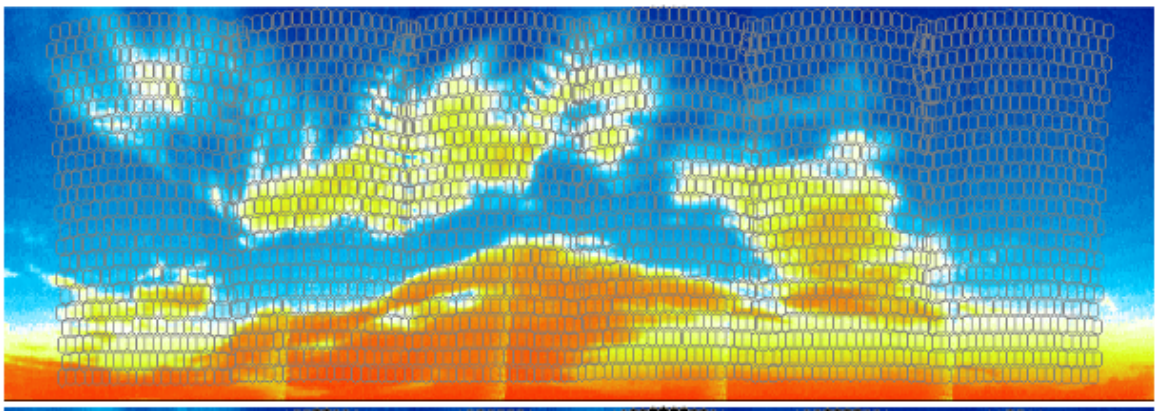

Figure 5.8: The 440 pixels of each of the 6 telescopes at one FD site mapped onto the sky. Telescopes are numbered 1 to 6 from right to left (lowest azimuth to highest azimuth angle). Credit: Patrick Van Bodegom [48]

Images taken by the cloud cameras at Auger (see section 3.5.8) are taken in sequence every 5 minutes, with two different sets of images taken. The first sequence is known as a Field of View (FOV) scan - this is a series of 5 images taken consecutively in the field of view of the fluorescence detectors (see section 3.1). Each individual image covers 40° in azimuth and 30° in zenith. The second series of scans is known as a full sky scan. For a full sky scan, a series of nine images are taken which covers the full 360° in azimuthal angle (and 60-90° in zenith angle). After this, the camera shifts down 30° in zenith and repeats the nine image sequence (for zenith angles between 30 and 60°). Finally, the camera shifts to point straight up before taking one single vertical image at 0° zenith. This sequence of 19 images is the full sky scan (FSS), which is done every 15 minutes for each cloud camera (there is one exception, which is discussed in section 5.4.1). An example of these two separate scans (FOV and FSS) are shown in Figures 5.9 and 5.10 respectively.

The camera field of view images with the overlapping FD camera pixels (and their calculated cloud fractions) are stitched together to create a map of the field of view of each FD, known as a mosaic [the right hand plot in Figure 5.11]. The cloud mask is the result of this mosaic, containing the cloud fraction information for each pixel. Cloud-affected reconstruction events are easy to pick out from this since one has a complete view of the portion of the atmosphere seen by the FDs during a data acquisition run.

In Figure 5.11, cloud fraction values calculated for each pixel are fed into the Auger cloud database (for those pixels in the field of view of the FDs). The vertical image in the full sky scan is used as the flat-fielding reference image (as described in section 5.2), since it is the flattest (in the infrared) of all images in the sequence.

The left panel in Figure 5.11 shows the brightness temperature for each pixel (each one of which is at a given zenith angle). The emissivity of the atmosphere increases with increasing zenith angle, due to the increased optical depths of water vapour and carbon dioxide, which are the main emitting gases in the atmosphere in the operational wavelengths of the cloud cameras (7-14 μm) [48]. Consequently, the atmospheric infrared flux received also increases with zenith angle. Due to the cloud cameras being sensitive to radiation at those wavelengths, it is difficult to distinguish a cloud's thermal emission from a clear sky background when viewing angles towards the horizon (since the atmosphere is opaque and bright) [48]. A simple equation has been empirically

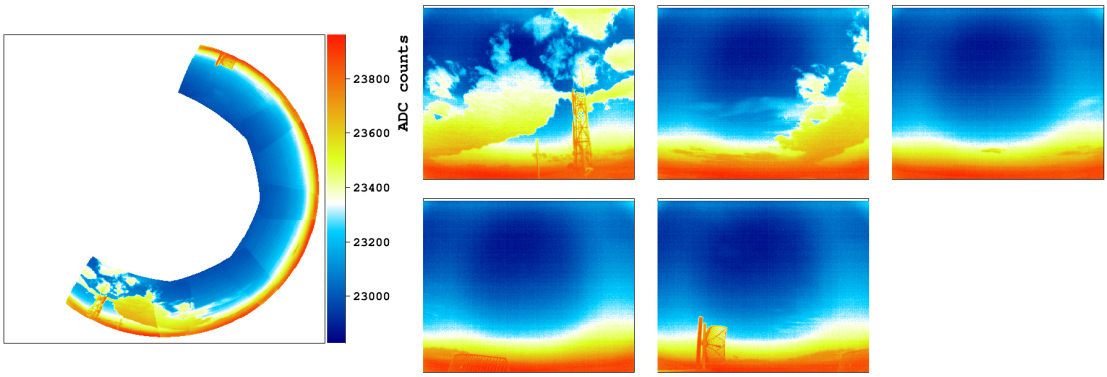

Figure 5.9: Field of View images from the cloud camera at the Coihueco FD site. Left panel: Overlapping the 5 images together to create a map (in the IR) of the field of view of the Coihueco FD. Right panel: Each FOV image is taken consecutively with a colour conversion in ADC counts. This is used to clearly show the location of cloud in the FOV. The 5 images are not in the order of the FOV scan.

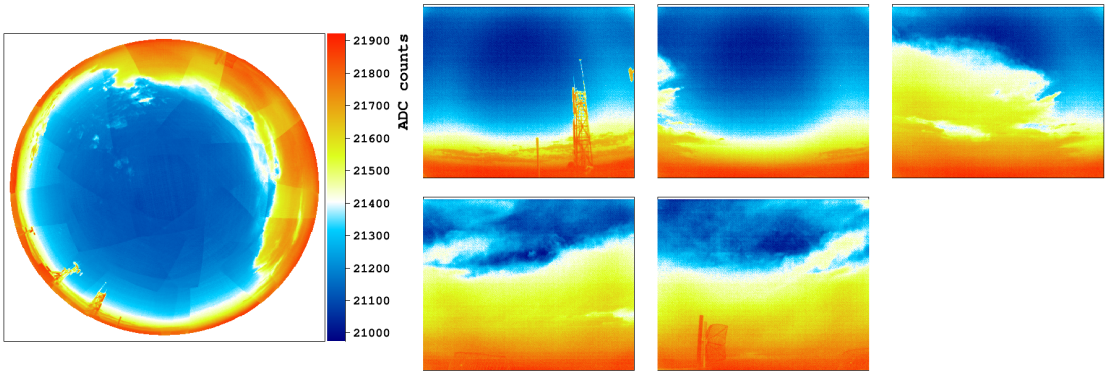

Figure 5.10: Full Sky images from the cloud camera at the Coihueco FD site. Left panel: Overlapping the 19 images together to create a map (in the IR) of the field of view of the Coihueco FD exclusively. Right panel: Each FOV image is taken consecutively with a colour conversion in ADC counts. This is used to clearly show the location of cloud in the FOV. Note that this scan was done at a later time than Figure 5.9.

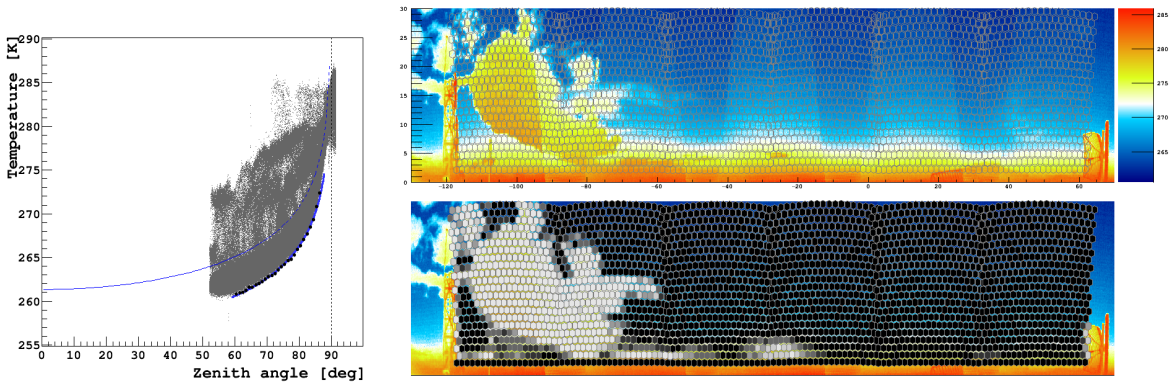

Figure 5.11: Example Cloud Mask from the cloud camera at the Coihueco FD site. Left panel: Temperature as a function of zenith angle of the images, with a fit to the clear sky and a threshold set for cloud. Right panel: Each image is mapped onto the FD FOV which creates a mosaic, with cloud fraction values calculated for each pixel (see section 6.3).

derived to describe how the effective brightness temperature changes as a function of zenith angle, as shown in equation 5.10.

$$T = A + B \ln sec\,\theta \tag{5.10}$$

where A and B are fitted parameters [48].

Equation 5.10 has been derived from measurements collected by single-pixel infrared radiometers and describes the relationship between brightness temperature and zenith angle quite well [48], particularly in the zenith angle range of the field of view of the FDs ($\sim 60-87°$). The radiometer used for this purpose has a similar spectral response to our cloud cameras, a fact also used in performing the temperature calibration (as described in Section 5.3.1).

The fit parameters A and B must be calculated in order to know the effective temperature of the atmosphere as a function of zenith angle, which was done for a catalogue of clear night skies [48]. Previous studies have shown that these two parameters depend on the local air temperature, as well as the relative humidity (i.e. the amount of water vapour present) [48]. With these known parameters (derived from clear night skies), a means of cloud detection is straightforward. This is due to cloud appearing warmer and brighter than the night sky, thus can be distinguished from a clear sky background due to this warmer temperature [48]. However, estimating the clear sky background using these parameters proved unsuccessful as the uncertainties associated with A and B were too large [48]. Thus, another method of detecting clouds in the camera's FOV is required.

As is shown in Figure 5.11, when plotting the sky temperature as a function of zenith angle, the warm bumps in the profile are associated with clouds. This is distinguishable from smooth clear sky profiles. Unless the sky is completely overcast, the coolest sky temperatures correspond to a clear sky. As a result, the clear sky background is best estimated directly from the data [48]. This background corresponds to the lower bound of pixels in the left panel of Figure 5.11. To estimate this background in the zenith angle range of the FDs we bin the pixels into zenith angle bins. For each bin, we then find the minimum pixel value, which is estimated to be the clear sky background [48].

Typically, a small region of clear sky is seen by the cloud cameras at a given zenith angle. A temperature threshold 3.5 K warmer than the clear sky background is also shown in Figure 5.11, indicated by the blue dashed line, and can be used to identify most clouds [48]. Those pixels with temperatures above this threshold are identified as viewing cloud, whereas pixels with values below the threshold are instead identified as viewing clear sky [48]. There are occasions when the threshold is lowered in order to increase the sensitivity to cloud detection, since in the case of a 3.5 K threshold, some pixels can still contain a cloud which is either very thin, or very away, resulting in those cloud containing pixels being < 3.5 K warmer than the clear sky background, which occurs quite frequently [48]. In other cases, the threshold must be raised to avoid false detections of cloud, which is most prevalent for the Coihueco cloud camera. Ultimately, the threshold is determined by the spread in the pixel values corresponding to the clear sky background, the values of which are typically adjusted by around ± 1 K [48].

The process for adjusting these thresholds is currently a manual process, which while it gives the best overall quality of data, is a tedious, time consuming process. However, there have been recent improvements to this process in order to speed up the production of cloud masks. This is outlined in Section 5.4.2.

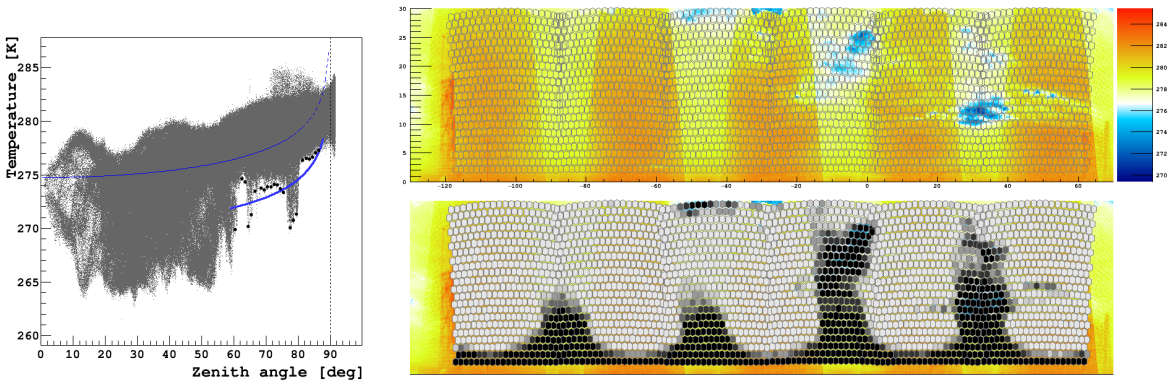

Figure 5.12: Field of View mosaic from the cloud camera at the Coihueco FD site in extremely overcast conditions. The curve produced from the fit is rather flat, corresponding to a very low B parameter.

Now, it is worth looking more in-depth into the fit function itself (given by equation 5.10); specifically, to better understand how the equation handles extreme cases such as a fully overcast sky. When the conditions are overcast, the $\ln \sec \theta$ curve flattens ($B \sim 0$) [see Figure 5.12]. When the sky is extremely overcast, extensive air shower analysis is not performed, due to the cloud quality cuts scrapping events during these conditions (see section 3.5.9 for more information).

The reason for this flattening of the curve is due to the emission properties of clouds in the infrared. When completely covered by cloud, the atmosphere will radiate at a single temperature; that is, the blackbody temperature of the cloud (see Section 4.2). As a result, the signal read in any image will be the same regardless of a pixel's zenith angle, resulting in the flat curve shown in Figure 5.12.

When the cloud database was first constructed, one only had access to four fluorescence detector sites for the cloud mask images. As it happens, the HEAT extension (see section 3.1.3) to Auger also covers a section of the sky close to the FOV of Coihueco. The full sky scan in the sequence covers HEAT's FOV during data acquisition, thus it is possible to include mosaics from HEAT in the analysis. As of January 2018, the program used to generate cloud masks was rewritten, allowing for this extension to happen. This is done by fitting to the $\ln \sec \theta$ curve over a different zenith angle range in the full sky sequence (30-60° Zenith), using the images generated by the Coihueco cloud camera.

5.4.1 Extension to HEAT

The program code to generate cloud masks, developed by Mat Cooper, Trent Grubb and Patrick van Bodegom, was originally written to use only the Fluorescence Detector FOVs. However, with the introduction of HEAT, and the analysis done on air shower events seen by both HEAT and Coihueco (see section 3.1.3), it was necessary to extend the program so that one could create cloud masks in the HEAT FOV (see section 3.1.3). To do this, changes were made; namely, modifying the number of pixels being looped over (as HEAT only has 3 telescopes as opposed to 6), and modifying the clear sky fit range to be from 30-60° zenith (as opposed to 60-90° zenith for the four fluorescence detectors). Figure 5.13 is an example of a generated mosaic for a full sky scan done at Coihueco, which covers HEAT's FOV.

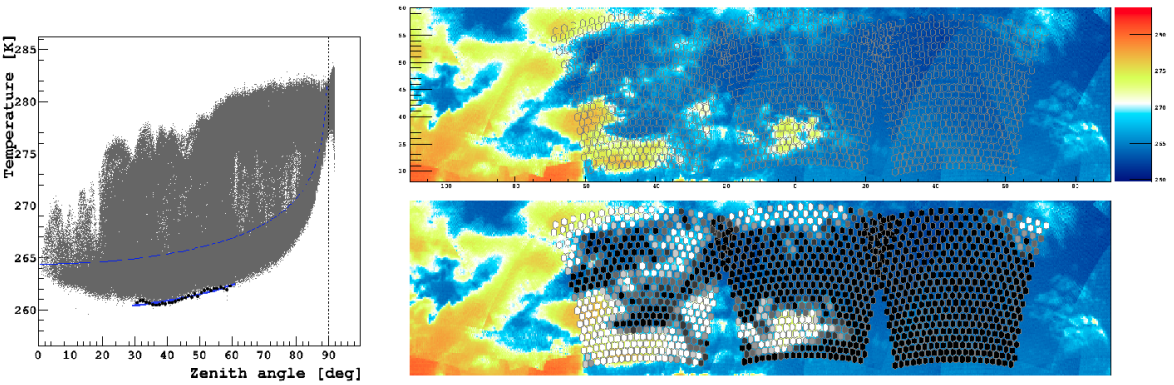

Figure 5.13: Field of View mosaic for HEAT from the cloud camera at Coihueco.

As can be seen by Figure 5.13, the integration of HEAT into the cloud mask analysis was seamless, allowing for higher quality reconstruction of lower energy events (for more information on this see section 3.1.3). More detail on the cloud database itself (and how it pertains to this extension) will be described in section 5.4.2.

5.4.2 Cloud Database

All cloud mask data is stored in the cloud camera database for use in cloud quality cuts as described in Chapter 3. Figure 5.14 shows the most up to date summary of the current cloud database, filled using data from the cloud mask production.

All the cloud mask data (see Section 5.4) that is generated is stored into a database, one of the many different databases in use by the Auger collaboration. For instance, there is a database for the LIDAR - see section 3.5.6, and for aerosols using data from the Central Laser Facility (see section 3.5.3). The Cloud Camera Database (DB) is the largest such database, due to the sheer number of cloud camera images that go into it, and as such requires a large degree of upkeep and maintenance. The biggest factor in reducing the size of the DB revolves around the concept of validity time. To discuss validity time, one first must understand important features of the data.

The data from cloud masks are made into cpd files, a text file which is split into columns with the GPS time of the image scan (the beginning of the image scan sequence is the recorded time), the ID of the cloud camera (1=Los Leones, 2=Los Morados, 3=Loma Amarilla, 4=Coihueco and 5=HEAT), and the cloud fraction value for all 440 pixels of each telescope. The typical time between scans is 5 minutes, though this can vary (as is shown later). As such, data has a 5 minute validity time placed on it. Validity time is the amount of time the data from a single scan is valid for. For example, if a scan starts at 00:00 UTC, then for any time up until 04:59 UTC, querying the cloud database will return the values found at 00:00 UTC. In the case where a given pixel has the exact same cloud fraction value for consecutive scans, it isn't a requirement to overwrite that pixel's value again. In this case, the validity time of the previous scan is extended by another 5 minutes for each pixel that remains unchanged.

The sole exception on this validity time rule is at HEAT, since, initially, data was only accessible every 15 minutes (since HEAT masks could only be done on full sky scans). HEAT had an initial validity time of 15 minutes, however, this presented a conundrum since a 15 minute validity time is too long for this purpose. This is due to the fact that atmospheric conditions can change quite rapidly within a 15 minute

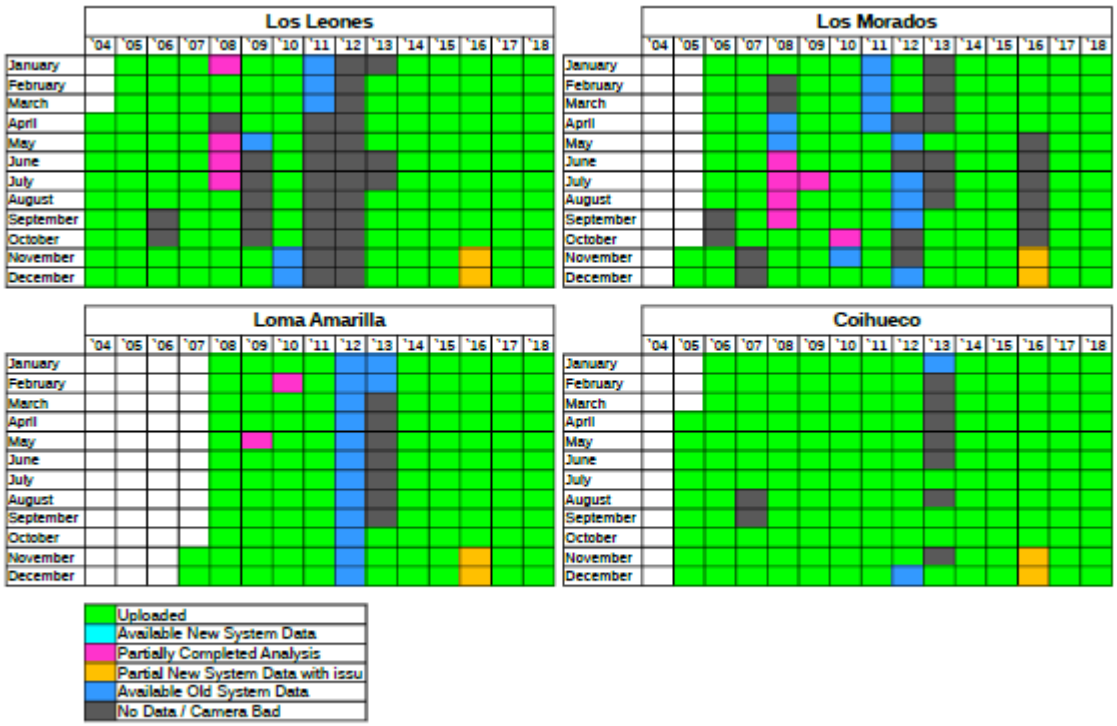

Figure 5.14: A table of completed cloud mask analysis done for each of the four FD sites. Green corresponds to a completed month of processed masks that is stored in the cloud database.

period. This results in inaccurate descriptions of the true atmosphere for the time period between consecutive scans. On the other hand, changing the validity time to 5 minutes (as for the 4 FD sites), is also not feasible since that creates large periods of time with gaps in the database. Essentially, if someone wishes to use the cloud camera data for HECo events (see section 3.1.3), then cloud mask data from HEAT is required as frequently as Coihueco. If the shower event occurs during a time where there is no HEAT data available, then the necessary cloud cuts are unable to be done, resulting in that particular event being rejected. The alternative is that cloud is not detected, meaning that the event is blind to potential uncertainties on the reconstruction due to the presence of cloud.

The only viable solution left was to change the pattern of scans for Coihueco such that a full sky scan is completed every 5 minutes, rather than a FOV scan. This was implemented in September of 2018 - Coihueco now performs a full sky scan every 5 minutes. This way HEAT will have a 5 minute validity time, rather than the 15 minute validity time it had before September 2018.

Validity Time Cross-Check

The only issue now is in the performance of the camera - constantly running full sky scans can place a lot of stress on the camera and as such, Coihueco now needs to be monitored much more closely than the other sites. However, this scan sequence change does allow for a 5 minute validity time to be applied to HEAT data, without unnecessary gaps in the database being created. Figure 5.15 shows distributions of the time (in seconds) between consecutive cloud camera scans, which ideally should be 300 seconds. This is to identify a couple of key points; mainly if the 5 minute validity time is a good choice, and also if there are any major gaps in the data (that might become problematic for analysis purposes).

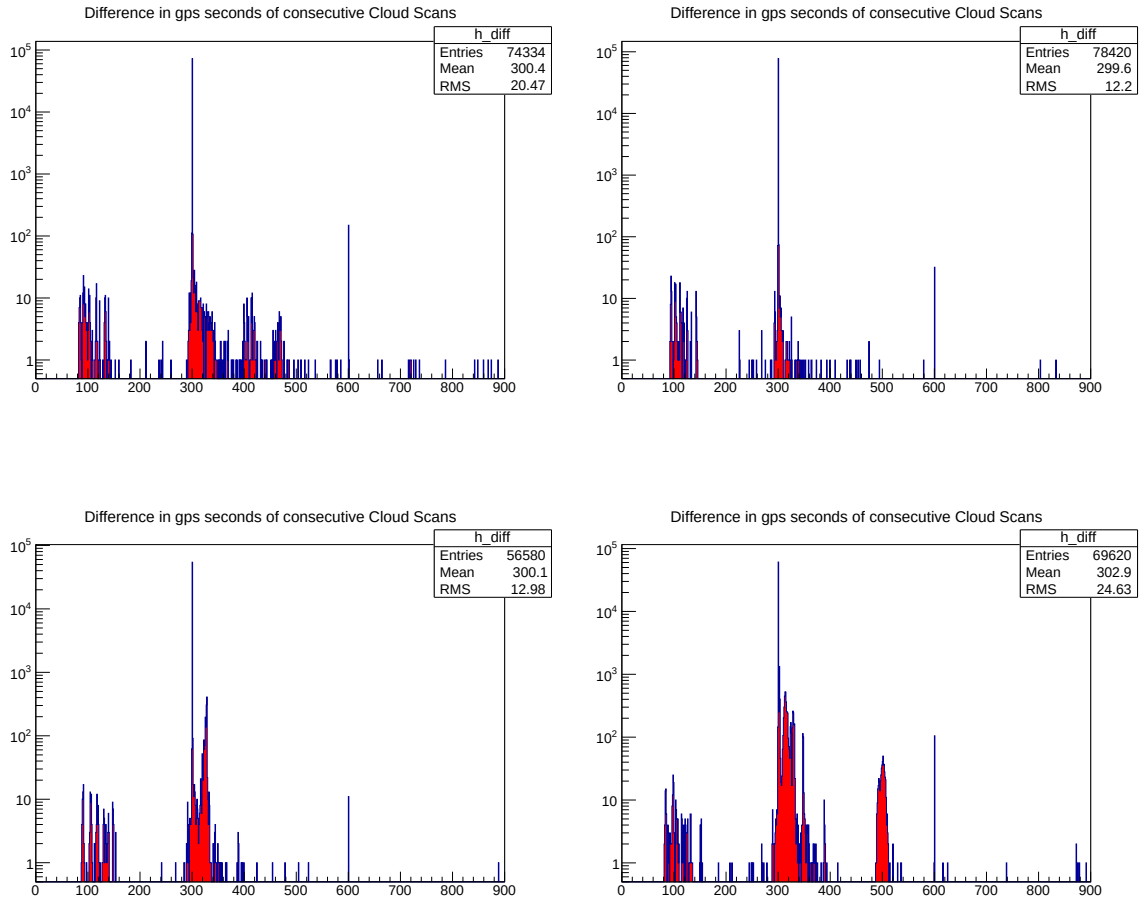

Figure 5.15: Distributions of time between consecutive cloud camera scans at each FD site (in seconds) between 2015 and 2018. All mean values of this difference are consistent with 300 seconds (5 minutes). Top left is Coihueco, top right is Los Leones, bottom left is Los Morados and bottom right is Loma Amarilla. Note the logarithmic scale on the vertical axis.

Finally, we can also look at the HEAT data, which should ideally have a gap of 15 minutes or 900 seconds (as this is the time between consecutive full sky scans before September 2018). Figure 5.16 illustrates this distribution.

The results from Figure 5.15 shows that for each of the four Fluorescence Detector sites, the 5 minute intervals between scans are mostly upheld (most sites have a mean difference close to 300 seconds). For reference, the proportions of how often this difference is 5 minutes is 99.1% for Los Leones, 97.7% for Coihueco, 96.1% for Los Morados, and 87.6% for Loma Amarilla. In each case, there are small proportions present at roughly 10 and 15 minute intervals (600 and 900 seconds respectively), however this is acceptable since that typically means scans were performed every 5 minutes, with one sequence skipped over due to bad data (usually due to an issue with the camera). Thus, this is not a problem since the validity time from the scan done previous to that can be extended to cover that gap.

The only remaining gap that is not well understood is the small proportion (around 8%) of data (between all 4 sites) with a time difference between 300 and 400 seconds. While it is unclear why there are many instances of this occurring (the actual difference

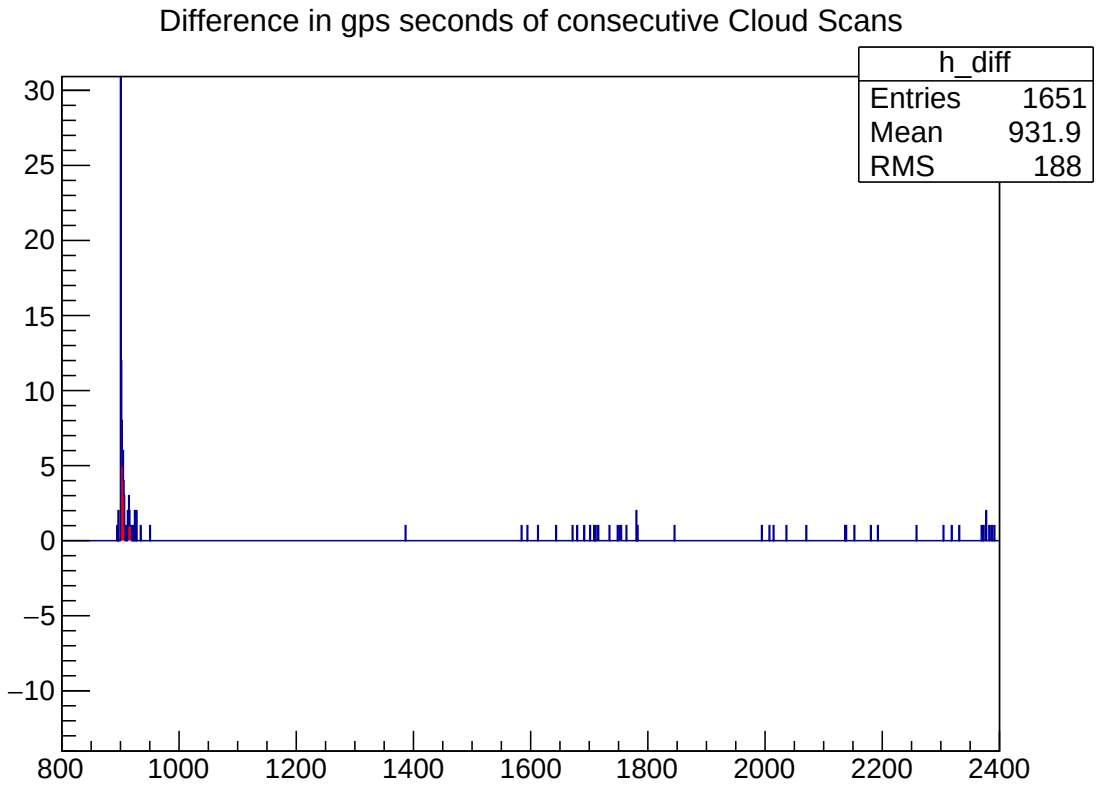

Figure 5.16: Distributions of time between consecutive cloud camera scans at the HEAT site. Once again, good consistency is shown as 94% of values are 15 minutes apart. This distribution was validated before the change to a 5 minute full sky scan at Coihueco was implemented, hence the 15 minute gap instead of 5 minutes.

is never a fixed amount), the gaps created in the database by this can be easily rectified by adjusting the validity time to be 6 minutes. Over the course of 1 minute, the atmosphere (in particular aerosols and clouds) will not change dramatically enough to affect the reconstruction analysis in any way. As a result, adjusting the validity time is able to fill in these unwanted gaps in the database without affecting analysis.

There is also a small proportion of data at the Loma Amarilla FD site specifically (the bottom right plot in Figure 5.15, with a time difference of around 500 seconds. Again, this is simply due to times where the cloud camera is not performing a scan every 5 minutes and there is a gap - this causes a small gap of around 2-3 minutes in the database with no data. Once again, extending the validity time by this extra 2-3 minutes can rectify this so that there are no gaps in the cloud database.

The reason gaps are not desirable in the database is so that cloud quality cuts (as discussed in section 3.5.9) can be applied to as many air shower events as possible. If there were no cloud cuts for a particular event, it could not be determined if cloud played a role in creating an 'exotic' event (see section 3.1), thus reducing the reliability of that event in the subsequent analysis done at Auger.

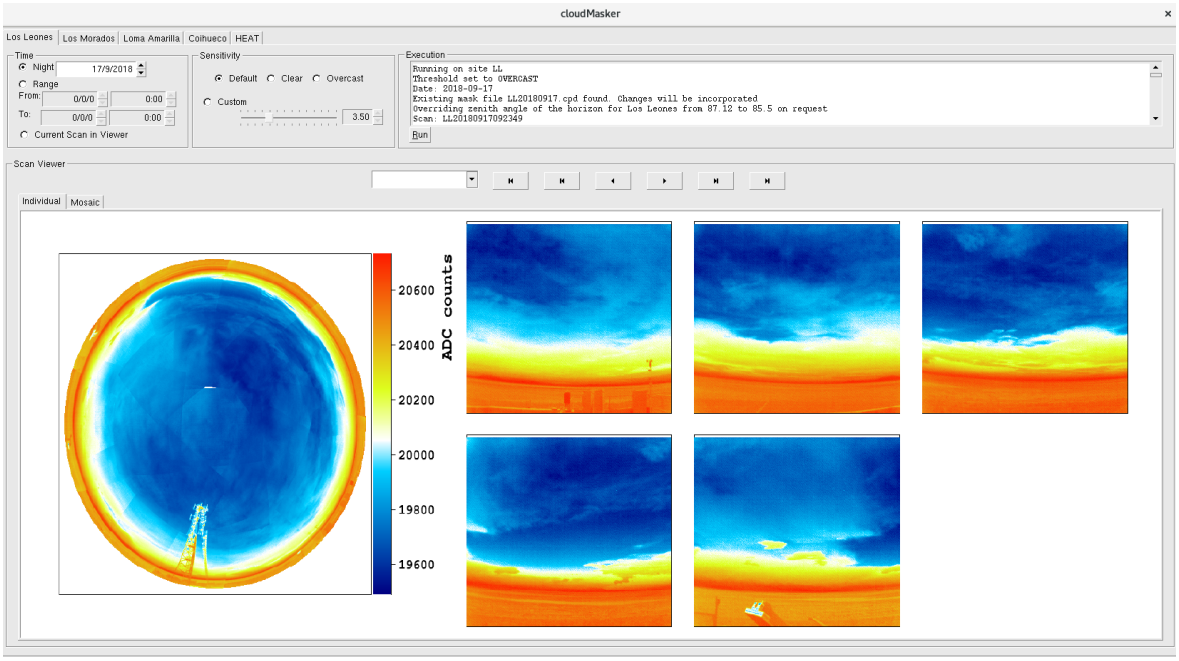

Figure 5.17: Screenshot of the Graphical User Interface for the Cloud Mask program.

Cloud Mask GUI

A recent change has also been made to the software used to generate the cloud mask images/database values. This change has implemented the previous code into a GUI (Graphical User Interface) (programmed by colleagues Bradley Manning and Violet Harvey) to improve the quality of cloud mask generation. A GUI is simply another way of displaying the graphics of the mosaics to more easily identify and troubleshoot issues with the cloud detection [85]. This allows for much faster access to the images and a much quicker way of cross-checking the threshold adjustments that are set. Figure 5.17 shows the output of the cloud mask mosaic software, displaying the results on a GUI.

As can be seen from Figure 5.17, applying a GUI to display the cloud camera images allows for ease of data analysis. In the left corner there are adjustable sliders/buttons for the cloud threshold mentioned earlier. As an example, there is a button to set an entire image to "clear" by adjusting the threshold to a very high value (as discussed earlier). This means that no pixel in the image can pass the cloud threshold, and will all be set to a cloud fraction of 0. After this button is clicked, the "Run" button is then clicked to implement this change. Figure 5.18 shows the brightness temperature as a function of zenith angle when the cloud threshold is set to clear (9999 K above the clear sky minimum fit to the brightness temperature). Figure 5.18 shows that there are no pixels that lie above the cloud threshold.

This means that all pixels in this particular cloud camera image will be marked as clear (with a cloud index of 0). These will show up as black, as illustrated in Figure 5.19.

Occasionally, there are a few scattered pixels well above the clear sky background (between 70 and 80° zenith), which would typically be marked as cloud. However, these warmer pixels result from the radio towers at each FD site coming into the camera FOV. This can be seen in Figure 5.18 as those pixels above the solid gray curve between 70 and 80° zenith. This is a known source of brightness which is removed in

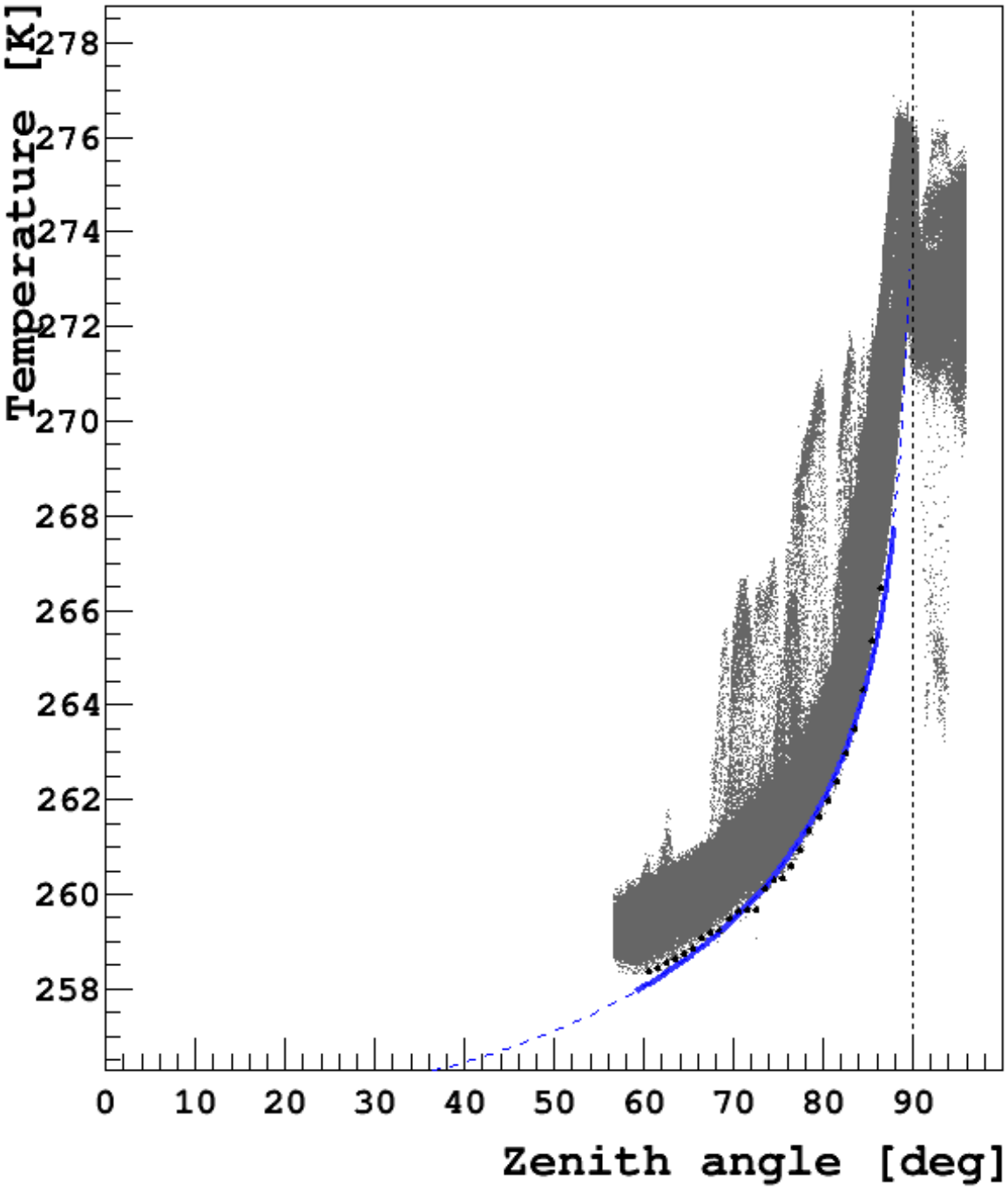

Figure 5.18: Brightness temeprature as a function of zenith angle with a cloud threshold of 9999 K above the clear sky minimum (dashed blue curve). As can be seen, there are no pixels above the cloud threshold, which means every pixel here is defined to be clear.

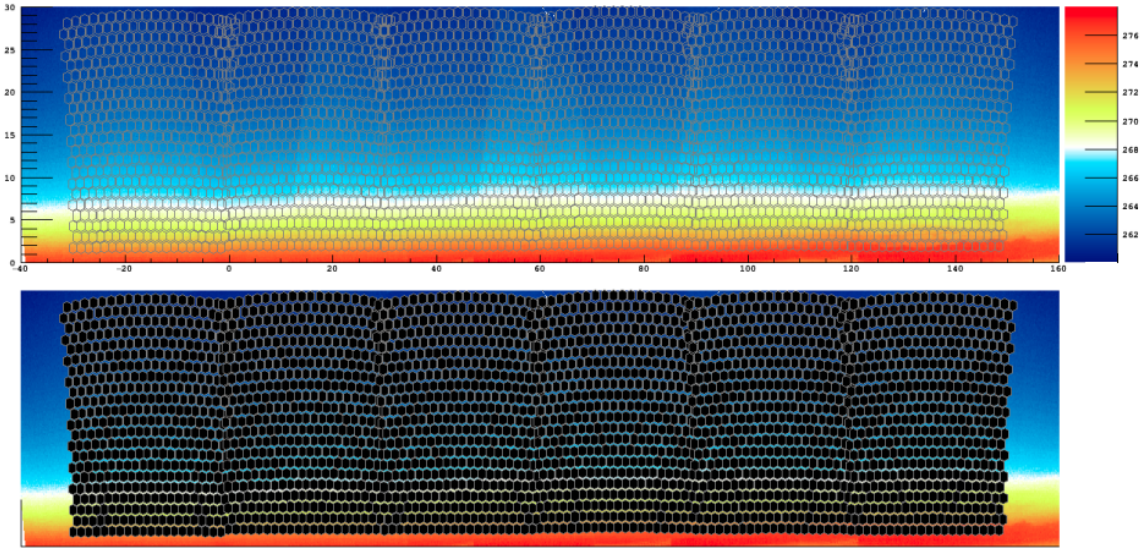

Figure 5.19: An example of the output of the cloud mask program when the cloud threshold is not defined. As a result, every pixel is set as "clear".

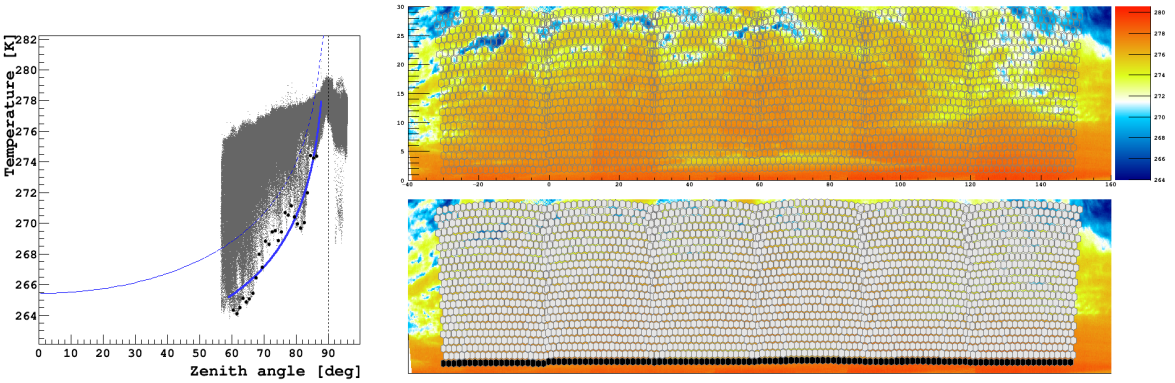

Figure 5.20: An example of the output when setting the Overcast value to True.

cloud mask production.

Additionally, in the cloud threshold options of the GUI, there is another button for the overcast option. Clicking this button will set the boolean "OVER" (or overcast) to true, making all pixels have a value of 1 once the run button is clicked. The bottom row of pixels in the cloud mosaic are still 0 due to the cut on elevation angles below 3° in the cloud mask software (as discussed in section 5.4). The reason for this cut is that there is a high concentration of water vapour very close to ground level, which means that the cloud camera quite frequently misinterprets the image and falsely identifies cloud. Setting these pixels to clear mitigates this false detection. An overcast example is shown in Figure 5.20.

Instructions on how to implement the Cloud Mask GUI are discussed in Appendix D.

5.5 Summary

Infrared Cloud Cameras installed at the Pierre Auger Observatory are used to detect cloud in the FD FOV and subsequently determine when air shower events detected by the Fluorescence Detectors are affected by cloud. Since the cloud can block the isotropically emitted nitrogen fluorescence from air showers, it is important to know where clouds are during FD data acquisition so that we know when air shower event data are affected by cloud. Knowing the presence and development of these clouds is a must in order to ensure the highest quality data in the Auger analysis is utilised. Improvements in our measurements of cloud have be seen over time, thus we have increased the effectiveness of the cloud camera to better study these clouds.

Chapter 6

Seasonal Dependence of Cloud Properties

As described in Chapter 4, there are various atmospheric parameters that change quite substantially with the seasons. An example of this would be the average water vapour content in the atmosphere; this value depends on the temperature and pressure, both of which fluctuate as a function of the time of year. Higher temperatures in summer means that water molecules are more likely to be in the vapour phase in the atmosphere, thus more water vapour is able to be held in the atmosphere (on average). Figure 6.1 illustrates the variation in ground temperature over the course of a year in Malargue.

Since the area of interest in this dissertation is clouds, it is desirable to study if they also vary as a function of the season. Most notably, one must understand how the properties of a cloud change between seasons; examples of such properties are the cloud base height (CBH) and cloud coverage (CC) [as discussed in Chapter 4]. If some correlation can be found, it may provide clues about the nature of the atmosphere given the variable conditions that exist when comparing the seasons. It is important to understand the nature and characteristics of clouds above the Pierre Auger observatory, in order to better understand the effect of clouds on observations of extensive air showers.

6.1 Cloud Base Height

During an FD run, LIDARs at the Auger site return a cloud base height every 15 minutes from the backscatter light profile (see section 3.5.6). However, the database for storing this information only records the lowest cloud base height for each LIDAR in a given hour, out of the sample of the four 15 minute scans within that hour. In the case where no cloud is found, a negative value is given - this is done to distinguish between a LIDAR reporting a clear sky, and not running at all (which would yield a CBH value of 0). In order to look for any seasonal fluctuations in the cloud base height (CBH), a histogram of all heights (as measured by the LIDARs) for each season has been done. The results are shown in Figure 6.2 for a 10 year period from 2005 to 2014.

Upon examination of Figure 6.2, it appears that the summer months have the lowest cloud base heights on average. After splitting the 10 years of LIDAR data in half for the summer months, some differences between the two time periods arise. Figure 6.3 shows the ten year period from 2005-2014 divided into 2005-2009 (left), and 2010-2014 (right).

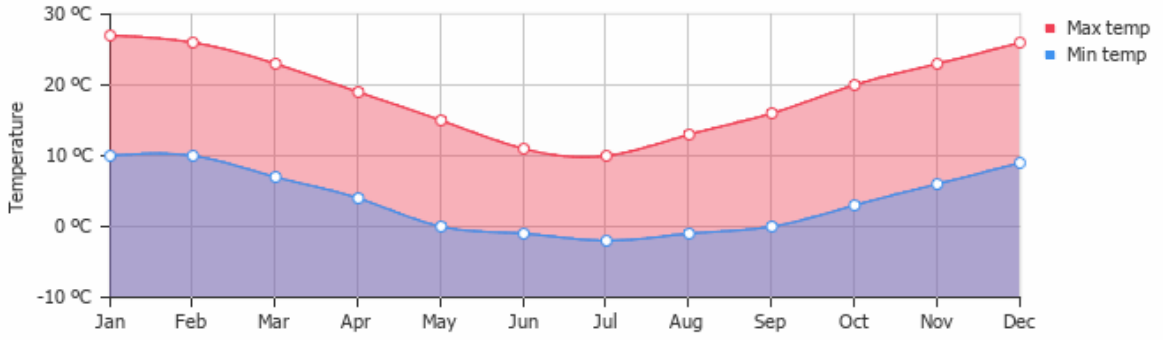

Figure 6.1: Average daily minimum and maximum Temperature in Malargue, Argentina for each month of the year. Clear periodic behaviour is present, where the peak of each curve is in February (Summer) and trough is in July (Winter). Source: <https://weather-and-climate.com/average-monthly-Rainfall-Temperature-Sunshine,malargue-mendoza-province-ar,Argentina>

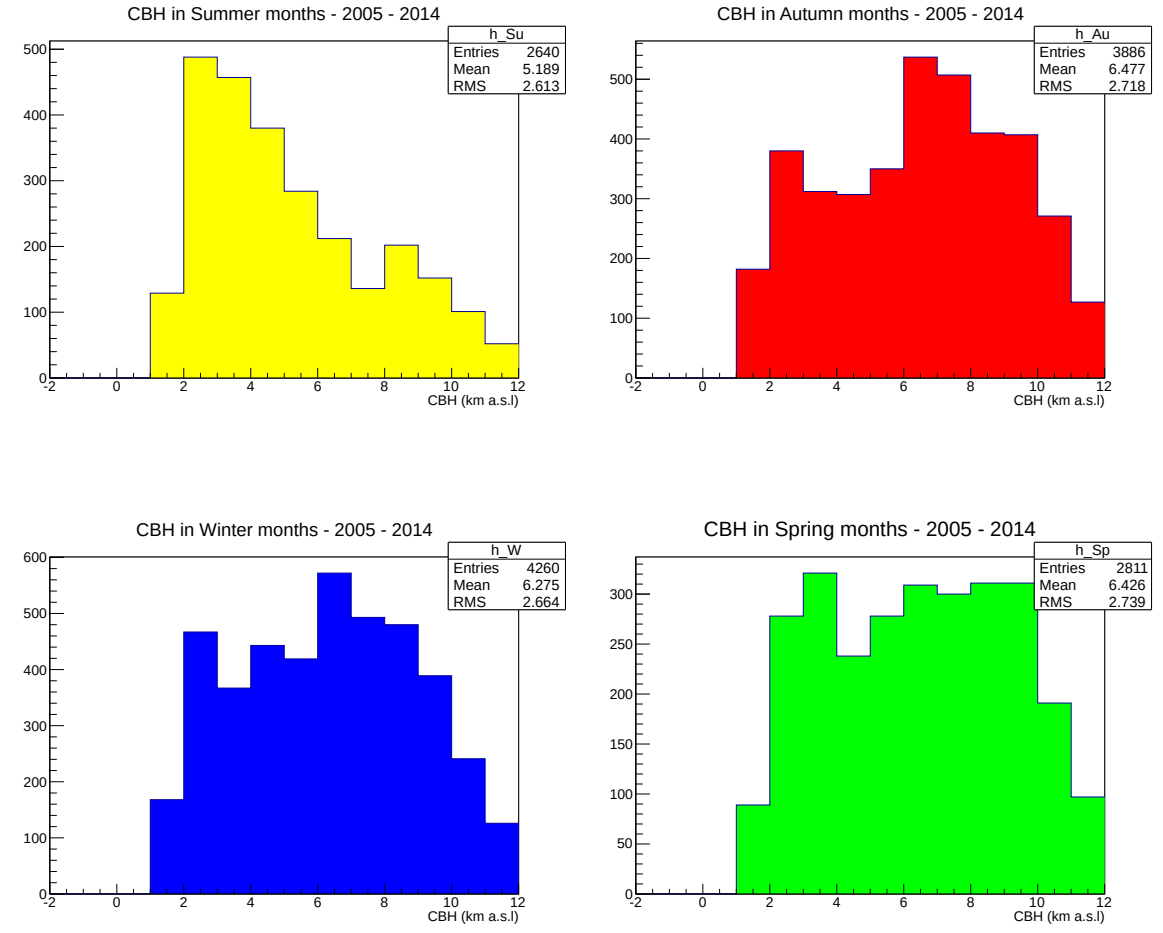

Figure 6.2: Cloud Base Height (above sea level) from all LIDARs, queried every hour during FD operation from 2005 to 2014. There are some distinctive features to investigate further. Here Summer is defined as December-February, Autumn is March-May, Winter is June-August and Spring is September-November. These definitions of the seasons will be the same throughout this chapter.

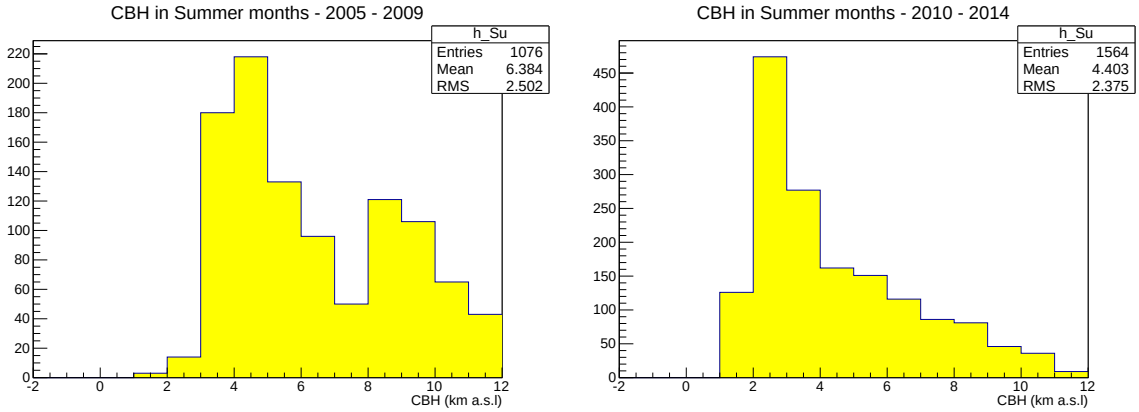

Figure 6.3: Cloud Base Height from all LIDARs, queried every hour from 2005 to 2009 (Left) and 2010-2014 (right) in the Summer months.

In looking at Figure 6.3, there is a noticeable difference between the two sets of data; however, this difference may not be significant. Whilst it may appear as though cloud heights are getting lower over time (on average), there are a number of reasons why the data could appear this way. A lack of statistics in the data can cause a skew or underestimation of the true nature of the data. Climate and weather events such as drought (which can last several years at a time) can influence the nature of cloud formation in the atmosphere, particularly in a dry region such as the location of the Pierre Auger Observatory in Argentina.

On the other hand, this apparent drift in cloud base heights could also be due to a decreased sensitivity of the LIDAR over time (as a result of the laser losing power). A study on the sensitivity of the LIDAR has been done to see if this general decrease is indeed caused by a systematic effect.

Firstly, the average cloud base height for each month of the year is investigated, to see if this decrease prevails over time. Looking at Figure 6.4, it appears that there is a slow drift in this mean CBH over time, consistent with Figure 6.3.

By splitting the data up from Figure 6.4, and isolating each FD site separately, the cause of this drift can be found. The site of most interest in this study is Coihueco. With this site being in closest proximity to the Andes mountain range, Coihueco is theorised to be the one with the most striking features, as discussed in Chapter 4. The average monthly cloud base height for Coihueco is shown by Figure 6.5.

Looking at this monthly average for Coihueco in Figure 6.5, this same drift that was noted earlier in the average cloud base height (see Figure 6.4) is still present. One wants to see if this same drift is also present at the other three sites. Figure 6.6 shows the average cloud base height over time for all four FD sites.

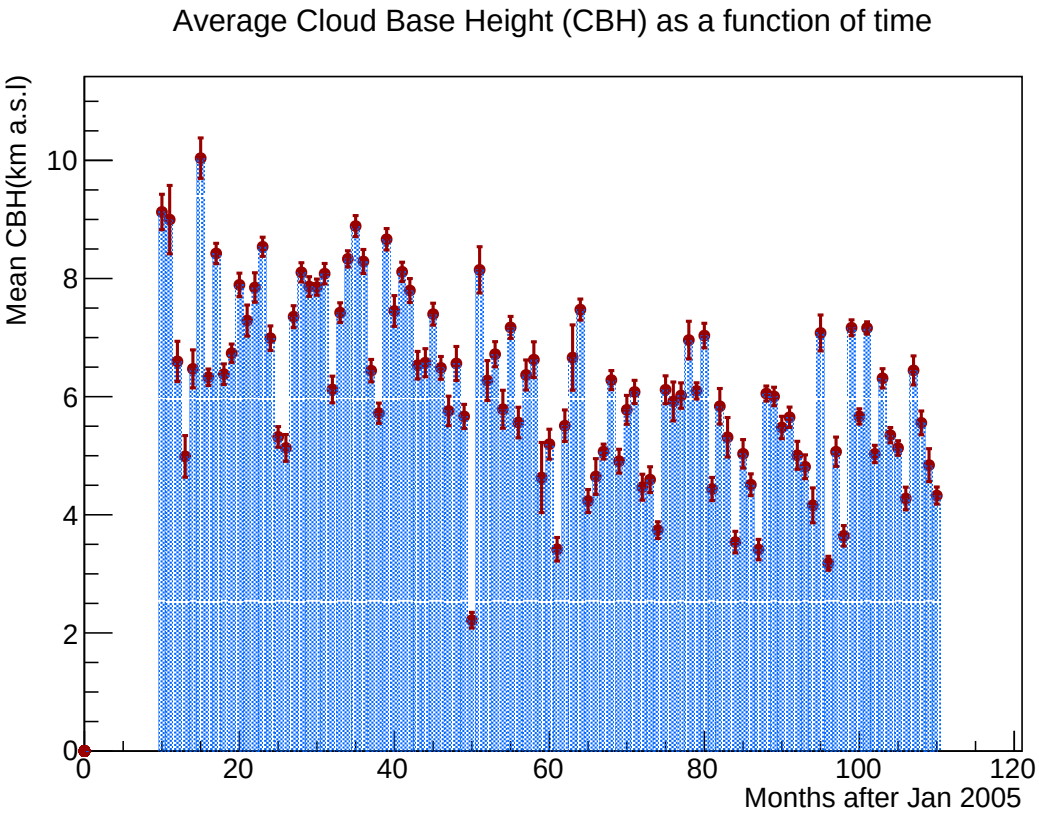

Figure 6.4: Average CBH for all sites, for each month between 2005 and 2014 based on LIDAR data. A slow drift in the average cloud base height calculated is noticeable here. One month (~ 50) exhibits a much lower CBH than the surrounding months, a systematic effect that needs to be explored further.

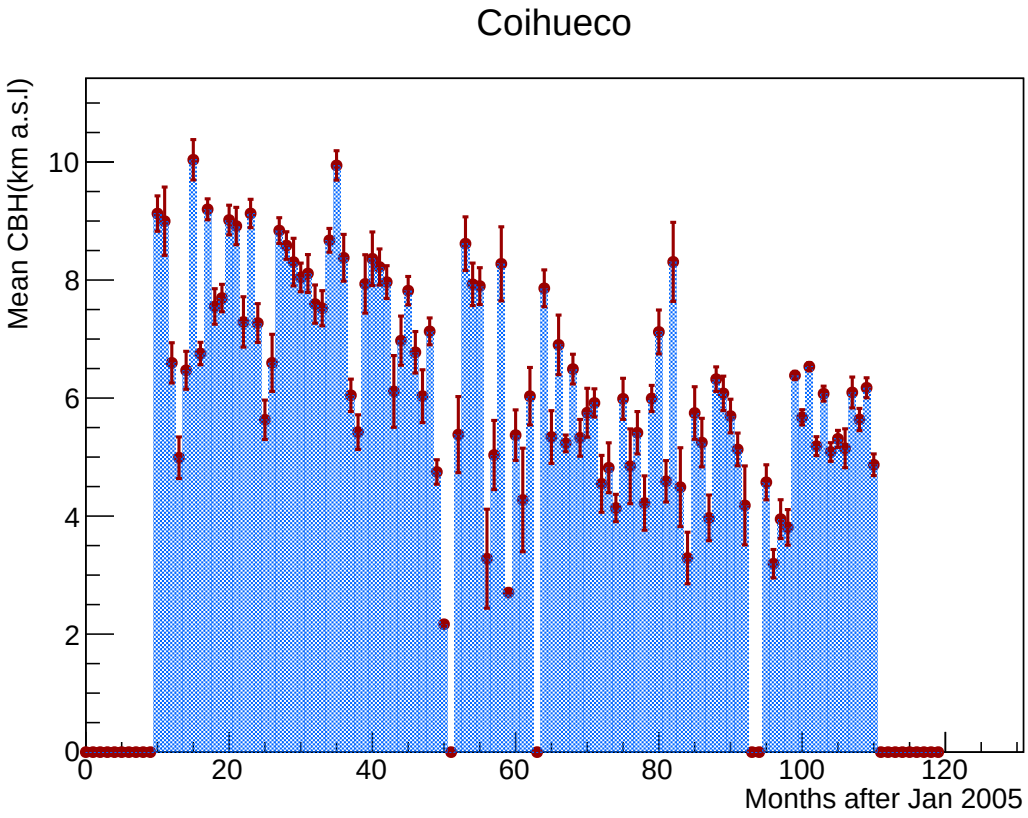

Figure 6.5: Average cloud base height (CBH) for the Coihueco site, for each month between 2005 and 2014 based on LIDAR data. The Coihueco LIDAR being unavailable in month 50 could explain the skew in the overall average CBH for the same month from Figure 6.4.

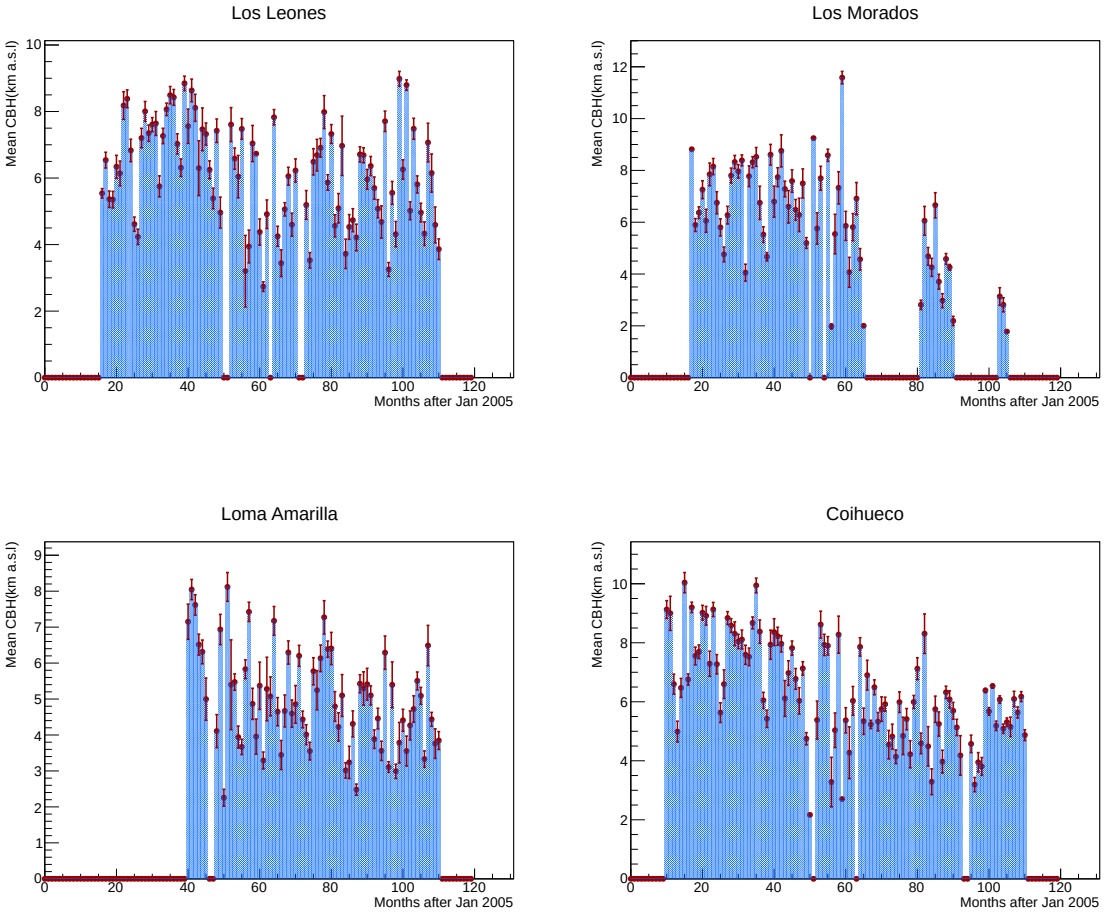

Figure 6.6: Average Cloud Base Height for each month for the LIDAR at Los Leones (top left), Los Morados (top right), Loma Amarilla (bottom left) and Coihueco (bottom right).

The results of Figure 6.6 illustrate that two of the FD sites (Los Leones, and Loma Amarilla) do not show this decrease in the average cloud base height over time as was seen at Coihueco. However, the same drift seen at Coihueco is also seen at Los Morados.

In order to check if the sensitivity of the LIDAR has affected these measurements, one must plot the standard deviation of the cloud base height measurements, searching for a similar increase/decrease. A decrease in the standard deviation over time suggests that there is a limit on the heights that can be measured by the laser, since the reduction in spread indicates that some heights are unable to be reached, due to a decrease in the laser power. Figure 6.7 shows how the spread evolves over time, specifically looking at the Los Morados and Coihueco FD sites (as these were the two sites that showed a drift in the cloud height).

Whilst the results of Figure 6.7 show no general increase or decrease in the standard deviation of cloud base height measurements over time, it is unable to be determined whether or not this is a noticeable atmospheric effect, or simply an effect of statistics or a random variation in the data.

Next, a check on the seasonal behaviour in the cloud coverage (CC) is required, which is discussed in Section 6.2.

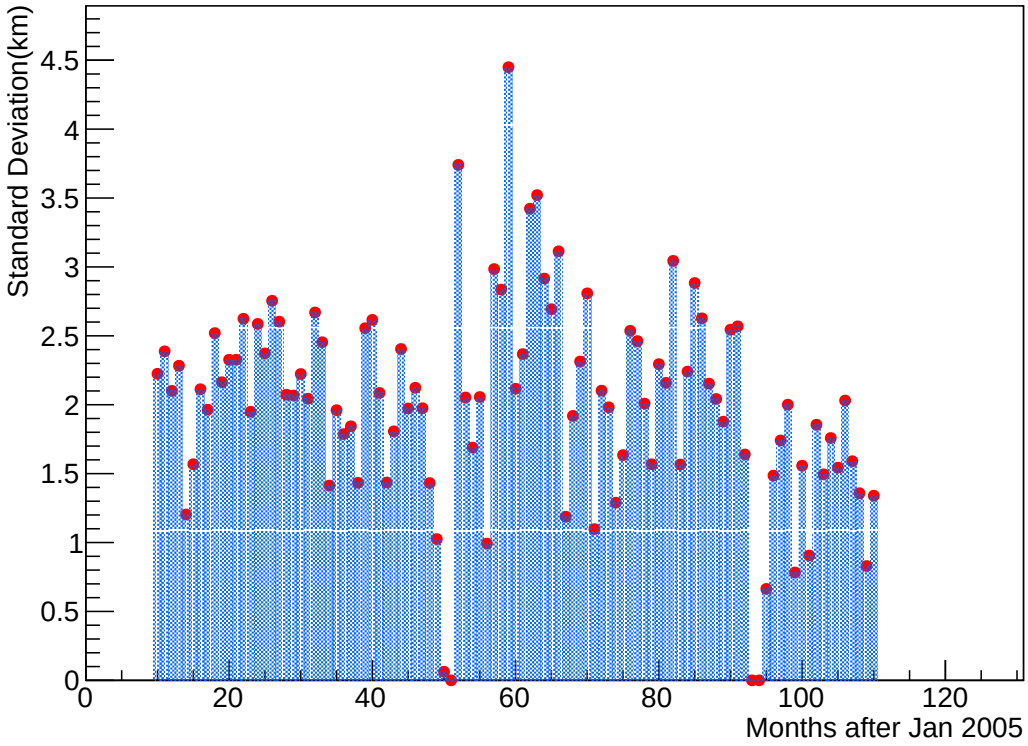

Figure 6.7: Standard deviation in CBH for each month between 2005 and 2014 for the Los Morados and Coihueco FD sites. No large increase or decrease in this spread is observed here.

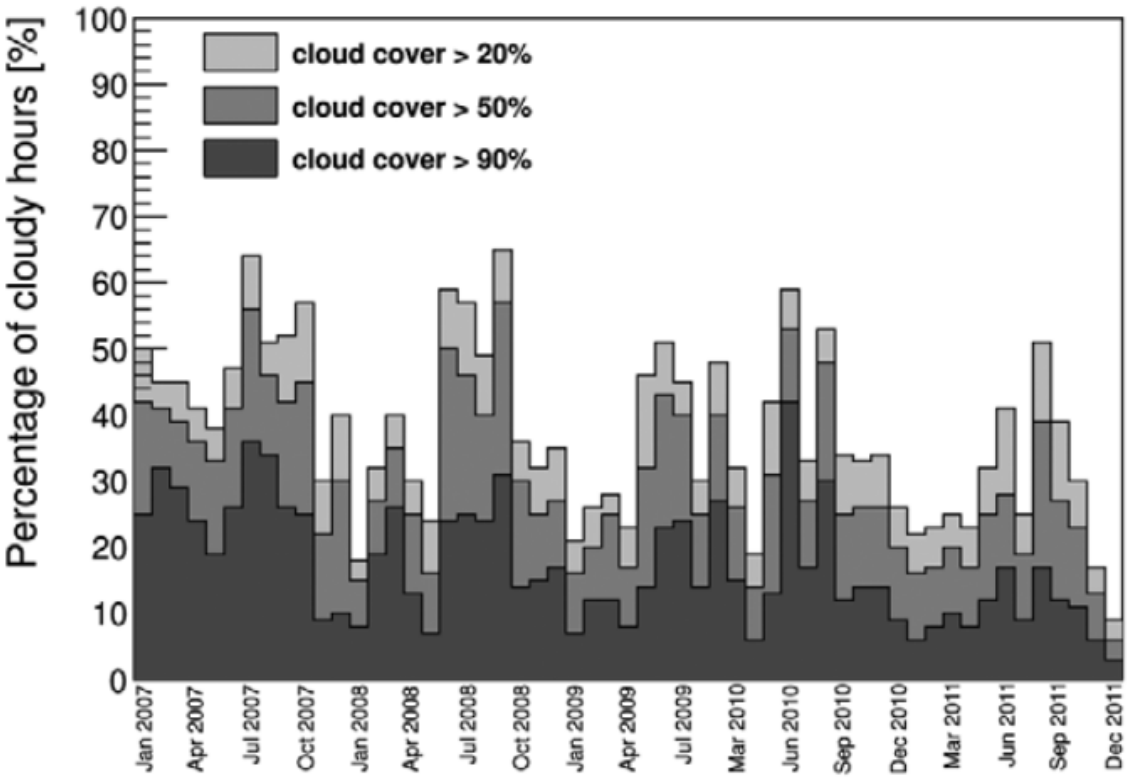

Figure 6.8: Monthly average cloud cover between 2007 and 2011 at Auger. Strong seasonal variations are observed, with cloud cover peaking in the winter months [43].

6.2 Cloud Coverage

In a similar way to the cloud base height described in section 6.1, the cloud coverage (in %) is recorded into the database for each LIDAR for each hour during data acquisition. Cloud coverage is defined as the fraction of sky that contains cloud, where the LIDAR performed the scan (see section 3.5.6 for more information). The scanning area of the laser forms a 45° inverted cone in the sky, and the entire volume is scanned over in a typical hour. Every 15 minute scan records a cloud coverage (0-100 %) and the database value for a given hour is the average cloud coverage of the four 15 minute scans. Cloud quality cuts on the FD reconstruction are done by using this calculated cloud coverage, namely that coverage values larger than 25 % for air shower events are rejected as the final reconstruction may have been influenced by the presence of a cloud. For more information on these cloud quality cuts see section 3.5.9.

A study by Rizi and Tonachini in 2011 [43] analysed the cloud coverage by splitting the data into three separate cloud coverage values ($> 20\%$, $> 50\%$ and $> 90\%$) for each month of the year between 2007 and 2011, using all four available LIDARs at Auger. Their results are shown in Figure 6.8.

As is shown in Figure 6.8, the highly overcast conditions (the black histogram in Figure 6.8) with a cloud coverage $> 90\%$, peak in the Winter time (around July/August). The goal is to replicate this study and see if the same pattern in the cloud coverage is observed. In addition, the study can be further extended in time (from 2005 up to 2014), to see if the same recognisable behaviour occurs over a longer period. To account for the "off" periods (i.e. when the Fluorescence Detectors were not running),

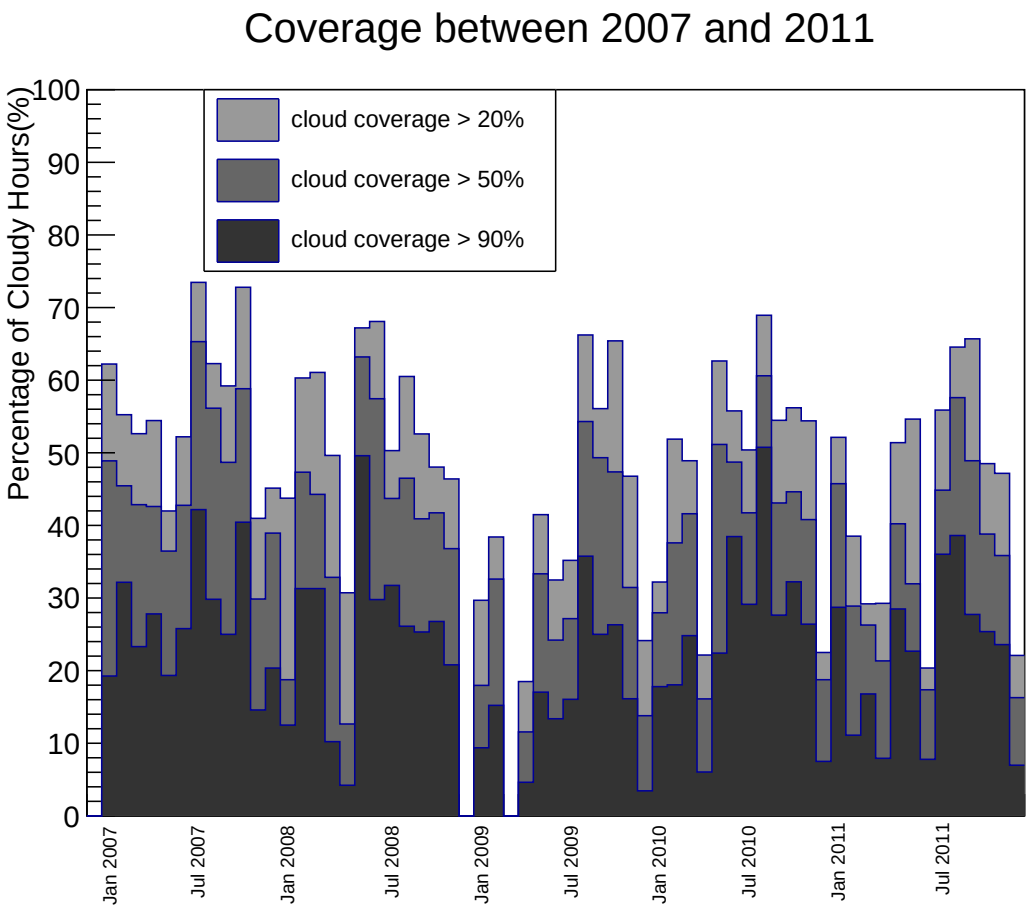

Figure 6.9: The same study on cloud coverage cuts (as Figure 6.8), repeated for the same time period (2007-2011). The results are similar to Figure 6.8, with the same periodic behaviour present (highly overcast conditions more frequent in Winter and less frequent in Summer). The January and July months are labelled to illustrate the difference between Summer and Winter.

each histogram is scaled by the total number of data acquisition hours. The histograms are filled based on whether the cloud coverage threshold was met in a particular hour, and scaled based on the number of times a cloud coverage was recorded (either the LIDAR reported clear conditions or detected a cloud).

The results of these are shown by Figures 6.9 and 6.10.

As can be seen from Figure 6.10, the extension of the study to later and earlier periods has shown the same prevailing pattern as found by Rizi and Tonachini [43]. There is consistency in the cloud coverage fluctuations between Figures 6.8 and 6.9.

As well as the seasonal variations in cloud cover, it is also interesting to look into the variations of cloud over the course of a single night. Of particular interest is the idea that cloud slowly builds up over time as the night grows colder, before dissipating towards the end of the night.

The same cloud cover cuts as done previously ($> 20\%$, $> 50\%$ and $> 90\%$) are now applied for each hour through the night for the Winter and Summer months. These results are shown in Figure 6.11. Note the use of UTC time as the reference time (local time at the Pierre Auger observatory is UTC -3 hours).

The winter months show a higher proportion of cloud when compared to the summer months.

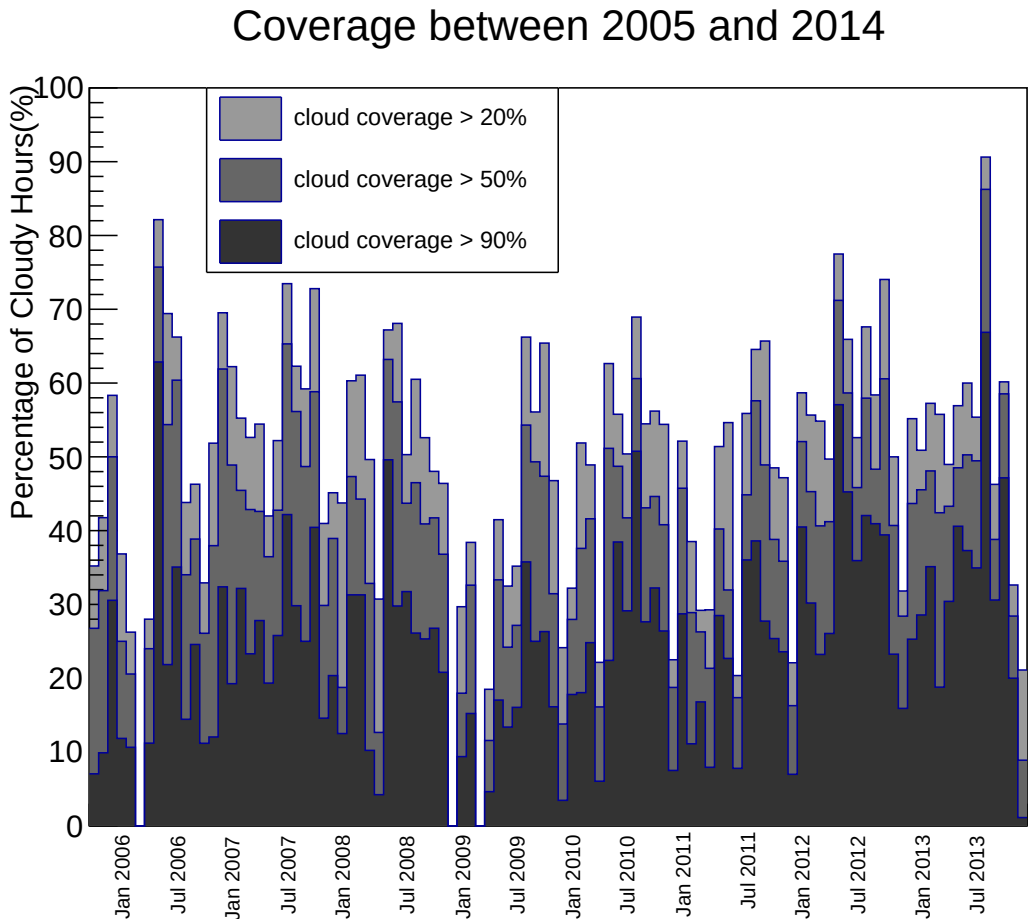

Figure 6.10: Cloud Coverage cuts study extended to range between 2005 and 2014.

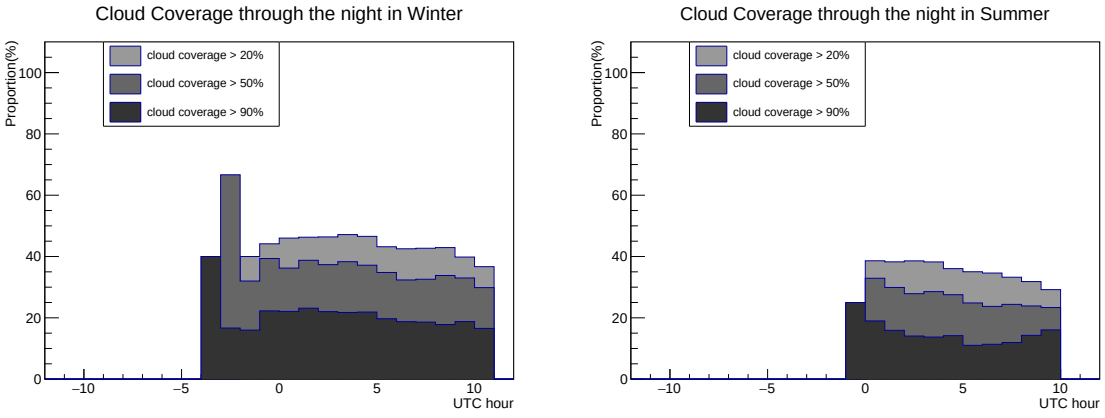

Figure 6.11: Cloud Coverage values for each hour of the night for the Winter months (left) and Summer months (right) between 2005 and 2014 for all LIDARs. Again, the histogram is scaled by the total run time of the LIDARs.

By looking at the winter months in Figure 6.11, the amount of cloud present at any particular time during the night appears to be constant. There is a spike seen in the first two bins for the winter months, however this is simply due to very low statistics during those hours, and is not a significant peak in the cloud coverage in winter. In the summer months, there does appear to be a slightly higher proportion of cloud earlier in the night. This can be investigated by comparing the clear times during the night in the summer months.

Figure 6.12 shows the clear times for each site during the summer months (December-February) over the course of the night, which appears to be consistently clear around $\sim 60\%$ of the time for all hours during FD data acquisition.

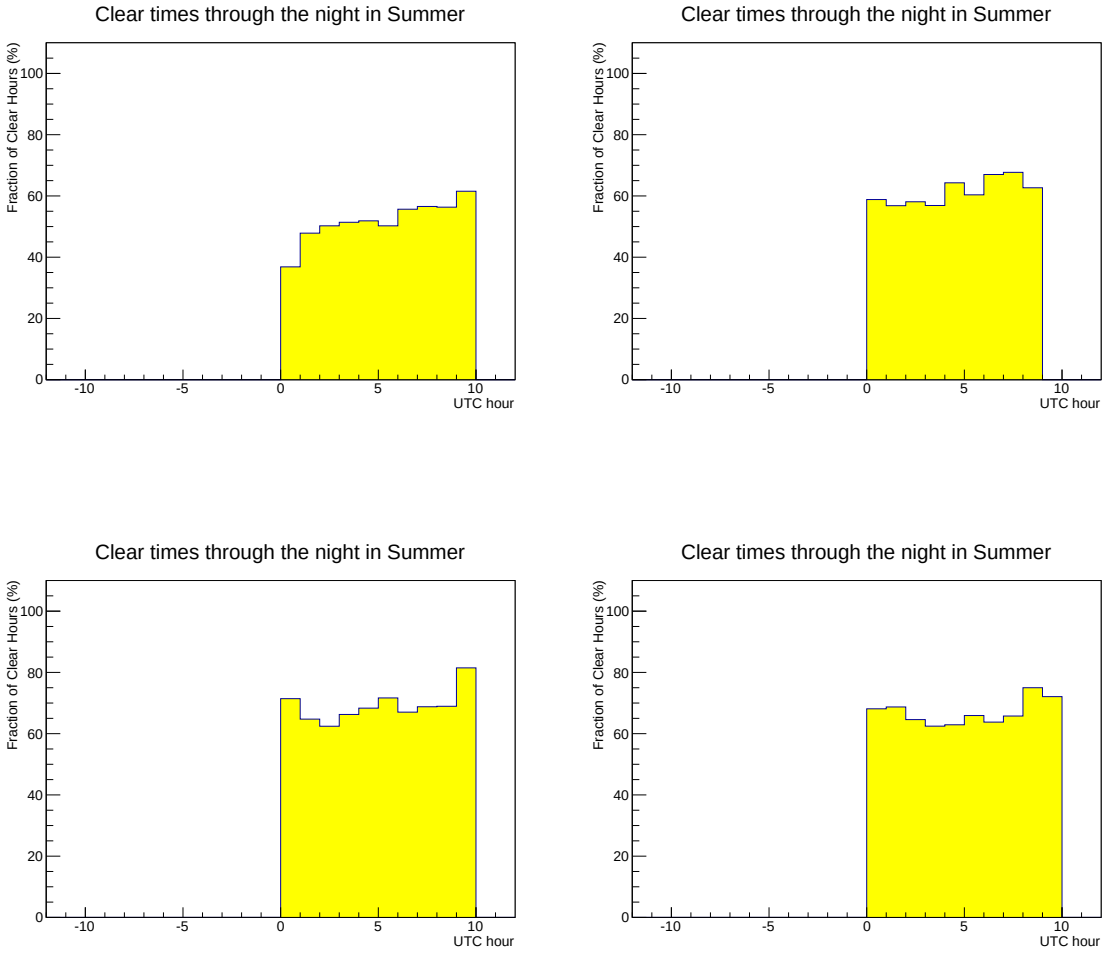

Figure 6.12: Clear hours through the night in the Summer months from 2005 to 2014 for the LIDARs at Los Leones (top left), Los Morados (top right), Loma Amarilla (bottom left) and Coihueco (bottom right).

Figure 6.12 shows that the level of cloud during the course of a typical night of FD operation will remain consistent. Since the atmosphere gets colder over the course of a night, one would expect that the temperature of the atmosphere should approach the dew point temperature, which would increase the likelihood of cloud formation. However, the data from Figures 6.11 and 6.12 do not reflect this fact.

One final check that one can do is to check the stability of the cloud coverage over time. To do so, each year is divided up into 12 months, with the average cloud

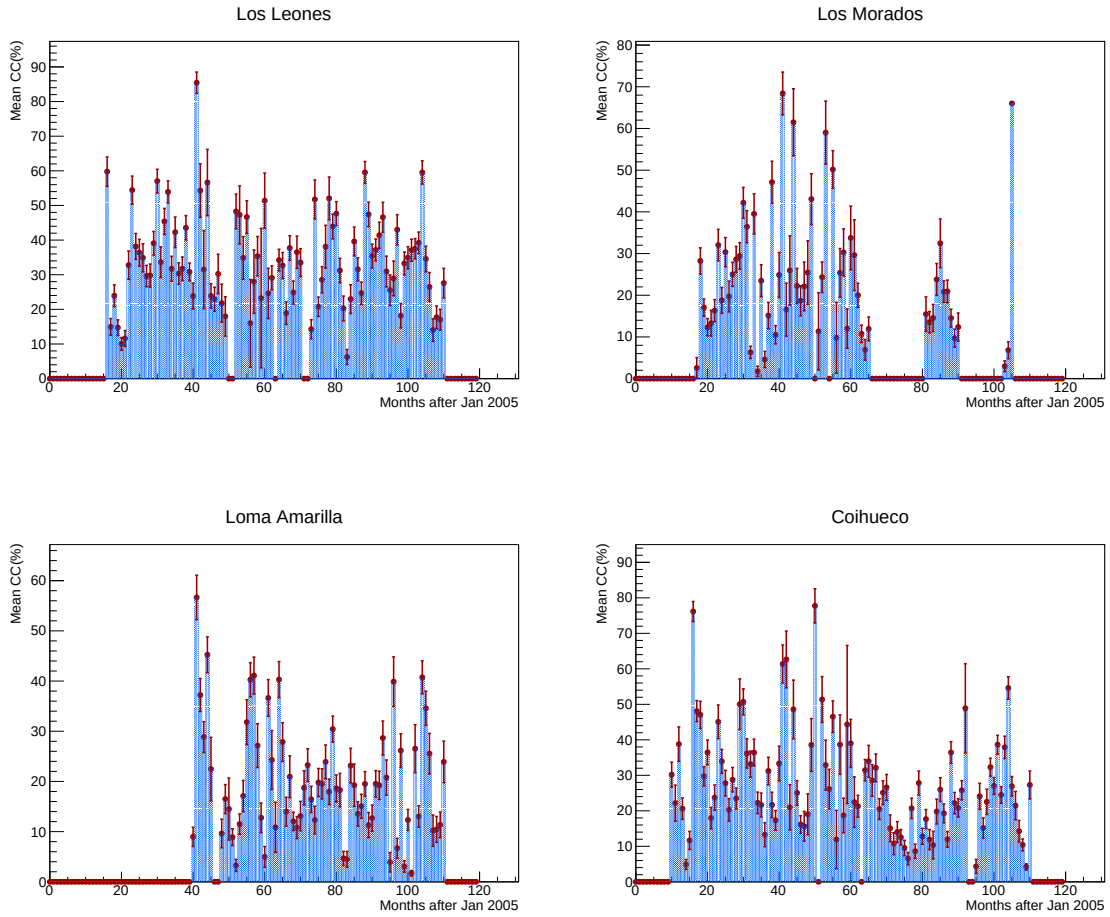

Figure 6.13: LIDAR Cloud Coverage for each month from 2005 to 2014 - to check for stability over time. Top left: Los Leones, Top right: Los Morados, Bottom left: Loma Amarilla and Bottom right: Coihueco.

coverage over each month taken and plotted as a function of time. The results are given in Figure 6.13 for each FD site.

The results of Figure 6.13 do not appear to follow with previous studies done (as far as seasonal variations go). However by zooming in further, the trends that are expected become more noticeable. Figure 6.14 shows a small portion of the LIDAR monthly cloud coverage at Coihueco.

When inspecting the results of Figure 6.14, one can see that even though climate properties can wildly fluctuate year-to-year, the underlying trend in cloud cover is present. A higher average cloud cover value in the Winter months agrees with the study done previously (Figure 6.18). The peak in the Summer of 2007 (around month 24) does not appear consistent with this general trend. A possible reason for this is due to there being more water vapour available in the atmosphere during Summer (due to the warmer temperatures); thus, there is a higher probability of cloud formation as there is a higher chance of condensation forming droplets and coalescing into cloud. This behaviour doesn't repeat in either adjacent Summer however, suggesting another potential cause. It is possible that 2007 was a particularly cloudy year, as atmospheric conditions can be difficult to predict and are generally unstable, which does not form patterns over time (as shown by Figure 6.14).

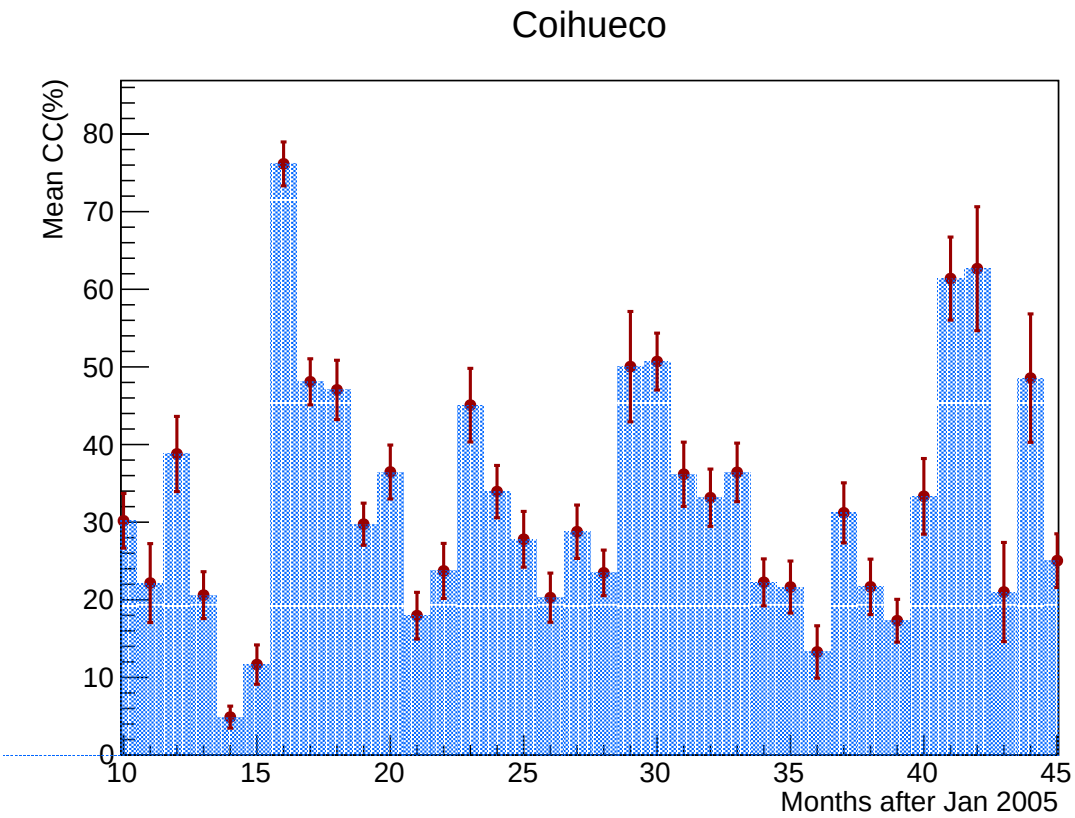

Figure 6.14: Zoomed picture of Coihueco results from Figure 6.13. Three distinct peaks in the mean cloud cover are present, each roughly 12 months apart and all corresponding to winter months, as expected. There is also a significant peak at month 24, which corresponds to summer 2007.

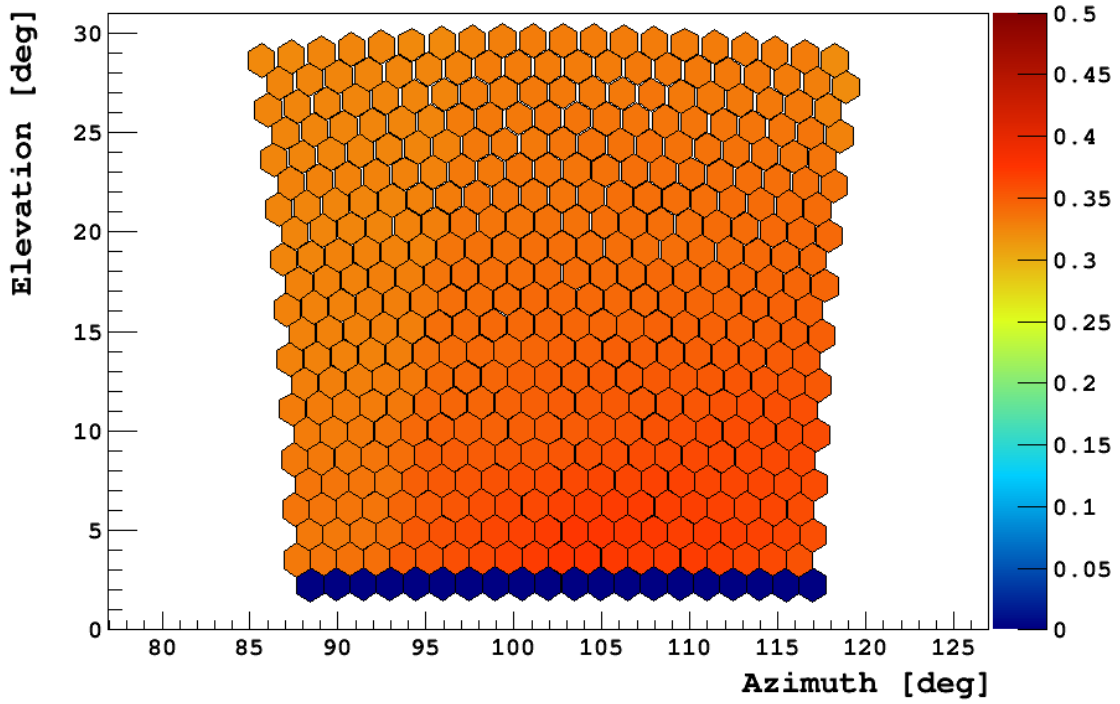

Figure 6.15: Cloud Fraction in Mirror 5 at Coihueco averaged for 2017

Another interesting property that can be studied is known as the cloud fraction (see section 5.4), which is calculated using the IR cloud cameras instead of the LIDAR. The seasonal study on cloud fraction is discussed in section 6.3.

6.3 Cloud Fraction

The cloud fraction value for each cloud camera pixel (described in chapter 5) is calculated for every 5 minute cloud camera scan. The aim here is to investigate the seasonal dependence of this cloud fraction over a long period of time. There are known issues with the bottom row of the FD camera pixels due to radiation from the ground creating a warm horizon, with water vapour absorption and emission close to the ground creating an optically thick atmosphere (as the water vapour concentration is highest close to the ground as mentioned in Chapter 3). Thus, warm clouds cannot be distinguished when close to the horizon. As a result, this bottom row is set to be clear in order to avoid false detection of cloud.

The analysis on the cloud camera is then done by extracting the cloud fraction value in each pixel for each telescope of each FD, before averaging the total cloud fraction for each FD pixel. The goal is to analyse the trend in this average cloud fraction over time, to see what behaviours are present.

Looking at the average cloud fraction per year for selected sites, telescopes, and years, some interesting patterns emerge - as an example see (Figure 6.15), the average cloud fraction in telescope 5 at Coihueco for 2017.

The most obvious behaviour observed in Figure 6.15 is that the average fraction slowly increases with decreasing elevation angle (the bottom section of the camera is

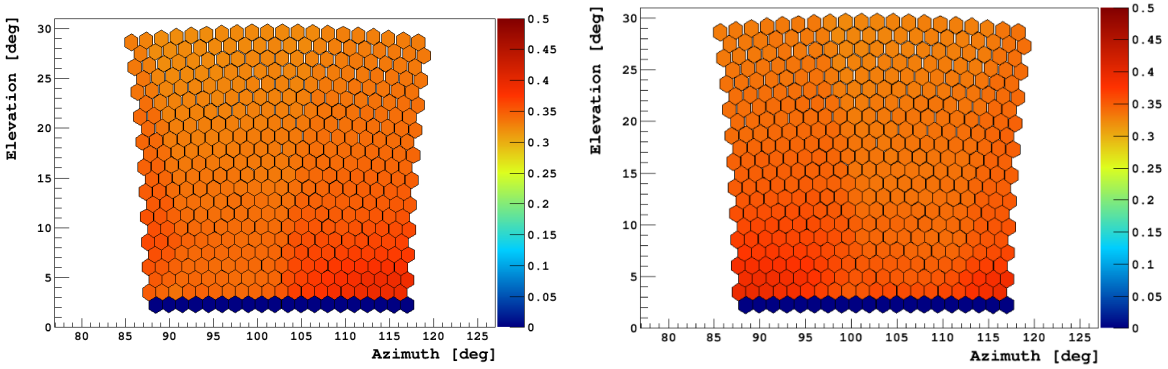

Figure 6.16: Cloud Fraction for each pixel of Mirror 4 (left) and Mirror 3 (right) at Coihueco averaged for 2017.

more obscured than the top). Looking at several different telescopes the same patterns emerge - see Figure 6.16. These are telescopes 3 and 4 from Coihueco during the same year (2017).

As can be seen from Figure 6.16, the overall trend leans towards a higher amount of cloud in the bottom rows of the camera. This trend does go according to expectations. Since the camera is looking through a larger portion of the atmosphere at lower angles, the probability of seeing a cloud at these lower angles is much higher. As a result, it is reasonable to expect that more cloud is seen at lower elevations.

Another interesting pattern is observed in the data in Figures 6.15 and 6.16. There is an asymmetry from left to right in the average cloud fraction values, due to the overlap between different images of the sky taken in sequence by the cloud camera. Since there are 5 images taken in the FOV scan (with 6 FD telescopes), each individual image spans more than an individual telescope. This results in a column of the FD pixels forming the boundary between two consecutive images in the cloud camera scan - which are calibrated differently [48]. The calibration of each cloud camera (known as "flat-fielding" (see section 5.2)) is different since they are independent of each other, thus each image will have a different correction applied to it [48].

As a result, the proportion of cloud present in each image will be slightly different, leading to the slight variations in cloud fraction between images from left to right as observed in Figures 6.15 and 6.16.

Lastly, the seasonal variations in this cloud fraction value are investigated. This is done by calculating the average pixel value for each FD telescope at a given time (for example, if half of all pixels were cloudy it would return a cloud fraction of 50%). A comparison is also done with the cloud coverage study given by the LIDAR (see section 6.2) - to see if similar seasonal patterns are observed with the cloud fractions, as was observed in the cloud coverage. Histograms showing every average cloud fraction value for every season in 2016/2017 are shown in Figure 6.17 for Coihueco.

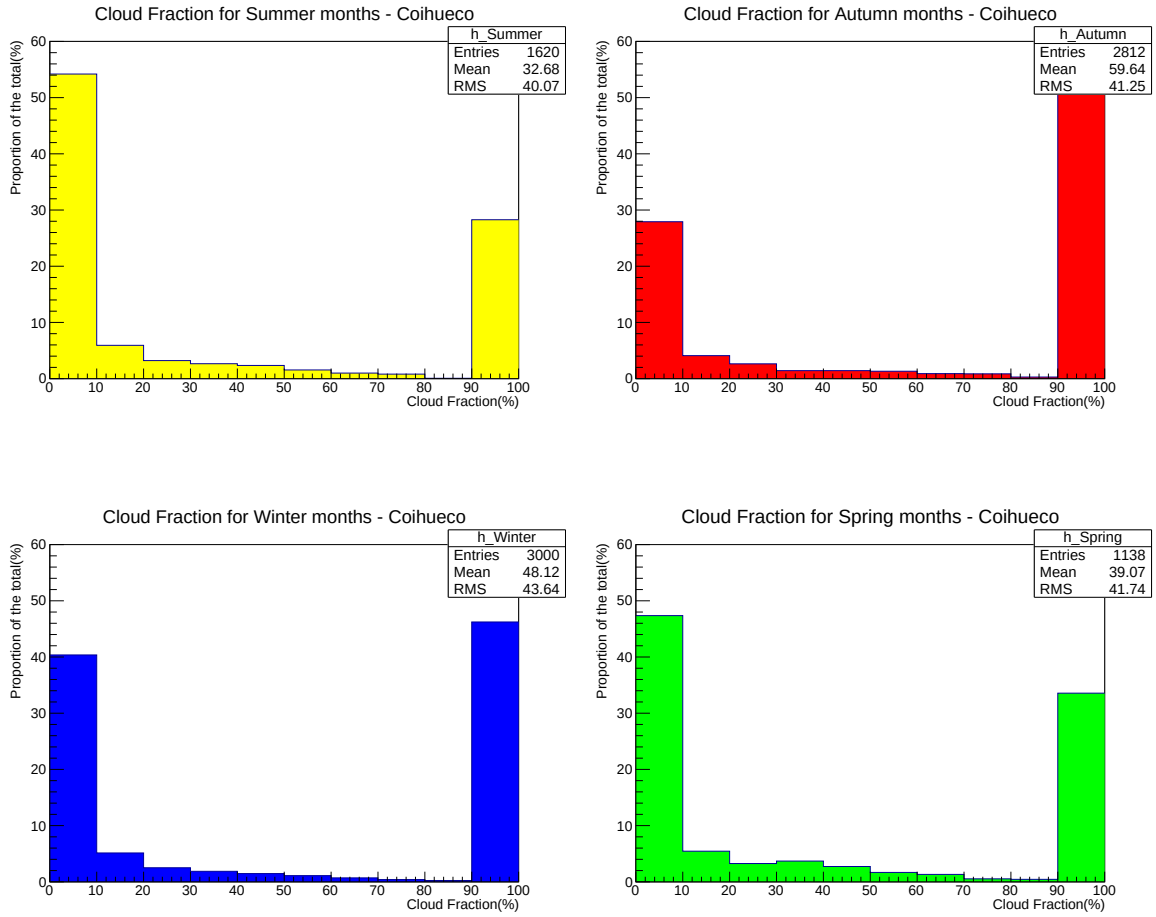

Figure 6.17: Proportion of the average cloud fraction for each season in 2016/2017 - as viewed by the Coihueco Cloud Camera. Each graph is normalised such that each bin makes up a percentage of the total.

Upon looking at Figure 6.17, summer is the clearest season, as expected (over 50% of all times are clear with less than 10% cloud cover), with Autumn being the cloudiest season. Secondly, the data shows that conditions are preferentially either totally clear (cloud fraction $< 10\%$ or totally cloudy (cloud fraction $> 90\%$). The proportion of time where these conditions are met is around 70-80% for each season.

Upon comparing summer with winter, the average cloud fraction overall is around 15% higher in winter. In particular, the proportion of overcast conditions is higher in winter compared to summer - this matches the study done with the LIDAR (see Section 3.5.6). A comparison can be made between the cloud fraction as seen by the IR cameras, and the cloud coverage given by the LIDAR. An analysis on the cloud coverage for every hour given by the LIDAR (for each season) is shown in Figure 6.18.

As can be seen by Figure 6.18, winter months produce the most overall cloud (with 30% cloud cover on average) whereas summer months produce the least cloud (21% cloud cover on average). This is what is expected from previous studies, and is consistent between the two separate sets of data. Interestingly, for this study, the average difference in cloud coverage between summer and winter is around 9% (compared to 15% for the cloud camera). This discrepancy in the proportions is most likely due to

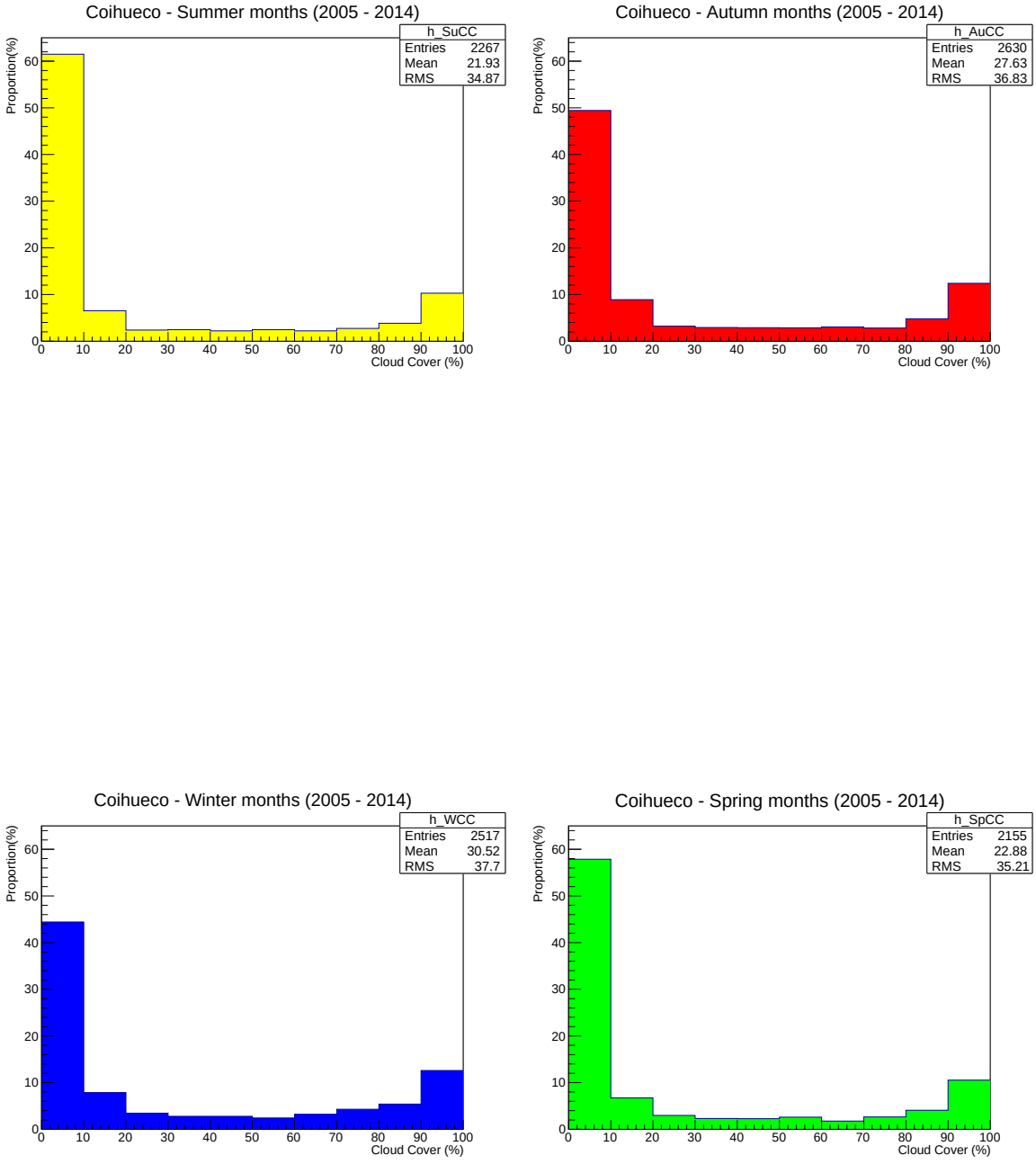

Figure 6.18: LIDAR Cloud Coverage for each season from 2005 to 2014 - as viewed by Coihueco. Each graph is normalised such that the sum of all bins gives 100% proportion.

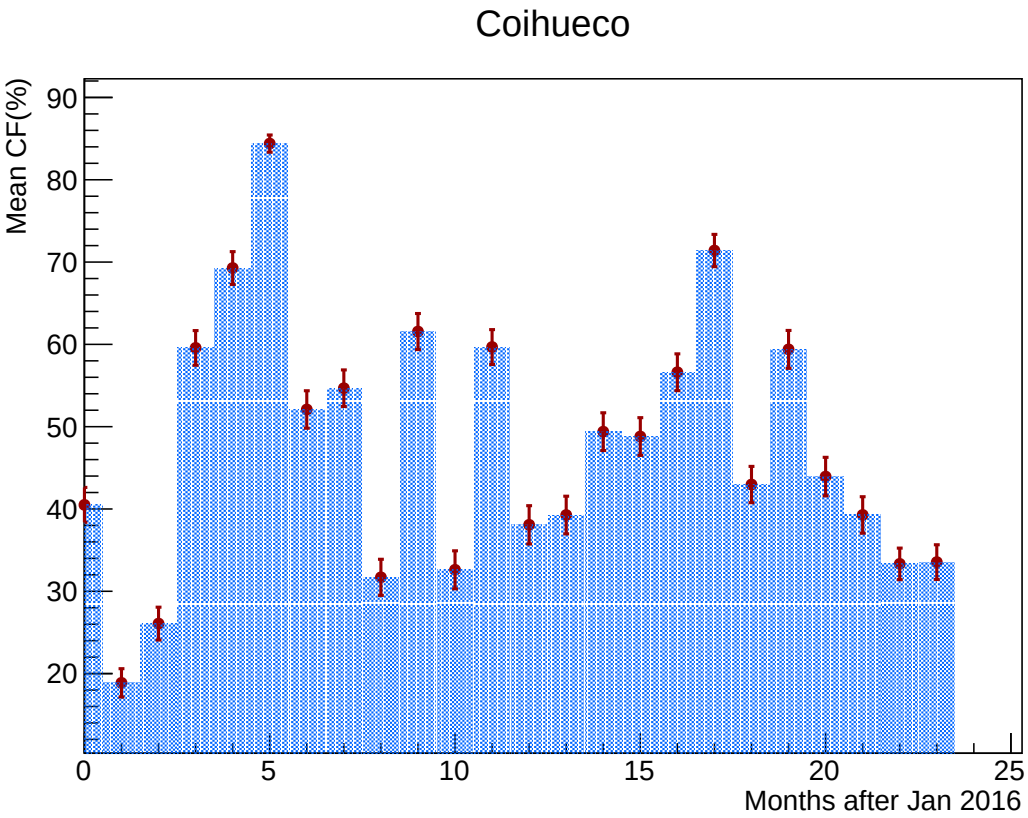

Figure 6.19: Cloud Fraction for 2016/2017 month by month at Coihueco (from the Cloud Camera). Two clear peaks are present at around 5 months and 17 months, which corresponds to the beginning of Winter for each year.

the difference in statistics between the two data sets. LIDAR analysis is over ten years (2005-2014), compared to the cloud camera analysis which is only over two years (2016 and 2017).

Another major difference between the cloud camera and the LIDAR is the projection of the atmosphere (i.e. the amount of atmosphere seen by each instrument). The cloud camera field of view is out over the array (compared to the LIDAR which scans vertically), meaning that for a typical cloud camera scan, more of the atmosphere is seen. As a result, the camera is more likely to see a cloud, which is shown by the higher cloud fraction values (on average) in comparison to the LIDAR.

The final check in order to round out this study is to investigate the stability of the cloud camera over time. Using the same method as described above, one can take the average cloud fraction over all pixels in each FD telescope per month as a function of time to check this stability. The results are shown in Figure 6.19.

As was shown by Figure 6.14, Figure 6.19 shows the average cloud fraction is stable and peaking around the winter months. This is exactly what was expected and rounds out the seasonal cloud studies.

6.4 Summary

The seasonal variation in cloud properties over time was investigated. It was able to be shown that there is a clear distinction between the conditions experienced at Auger in Summer compared to Winter, with the study revealing that clouds are more prevalent in the Winter (and Autumn) months. However, when looking over the course of a night of FD operation, the amount of cloud cover was shown to be consistent over the entire night, with no real change observed between the beginning and end of a night. It was expected that as the night gets colder, air temperature would more closely match the dew point temperature, thus leading to a higher likelihood of cloud formation towards the end of the night. However, the observed data did not reflect this.

Chapter 7

Simulating the Cloud Camera Response and the effect of Cirrus Cloud in MODTRAN

The IR cloud cameras installed at the Auger site have a spectral response used to determine the location of clouds within the field of view of the FDs. Thick cloud in the infrared is optically thick (acting as a blackbody at temperatures of $\sim 300\,\mathrm{K}$); therefore it is easily imaged by the cloud camera. However, a cirrus cloud is optically thin in the infrared, meaning the cloud camera has difficulty imaging them.

As a result, one wishes to simulate the response of the cloud camera to cirrus cloud, in order to better understand why there are difficulties in imaging this sort of cloud, as well as the effects they have on the analysis of extensive air showers. This is done using MODTRAN. The basics of MODTRAN will be described in section 7.1, before a discussion on the response of the IR cloud camera to cirrus cloud.

7.1 MODTRAN

MODTRAN (MODerate resolution atmospheric TRANsmission) is a software program designed to model atmospheric radiation over a wide spectral range. The program was developed and continues to be maintained through a collaboration between Spectral Sciences, Inc. and the U.S. Air Force [48]. MODTRAN is adopted in government agencies, educational institutions and other commercial organisations worldwide to predict and analyse optical measurements through the atmosphere. The main purpose of MODTRAN is for use in remote sensed multi-spectral and hyperspectral imaging (MSI/HSI). For the purposes of this thesis, MODTRAN version 5.2 is the model that is used [86]. In MODTRAN, line of sight (LOS) spectral radiance and transmittance are calculated in the wavenumber range $0 - 50,000\,\mathrm{cm}^{-1}$, which covers the ultraviolet to longwave infrared wavelength range of the electromagnetic spectrum [87] [88]. Radiation Transport physics within MODTRAN is accurate in modelling horizontally homogeneous atmospheres [86].

MODTRAN is written in the FORTRAN computing language and is structured by putting all variables into "cards". Card 1 sets the model atmosphere, as well as the output (i.e. in either radiance or transmittance) [89]. There are also other atmospheric parameters that can be set, including multiple scattering (on or off). The main function (at least for the purposes of this dissertation) of Card 1A is to set the ground level (in

```
1  M   2    3    1    0    0    0    0    0    0    0    0    0    0    .000    .00
2  f   8    0   385.000  1.000000  1.00000  f t f
3  01_2008
4     -1    0    0    0   18    0       .000       .000       .000       .000      1.420
5     1.000   4.00  0
6        .000                75.000
7              2         50          1          2rm         mlaa
8     0
```

Figure 7.1: Example of a set of cards in MODTRAN. Each line in the file corresponds to a different card.

km above sea level - a.s.l), which is 0 km by default but is also adjusted to the altitude of the Auger observatory (which is at 1420 m a.s.l on average).

Card 2 sets the aerosol and/or cloud model that you wish to use. Positive integers here will cycle through a few options for aerosol models, including Rural, Urban, Maritime (over ocean), Desert and Tropospheric [86]. Since the observatory is located in a rural environment, the most often picked aerosol models for the analysis in the rest of this chapter are the Rural - 5 km VIS and Rural - 23 km VIS options. VIS is short for visibility, i.e. the amount of atmosphere you can see through (in km) for the given aerosol distribution. In this case, 5 km VIS means a dense aerosol distribution, whilst 23 km VIS means low [86].

An example of the cards used in MODTRAN is shown in Figure 7.1.

For the most part, cards 1 and 1A are left unchanged as shown in Figure 7.1. The first integer (a 2 in this case) sets the atmosphere as Mid Latitude Summer - a 3 would indicate Mid Latitude Winter (45°). This mid latitude model is chosen as it is closest to the conditions that are present at the observatory in Malargue (38°S latitude).

Card 2 (line 4 in Figure 7.1) sets the aerosol and cloud options in the model. After setting the aerosol model in Card 2, the next number of interest is the 5th integer (18 in this case). This number sets the cloud/rain model that is being used. There are 6 main options for different types of cloud which are used for the purposes of this analysis. The first 5 are all preset clouds with a fixed altitude and thickness. Some examples of these clouds are stratus, cumulus and nimbostratus, all of which are described in chapter 4.

For a cirrus cloud, card 2A contains 3 values - these set the cloud thickness, cloud height (above ground) and extinction coefficient. The extinction coefficient is defined as the ratio of attenuated light to transmitted light at a given wavelength (which is set in MODTRAN). For the extinction, there are 2 options. One is to set your own value for the extinction, or by leaving the column as a 0 it gets preset to $0.14 * T$, where T is the cloud thickness (in km).

Now that we have the basis of MODTRAN, we can begin to do some cloud modelling with the tools available. The infrared cameras at Auger have difficulty in detecting cirrus cloud (as discussed at the beginning of this chapter). As a result, we turn to MODTRAN to model the camera response to cirrus cloud. Section 7.2 describes the study of this cirrus cloud in MODTRAN.

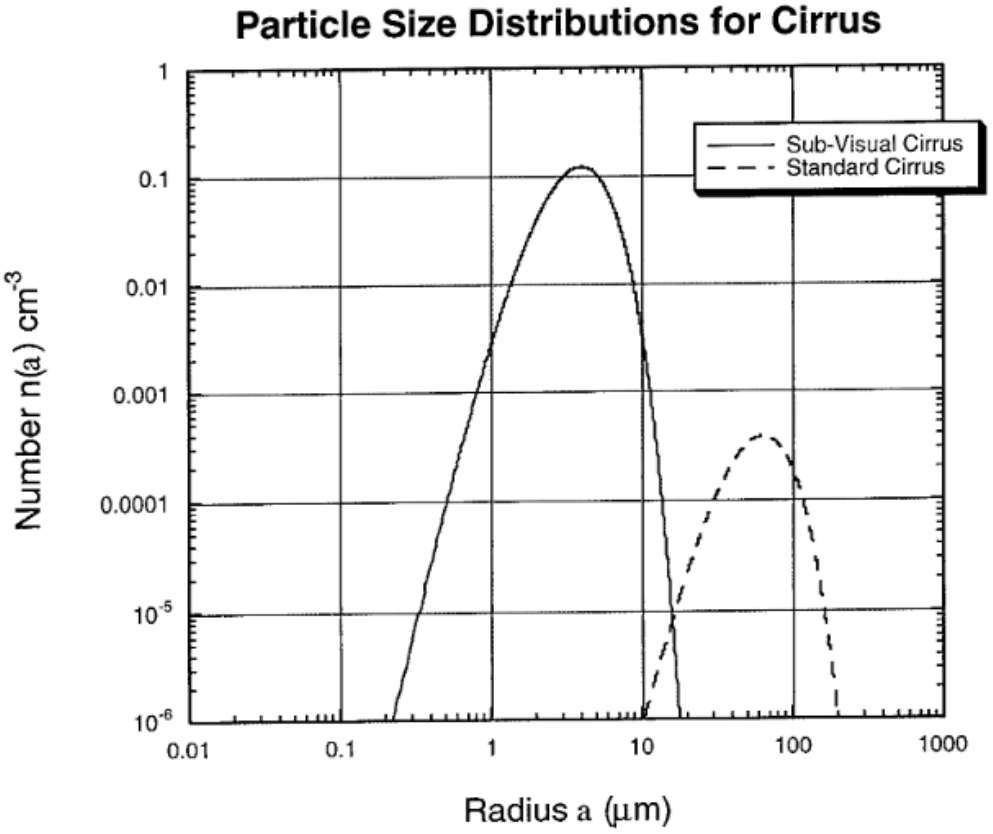

Figure 1. The particle size distribution for the subvisual model contains more particles cm^{-3}, and they are on average much smaller (a_{mode} = 4 µm) than for the standard model (a_{mode} = 64 µm).

Figure 7.2: Number Density of particles in the two Cirrus models in MODTRAN. [90]

7.2 Models of Cirrus Cloud

Cirrus clouds in the FOV of the Cloud Cameras are not as well understood. Since this type of cloud is high altitude, comprised mainly of ice crystals (see Section 4), it has different radiative properties to thick, low altitude cloud. Using MODTRAN, we can simulate the radiance emitted from such a cloud. There are two different models of Cirrus cloud in MODTRAN, which are defined as sub-visual and standard cirrus models. The only major differences between the two models are the mean ice particle sizes (64 μm for the standard and 4 μm for the sub-visual model) and the corresponding number densities of particles. The number density as a function of particle size is shown in Figure 7.2.

For this simulation, three parameters are set: Cloud altitude (above ground level - a.g.l in km), cloud thickness (in km) and extinction coefficient (in km^{-1}). For the extinction, MODTRAN employs two different settings. Either the value can be set to whatever is chosen, or a preset option can be used. In this preset option, the extinction coefficient is set to be a constant value C (where $C = 0.14\,km^{-2}$) multiplied by the cloud thickness, at a fixed wavelength of 550 nm. This comes from an NOAA study done in the 1980s, which related optical depth (τ), extinction coefficient (K_e) and cloud thickness (L) as in equation 7.1 [90]. Here, $\frac{1}{\tau} = 0.14$ for the reference wavelength of 550 nm. For the purposes of these simulations, the preset extinction will be used.

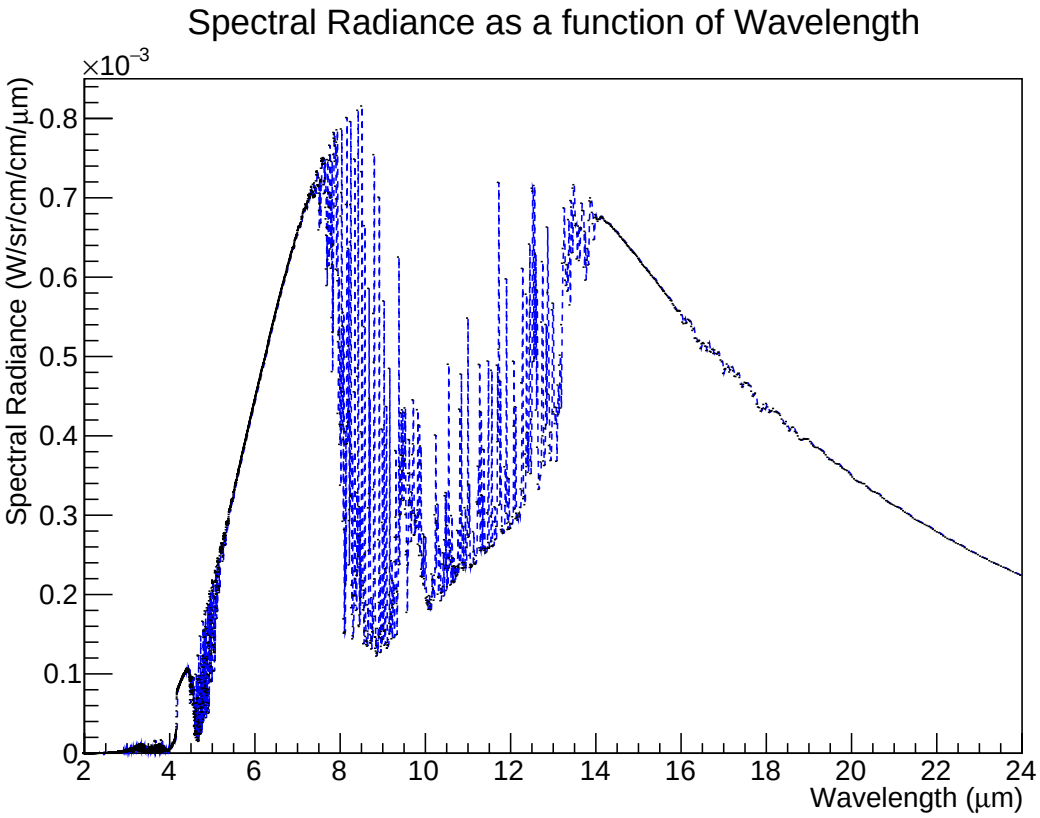

Figure 7.3: Spectral radiance as a function of wavelength for an atmosphere with a cirrus cloud overhead. The parameters of the cloud are: Altitude of 6 km, thickness of 0.4 km, and extinction coefficient (at 550 nm) of 0.14*(thickness). The atmospheric model used was Mid Latitude Summer which has a ground temperature of 294.2 K.

$$K_e = \left(\frac{1}{\tau}\right) L \tag{7.1}$$

Once the values have been set, we can plot out the radiance as a function of wavelength, as shown in Figure 7.3.

Given this spectral radiance curve, we can convert to a cirrus cloud brightness temperature by using the Stefan-Boltzmann Law.

$$R = \sigma T^4 \tag{7.2}$$

where σ is the Stefan-Boltzmann constant ($5.67 \times 10^{-8}\, W/m^2/K^4$), and R is the radiance for a given absolute temperature T. In order to get the brightness temperature of the object (in this case a cirrus cloud) we must integrate the radiance curve. These are shown in equations 7.3 and 7.4.

$$R_{tot} = \int_0^\infty R_\lambda d\lambda \tag{7.3}$$

$$T_{sky}^4 = \frac{\pi R_{tot}}{\sigma} \tag{7.4}$$

Since the Auger cloud cameras operate in the $7 - 14\,\mu$m range, the integration will be performed in that range. The radiance (from MODTRAN) is given per unit solid

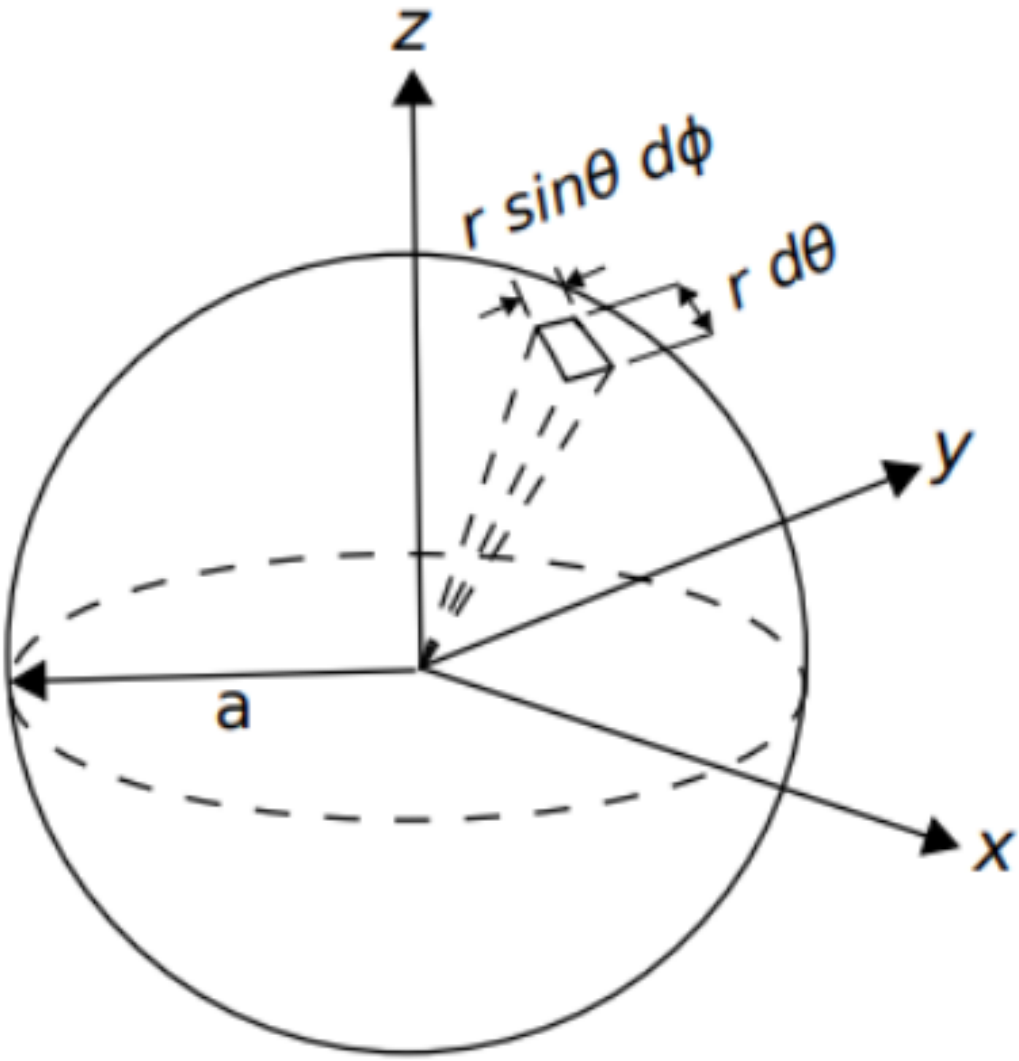

Figure 7.4: Element area dA of a sphere. Source: [91]

angle and per unit area. As a result, the integration must be done over the solid angle being subtended for an area element. For a hemisphere, the solid angle subtended is 2π.

One wants to integrate the radiance over the entire hemisphere, which is done by first calculating the element area of the sphere (denoted dA) given by equation 7.5

$$dA = r^2 \sin\theta \, d\theta \, d\phi \tag{7.5}$$

in spherical co-ordinates, as shown in Figure 7.4.

For this case, one can set $r = 1$ and perform the integration over zenith angle $\theta \in [0, \frac{\pi}{2}]$ and azimuth angle $\phi \in [0, 2\pi]$. This is the area A on the sphere. Assuming the detector has an area $A^{'}$, then from the zenith the detector will look like it has an area $A^{'}$. However, viewed at a zenith angle of θ, this detector area will be $A^{'} \cos\theta$ (where $\cos\theta = 1$ at the zenith). As a result, integrating over the detector area over these angles gives the result shown in equation 7.6 [92].

$$\int A^{'} = \int \cos\,\theta \, dA = \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \sin\theta \cos\theta \, d\theta \, d\phi = \pi \tag{7.6}$$

This is where the factor of π (from equation 7.4) comes from, to calculate the radiance in units of W/m^2 and hence, get the correct units for the brightness temperature.

Once this brightness temperature is calculated, it can be plotted as a function of the observer zenith angle. The last step is then to calibrate the results from MODTRAN by equation 7.7.

$$\tau(T_{MOD}^4) + (1-\tau)(T_{ground}^4) = T_{Cam}^4 \tag{7.7}$$

Here τ is the transmission of the camera (defined as the fraction of light transmitted through the lens and window), T_{ground} is the ground temperature of the MODTRAN model, and T_{MOD} is the sky temperature for a given cloud camera temperature T_{Cam}.

The cloud cameras at Auger measure radiation from two sources; one being the radiation from the sky, and the other being the radiation from the optics of the camera itself. The first term on the left hand side in equation 7.7 is the radiation received from the atmosphere by the camera (which has transmission factor τ), given as σT_{MOD}^4. The second term is the radiation absorbed and re-emitted by the camera optics (which radiates at the ground temperature), given as σT_{ground}^4. The sum of these two is the measure of the total radiation seen by the cloud camera (σT_{cam}^4). The Stefan-Boltzmann constant σ cancels out in all terms which leaves equation 7.7 as shown. The result is a simulated atmospheric temperature in the bandpass of the cloud camera with the given atmospheric conditions.

Given we don't know the camera transmission initially, we can rearrange and solve equation 7.7 for the transmission factor τ as;

$$\tau(T_{MOD}^4 - (T_{ground})^4) + (T_{ground}^4) = T_{Cam}^4 \tag{7.8}$$

$$T_{ground}^4 - T_{Cam}^4 = \tau(T_{ground}^4 - T_{MOD}^4) \tag{7.9}$$

$$\tau = \frac{(T_{ground}^4 - T_{MOD}^4)}{T_{ground}^4 - T_{Cam}^4} \tag{7.10}$$

For a vertical image of clear sky from the cloud camera, the brightness temperature calculated from the camera signal is typically around 255 K. Simulating clear sky with MODTRAN for a vertical image and integrating within the wavelength range of the camera yields a T_{MOD} value of 159.3 K, and the ground temperature of the atmospheric model used was 294.2 K. Plugging these into equation 7.10 yields;

$$\tau = \frac{(294.2^4 - 159.3^4)}{294.2^4 - 255^4} = 0.48 \tag{7.11}$$

Given this transmission factor, the calibrated temperature T_{cam} can be calculated for a given observer zenith angle. The results are shown in Figure 7.5. Figure 7.5 shows a small difference in brightness temperature between a clear sky and a cirrus cloud (of the order of ~ 1 K), due to From Figure 7.5, it appears that the brightness temperature follows the $\ln sec\,\theta$ curve (see chapter 5). The easiest way to cross-check this is to plot temperature as a function of $\ln sec\,\theta$, which should yield a straight line with slope corresponding to the B parameter with units of K (see chapter 5). The result of this (for a clear sky) is shown in Figures 7.5 and 7.6.

From Figure 7.6, the slope of the fitted line is 3.27 K. This is comparable with typical results that are obtained with the cloud cameras at Auger.

Now, the process can be repeated, but with a cirrus cloud placed 4 km above ground level, with a thickness of 1 km and extinction coefficient of 0.14. The results of this are shown in Figure 7.7.

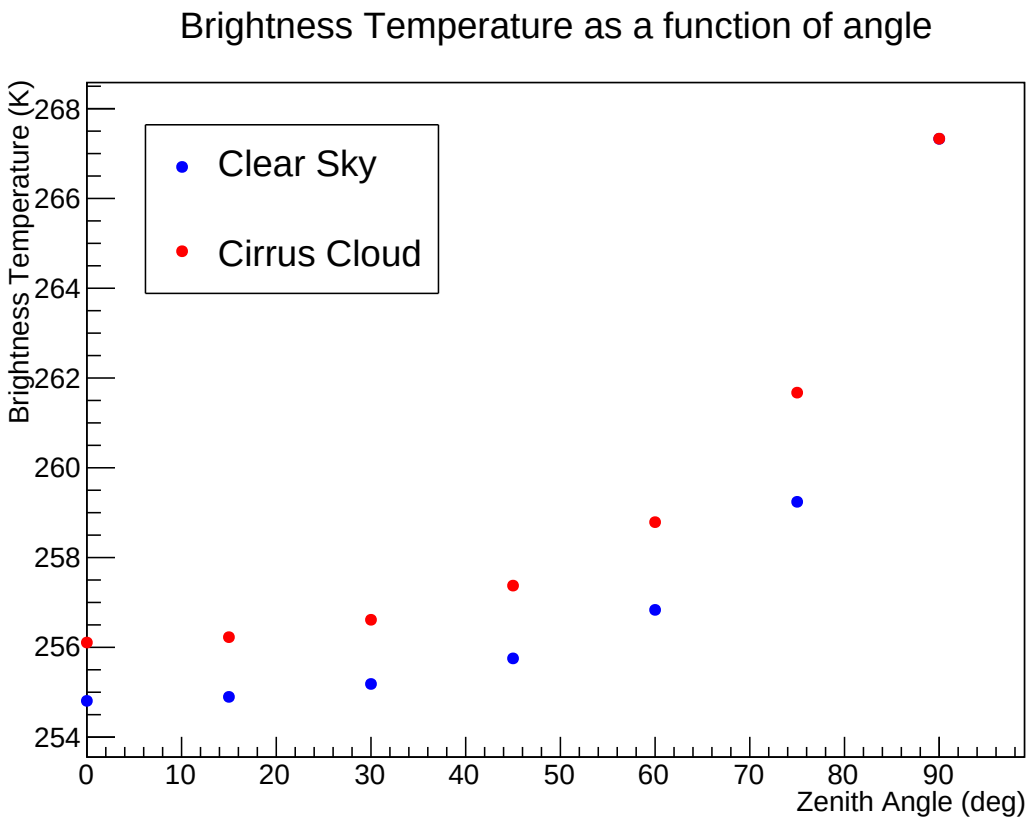

Figure 7.5: Simulated sky brightness temperature as a function of zenith angle. The different colours represent a clear sky (blue) and a cirrus cloud present overhead (red) for a Mid-Latitude Winter. The parameters of the cirrus cloud are: 4 km altitude, 1 km thickness and extinction coefficient (at 550m nm) of 0.14.

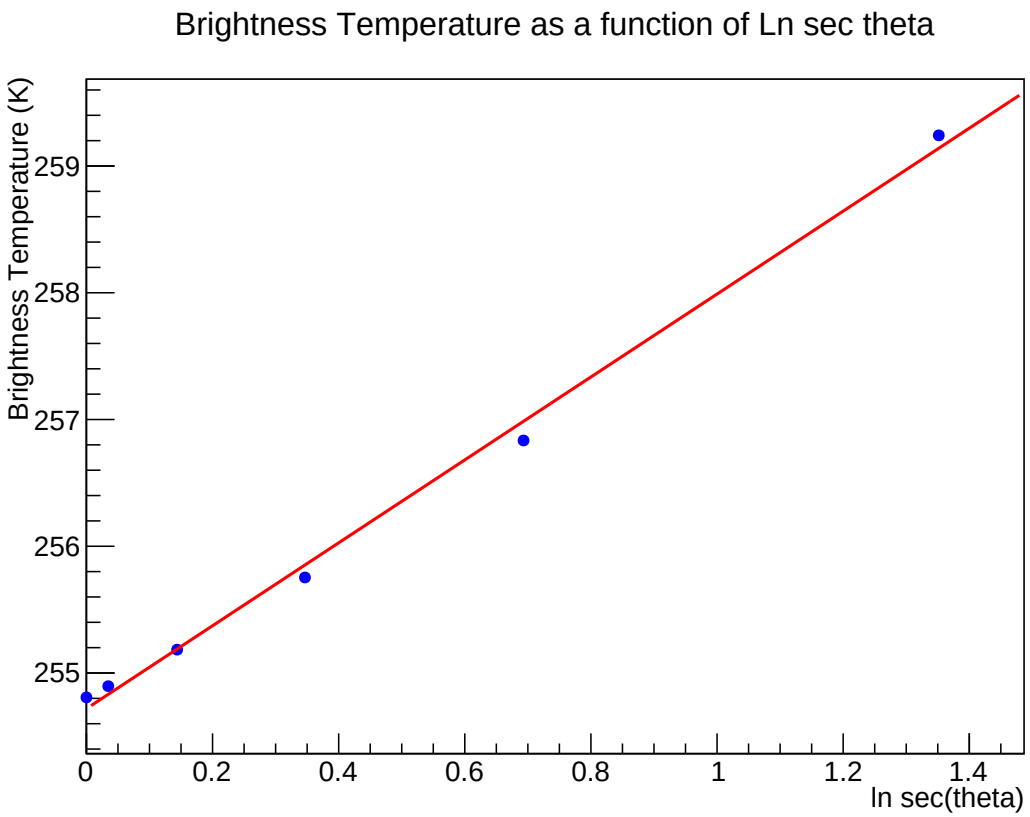

Figure 7.6: Simulated brightness temperature as a function of $\ln sec\,\theta$ for a clear sky in Mid Latitude Winter. Performing a linear fit yields a slope of 3.27 K per unit $\ln sec\,\theta$.

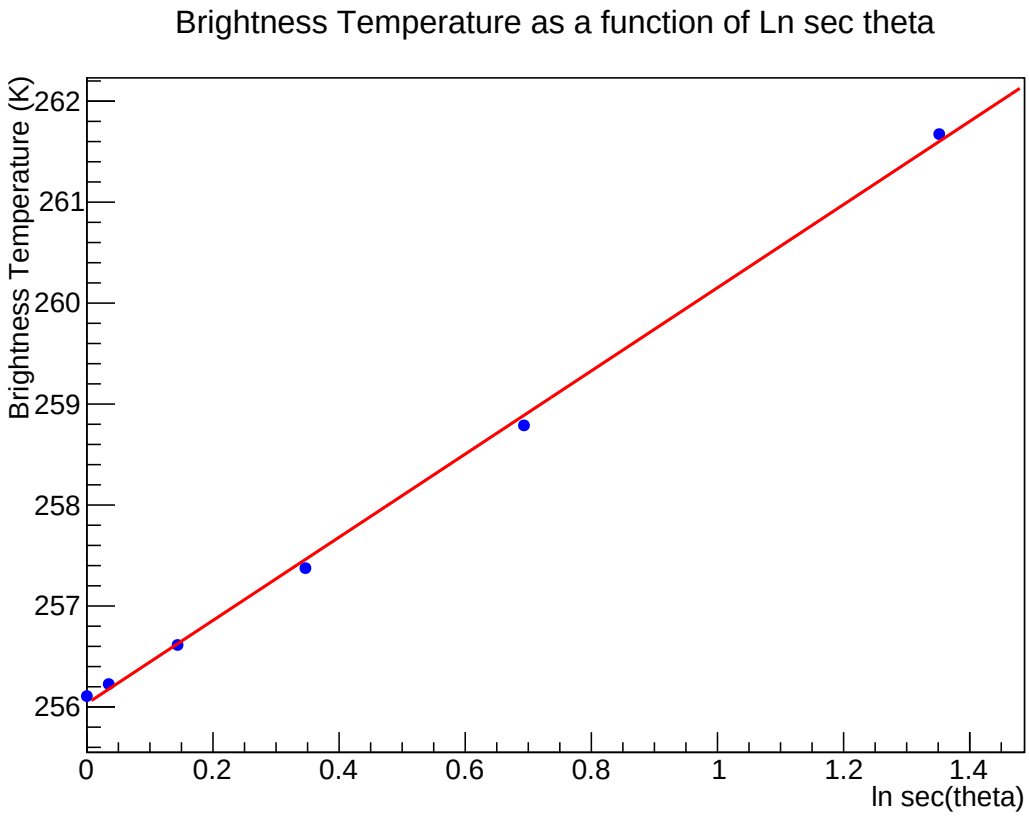

Figure 7.7: Simulated brightness temperature as a function of $\ln sec\,\theta$ for a Mid Latitude Winter, with a cirrus cloud present. Performing a linear fit yields a slope of 4.11 K per unit $\ln sec\,\theta$.

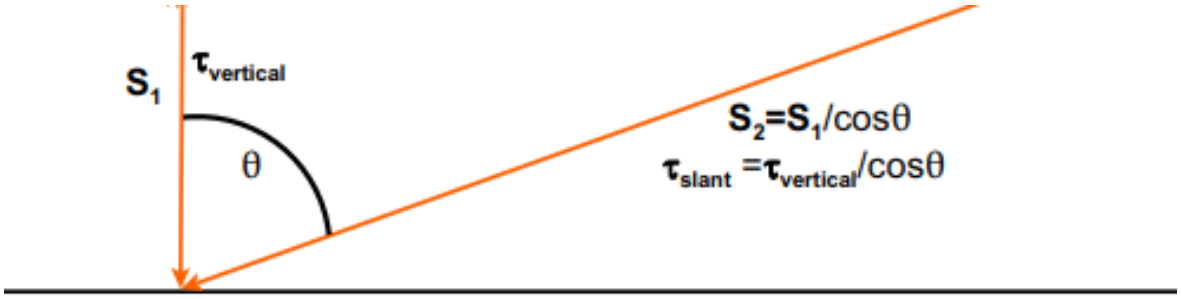

Figure 7.8: Optical depths at the vertical (τ_{vert}) and at a given zenith angle θ (τ_{slant}) for given radiances S_1 and S_2 respectively. Source: [93].

From Figure 7.7, the slope has increased by adding in a cirrus cloud. Due to the optical depth of the atmosphere increasing with increasing zenith angle, the brightness temperature increases also. When a cirrus cloud is present in the atmosphere, the vertical optical depth of the atmosphere is higher than it is for a clear sky. Optical depth at a given zenith angle depends upon the vertical optical depth. In a simplified plane parallel model of the atmosphere (as assumed in MODTRAN), the optical depth at any zenith angle (τ_{slant}) can be calculated from Figure 7.8, with vertical optical depth τ_{vert} as shown. As a result, we find that

$$\cos\,\theta = \frac{\tau_{vert}}{\tau_{slant}} \tag{7.12}$$

Hence;

$$\tau_{slant} = \frac{\tau_{vert}}{\cos\,\theta} = \sec\,\theta\ (\tau_{vert}) \tag{7.13}$$

Similarly, the slope of the optical depth as a function of zenith angle also depends on the vertical optical depth. This slope is equal to the derivative of 7.13 with respect to zenith angle, which results in;

$$\frac{d}{d\theta}(\tau_{slant}) = \frac{d}{d\theta}(\sec\,\theta\ (\tau_{vert})) = \tau_{vert}\,\tan\,\theta\,\sec\,\theta \tag{7.14}$$

From equation 7.14, we can see that the slope of the optical depth increases as the vertical optical depth increases. Due to these relationships, the slope of the brightness temperature curve is larger when a cirrus cloud is present.

The camera transmission factor of 0.48 (from equation 7.11) is a reasonable first order approximation to the true value for the cloud cameras at Auger, but it seems to be an underestimation. This is due to the cameras at Auger being designed (and calibrated) to be highly transmissible in the IR window. As a result, the calculation of this transmission must be cross-checked. This is shown in section 7.3.

7.3 Second attempt at Calibration

There is another method of calibration that can be done on the modelled data. At the Pierre Auger Observatory, the infrared cloud cameras are calibrated by pointing at a known blackbody (with known temperature T) [48]. This idea can be used in order to calibrate the brightness temperatures calculated from MODTRAN. To do so, one can

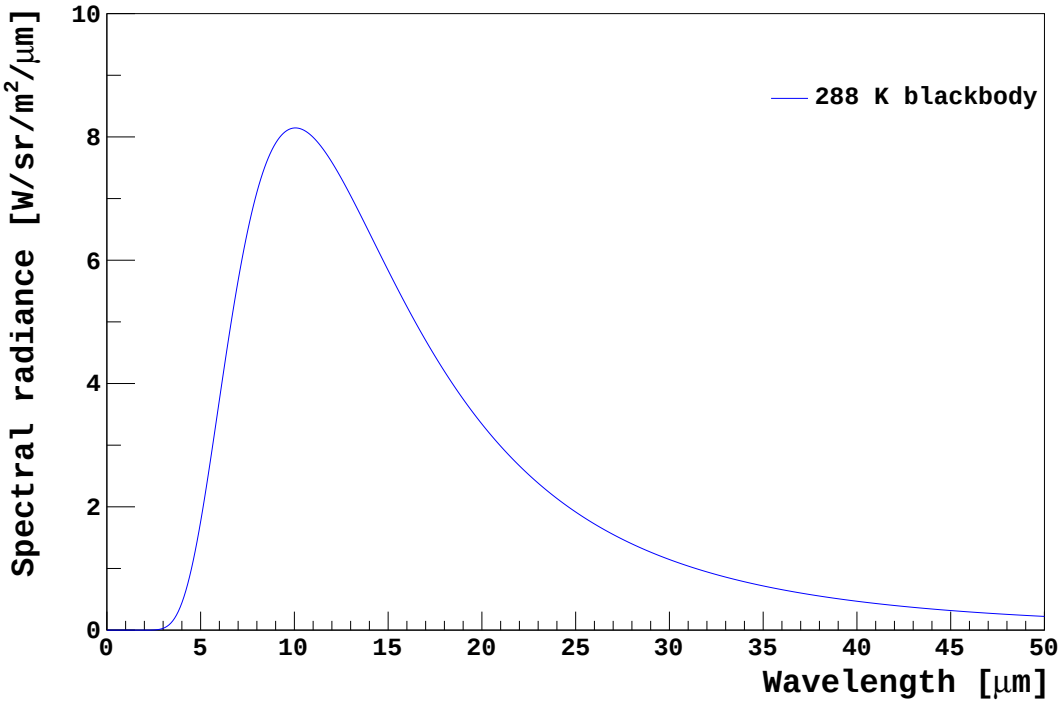

Figure 7.9: Planck distribution for a 288 K blackbody object. A very distinct radiance peak is present, at a wavelength given by Wien's Law. [48]

calculate the Planck distribution for a given blackbody at temperature T via equation 7.15 [94].

$$B_\lambda(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \tag{7.15}$$

Another important mathematical equation is known as Wien's displacement law (equation 7.16), which calculates the wavelength of the peak of the Planck distribution at temperature T. [94]

$$\lambda_{max} = \frac{2.898 \times 10^{-3}}{T} [\mu m] \tag{7.16}$$

Figure 7.9 shows the Planck distribution for an object at a temperature of 288 K, and Figure 7.10 shows the Planck distribution for blackbody temperatures ranging from 200-300 K (with corresponding wavelength peaks as given by Wien's Law).

From here, we calculate the integrated radiance in the operational wavelength range of the cloud cameras ($7-14\,\mu$m) to get the band radiance, R_{band}, and plot as a function of the blackbody temperature. This is shown in Figure 7.11 for temperatures ranging from 200 to 300 K. Temperature is plotted as a function of radiance since the blackbody temperature is known and one wishes to see how the temperature depends on this band radiance as that is the wavelength range that the cloud cameras at Auger are sensitive to. Figure 7.11 plots temperature vs radiance since the radiance is calculated by MODTRAN and one wishes to know the corresponding blackbody temperature for the calculated radiance. This temperature is used for all subsequent analysis, rather than the radiance.

The next step in the calibration is to fit a function to the resulting curve. Since the radiance of a blackbody is proportional to the temperature to the fourth power

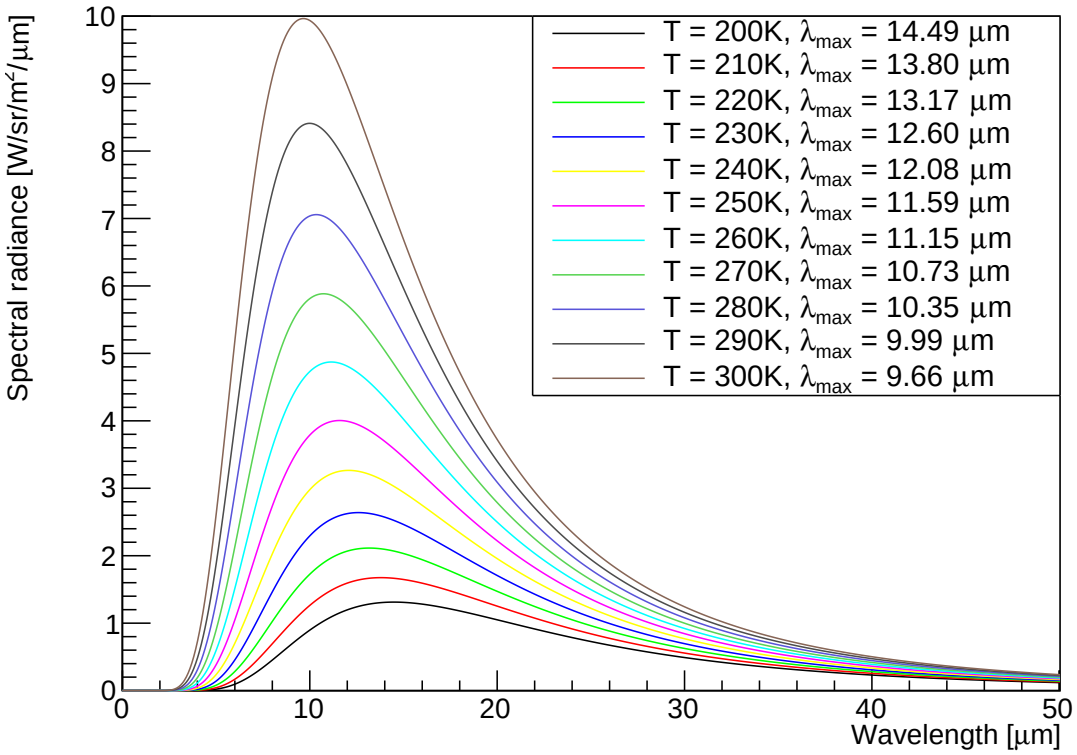

Figure 7.10: Planck distributions for objects of various temperatures, ranging from 200 K to 300 K.

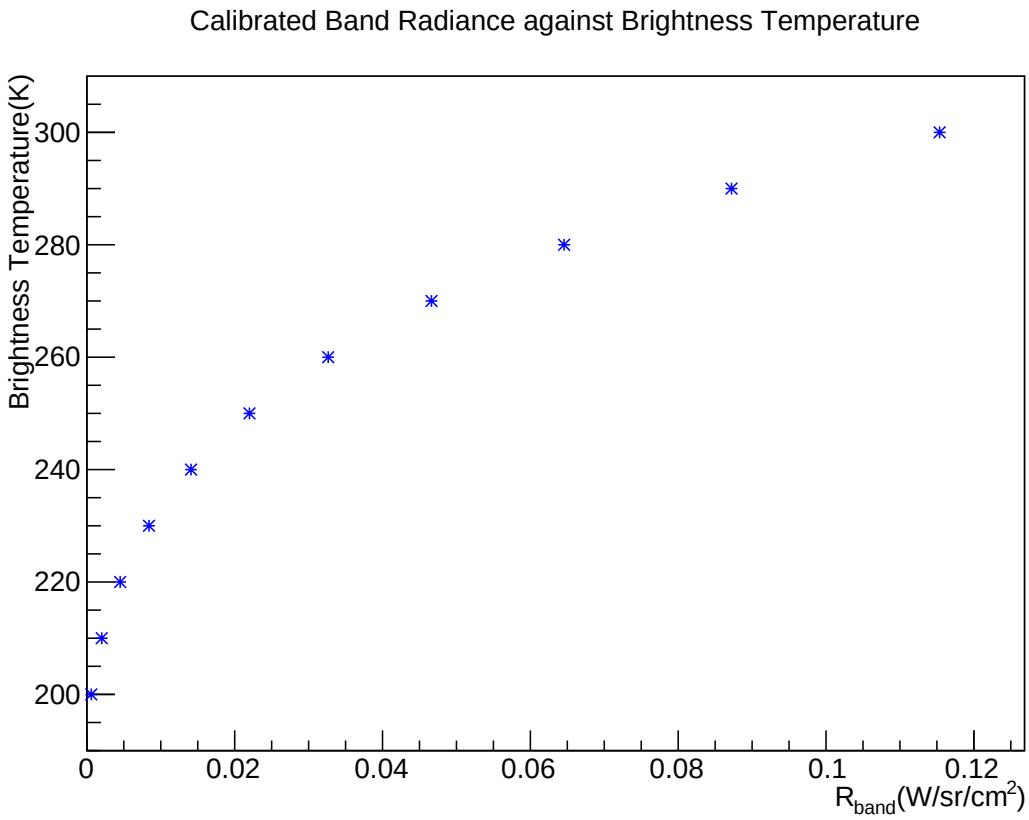

Figure 7.11: Blackbody temperature as a function of integrated band radiance over 7-14 μm.

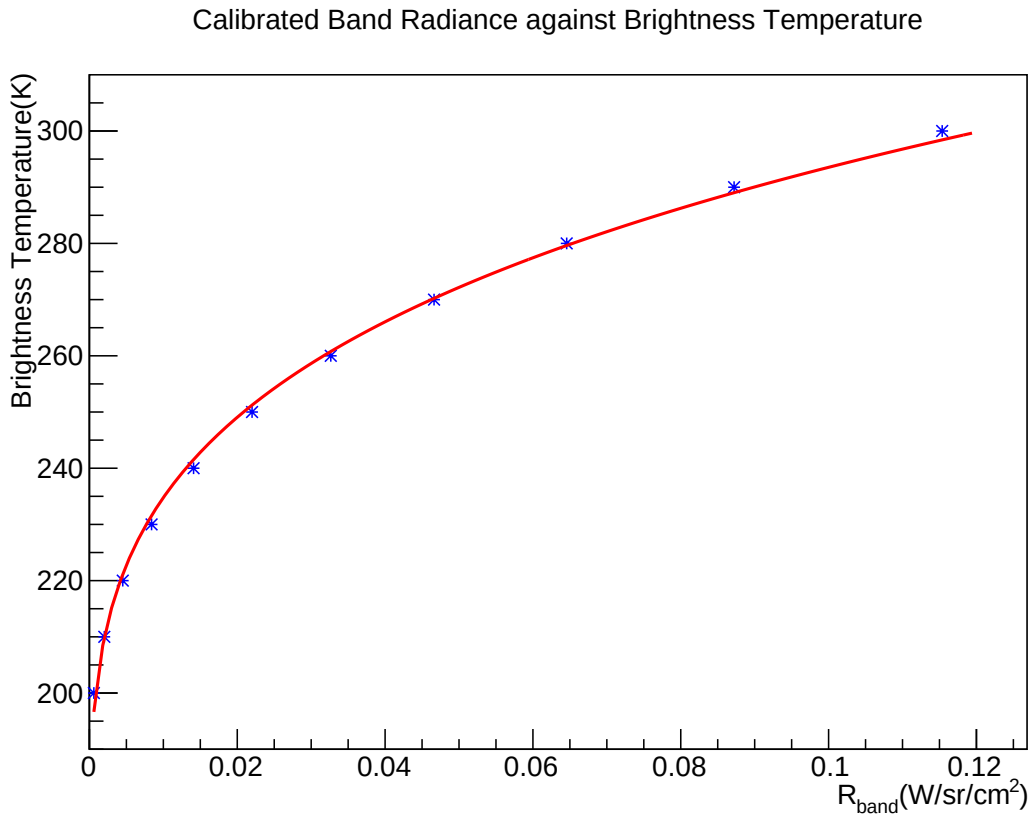

Figure 7.12: Brightness Temperature as a function of band radiance with a fit given by equation 7.17. The parameters of the resulting fit are a = 238.7 $\mathrm{K\,cm^2W^{-1}sr^{-1}}$ and b = 159.3 K.

(Stefan-Boltzmann Law - equation 7.2), this will become the first attempt at a fitted function, given by equation 7.17 for fit parameters a and b.

$$T = aR_{band}^{\frac{1}{4}} + b \tag{7.17}$$

The results of this quartic fit are shown in Figure 7.12 with the given fit parameters.

As a result, when cirrus cloud is modelled in the atmosphere, the calculated band radiance can be used to calculate what the brightness temperature must be. The resulting fit becomes equation 7.18.

$$T = (238.7)R_{band}^{\frac{1}{4}} + 159.3 \tag{7.18}$$

Now that the calibration has been done, the next step is to account for the transmission of the camera, which is done by equation 7.19 for given transmission factor τ.

$$\tau(T_{MOD}^4) + (1-\tau)(T_{ground}^4) = T_{Cam}^4 \tag{7.19}$$

In equation 7.19, T_{MOD}^4 corresponds to the calibrated brightness temperature from MODTRAN, T_{ground} corresponds to the ground temperature, and T_{Cam}^4 is the IR cloud camera brightness temperature calculated at Auger (see section 3.5.8). For a vertical image of the clear sky using MODTRAN, the integrated band radiance is given as $\mathrm{R_{band}} = 1.16 \times 10^{-3}\,\mathrm{W/cm^2}$. As a result, the calibrated temperature is given by equation 7.20.

$$T = (238.7)(1.16 \times 10^{-3})^{\frac{1}{4}} + 159.3 = 203.3\,\mathrm{K} \tag{7.20}$$

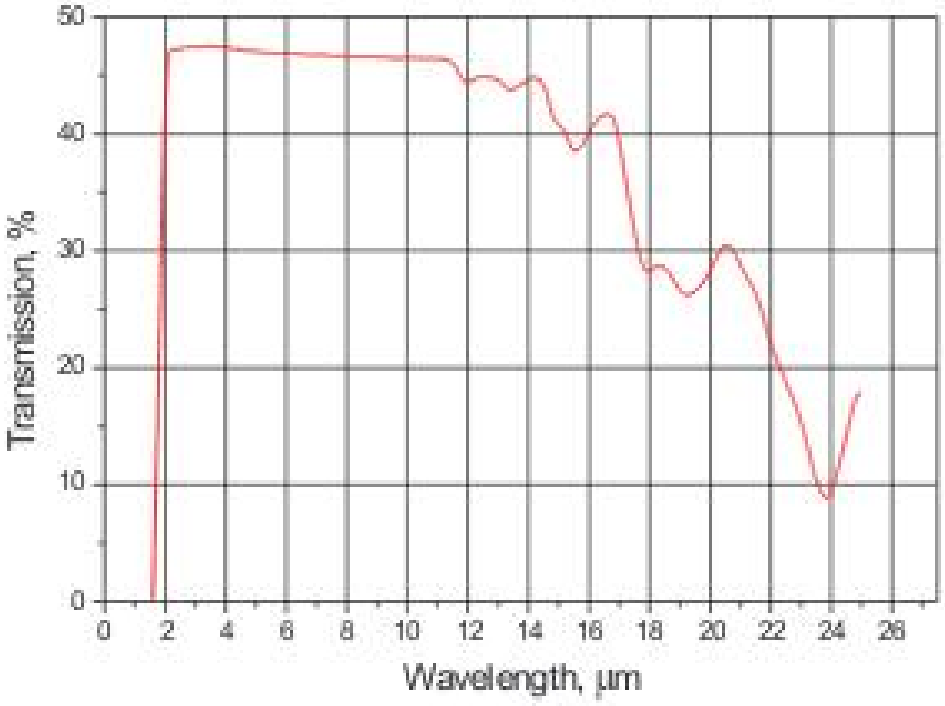

Figure 7.13: Transmission of a 2 mm sample of Germanium in the $2 - 26\,\mu m$ range. Source: [95].

As a result, $T_{MOD} = 203.3$ K. This was simulated for a clear sky in the vertical direction using the Mid Latitude Summer model, and thus in order to calculate the transmission one needs a real time measurement of the cloud camera vertical temperature and the ground temperature. Picking a clear night in the Summer of 2017 gives a ground temperature of 286.9 K and a vertical sky temperature of 257.6 K. Plugging these values into equation 7.10 gives the transmission as

$$\tau = \frac{(286.9^4 - 203.3^4)}{286.9^4 - 257.6^4} = 0.47 \tag{7.21}$$

This transmission is similar to what was obtained previously, which seems low. The cameras installed at Auger use a window made of Germanium material, which has a thickness of 2.05 mm. The transmission of a 2 mm thick Germanium window is shown in Figure 7.13.

As can be seen by Figure 7.13, the transmission of 2 mm thick Germanium in the atmospheric window ($7 - 14\,\mu$m) is between 0.45 and 0.5. This is consistent with the calibrated values calculated in equations 7.21 and 7.11. Thus the model is sufficient in simulating the transmission of the cameras at Auger.

Now, one can simulate the effect that cirrus cloud in the atmosphere has on temperature measurements. Firstly, one can compare the temperature curve of the brightness temperature (T_{cam}) as a function of zenith angle for both a clear sky and a cirrus cloud, as was done in section 7.2. This is shown in Figure 7.14.

The results of Figure 7.14 show that the brightness temperature increases as a function of $\ln sec\,\theta$. One can compare the brightness temperature fits as a function of $\ln sec\,\theta$ for various cirrus cloud heights, which are shown in Figures 7.15 and 7.16

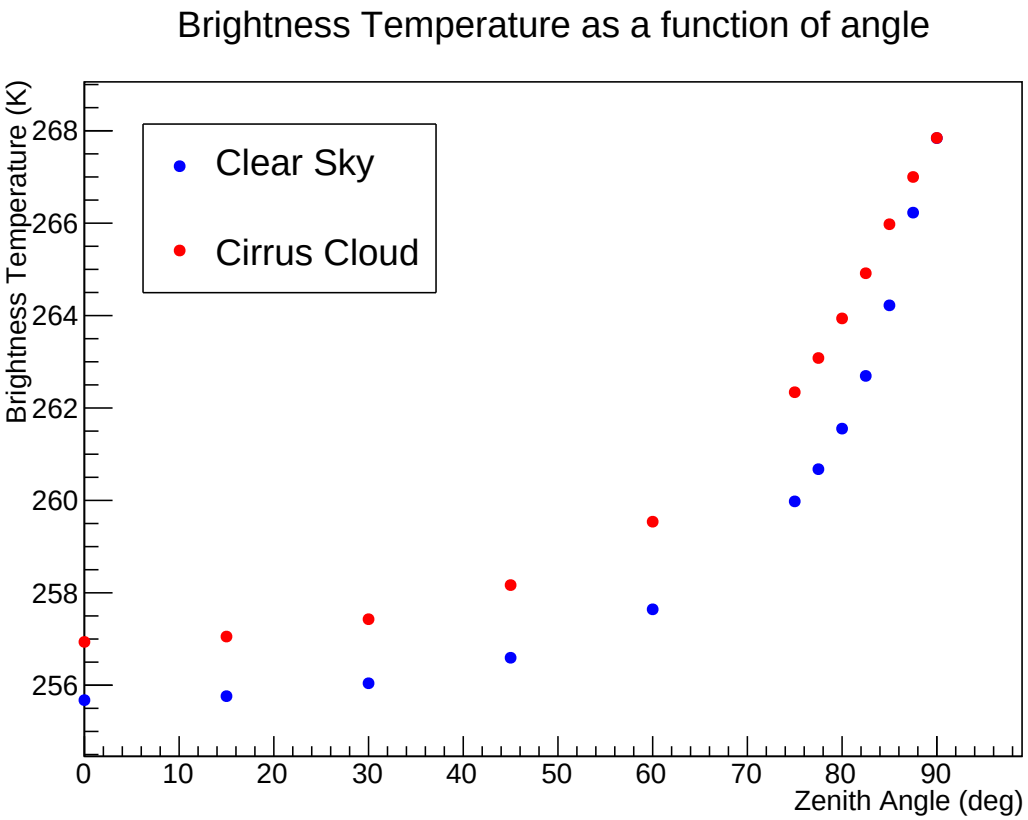

Figure 7.14: Brightness temperature (T_{cam}) as a function of Zenith angle for a clear sky (blue dots) and with a cirrus cloud 4 km above ground level (a.g.l). The model used was Mid Latitude Winter.

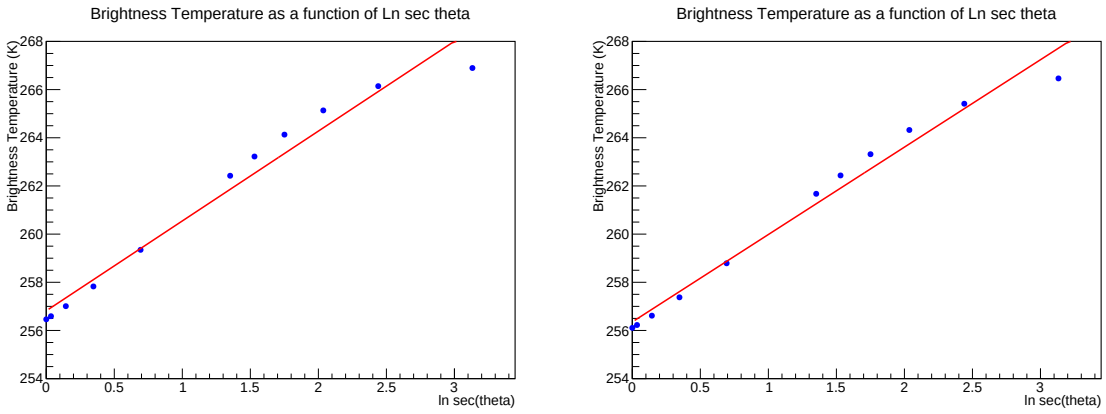

Figure 7.15: Brightness Temperature as a function of angle with Cirrus cloud in the atmosphere located at 2 km (left panel) and 4 km (right panel). The slope of the fitted line is 3.73 K (left) and 3.62 K (right). The model used was Mid Latitude Winter.

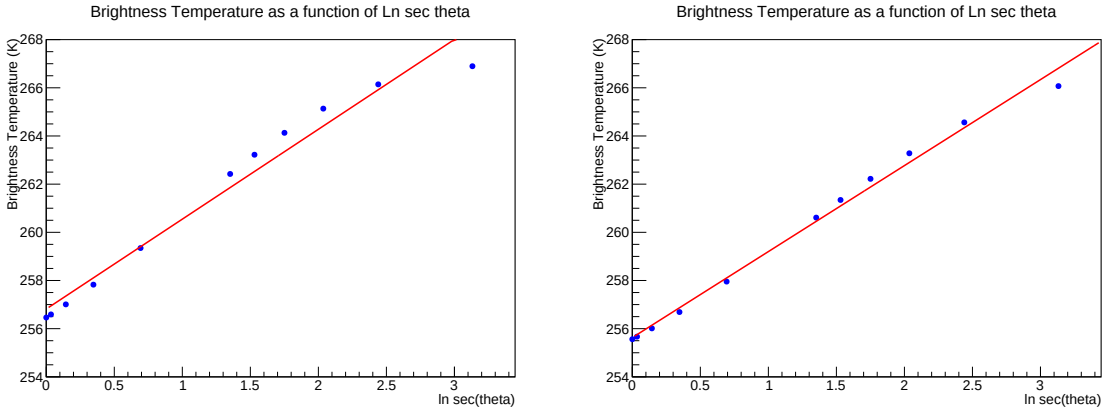

Figure 7.16: Brightness Temperature as a function of angle with Cirrus cloud in the atmosphere located at 2 km (left panel) and 8 km (right panel). The slope of the fitted line is 3.73 K (left) and 3.56 K (right). The model used was Mid Latitude Winter.

comparing a cirrus cloud at 2 km to cirrus clouds at 4 km and 8 km respectively, for the same values of cloud thickness (1 km) and extinction coefficient (0.14) in all cases.

Figures 7.15 and 7.16 show that the slope of the fit decreases as a function of cirrus cloud altitude. However, this decrease is small ($\sim$ 0.1 K), indicating that the height/temperature of cirrus cloud in the atmosphere is largely independent of its brightness in the infrared.

The last question to ask is at what altitude do these high level cirrus clouds become invisible in the infrared? To answer this question, one can compare this $\ln sec\,\theta$ fit for a high altitude cirrus cloud and a clear sky, to see when the two fits become almost indiscernible. Figure 7.17 shows the fit for a clear sky compared to the fit result with a cirrus cloud located 8 km above ground level (a.g.l).

Figure 7.17 shows similar behaviour when comparing the atmosphere for a clear sky to an atmosphere with a cirrus cloud at high altitude. Since the slope of the fit is within 0.01 in both cases, it means that there is no discernible difference between the two atmospheres. What this demonstrates is that a high altitude cloud behaves in

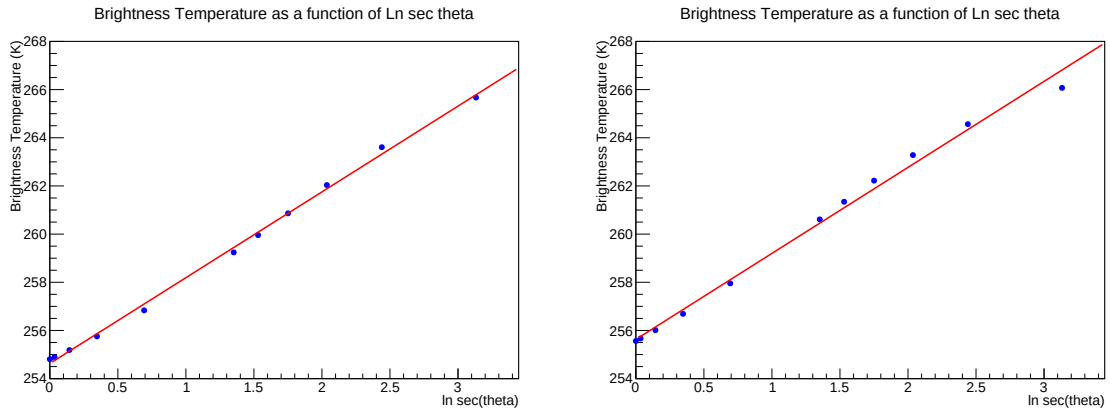

Figure 7.17: Brightness Temperature as a function of angle for a clear sky (left panel) and cirrus cloud located at 8 km (right panel). The slope of the fitted line is 3.56 K in both cases. The model used was Mid Latitude Winter.

much the same way as a clear sky, with a similar infrared response. Effectively, a cloud at such an altitude is unable to be detected with cloud cameras, like the ones used at the Auger observatory. This is why some cloud masks are still unable to detect faint cloud that is visible to the naked eye (see section 5.4). This is also due to the fact that ice has a lower emissivity than water vapour [96], which means that in the infrared wavelength band ice crystals will appear less bright than water droplets. From this fact alone, it is harder to detect cirrus cloud in the IR, since it more closely approximates the brightness of a clear sky.

It is also worth checking the effects that other types of cloud have on the infrared response of the camera. It is well established that overcast conditions cause the $\ln sec\,\theta$ curve to flatten, thus reading a very warm temperature (for more on this, see section 5.4). This is also shown in Figure 7.18 - the brightness temperature as a function of zenith angle for a thick cloud at low altitude.

From Figure 7.18, it is clear that having a thicker cloud in the atmosphere increases the optical depth. A high value of optical depth means that the temperature of the atmosphere increases very slowly as the zenith angle increases (shown by a slope of 0.51 K from Figure 7.18). The reason for this is due to the thick cloud present acting as an almost perfect blackbody (see Section 4.2), meaning that the atmosphere will radiate at a single temperature (the temperature of the thick cloud) which is not dependent upon zenith angle. This is consistent with the results from section 5.4.

Next, the response study was repeated for a mid latitude winter model, to see if the pattern is consistent. Figures 7.19, 7.20, and 7.21 show the same study of cirrus cloud as done previously (Figures 7.15, 7.16 and 7.17 respectively).

As can be seen by Figure 7.20, there is a discernible difference between a cirrus cloud located at 2 km and 8 km respectively. The lower the altitude of cloud, the larger effect it has on our infrared response, which is irrespective of the time of year. Another distinct feature found from the MODTRAN simulations done in a mid latitude summer model is that the difference between a clear sky and an atmosphere with a high altitude cirrus cloud is minimal. The question is, does this also hold for a winter model? This is shown in Figure 7.21.

The results of Figure 7.21 shows that there is a large difference between a clear sky

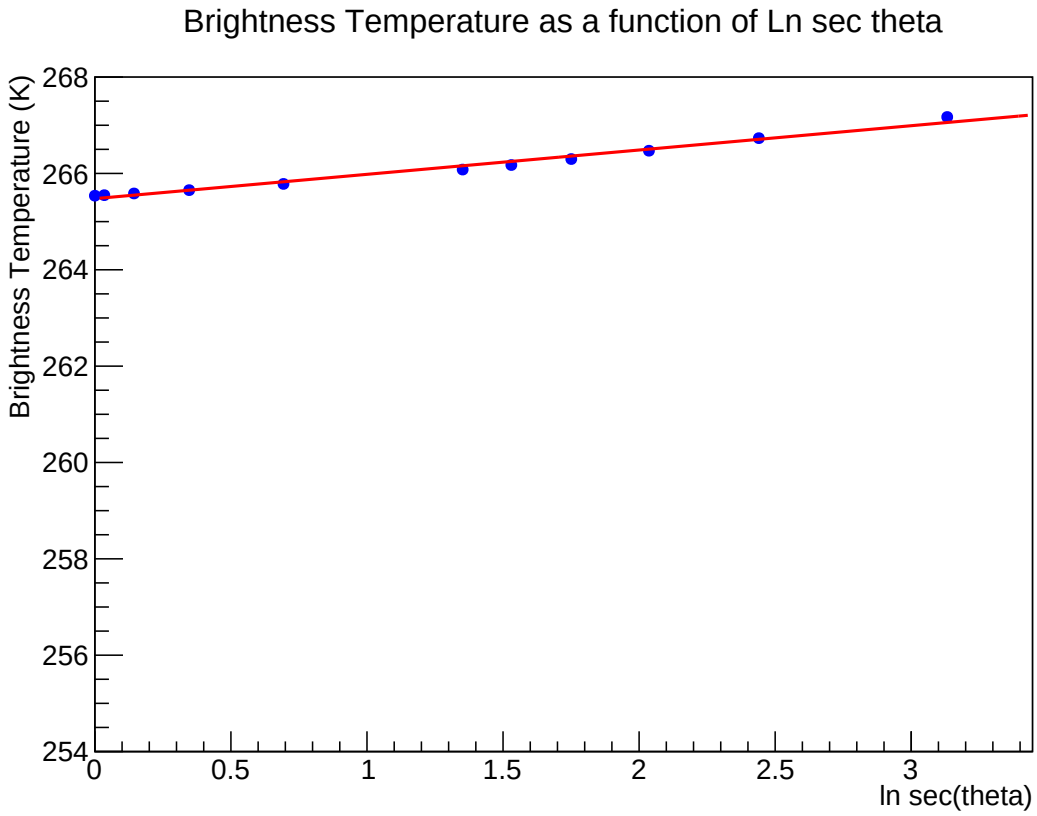

Figure 7.18: Brightness temperature as a function of zenith angle with an Altostratus cloud located at 2.4 km. The slope of the fitted line in this case is 0.51 K. The model used was Mid Latitude Winter.

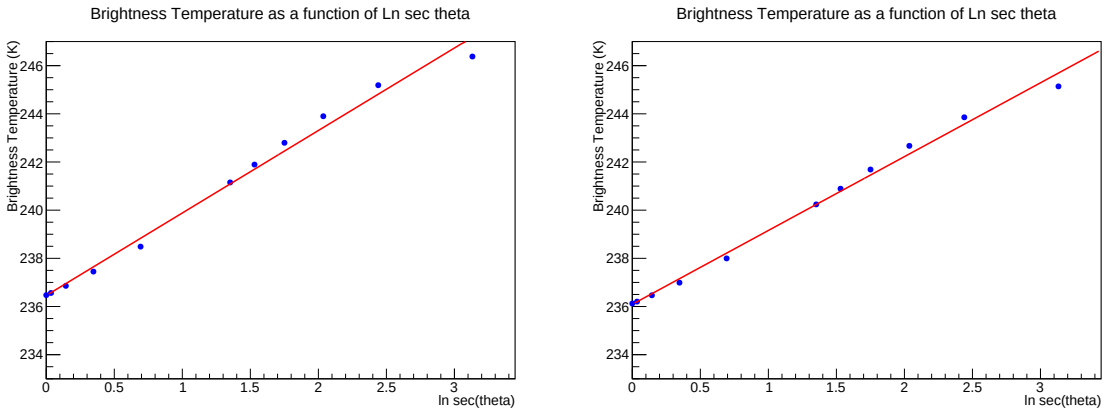

Figure 7.19: Brightness Temperature as a function of angle with Cirrus cloud in the atmosphere located at 2 km (left panel) and 4 km (right panel). The slope of the fitted line is 3.42 K (left) and 3.06 K (right). The model used was Mid Latitude Winter.

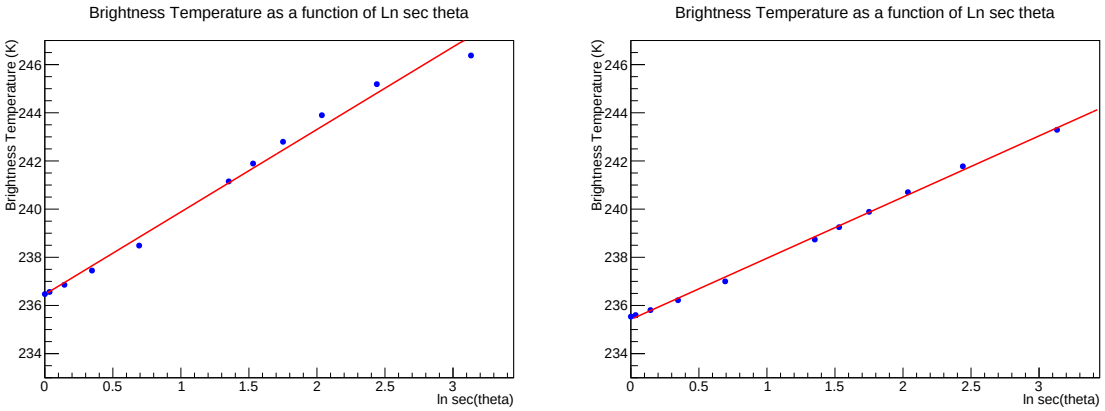

Figure 7.20: Brightness Temperature as a function of angle with Cirrus cloud in the atmosphere located at 2 km (left panel) and 8 km (right panel). The slope of the fitted line is 3.42 (left) and 2.53 (right). The model used was Mid Latitude Winter.

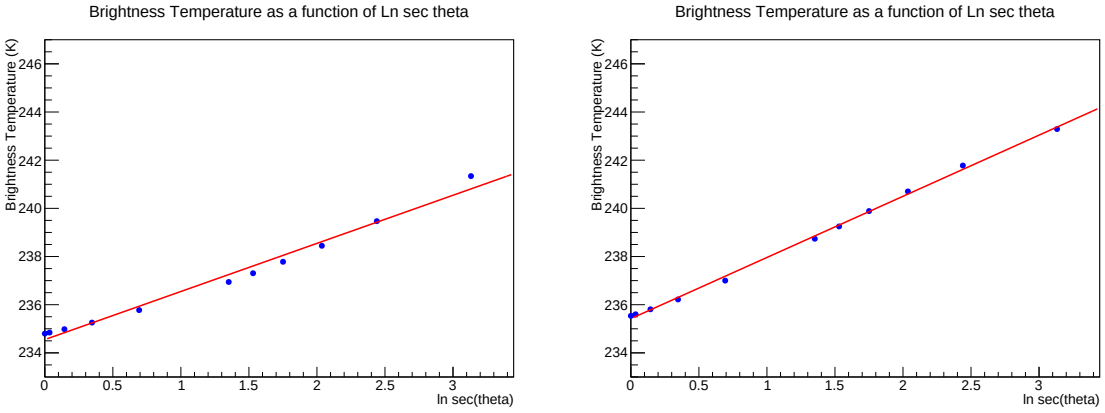

Figure 7.21: Brightness Temperature as a function of angle for a clear sky (left panel) and cirrus cloud located at 8 km (right panel) using the Mid Latitude Winter MODTRAN model. The slope of the fitted line is 1.99 (left) and 2.53 (right).

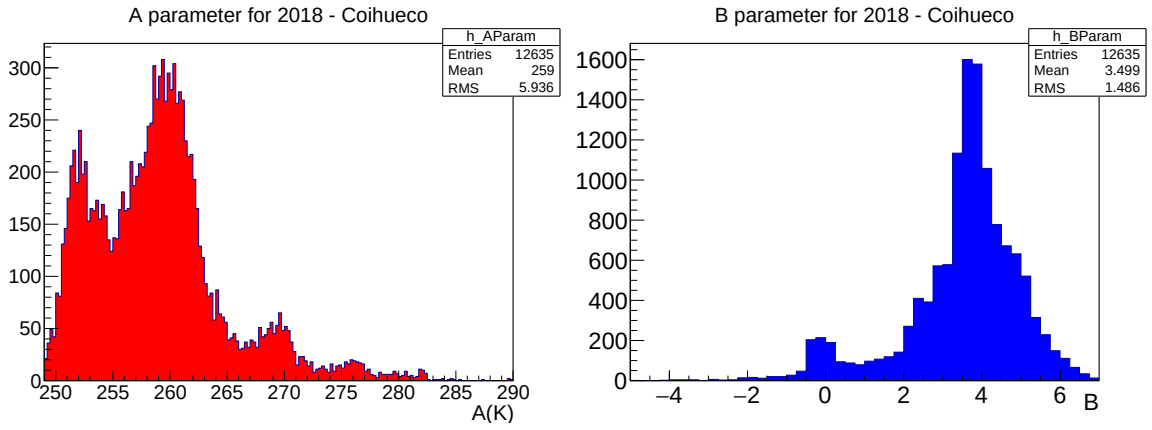

Figure 7.22: Distributions of the A parameter (left) and B parameter (right) for all nights in 2018 at Coihueco. The A parameter is the vertical sky temperature whilst the B parameter is the slope of the fitted $\ln \sec \theta$ curve.

and a cirrus cloud for a winter model. This suggests that it is easier to detect cirrus cloud when the atmosphere is colder.

Additionally, the results of Figure 7.21 illustrate that the slope of the brightness temperature vs zenith angle curve is lower in winter, compared to summer. This is due to the vertical optical depth being lower in winter, as there is less water vapour present in the atmosphere (which is the largest absorber of infrared radiation in the atmospheric window). Water vapour (and other constituent atmospheric molecules) determine the vertical optical depth due to their absorption properties, thus a lower concentration (as one gets in winter) means a lower vertical optical depth. As was shown earlier in this chapter, the slope of the curve depends upon the vertical optical depth, thus a lower vertical optical depth corresponds to a lower slope.

A way of validating these results from MODTRAN is to examine the distributions of the fit parameters from the cloud cameras at Auger, as shown in equation 7.22. Figure 7.22 shows the distributions of the fit parameters at Coihueco for a single year.

$$T = A + B \ln sec\,\theta \tag{7.22}$$

Another comparison of these parameters is to look at these distributions over Summer and Winter on clear nights. Figure 7.23 splits the A parameter distributions into the summer and winter months of 2018.

Figure 7.23 shows a much cooler vertical temperature in winter compared to summer, as expected. It is also worth comparing the B parameter for those same clear nights, in both summer and winter, as shown by Figure 7.24.

From Figure 7.24, it is clear that the B parameter is lower in winter (on average). This is due to the fact that a smaller B value corresponds to a higher proportion of cloud, which is expected for the winter months as was shown in chapter 6 The bimodal behaviour of the distribution in the Winter months can be explained by a higher prevalence of either totally overcast conditions in Winter (the left peak of the distribution) or totally clear conditions (right hand peak).

Now, one can show the distributions of these parameters for all four FD sites at Auger, to check for consistency. Figures 7.25 and 7.26 show the distributions of A and

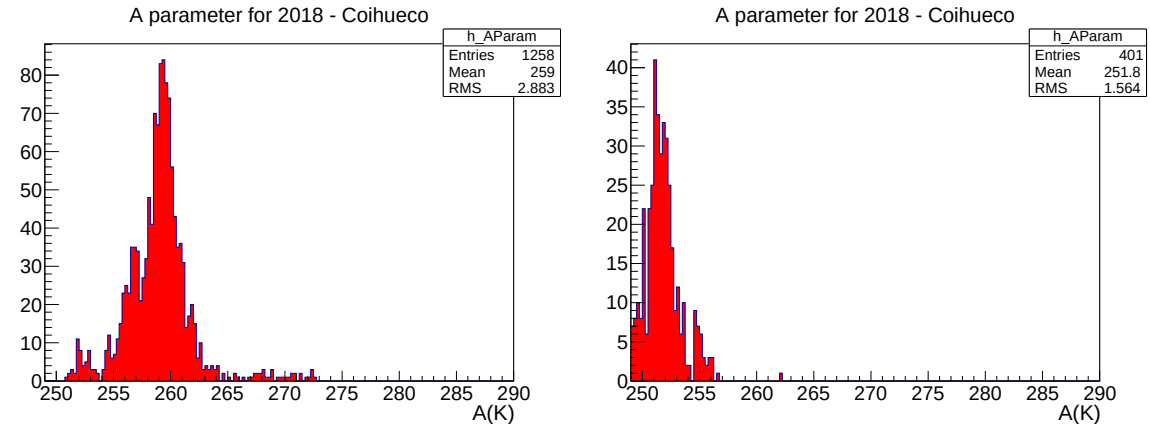

Figure 7.23: Parameter A distributions during clear nights for the summer (left) and winter (right) months at Coihueco in 2018.

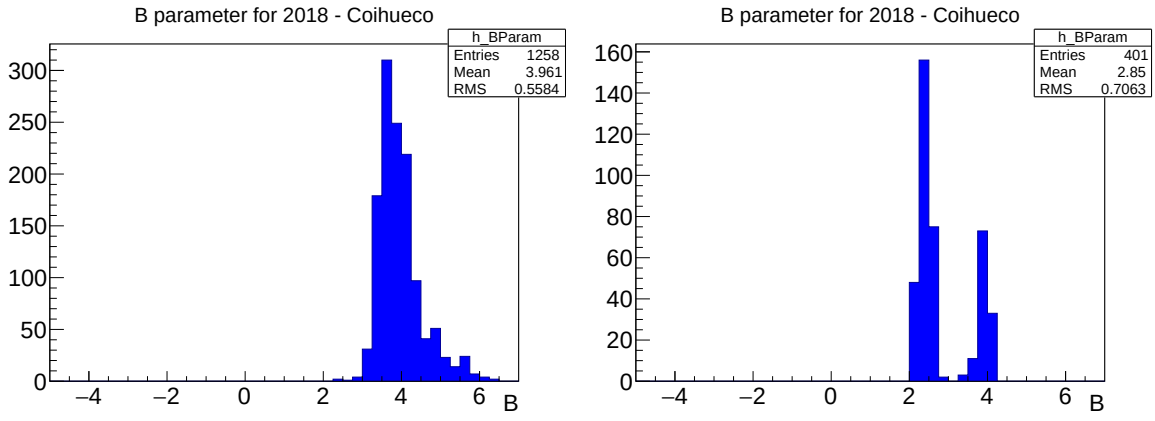

Figure 7.24: Parameter B distributions during clear nights for the summer (left) and winter (right) months at Coihueco in 2018.

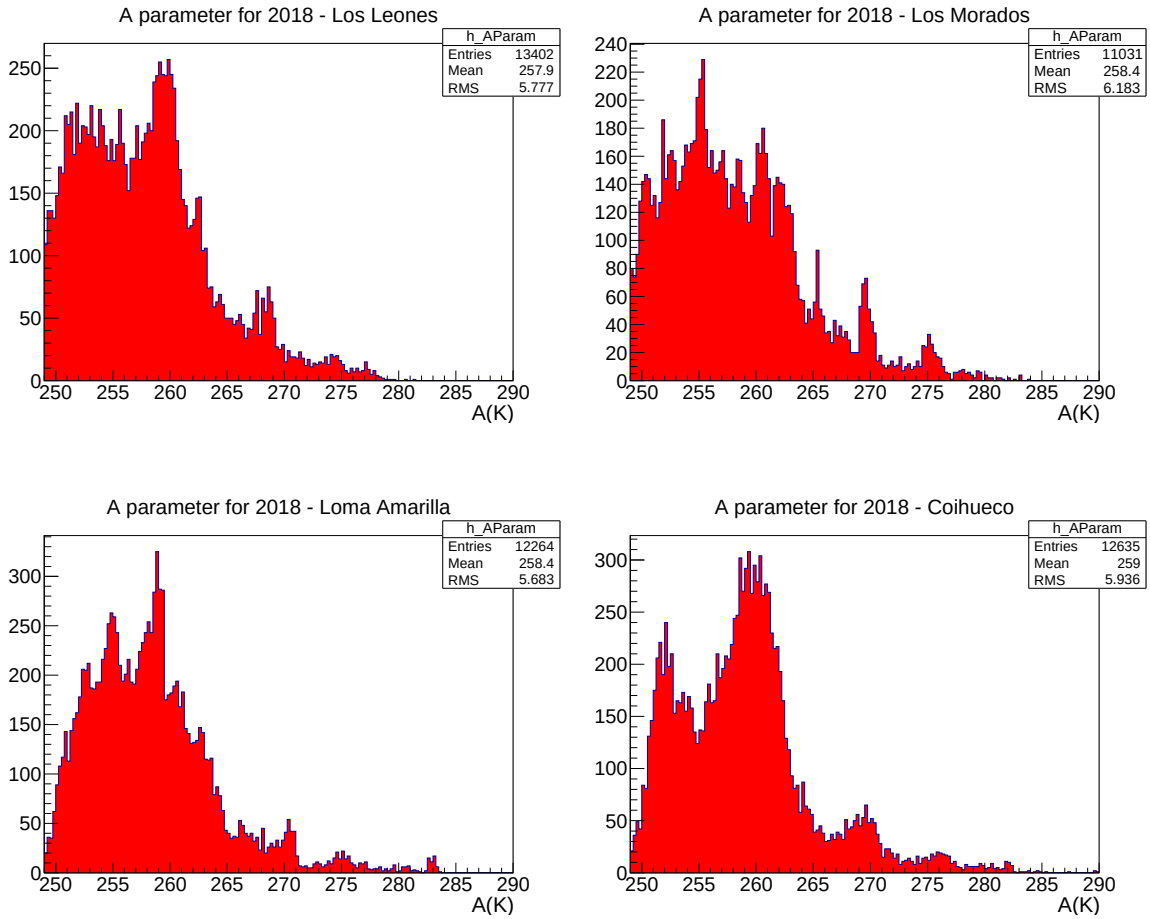

Figure 7.25: A parameter distributions for Los Leones (top left), Los Morados (top right), Loma Amarilla (bottom left) and Coihueco (bottom right) for all nights in 2018.

B respectively for each site in 2018, for all nights.

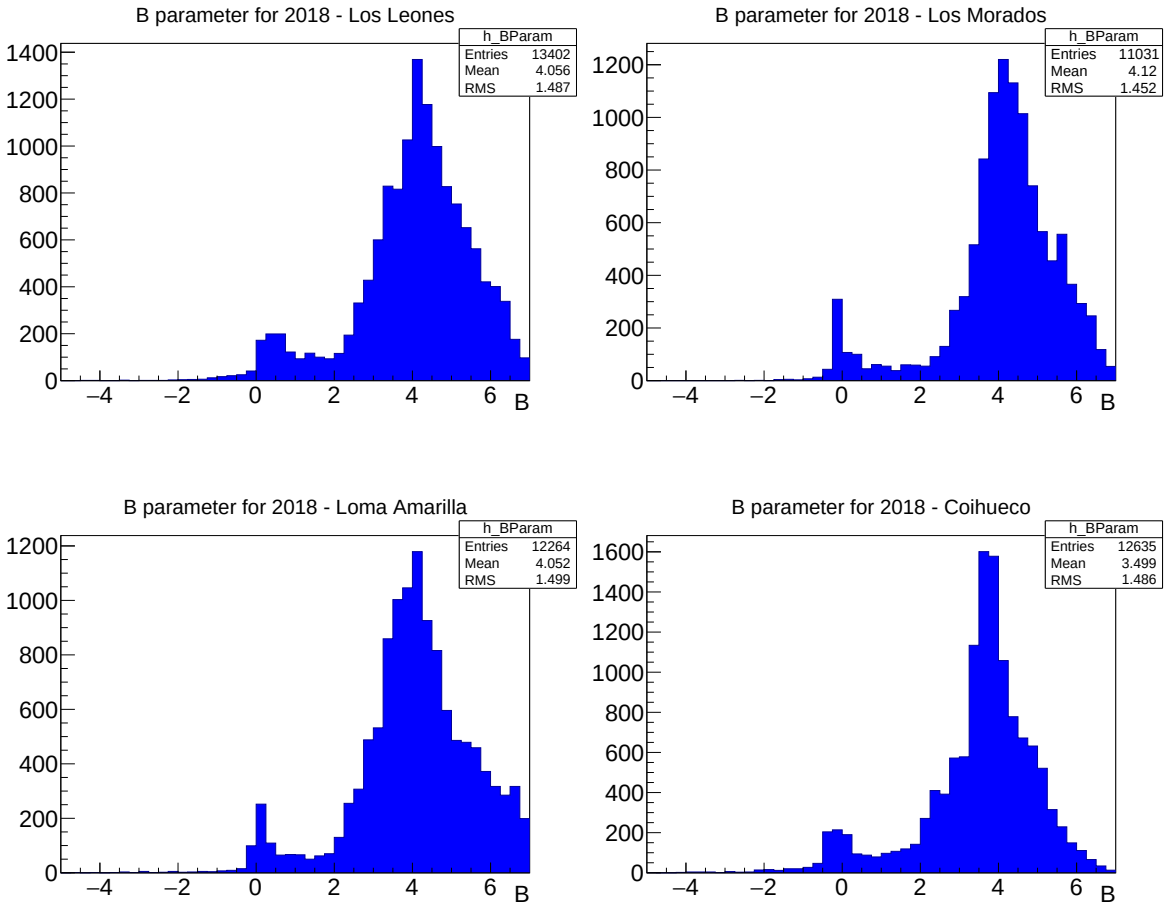

Figure 7.26: B parameter distributions for Los Leones (top left), Los Morados (top right), Loma Amarilla (bottom left) and Coihueco (bottom right) for all nights in 2018.

From Figures 7.25 and 7.26, the four FD sites all show consistency in their distributions - showing that the $\ln sec\,\theta$ fit performs well for all IR cameras.

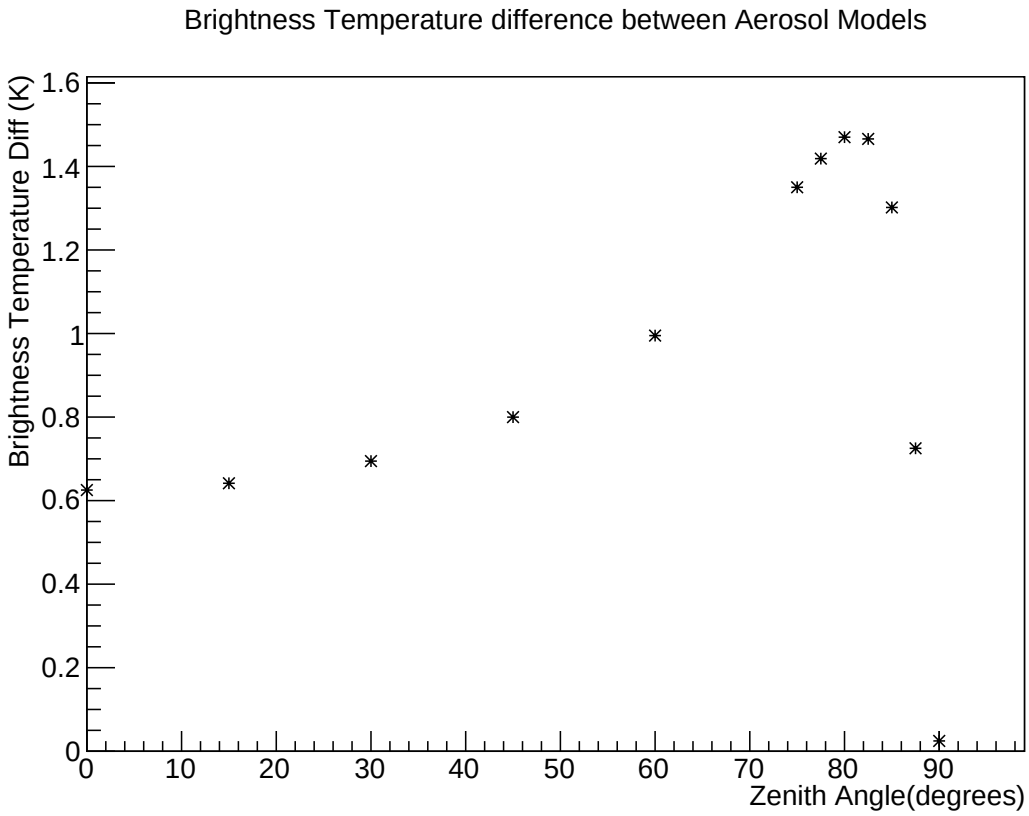

Figure 7.27: Difference in Brightness Temperatures at each Zenith Angle for the High and Low Aerosol content MODTRAN models. The difference is defined as the brightness temperature (high aerosol distribution) - brightness temperature (low aerosol distribution).

Effect of Aerosols on brightness temperature

Aerosols in the atmosphere are linked with cloud formation as discussed in Chapter 4. Aerosols also effect the net radiative forcing of the atmosphere as they scatter and absorb infrared radiation [72]. They are an important constituent of the atmosphere, and analysis is done extensively on their effects on the transmission of light and on extensive air shower reconstruction [39].

As a result, one wishes to check the difference in brightness temperature for each zenith angle when comparing different aerosol and cirrus cloud distributions, to see the effects of aerosols on this radiative signature. Figure 7.27 shows the difference in brightness temperature for a high (aerosol vertical optical depth of 0.18) and low (aerosol vertical optical depth of 0.05) aerosol distributions in clear sky conditions.

Figure 7.27 illustrates that a higher aerosol distribution atmosphere is warmer than a low aerosol distribution atmosphere. Additionally, Figure 7.27 shows an increase in the brightness temperature difference at higher zenith angles (peaking at $\sim 80°$). The decrease at zenith angles close to the horizon is due to the atmosphere becoming optically thick, thus all radiative models read the same infrared temperature. It also shows that changing the aerosol content in the atmosphere has a minimal effect on the radiance detected at ground level.

The next step is to look at this difference again, but with a cirrus cloud compared to a high level of aerosols. This is shown in Figure 7.28.

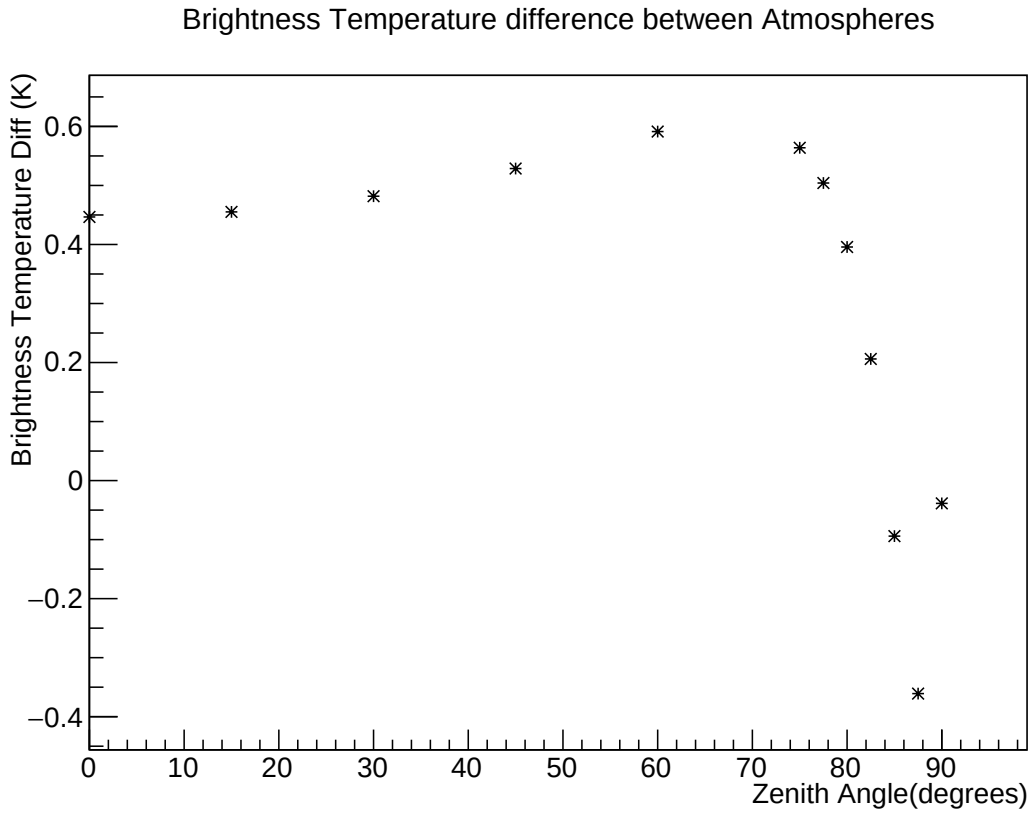

Figure 7.28: Difference in Brightness Temperatures at each Zenith Angle for a 4 km cirrus cloud and a high concentration of aerosols. The difference is illustrated as Brightness temperature (cloud) - brightness temperature (aerosols). The aerosol atmosphere does not contain a cirrus cloud, and the cirrus cloud atmosphere does not contain aerosols.

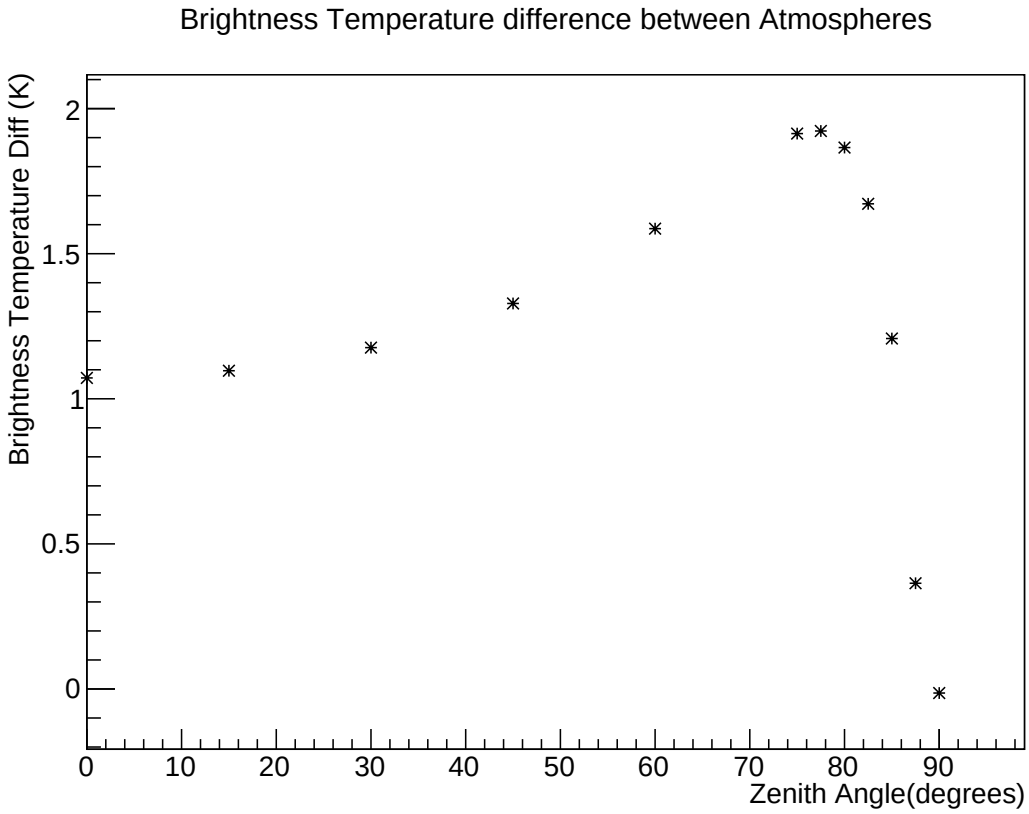

Figure 7.29: Difference in brightness temperatures as a function of zenith angle for a 4 km cirrus cloud and the low aerosol model. Once again, the difference is illustrated as Brightness temperature (cloud) - brightness temperature (aerosols but no cirrus).

The brightness temperature difference in Figure 7.28 is small, indicating that even a high aerosol concentration in the atmosphere is leaving the radiative signature largely unchanged (since cirrus cloud itself makes very little difference as described in previous sections).

Figure 7.29 shows similar behaviour to that of Figure 7.28, in that aerosols present in the atmosphere are having a minimal effect on the radiance of the atmosphere. Additionally, Figure 7.29 further indicates that cirrus cloud is having a larger effect on the radiation/absorption in the infrared band than the aerosols are.

7.4 Summary

In summary, cirrus cloud present in the field of view of the IR cloud cameras at Auger is difficult to detect due to its differing radiative properties (from thick cloud). Due to cirrus cloud existing at such high altitude, extensive air showers (which develop below the height of the cloud) are largely unaffected by their presence, thus cirrus cloud can be ignored for the purposes of air shower reconstruction. The presence of aerosols in the atmosphere was also shown to minimally affect the brightness temperature of the atmosphere.

Chapter 8

Estimation of Cloud Base Height using the Infrared Cloud Cameras

There have been previous attempts to use the information from the IR cloud cameras at Auger (see Chapter 5) in order to estimate the height of clouds in the atmosphere, however they have been unsuccessful. The camera itself can detect clouds in the field of view, but there was (until now) no way of knowing how high up, or how far away, the cloud is. Knowing these parameters would be a huge positive for the reconstruction of air showers since the cloud cuts applied could be relaxed if the shower is known to develop below the base of the cloud (see section 3.5.9).

This chapter outlines a technique in development which allows the base height of a cloud in the camera's Field of View (FOV) to be estimated, by first estimating the temperature of the cloud. The calculated cloud height using this technique will be compared to other cloud height measurements already done at Auger, to test its validity. The outline of this technique is described in section 8.1.

8.1 Cloud Height Estimation Technique

There has been a previous attempt to measure cloud temperatures in the infrared [97]. This study measured cloud height and sky temperature using a single pixel radiometer with a bandpass centred at 10 μm [97]. The estimated sky temperature T_{sky} from the radiometer (for totally overcast conditions) was then compared to the calculated sky temperature from the ceilometer using equation 8.1 [97].

$$T_{sky} = T_{ground} - \Gamma \times h \tag{8.1}$$

T_{sky} was estimated by using the measured ground temperature T_{ground} and cloud base height h from the ceilometer, and assuming an average lapse rate Γ of 2.2 K/300 m [97]. The results of comparing these two temperatures are shown in Figure 8.1 for cloud base heights < 750 m, and in Figure 8.2 for cloud base heights > 1200 m [97].

The results from Riordan et al. [2005] showed agreement between the cloud temperature and sky temperature for low clouds (base height < 750 m - as shown in Figure 8.1), but the agreement appeared to get worse for higher altitudes, as can be seen by Figure 8.2 [97]. For higher altitudes, the measured temperature from the radiometer was always warmer than the estimated temperature from the lapse rate. Two effects were thought to be the cause of the warmer measured temperature, one being the effect

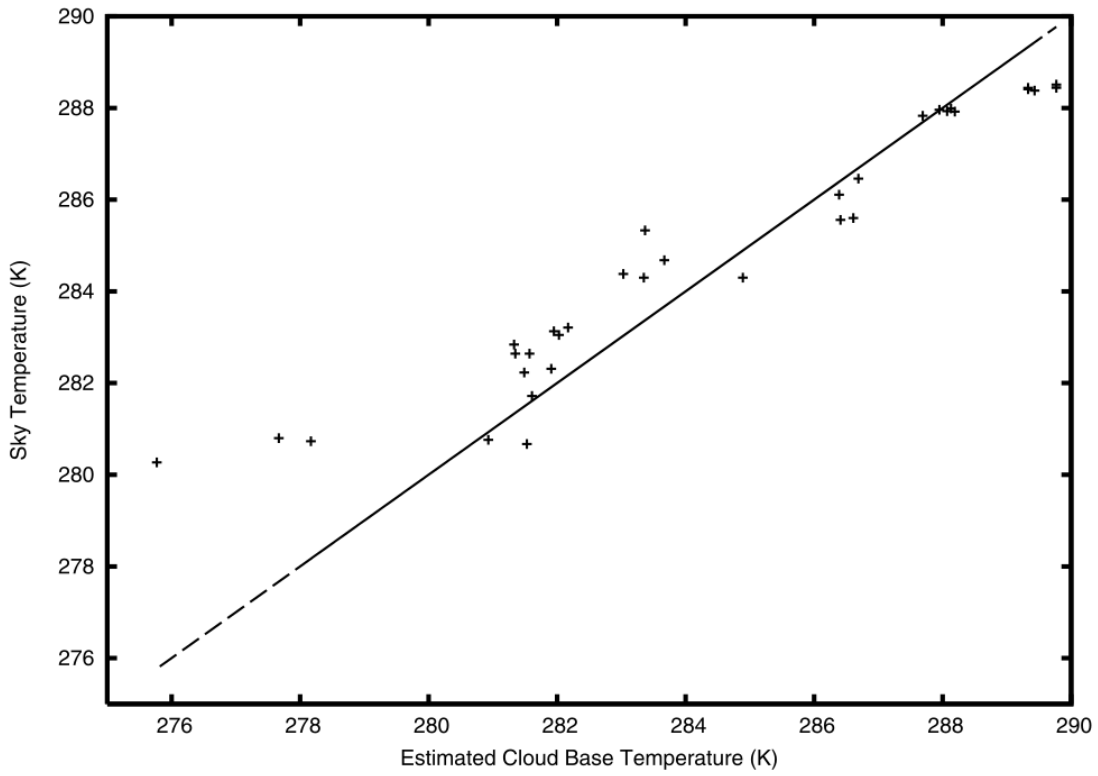

Figure 8.1: A comparison of cloud base temperature from the cloud height and an average lapse rate (from the ceilometer - horizontal axis) to the measured sky temperature (from the radiometer - vertical axis) for cloud base heights $< 750\,\mathrm{m}$. Source: [97].

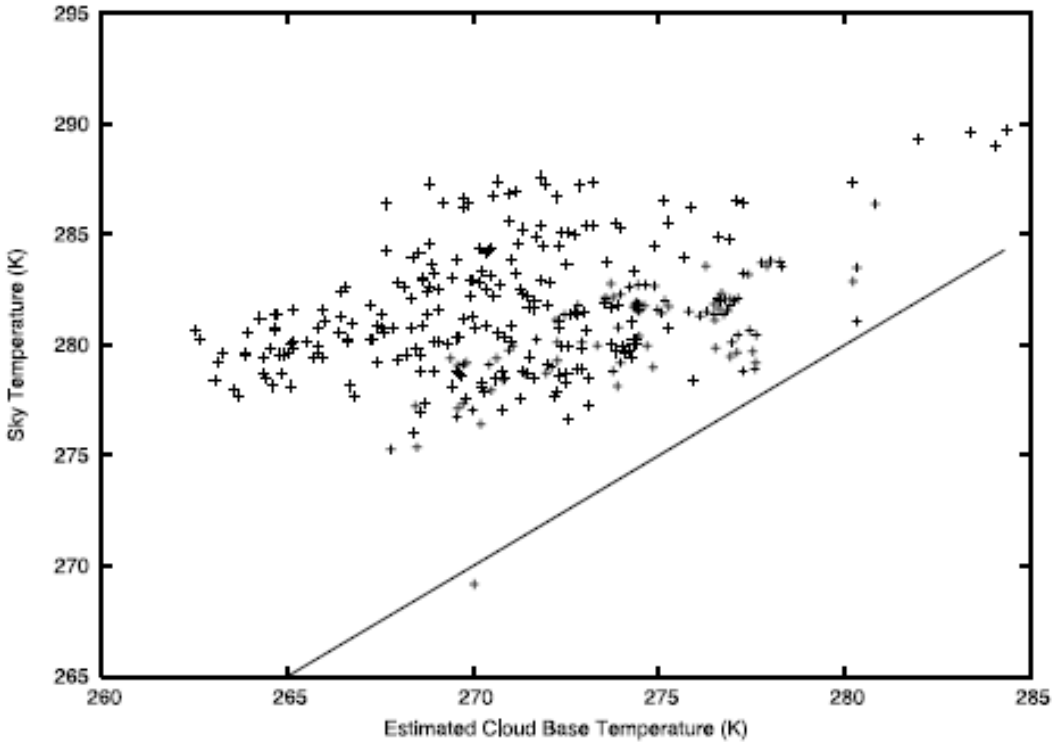

Figure 8.2: A comparison of cloud base temperature estimated from the cloud height and an average lapse rate (from the ceilometer - horizontal axis) to the measured sky temperature (from the radiometer - vertical axis) for cloud base heights $>$ 1200 m. Source: [97].

of intervening water vapour between the ground and the cloud [97], and the other effect due to the reflection of ground heat by the cloud (which occurs when a cloud is not a perfect blackbody radiator).

By utilising sky temperature measurements from the cloud cameras at Auger, the aim is to improve upon these results found previously, by improving the estimation of cloud base heights and by better understanding the effects causing deviations in the estimated cloud base height. To do this, there are two methods that will be presented in this chapter, both of which use the measured sky temperature from the cloud cameras, and use the GDAS temperature profiles (see section 3.5.2) to estimate the height of clouds in the field of view of the cloud cameras. The first method measures the sky temperature T_{sky} and finds the height in the GDAS profile which has a temperature that matches T_{sky}. This will be referred to as Method A in this chapter. To get the correct temperature to match the GDAS profile, the difference between ground temperature of GDAS and the local weather station is found, before the GDAS height profile is adjusted by that difference. This allows for the most accurate value of temperature to compare the measured sky temperature to. The second method, denoted Method B, uses the measured sky temperature T_{sky} from the cameras, the measured ground temperature T_{ground} from the local weather station and a fitted lapse rate Γ to the GDAS temperature profile (which has its ground temperature matched to the weather station as it is in method A). The cloud base height is then estimated by equation 8.2.

$$h = \frac{T_{ground} - T_{sky}}{\Gamma} \tag{8.2}$$

The fit region of the GDAS profile is chosen to be from ground level up to 5 km above ground level (A.G.L). The reason for this is discussed in section 8.3.3.

To measure cloud temperature from the cloud camera, one finds times when the vertical image from a full sky scan (see section 5.4) contains cloud. There are two possible scenarios for this. First, if the image is totally overcast, we measure the sky temperature as the mean temperature value of all pixels in the vertical image. If the image is only partially cloudy, we select only those pixels that contain cloud and calculate the average temperature of those cloudy pixels. The illustration of this measured sky temperature is discussed in section 8.3.4.

In order to test the validity of this measured height, another instrument is required to compare to, thus making sure that the cloud base height estimations from Methods A and B are accurate. The cloud height given by each FD viewing the CLF (see section 3.5.3) is used for this purpose. The central laser facility measures cloud base height every 15 minutes. However, the cloud height that is input to the database (to compare to the height given by equation 8.2) is the lowest cloud base height for an entire hour (out of the four 15 minute laser shots), which may lead to uncertainties in the calculated cloud height.

One key assumption that is made in estimating cloud base heights is that cloud viewed by the cloud camera radiates as a blackbody with an emissivity of 1. The potential issues with this assumption will be discussed in section 8.3.1. Section 8.2 describes the results of testing the two cloud height estimation methods.

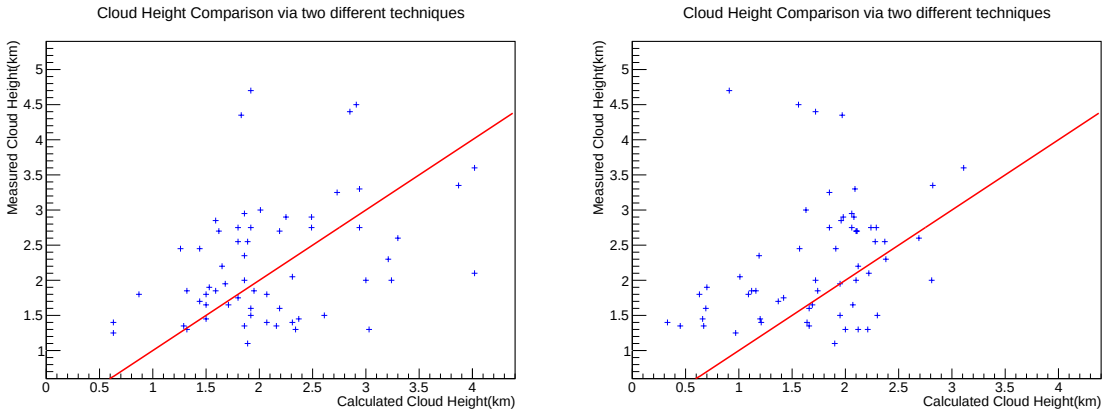

Figure 8.3: Left: Method A comparison of calculated height (horizontal axis) [found by matching the temperature in GDAS to the sky temperature from the cloud camera] with the height given by the CLF laser shots (vertical axis). A red 1:1 line is drawn for the sake of comparison. Right: Method B comparison of calculated height (horizontal axis) using equation 8.2 with the measured height from the CLF (vertical axis). The red line corresponds to equality of the two heights.

8.2 Testing of the Technique

The initial test of both Method A (left) and B (right) are shown in Figure 8.3, with all initial ground temperature, lapse rate and sky temperature corrections applied as described in the previous section. The selected nights were chosen over 2 years (2017 and 2018) where the CLF recorded a cloud base height below 5 km and the cloud camera imaged a cloud in the vertical image (see section 5.4). One good way of representing how well each method performed is to look at the residual difference between the heights via each method compared to the CLF. These are shown in Figure 8.4 for method A (left) and method B (right).

Figures 8.3 and 8.4 show that method A (using the GDAS profile to match the temperature to that of the cloud camera) provides a better estimate of cloud height than method B, however there is a large degree of spread in the estimation of cloud base height. The sources of uncertainty that could contribute to this spread will be discussed in section 8.3.

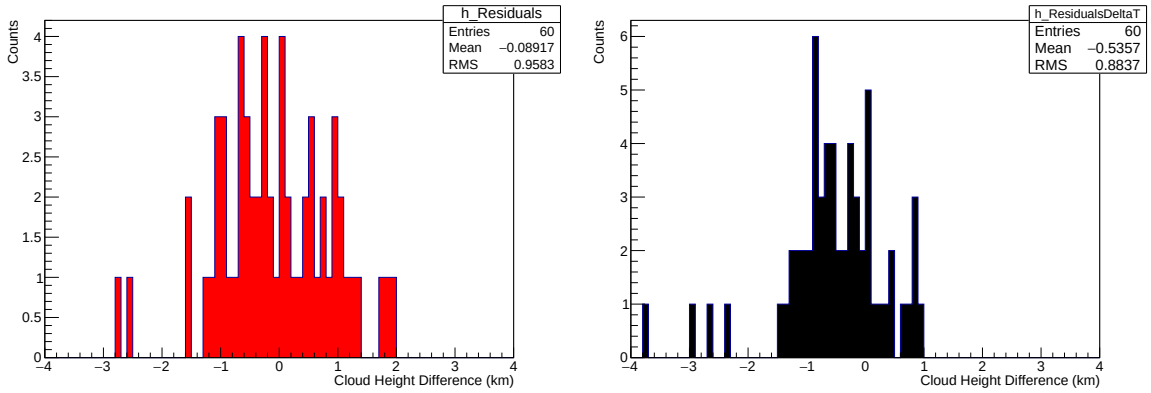

Figure 8.4: Residuals between the cloud base height estimated via method A (left) and method B (right) with the cloud base height measured by the CLF without corrections applied.

8.3 Sources of Error in the Technique

There are various potential sources of error when estimating the height of clouds in the atmosphere. As found in previous studies, the effect of intervening water vapour (and potentially other molecules) can cause an overestimate of sky temperature, which will be further explored in section 8.4 [97]. Another potential source of error in the estimation of cloud base heights is the emissivity of a cloud, which, when not equal to 1, can cause a difference in a cloud's measured temperature to its true temperature.

8.3.1 Emissivity

Firstly, sky temperature in the infrared depends on an object's emissivity, which is lower than one for optically thin clouds. The Stefan-Boltzmann law shows how the radiance R depends upon the emissivity ϵ and temperature T for a blackbody.

$$R = \epsilon \sigma T^4 \tag{8.3}$$

Thick clouds that act as blackbodies will radiate as shown by equation 7.2. If the emissivity is slightly less than 1 [typically between 0.9 and 1], this means the cloud will reflect incident radiation from the ground (the reflectivity is given by $1 - \epsilon$), causing a warmer sky temperature to be measured than what is expected for the cloud. Asa result, the cloud height estimated will be lower than expected for the cloud. In this scenario, the radiation is either reflected off the cloud or absorbed by the cloud. This has also been shown in previous results [97].

On the other hand, thinner clouds do not follow the Stefan-Boltzmann Law, due to their transparency. For clouds with an emissivity much smaller than 1, the radiance R (for a given temperature T) is also much smaller. This makes the cloud appear much cooler than expected, leading to an estimate in the cloud base height that is much higher. These two emissivity effects compete with each other and can cause deviations in the cloud height both above and below what is expected, as shown in Figure 8.3, depending on the type of cloud present.

There are other smaller effects that must also be considered for the purposes of this study. For instance, the full sky scan from the cloud camera only takes place every 15 minutes, and in that time frame, cloud can move, change and morph substantially due to atmospheric circulation. This can happen over as short a period as 5 minutes. Thus, each instrument monitoring the atmosphere will be potentially looking at a different patch of cloud. As a result, it would not be possible for the true height to be estimated. As is the case for previous studies utilising GDAS temperature profiles to estimate cloud base heights (see section 4.4.1), there are some small issues which could potentially affect the accuracy of the lapse rate calculation. The data are only modelled every 3 hours (at 00, 03, 06 and 09 UTC over the course of a night). Therefore, times being tested that are far away from this time (for example at 01:30, 04:30 and 07:30 UTC) will most likely be unable to accurately model the atmospheric conditions at that time. More on the uncertainties associated with the lapse rate can be found in section 8.3.3. Next, the uncertainty on ground temperature will be discussed.

8.3.2 Ground Temperature

The first variable to consider from equation 8.2 is the ground temperature T_{ground}. There are two different instruments one can use for the ground temperature measurement. The standard one used in this technique is the local weather station, however one can also extract ground temperature from GDAS given the known altitude of each individual FD. Ideally, these two ground temperatures will agree, and one can plot the difference of the two in a histogram. The result of this is shown in Figure 8.5 for the data used in Figure 8.3.

Figure 8.5 shows that there is an average difference of 2 K between the two ground temperatures from the two instruments, meaning that GDAS reads a slightly higher temperature (on average). However, this difference between the two ground temperatures can be substantial (due to a large 3 K RMS spread), suggesting that GDAS cannot always accurately model the true ground temperature. Furthermore, one can check how the difference between the ground temperatures varies for each three hour period over a night. Figure 8.6 shows the difference in ground temperature between GDAS and the weather stations for each three hour interval over a 24 hour period.

Figure 8.6 shows that there is a clear diurnal variation in the difference between the two ground temperatures. On average, the best agreement occurs at 0 and 12 UTC [48]. This is due to the nature of the data assimilation model. GDAS uses data from radiosonde flights, which are typically performed at 0 and 12 UTC [98]. For these hours, GDAS has a much more accurate handle on the true nature of the atmosphere since radiosonde data fed into the model provides the best representation of the atmosphere at that time [98].

As a result, the best choice for ground temperature in the lapse rate method (method B) is to use the local weather station ground temperature. Due to the large spread in the two ground temperatures, the best way to handle ground temperature for method A is to calculate the ground temperature difference between GDAS and the weather station, and subsequently adjust the entire GDAS profile by this difference. This way, one can estimate the best temperature from the GDAS profile that matches the measured sky temperature from the cloud camera. These are the corrections to the GDAS profile for both methods that are employed in Figure 8.3.

The next parameter to check for uncertainty is the lapse rate, Γ.

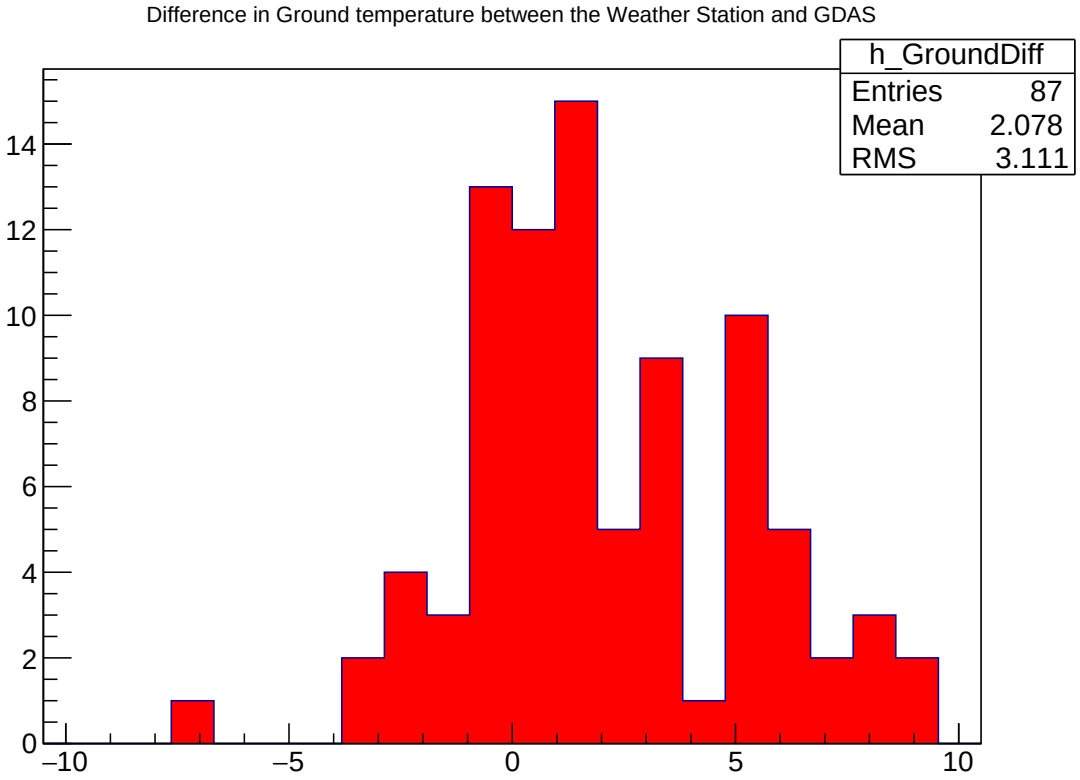

Figure 8.5: Histogram of the difference between the ground temperature modelled by GDAS and the ground temperature read by the local weather station for the cloud camera scans chosen in Figure 8.3 [in units of K]. The difference is read as Ground Temperature (GDAS) - Ground Temperature (Weather Station).

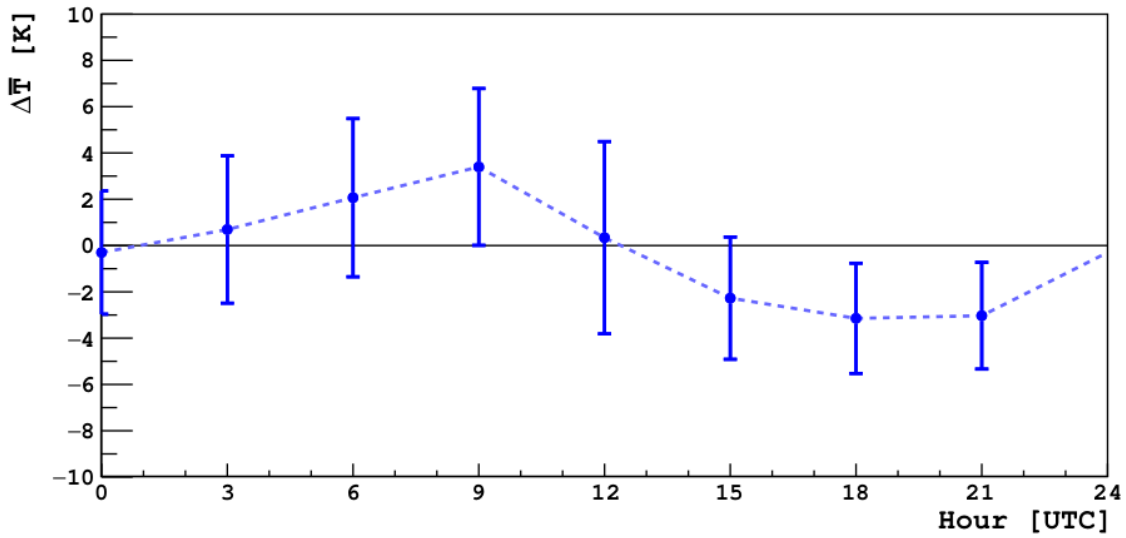

Figure 8.6: Comparison of the difference between the GDAS model and weather station ground temperatures for each 3 hour period over a day [48]. Figure 8.5 corresponds to 0-9 UTC, which is preferentially positive and agrees with the results shown here. Error bars indicate standard error on the mean.

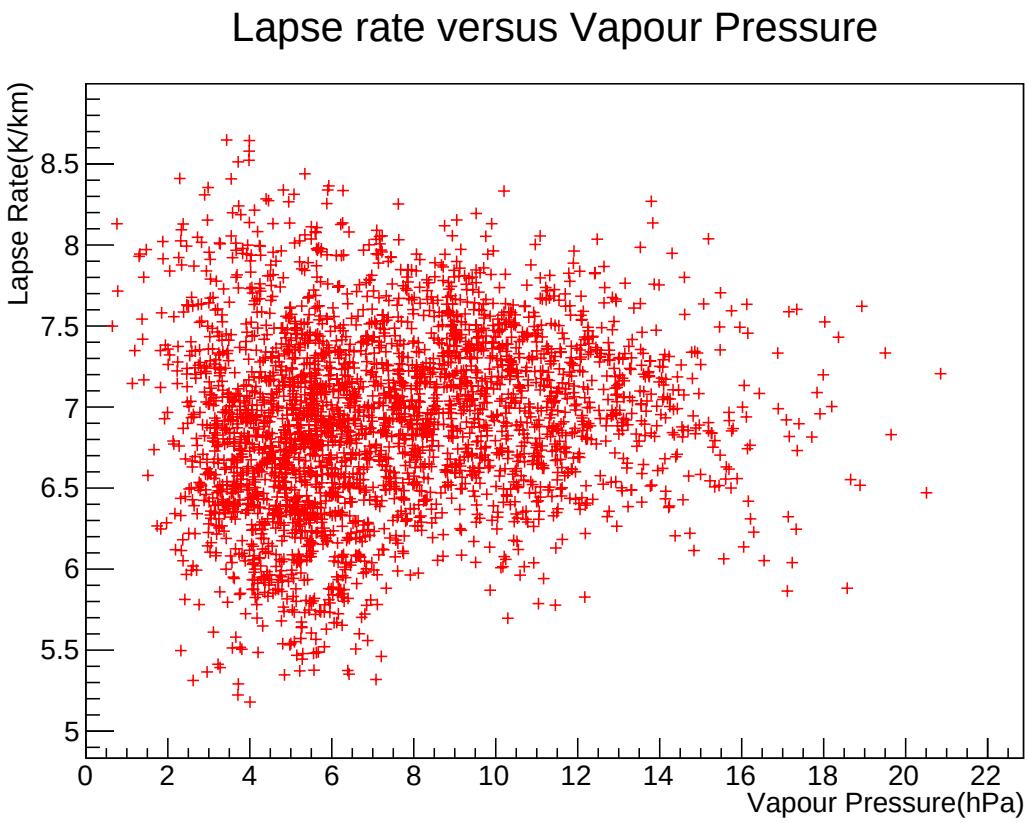

Figure 8.7: Lapse rate (fit between 4 and 10 km above ground level) of the atmosphere as a function of the vapour pressure (at ground level) - taken every 3 hours from GDAS.

8.3.3 Lapse Rate

Looking at the lapse rate, which one can study as a function of time, leads to some interesting outcomes. Firstly, the linear fit to the GDAS profile to calculate the lapse rate has been taken from 4-10 km (above ground level) which matches what was done from the previous lifting condensation level study (see Chapter 4). For the purposes of this analysis, since the majority of clouds found are below 5 km, the lapse rate fit will only be taken from ground level up to 5 km. The reason for this is shown later on in this section. This lapse rate (fit from ground level to 5 km above ground level) is applied to method B in Figure 8.3, and will hopefully allow for a more accurate calculation of the lapse rate.

The lapse rate depends on a number of atmospheric factors such as humidity, pressure and density. One can plot these parameters against the lapse rate to see if there are other determining factors. Firstly, Figure 8.7 shows the lapse rate as a function of the vapour pressure given by GDAS.

Figure 8.7 shows a correlation between the lapse rate and the vapour pressure. This is due to the process of adiabatic cooling. When a parcel of air rises in the atmosphere, water vapour inside the parcel will condense and release latent heat. The higher the vapour pressure at ground level (i.e. the more water vapour there is), the more latent heat is released and the more slowly a parcel will cool as it rises. This is why lapse rate is dependent upon the vapour pressure.

However, a large spread is present in Figure 8.7, thus fitting a functional form to the lapse rate as a function of vapour pressure is not straightforward, as there are other

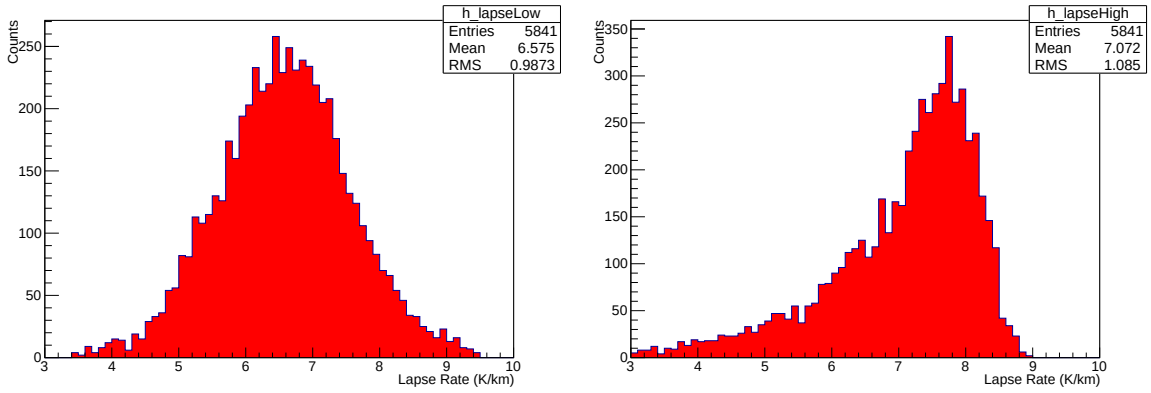

Figure 8.8: Distribution of Lapse Rates over 2 years fitting between the ground and 5 km above ground level (left), and between 5-10 km above ground level (right) for all times of the day.

factors contributing to this spread. It is worth noting that this correlation exists, as it may be able to explain some of the behaviour shown in the results (see Figure 8.3).

The next step is to find the best region of the atmosphere to fit the temperature profile from GDAS, in order to get the best possible lapse rate to use for cloud height estimation. The analysis of the LCL calculations from Chapter 4 fitted the temperature profile in the 4-10 km height range, due to this portion of the atmosphere being known as the "free" troposphere. However, for the purposes of cloud height estimation, this may not be the best region to choose due to the fact that he majority of clouds being investigated are below this height range. As a result, the fit may perform better over a lower height range. Figure 8.8 shows the distributions of lapse rate values in 2017 and 2018 by fitting over two different height ranges (0-5 km above ground and 5-10 km above ground).

The results of Figure 8.8 shows a more consistent fit for the 0-5 km region due to the lower overall spread compared to the 5-10 km region (RMS values of 0.99 K/km compared to 1.09 K/km). It is also interesting to note that the mean value of lapse rate for the 0-5 km is 6.575 K/km, close to the standard value of 6.5 K/km chosen earlier. This fit appears to be the better overall choice in order to calculate a lapse rate, however this requires cross-checking of the fit routine.

One can test the goodness of fit in the two fit regions by looking at the RMS (root mean squared) values for each fit to the GDAS temperature profile. The RMS is calculated by equation 8.4.

$$RMS = \frac{1}{N}\sqrt{\sum_i (T_i - T_{expected})^2} \tag{8.4}$$

where T_i is the temperature of the GDAS profile at a given height, and $T_{expected}$ is the temperature one would expect for a given height (above ground level) assuming a linear lapse rate Γ over a given fit region. This is shown by equation 8.5. If the linear fit to the lapse rate is accurate, then T_i and $T_{expected}$ will match, as the expected temperature from the fit will match the true temperature T_i.

$$T_{expected} = T_{ground} - (\Gamma h) \tag{8.5}$$

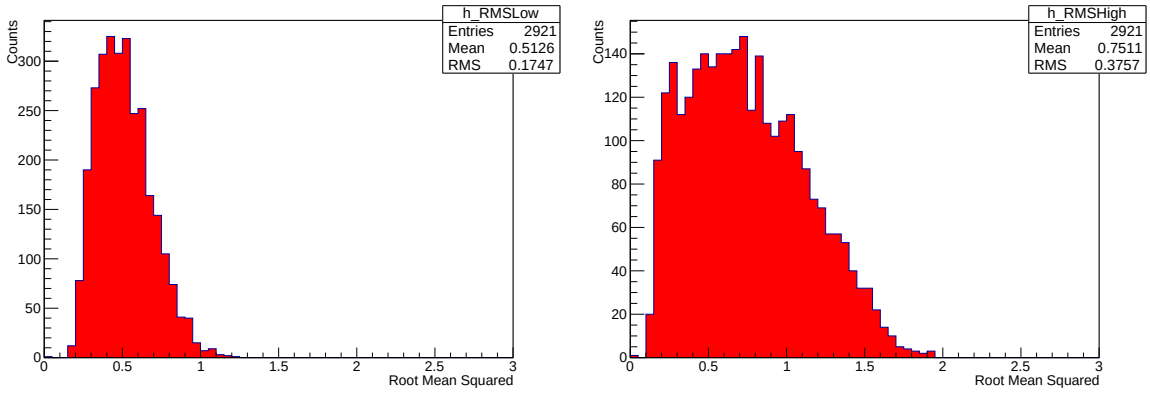

Figure 8.9: Root Mean Squared (in units of K) fit to the lapse rate over 2 years (2017/2018) fitting between the ground and 5 km (left) and between 5-10 km A.G.L (right).

As a result, for a perfect straight line fit this RMS would be zero. Figure 8.9 shows the results of the calculation of RMS for each fitted GDAS profile over the two height regions.

Figure 8.9 shows that the 0-5 km fit had a lower average RMS value than the fit over the 5-10 km region. This indicates a more accurate linear fit from 0-5 km. The lapse rate being fit over the 0-5 km region is still the best choice, due to all clouds being considered belonging to this height range. This is why the linear fit region for the lapse rate was chosen to be 0-5 km above ground level for the application of equation 8.2 in Figure 8.3.

The lapse rate value is not causing any systematic differences between the calculated cloud height and the measured cloud base height from the CLF, as the best possible choice has already been applied. Thus, the spread in the cloud base height is due to another factor. The final parameter to check (from equation 8.2) for sources of uncertainty is the sky temperature measured by the cloud camera.

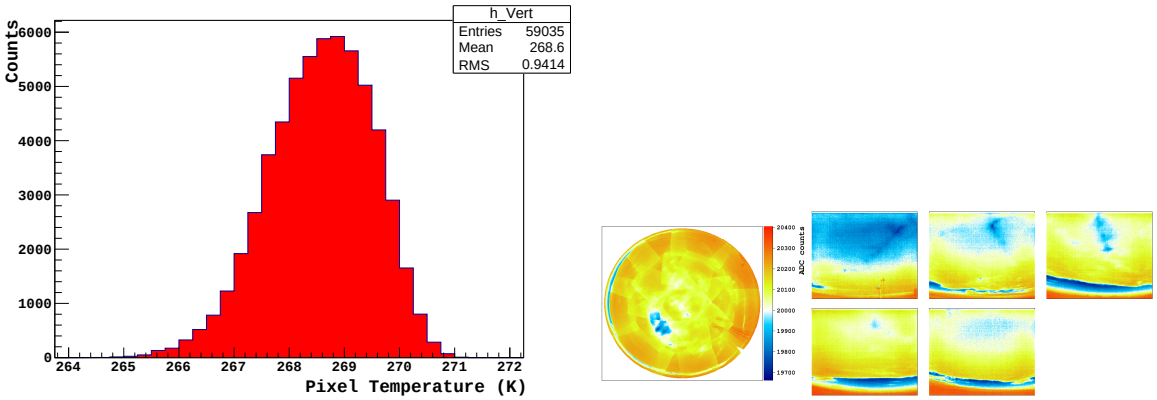

Figure 8.10: Left: Distribution of Brightness Temperature of every pixel in the vertical image that is totally overcast. Vertical pixels are defined as those pixels with a pointing zenith angle between 0 and 30°. Right: Cloud Camera field of view for this image, showing the entire vertical image filled with cloud [left image], also showing the cloud camera field of view that matches the local FD [right image]. Image is taken from Los Morados on July 24th, 2017.

8.3.4 Sky Temperature

The final variable to consider is the sky temperature detected by the cloud camera. A measure of uncertainty on the sky temperature is the distribution of temperatures across the entire vertical image, for each pixel. Figure 8.10 shows the distribution of temperatures across each pixel in the vertical image (left) which is completely filled with cloud (right).

Figure 8.10 shows good consistency of temperature over each pixel. There is a spread of 0.94 K, which would correspond to a spread in cloud heights of $\sim$ 140 m, which illustrates that the mean temperature value of cloudy pixels is a good estimate of the true temperature of the cloud.

Next, one wishes to compare the pixel temperatures for an image that is partially cloudy, shown in Figure 8.11. This is expected to have a much larger spread.

As expected, Figure 8.11 shows a much larger spread than the completely overcast image. It also displays a distinct double peak pattern. This behaviour is well understood since the cloudy portions of the image are considerably warmer than the background clear sky. The warmer right hand peak (corresponding to the cloudy portions of the image) also has a small spread, similar to Figure 8.10. In this case, taking the value of the warmer peak (as opposed to the mean of the entire image) will give the best estimate to the cloud brightness temperature.

Lastly, one can look at the temperature distribution from a clear vertical image. This is displayed in Figure 8.12.

Figure 8.12 shows a similar distribution to the overcast image, namely good consistency in the spread (being low). Interestingly, in the case of the clear image, the spread is much smaller than in the overcast image. There is a reason why this spread is present, even for a clear image.

As discussed in Chapter 5, there are pixel-by-pixel variations in the infrared signal received by each cloud camera. The flat-field correction and temperature calibrations of the cloud cameras at Auger (see chapter 5) attempt to reconcile these variations,

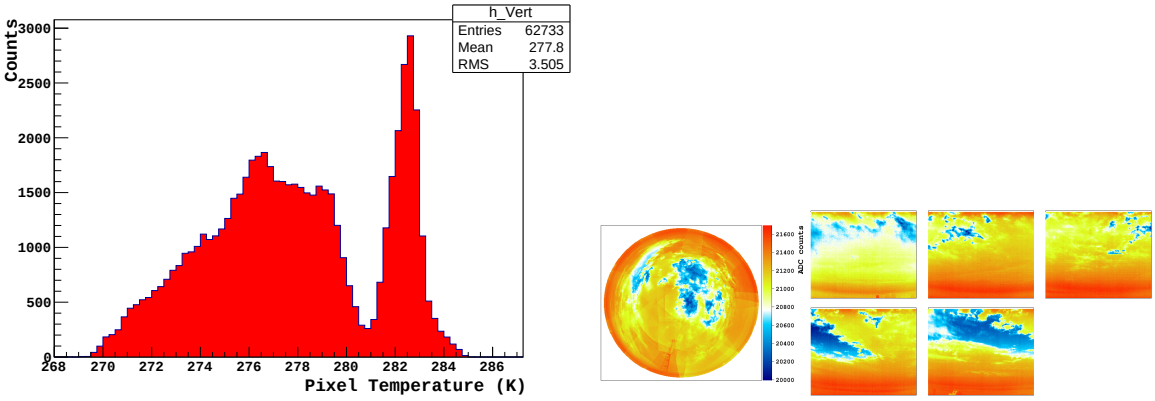

Figure 8.11: Distribution of Brightness Temperature of every pixel in the Vertical Image that is partially overcast. Image taken from Los Leones on February 26th, 2017.

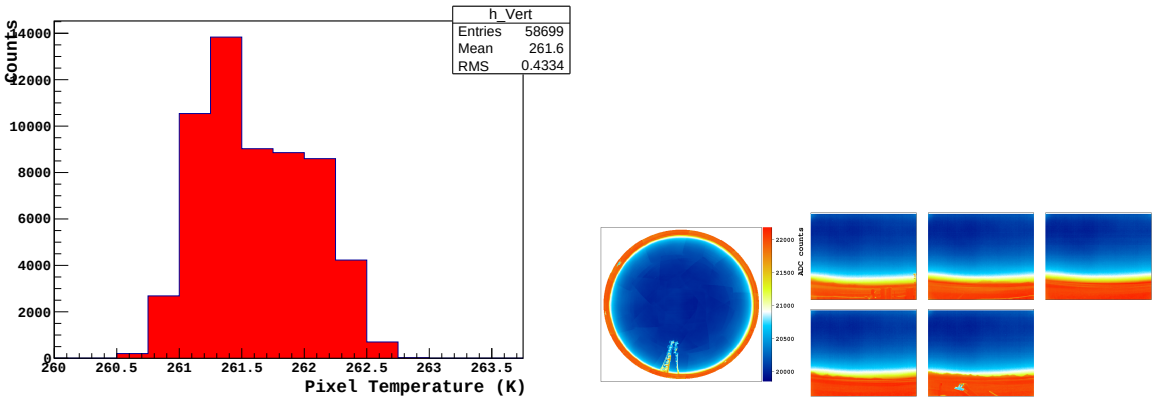

Figure 8.12: Distribution of brightness temperature of every pixel in the vertical image that is completely clear. Image taken from Los Leones on the 29th of December, 2018.

however there are still small systematic differences in the gain of each cloud camera pixel (even after applying these corrections), which would mean each one reads a slightly different brightness temperature. The best way of reducing this uncertainty is to use the average temperature of the peak of the distribution, since this is the best measure of temperature of the majority of cloudy pixels in the image. This is the correction made for sky temperature in both cloud base height estimation methods shown by Figure 8.3.

Of course, the main question that arises from this is the following: how accurate is this measured sky temperature (from the cloud camera) to the true temperature of the cloud in the atmosphere?

The remaining sections describe the study done to answer this question, using MODTRAN (see section 7.1).

8.4 Studies with Blackbody Curves

As discussed in section 4.2, clouds in the atmosphere emit their own infrared blackbody radiation B_λ (the amount of energy emitted by an object per unit area and per unit wavelength), which is a continuous spectrum over a range of wavelengths, and purely depends on the temperature of the cloud, as given by Planck's Law:

$$B_\lambda = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \tag{8.6}$$

where λ is the wavelength of the emitted radiation, h is the Planck constant, k is the Boltzmann constant and c is the speed of light in a vacuum. The wavelength λ_{max} where the blackbody spectrum peaks is given by Wien's Displacement Law as:

$$\lambda_{max} = \frac{2.9 \times 10^{-3}}{T} \tag{8.7}$$

One can then calculate the radiance of any blackbody at a given temperature T as given by the Stefan-Boltzmann Law:

$$R = \sigma T^4 \tag{8.8}$$

The radiance R is calculated as the integral of the Planck curve (equation 8.6) as in equation 8.9.

$$R = \int_0^\infty B(\lambda) d\lambda \tag{8.9}$$

It is this radiance R that is measured by the infrared cloud cameras at Auger, which is used to infer the temperature of a cloud in the cloud camera's field of view. Assuming a blackbody radiator, and in the absence of any other forms of radiation, the cloud would radiate at a given temperature which would be accurate to the true temperature of the cloud as calculated by equation 8.9.

One requires a method of correctly checking the output from MODTRAN in terms of this radiance, when clouds are present in the atmosphere. As was shown by Riordan et al. [2005], the measured sky temperature estimated from the radiometer was warmer than is expected given a cloud base height and lapse rate [97].

The cloud temperature being warmer than expected is due to the presence of another form of radiation, which comes in two parts. As described in section 4.2.1, the

Earth also radiates as a blackbody (with an approximate temperature of $\sim 288\,\mathrm{K}$ and peak wavelength of $\sim 10.1\,\mu\mathrm{m}$), and this radiation can be absorbed by the cloud, heating it up. This causes a warmer temperature of the cloud to be measured by our infrared detectors than what is expected. This radiation from the Earth can also be reflected by the cloud, which contributes to the warmer temperature that is measured. Similarly, there is another source of heat that also comes from the Earth's blackbody radiation. Various molecular gases (such as water vapour, carbon dioxde, methane, nitrous oxide and ozone) can absorb the Earth's blackbody radiation [68]. Infrared radiation is absorbed by these molecules when their electric dipole moment changes as they rotate or vibrate [94]. As long as the frequency of radiation matches the resonant frequencies of the given molecule, the radiation can be absorbed. All gases mentioned previously are able to absorb this infrared radiation. Once absorbed, it is then re-emitted at the same characteristic wavelength, adding an extra source of radiation into the atmosphere which is detected by the cloud cameras. This extra radiation may also be absorbed or reflected by the cloud.

The cloud cameras are designed to be sensitive to radiation in the atmospheric window ($\sim 7-14\,\mu\mathrm{m}$) since this is the wavelength range where radiation is least affected by the constituents of the atmosphere. Infrared radiation outside of this wavelength range is also absorbed by molecular gases (mainly water vapour and carbon dioxide). However, integrating the radiance curve (as in equation 8.9) at these wavelengths is not an accurate representation of the cloud temperature since a portion of this radiance is due to these atmospheric molecules absorbing and re-emitting radiation between the ground and the cloud [99]. This biases the temperature, making it an inaccurate portrayal of the temperature of the cloud.

The combination of these two sources of extra radiation (due to the Earth) within the atmospheric window is what contributes to the overestimate in temperature seen previously, which was unaccounted for [97]. One wishes to look at each of these effects separately, and determine their significance to the difference in the measurement of cloud temperature. As described in section 4.2.1, the release of latent heat when water vapour condenses will also cause the cloud to warm up, thus reading a warmer temperature than expected. This effect, combined with the effects due to terrestrial radiation, must be accounted for in order to measure an accurate cloud temperature.

Figure 8.13 illustrates the difference in radiance for an atmosphere that contains a cloud, compared to the expected radiance of the cloud given it is at a known height (hence a known temperature), assuming it is a blackbody. Placing a cloud at 8 km altitude produces the blue curve in Figure 8.13. The US standard atmosphere model used in MODTRAN has a ground temperature of 288.2 K, and a temperature of 236.2 K at 8 km altitude. The blackbody curves for these temperatures are the red (ground temperature) and black (temperature at 8 km altitude) curves in Figure 8.13. The radiance of the cloud is calculated assuming that the observer is located at ground level.

Two peaks in the radiance curve from Figure 8.13 are present, at ~ 7 and $14\,\mu\mathrm{m}$. The first of these (left hand peak) is due to water vapour close to the ground absorbing radiation from the ground and reaching thermal equilibrium with the Earth, and the other (right hand peak) is due to the same phenomenon exhibited by carbon dioxide. This effect extends beyond $14\,\mu\mathrm{m}$ as both carbon dioxide and water vapour reach a thermal equilibrium with the Earth at these longer wavelengths. This is why the temperature at these longer wavelengths matches the ground temperature (the red

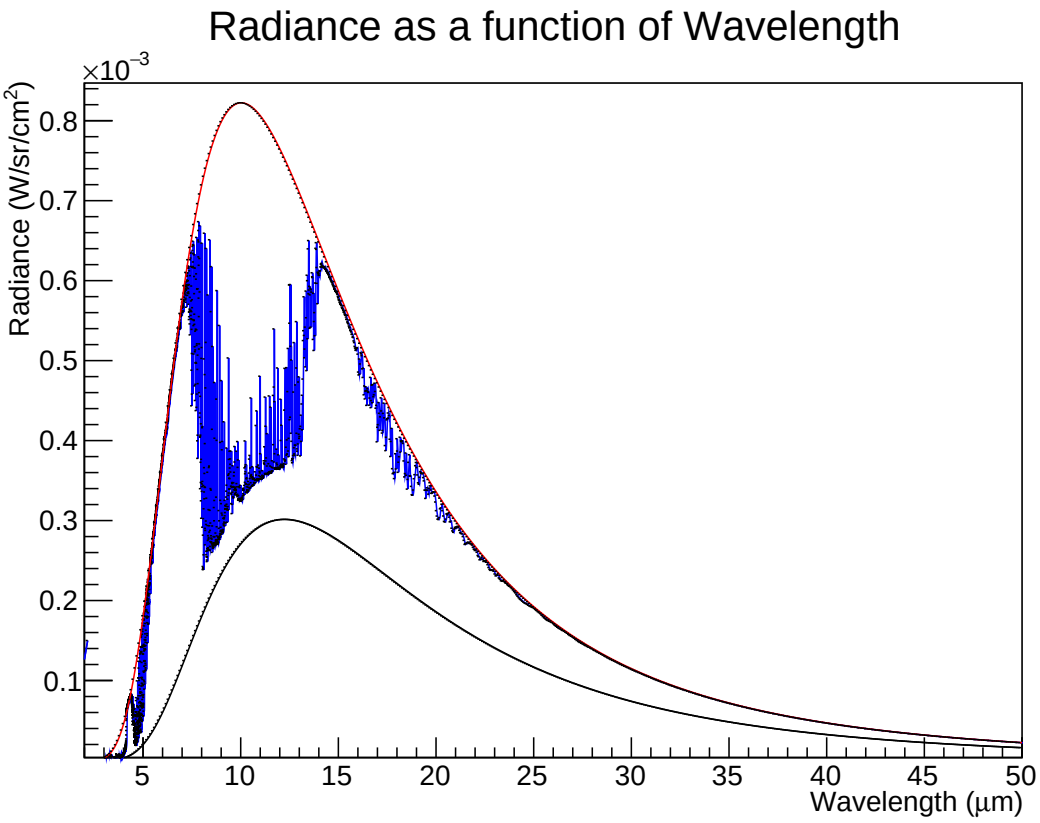

Figure 8.13: Radiance output for a 288.2 K blackbody (Red), a 236.2 K blackbody (black) and for a cumulus cloud (blue) at 8 km height in the atmosphere. The red curve is for a ground level blackbody, and the black curve is for a blackbody at the height of the cloud. The cloud radiance sits in between the two blackbody curves, suggesting a combination of absorption features that increase the temperature of the cloud from what is expected (black).

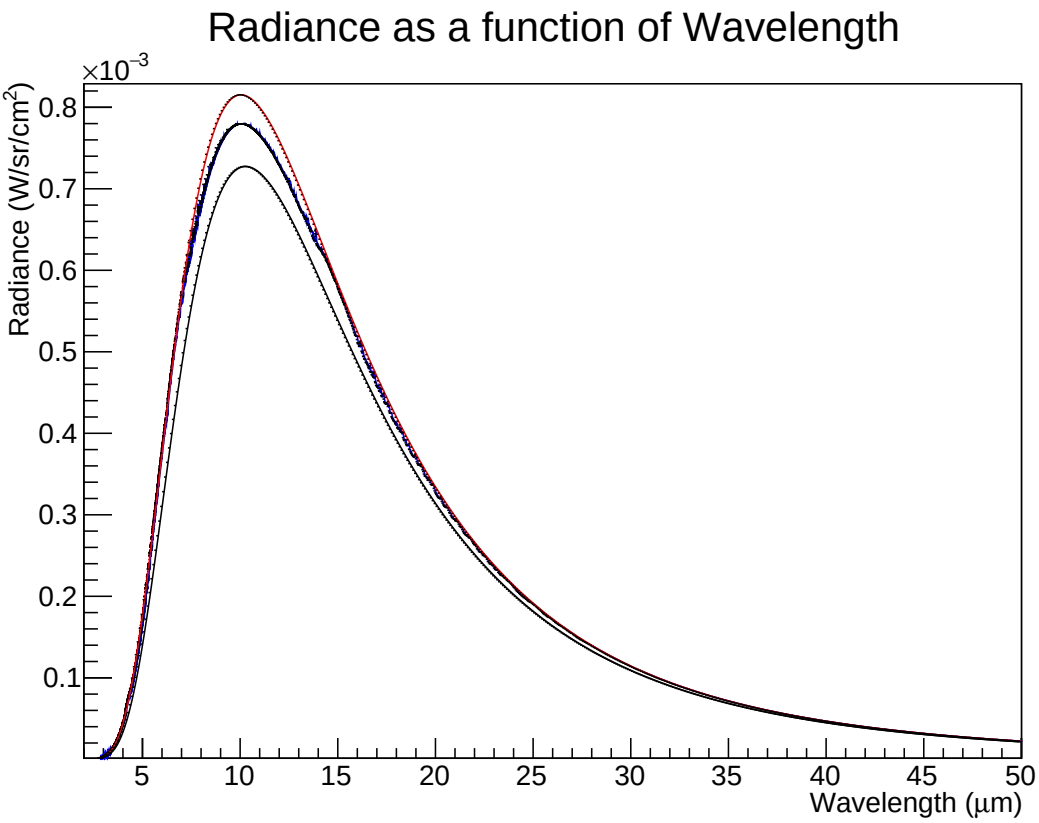

Figure 8.14: Radiance output for a 288.2 K blackbody (Red), a 281.7 K blackbody (black) and for a cumulus cloud (blue) at 1 km height in the atmosphere. The differences in radiance between all three curves are much smaller than it is for those curves in Figure 8.13.

curve in Figure 8.13).

The results of Figure 8.13 also show that the radiance output from MODTRAN is a combination of the cloud's blackbody radiation, the latent heat released via the condensation of water vapour, the absorption, re-emission and reflection of Earth's radiation by the cloud, and the absorption/re-emission of Earth's radiation by constituent molecules between the ground and the cloud. This is due to the cloud radiance curve (blue) lying above what would be expected for a blackbody at the cloud height (8 km), suggesting that the extra effects are indeed contributing to the overall radiance (thus warmer temperature) of the cloud. This must be corrected for, such that the radiance calculated is an accurate value corresponding to the true cloud temperature.

This difference in radiance is less pronounced as the cloud moves lower in the atmosphere. Figure 8.14 shows the comparison between two blackbody curves, at ground temperature (red) and 281.7 K (black) [the temperature at 1 km altitude] with the radiance output from MODTRAN for a cumulus cloud at 1 km altitude (in blue).

The results of Figure 8.14 suggest that the effect of the extra sources of radiation is not as significant for low altitude cloud as it is for cloud at higher altitude. This is due to the integral number of molecules (between the ground and the cloud) being smaller, meaning that there are less molecules available for absorption and re-emission of the Earth's radiation. This also provides a clue that the effect of cloud absorbing Earth's radiation is independent of the height of the cloud in the atmosphere.

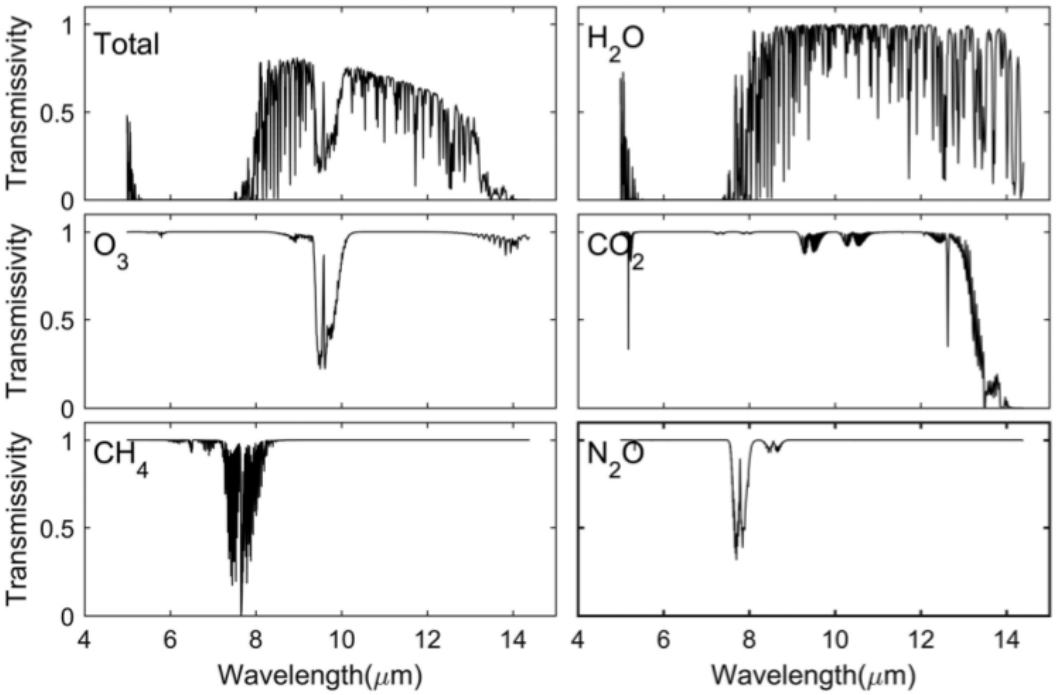

Figure 8.15: Absorption bands in the infrared band (5-14 μm) for various constituent molecules in the atmosphere. Molecules shown are water vapour, carbon dioxide, nitrous oxide, ozone and methane. Vertical transmissivities are shown for a particular set of molecular densities [100].

One can look at the effect on the radiance that is caused by the absorption and re-emission of radiation by various constituent molecules in the atmosphere that exist between the ground and the cloud. Figure 8.15 shows the transmission and absorption of various molecules in the atmosphere between 5 and 14 μm.

Figure 8.15 shows that each molecule has emission features in the atmospheric window (7-14 μm), suggesting that each one is adding extra radiance by emitting Earth's radiation at characteristic wavelengths (some examples are 6.3 μm for water vapour, 7.5 μm for methane, and 14.1 μm for carbon dioxide) [100]. The results of Figures 8.13 and 8.14 imply that this small, measurable effect on the radiance due to these molecules causes a warmer temperature to be read than what is expected for the cloud at a given height, due to the re-emission of absorbed infrared radiation from the Earth. This effect increases as the cloud moves higher in the atmosphere, simply due to there being a higher total number of trace gases present between the ground and the cloud (therefore, a higher amount of infrared radiation being emitted).

In order to understand this effect, one must integrate the radiance (the blue curve in Figure 8.14) in the atmospheric window and correct for this bias in the temperature. To do so, we define a quantity I, given by equation 8.10, which is the integral of a blackbody curve (in the atmospheric window) at a known blackbody temperature T_{BB}.

$$I = \int_{7}^{14} R\, d\lambda \tag{8.10}$$

Since the cloud temperature is warmer than expected, the difference in the value of I between the blue and black curves in Figures 8.13 and 8.14 will correspond to

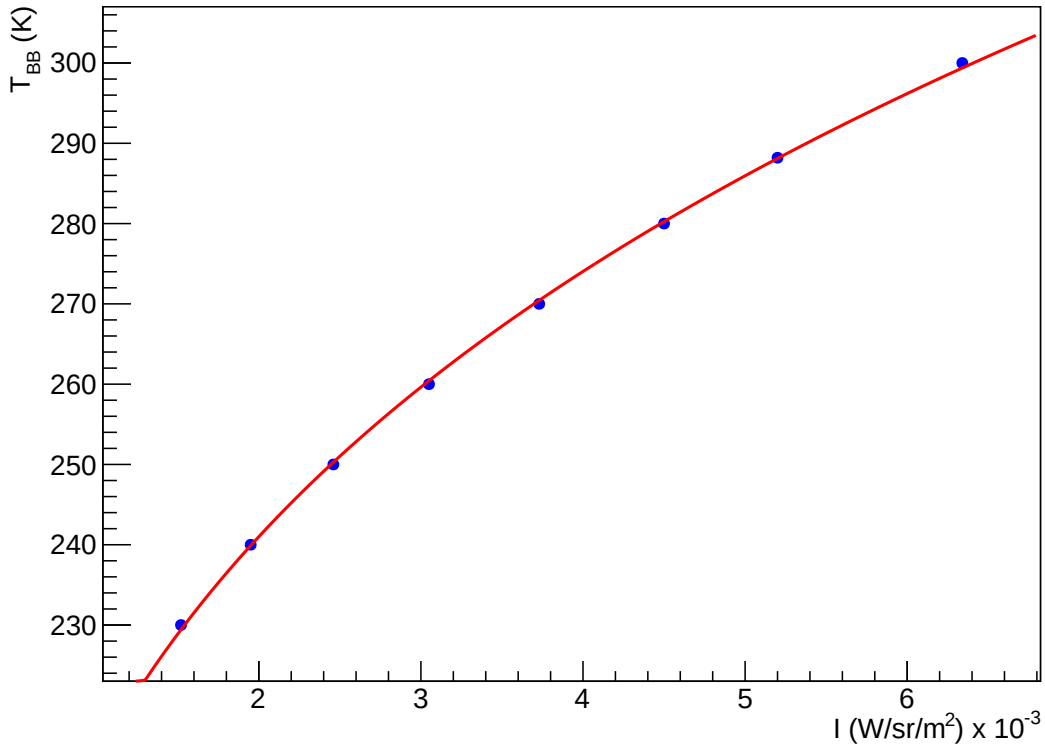

Figure 8.16: Integral of a blackbody curve (equation 8.10) in the atmospheric window as a function of its blackbody temperature T_{BB}. The atmospheric window is defined to be between 7 and 14 μm.

a difference in the blackbody temperature of the cloud compared to the atmosphere. Thus, one can measure the temperature bias caused by the presence of constituent molecules.

Figure 8.16 shows this integral quantity I of a blackbody curve in the atmospheric window against its corresponding blackbody temperature (which is known from the Planck curve). The red curve is a fit to this data, which best matches the relationship between the two parameters.

From Figure 8.16, one can fit the relationship between the integration in the atmospheric window (from equation 8.10) and the blackbody temperature T_{BB}, which is given by equation 8.11.

$$T_{BB} = A(I^{\frac{1}{4}}) + B \tag{8.11}$$

The fit to equation 8.11 gives A = 146.8 K sr m^2 W^{-1} and B = 66.4 K. Now, for any cloud radiance curve (for example, the blue curve in Figures 8.13 and 8.14), one can calculate the integral in the atmospheric window, and using equation 8.11, calculate the temperature of the cloud. One can then compare this temperature to the temperature of the atmosphere (given by the black curves in Figures 8.13 and 8.14) at the height of the cloud. This is shown in Figure 8.17 for a US standard atmosphere with a thick cumulus cloud ranging in height from 1 km to 8 km [in 1 km intervals].

The results of Figure 8.17 confirms that the temperature calculated from the radiance curve is larger than what would be expected for a cloud given its height and temperature. What this implies is that the radiance being measured here is a combina-

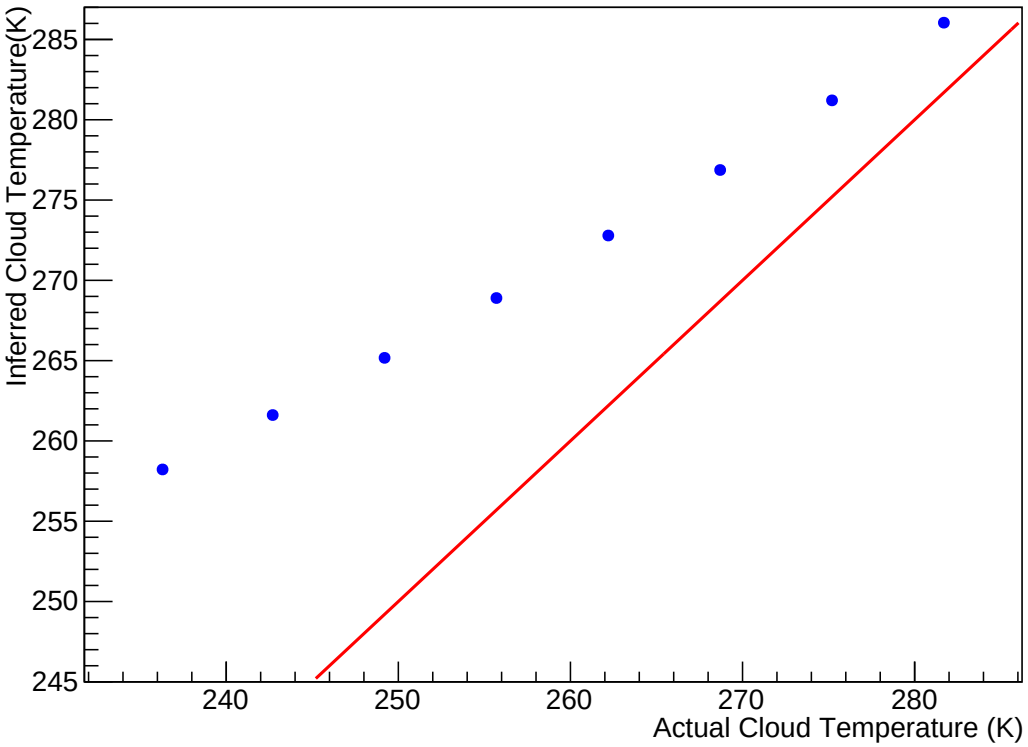

Figure 8.17: Comparison of cumulus cloud temperature (vertical axis) versus the expectation for a blackbody at the same height in the atmosphere (horizontal axis). The blue points are the cloud temperatures measured in MODTRAN by the use of equation 8.11, and the red line is the atmospheric temperature at the height of the cloud. Cloud heights shown by the blue points are 1 km - 8 km in 1 km intervals.

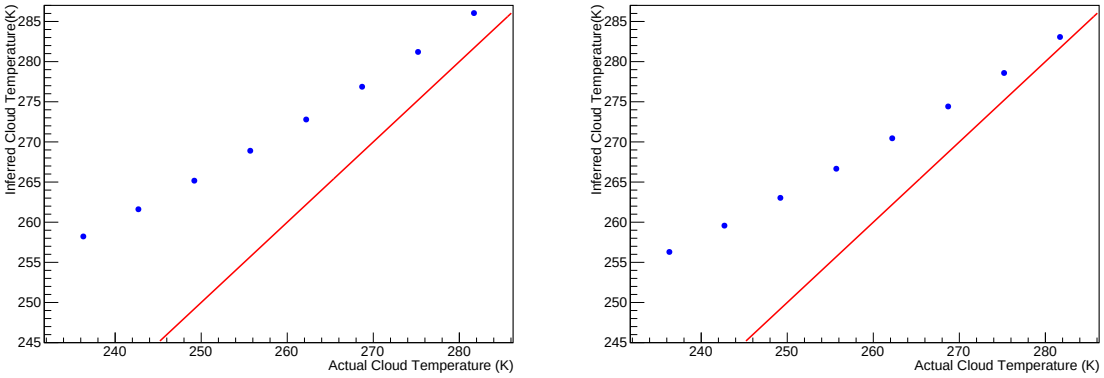

Figure 8.18: Comparison of the difference between the Cloud Temperature and Atmospheric Temperature for thick (thickness of 1 km) cumulus cloud (left) and thin (thickness of 100 m) cumulus cloud (right).

tion of the radiance due to the thick cloud in the atmosphere, as well as the radiance due to atmospheric molecules between the observer at ground level, and the cloud. There is a small effect due to the latent heat that is released within a cloud, something that is discussed in more detail in section 8.6. The biasing effect here is caused by the absorption and re-emission of Earth's radiation by atmospheric molecules which include; water vapour, carbon dioxide, ozone, methane and nitrous oxide. The effect increases in severity as cloud goes higher in the atmosphere since there are more molecules available for the absorption and re-emission of Earth's radiation, as demonstrated by Figure 8.17.

Another check that can be done, is to see the effect of the cloud thickness on this bias. The study from Figure 8.17 is repeated, adjusting the cloud thickness to 100 m (rather than 1 km). The results are shown in Figure 8.18.

Figure 8.18 shows that decreased thickness of the cloud creates a systematic shift in the bias (about 3 K cooler at all altitudes), which lowers but does not remove the temperature biasing effect. This is due to the thinner cloud being less efficient at absorbing the infrared radiation from Earth, and also reflecting a portion of radiation from the Earth. However, this means it does not lower the effect caused by the atmospheric molecules. As a result, using a fixed average thickness of 1 km will be able to best study the effect of the temperature bias.

It is also worth inputting an atmosphere in MODTRAN (taken from the GDAS gridpoint at Auger [see section 3.5.2]), to see how big of an the effect the constituent molecules have on the cloud temperature measurements with the infrared cameras. The GDAS atmosphere is defined by temperature, pressure, and relative humidity as a function of height. All other atmospheric molecules have standard values defined by the US Standard Atmosphere (see Figure 8.23). The idea of this is to control the amount of water vapour in the atmosphere and understand its effect on the bias, since absorption and re-emission of infrared radiation from the Earth by water vapour is a large factor in calculating a larger radiance, causing the temperature of the cloud to be overestimated. This has been shown in previous studies on Earth's radiation budget, with an estimated 60% of the total absorption and re-emission of terrestrial radiation being done by water vapour [71]. However, one wishes to check how much of an effect

```
M   7    3    1    0    0    0    0    0    0    0    0    1    0    .000    .00
f   8    0   385.000  1.000000  1.00000  f t f
01_2008
   -1    0    0    0    2    0       .000       .000       .000       .000     0.000
 1.000   8.000  -9.000  -9  -9  -9.000  -9.000  -9.000  -9.000  -9.000  -9.000
   31    0    0
     0.000   865.121     9.491    43.597     0.000     0.000ABH2222222222222
     0.400   825.628     8.744    44.775     0.000     0.000ABH2222222222222
     0.800   786.565     6.734    47.727     0.000     0.000ABH2222222222222
     1.200   749.130     4.251    42.550     0.000     0.000ABH2222222222222
     1.600   713.040     2.175    36.668     0.000     0.000ABH2222222222222
     2.000   678.469    -0.050    40.754     0.000     0.000ABH2222222222222
     2.400   645.381    -2.441    52.063     0.000     0.000ABH2222222222222
     2.800   613.515    -5.294    61.229     0.000     0.000ABH2222222222222
     3.200   582.948    -7.935    64.705     0.000     0.000ABH2222222222222
     3.600   553.699   -10.408    62.888     0.000     0.000ABH2222222222222
     4.000   525.516   -12.942    61.098     0.000     0.000ABH2222222222222
     4.400   498.686   -15.489    59.659     0.000     0.000ABH2222222222222
     4.800   472.780   -18.117    59.770     0.000     0.000ABH2222222222222
     5.200   448.179   -20.760    60.152     0.000     0.000ABH2222222222222
     5.600   424.375   -23.596    60.252     0.000     0.000ABH2222222222222
     6.000   401.835   -26.432    60.762     0.000     0.000ABH2222222222222
     6.400   379.976   -29.508    60.422     0.000     0.000ABH2222222222222
     6.800   359.262   -32.606    60.492     0.000     0.000ABH2222222222222
     7.200   339.324   -35.840    59.455     0.000     0.000ABH2222222222222
     7.600   320.201   -39.193    57.668     0.000     0.000ABH2222222222222
     8.000   302.157   -42.545    56.601     0.000     0.000ABH2222222222222
     8.400   284.539   -45.452    50.802     0.000     0.000ABH2222222222222
     8.800   267.871   -48.297    45.381     0.000     0.000ABH2222222222222
     9.200   252.179   -51.141    40.938     0.000     0.000ABH2222222222222
     9.600   237.086   -52.329    30.902     0.000     0.000ABH2222222222222
    10.000   222.846   -53.240    22.640     0.000     0.000ABH2222222222222
    10.400   209.462   -54.150    16.605     0.000     0.000ABH2222222222222
    10.800   196.830   -55.141    12.543     0.000     0.000ABH2222222222222
    11.200   184.820   -56.369    10.317     0.000     0.000ABH2222222222222
    11.600   173.543   -57.596     8.504     0.000     0.000ABH2222222222222
    12.000   162.954   -58.824     7.024     0.000     0.000ABH2222222222222
    00.000     0.000   000.000         0         0         0                  0
         2        50         1        2rm       mlaa
    0
```

Figure 8.19: The cards in MODTRAN that are used to create a user defined atmosphere. The columns of data from left to right are altitude (km), pressure (hPa), temperature (°C) and relative humidity (%).

other molecules have on this temperature bias as well, to see how their effect compares to water vapour.

Figure 8.19 shows the modified MODTRAN tape file for the user defined atmosphere, which adds a new card (card 2C) to define height, pressure, temperature, and relative humidity from a GDAS profile.

Figure 8.20 shows the difference in temperature between the atmosphere and the cloud for an example autumn GDAS profile (May 2016), with a ground temperature of 282.6 K and PWV (Precipitable Water Vapour) content of 12.5 mm. PWV is defined as the amount of water vapour above a unit area of the atmosphere if it all condensed and was placed into a unit area box (more on this PWV is discussed later).

Figure 8.20 shows that an atmosphere at Auger (output from GDAS) has the same temperature bias as was seen in previous figures. The rest of this chapter studies this bias, the cause of it, and how it can be corrected for in our cloud height estimation analysis.

An interesting change that can be made to the user defined atmospheres from GDAS (such as the one seen in Figure 8.20) is to remove all of the water vapour in

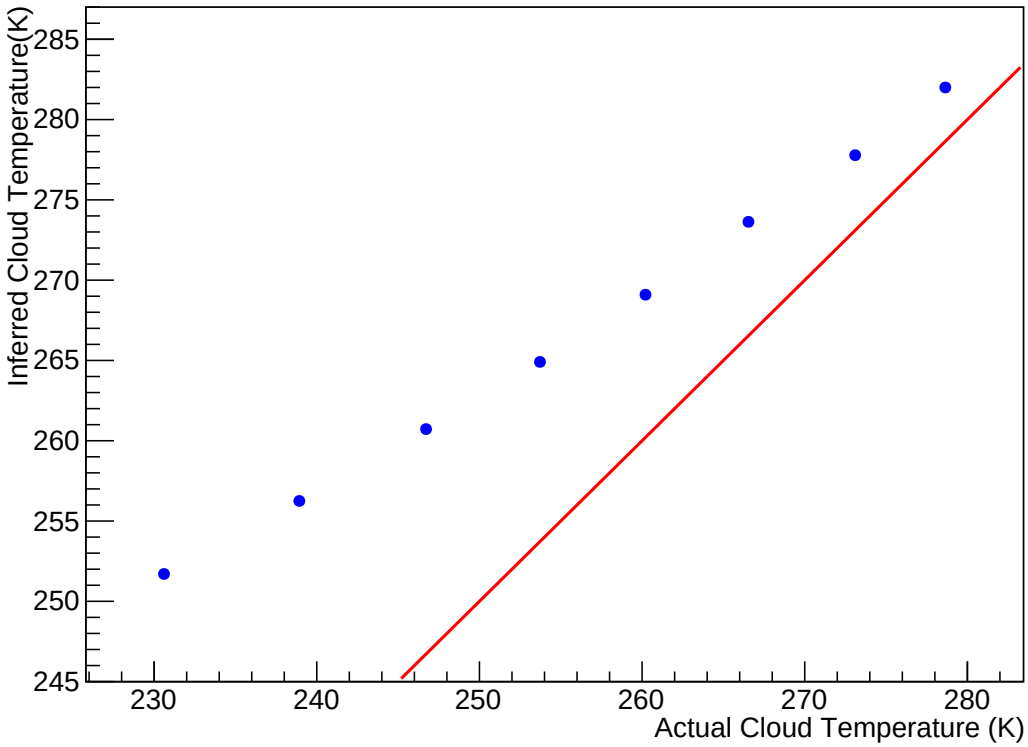

Figure 8.20: Comparison of sky temperature to atmospheric temperature output from MODTRAN from the GDAS gridpoint located at Auger.

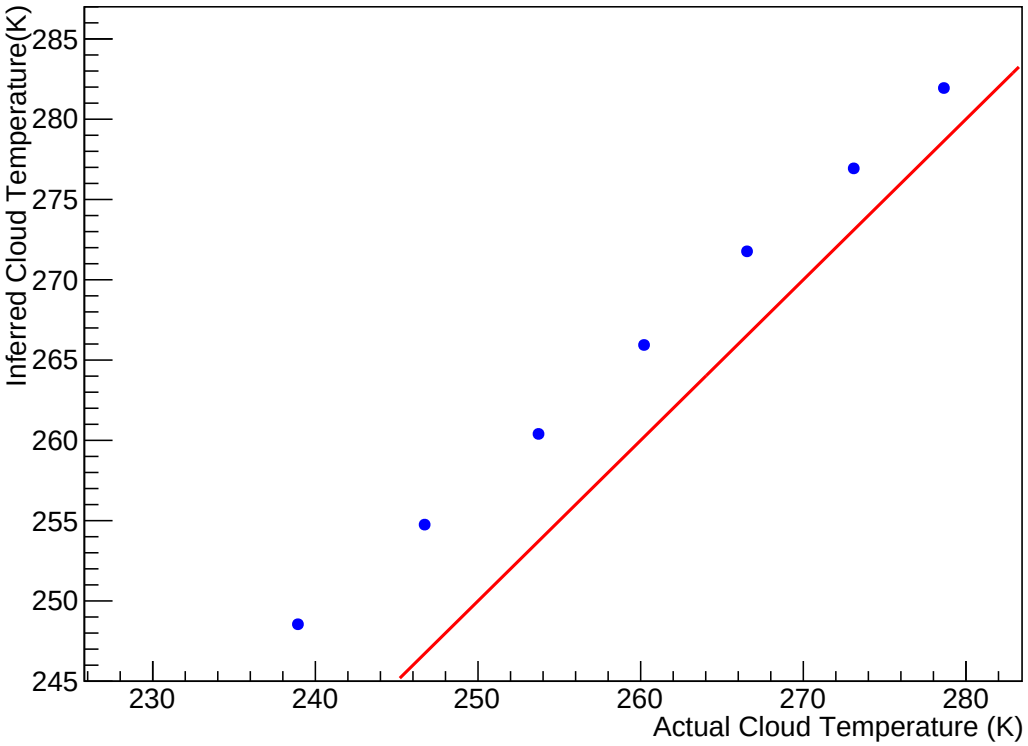

Figure 8.21: Comparison of Cloud Temperature to Atmospheric Temperature output from MODTRAN in an atmosphere with zero water vapour.

the atmosphere and check the effect that this has on the bias. One expects that since the water vapour creates the majority of the bias (water vapour is one of the largest absorbers of infrared radiation in the atmospheric window), that the temperatures of the cloud and the atmosphere should more closely match [71]. Figure 8.21 illustrates the effect with zero water vapour content in the atmosphere.

The results of Figure 8.21 show that the bias is still present. This is due to other gases present in the atmosphere, which also absorb and re-emit infrared radiation from the Earth. All of these gases (as described previously) have absorption/emission bands in the atmospheric window [99]. This shows that whilst water vapour in the atmospheric window is a factor in the temperature bias, the other gases such as carbon dioxide, nitrous oxide and methane also affect the temperature bias (a factor which will be considered further later on in this chapter).

Figure 8.22 shows the radiance curve for a clear atmosphere with no water vapour. The expectation, for no temperature bias (or cloud), is that the radiance curve will be flat in the atmospheric window, assuming that water vapour is the only absorber within this wavelength range (7-14 μm).

Clearly, the radiance curve in Figure 8.22 is not flat, and several characteristic peaks emerge. Figure 8.22 shows emission in the atmosphere by various gases. The peaks in the radiance curve (in blue) are 7-8 μm for CH_4 and N_2O, 9.6 μm and 14.1 μm for O_3, and 12.7 μm for CO_2. Due to this emission of infrared radiation, there will always be a bias in the sky temperature estimates, causing a systematic uncertainty on the cloud height (from section 8.1). However, this bias in the sky temperature is small for

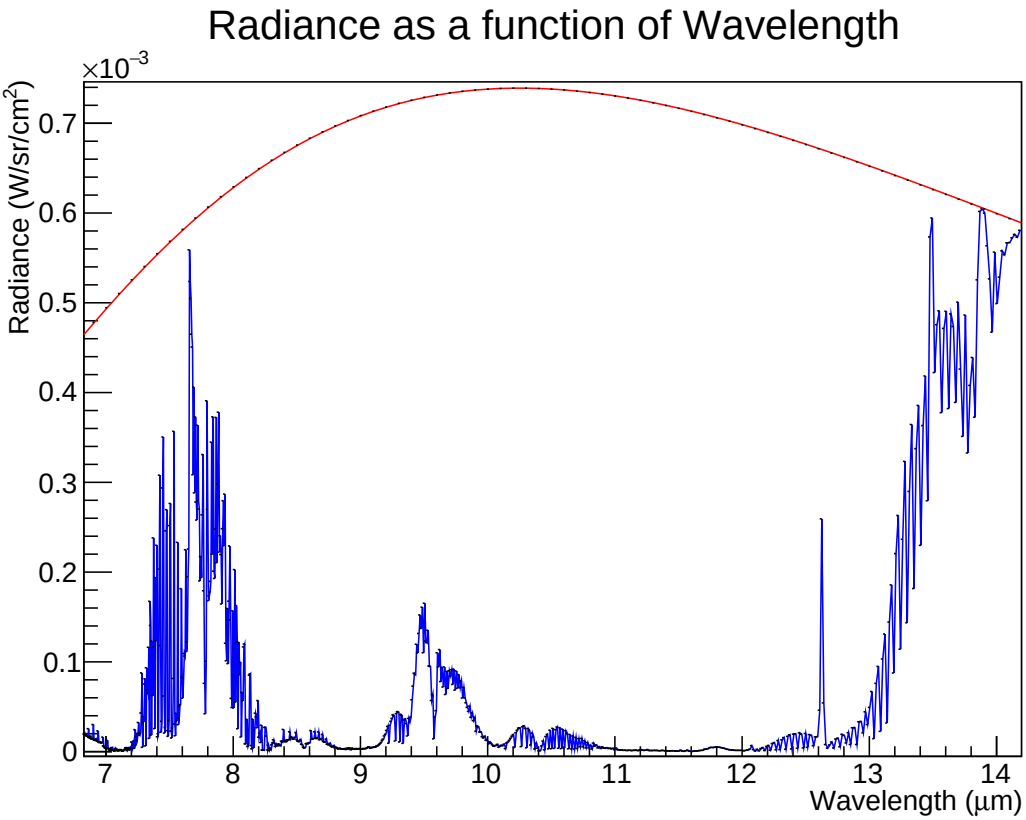

Figure 8.22: Radiance output for a ground temperature blackbody (red) and a clear atmosphere absent of water vapour (blue). The peaks in the atmospheric window correspond to various remaining gases in the atmosphere.

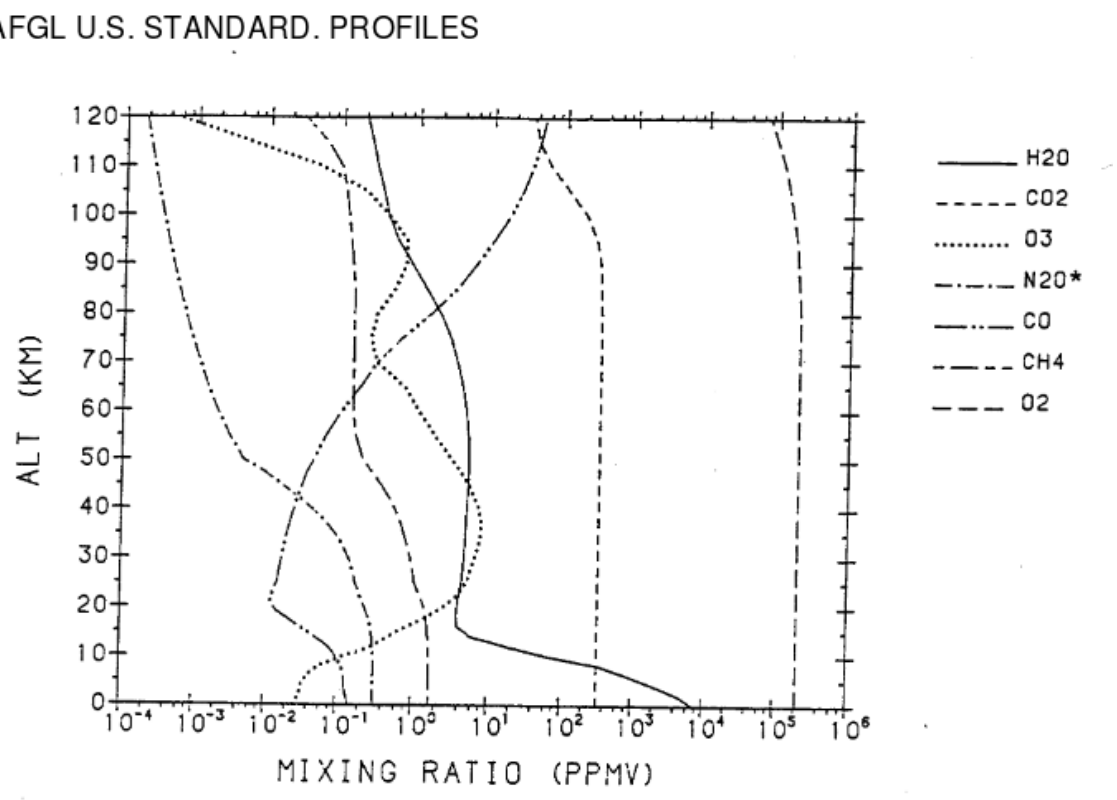

Figure 8.23: Mixing ratio (in ppmv) for each constituent molecule in the US Standard Atmosphere. Source: [101]

heights below 4 km. This height distinction is important since thick clouds generally don't develop above this altitude, where cirrus cloud is more prevalent (and difficult to image with the infrared cloud cameras as discussed in Chapter 7).

Figure 8.23 shows the mixing ratios for each of these constituent molecules in the atmosphere, which are the default values for the US standard atmosphere in MODTRAN [101]. For all user defined atmospheres in MODTRAN, one is allowed to specify the concentrations of any one particular molecule, however, if a mixing ratio is not defined, the concentrations revert back to the US standard atmosphere values, as depicted by Figure 8.23. This is what is used for the analyses in this chapter.

The integral number of each molecule in the atmosphere is steadily increasing with height for the region of interest (below 10 km), as shown in Figure 8.23. This suggests that the temperature bias (due to absorption and re-emission of the Earth's blackbody radiation by these atmospheric gases) does grow, slowly, as a function of altitude. The peak at 9.6 μm due to ozone has no significant effect on the bias since ozone predominantly exists above 10 km. The portion of emission due to ozone (from Figure 8.21) is estimated to be around 15% of the total emission (of all atmospheric gases not including water vapour) from Figure 8.22, which means the bias is actually about 15% less significant than originally estimated.

There are two other molecules present in Figure 8.23 that have not been discussed previously. These are carbon monoxide (CO) and diatomic oxygen (O_2). However, since CO has an absorption band at 4.7 μm, which is outside of the atmospheric window, it has no effect on the infrared absorption in our analysis here [102]. Similarly for the O_2

molecule, an absorption band exists between 15 and 16.5 μm, once again outside of the atmospheric window. Thus, this molecule also has no effect on the results.

Before discussing those corrections, one can discuss the effect of aerosols on the temperature measurements.

8.4.1 The effect of Aerosols

As well as the constituent molecules of the atmosphere, it is worth checking if aerosols present in the atmosphere will effect cloud temperature estimates. As discussed in section 3.5.3, the Central and Xtreme Laser Facilities at the centre of the Pierre Auger Observatory monitor aerosols present in the atmosphere. The typical measure of aerosol content for these instruments is VAOD (Vertical Aerosol Optical Depth), which for small optical depths is the proportion of incident radiation that is scattered or reflected by aerosol particles in the atmosphere. Typical values for the VAOD range from 0.01-0.1. A value of 0.1 is known as a "dirty" atmosphere and nights with aerosol concentration this high (and higher) are cut out in the air shower reconstruction as too much of the emitted fluorescence light is attenuated by aerosols (see section 3.1).

As mentioned in section 4.2, aerosols in the atmosphere scatter incoming solar radiation back into space, contributing a small net cooling effect on the atmosphere. It is worth checking how much of an effect aerosols have on our cloud temperature estimations by looking at the difference caused by aerosols present in the atmosphere. The results from a standard atmosphere have already been shown (see Figure 8.20). Next, one uses this same atmosphere, but also adding aerosols, to see how much more the temperature differs to expectation. This is done using a desert model atmosphere provided by MODTRAN, where the set wind speed determines the aerosol concentration. The desert model is chosen as it most closely matches the climate at the observatory. It is well established that wind speed correlates with the concentration of aerosols in the atmosphere for desert environments similar to the Auger site. To get the aerosol concentration into a comparable amount, one can calculate the vertical aerosol optical depth by computing the vertical transmission of the full atmosphere at a given reference wavelength. Since the CLF laser wavelength is 355 nm (see section 3.5.3), this is the wavelength that is used to compute the vertical transmission and hence VAOD for the simulated atmosphere in MODTRAN. For this calculation, define T_1 as the vertical transmission factor for a purely molecular atmosphere (with no aerosols present) and T_2 as the vertical transmission factor for the same atmosphere but with aerosols present. The chosen atmosphere utilises the desert model with a wind speed of 3 m/s (this is a typical average wind speed at the observatory in Argentina). For a purely molecular atmosphere the vertical transmission is given by;

$$T_{355nm} = 0.549 = T_1 \tag{8.12}$$

Next, for the aerosol atmosphere the vertical transmission becomes;

$$T_{355nm} = 0.456 = T_2 = T_1 \times T_a \tag{8.13}$$

In equation 8.13, T_a is the vertical transmission factor purely due to the aerosols present. Rearranging this equation one finds;

$$T_a = \frac{T_2}{T_1} = \frac{0.456}{0.549} = 0.83 \tag{8.14}$$

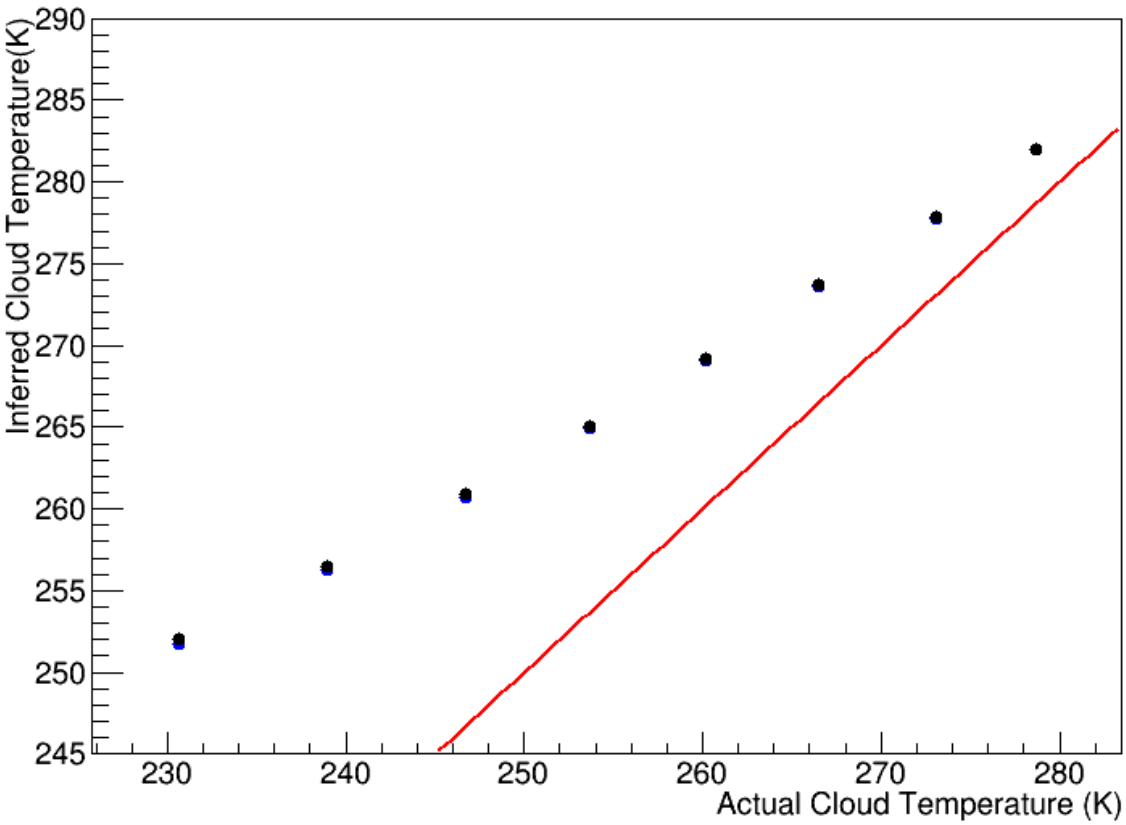

Figure 8.24: Temperature calculated by MODTRAN for the same atmosphere but with aerosols added in (black) compared to a purely molecular atmosphere (blue). High aerosol concentration appears to have no significant effect on the calculated radiance temperature of the cloud.

From this, one can calculate the Vertical Aerosol Optical Depth at the reference wavelength (labelled τ_a) as;

$$\tau_a = -ln(T_a) = -ln(0.83) = 0.185 \tag{8.15}$$

As can be seen by equation 8.15, the calculated VAOD is a much larger value than is typically measured by the atmospheric instruments at Auger. As a result, this is an appropriate test of the aerosol effect, since a high concentration might yield some discernible uncertainty to the cloud temperature estimations.

Figure 8.24 is the result of calculating the cloud temperature for a purely molecular atmosphere with no aerosols (blue dots) compared with the same atmosphere with aerosols added (at the VAOD calculated by equation 8.15).

The results of Figure 8.24 show that aerosols present in the atmosphere have no noticeable effect on the cloud temperature calculated from the radiance curve. The largest discrepancy in temperature between the molecular and aerosol atmospheres was 0.3 K, an effect that can be ignored for the purpose of cloud height estimation.

Since the effects of aerosols in the atmosphere does not seem to affect the temperature systematic, then the systematic must be largely caused by the trace gases in the atmosphere, and the cloud itself. Section 8.5 will discuss the corrections made for these two effects.

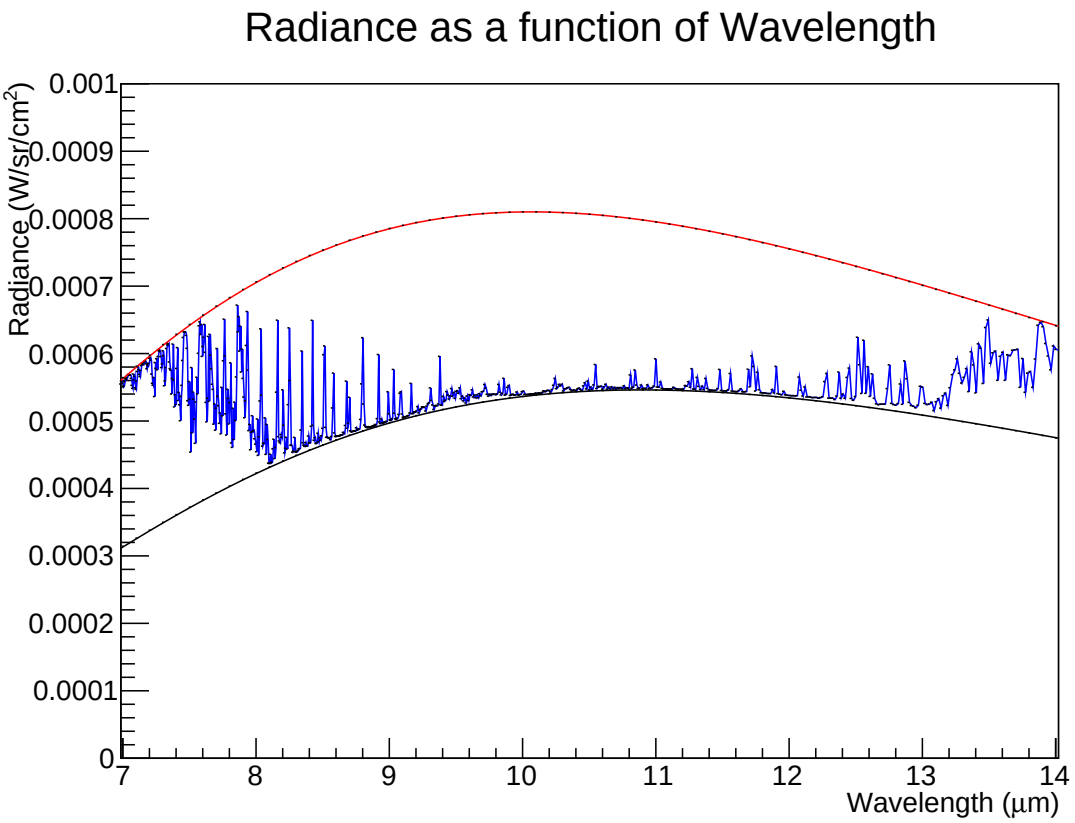

Figure 8.25: Radiance curve (blue) for a thick cloud at a height of 4 km in the atmosphere, viewed by a ground observer. The black curve is the expected result for a blackbody at that height, whilst the red curve corresponds to a blackbody at ground temperature. The resulting radiance is larger indicating that the cloud is warmer than a blackbody at the temperature of the cloud.

8.5 Using Radiance Curves to study the overestimation of Sky Temperature

As shown by Riordan et. al. [2005] (see Figure 8.2), the cloud temperature (in the infrared) was overestimated in the atmosphere [97]. The goal is to understand why this happens. This can be achieved using MODTRAN, by simulating atmospheres with thick cloud present, to see if this overestimation is seen.

Before showing the simulation results, it's important to refresh the blackbody properties of cloud. As shown in section 8.4, the radiance of a blackbody object is proportional to the temperature of the object to the fourth power. This is the key point in interpreting the results from MODTRAN, turning spectral radiance into a meaningful quantity (temperature) that can be compared to the expectation for a given atmosphere.

The radiance curve of a thick cloud at 4 km in the atmosphere is shown in Figure 8.25. The important point is to compare this radiance curve (shown in blue) to the radiance curve of a blackbody at the temperature expected for that height in the atmosphere (in this case 4 km).

The results of Figure 8.25 show that the temperature of the cloud is higher than

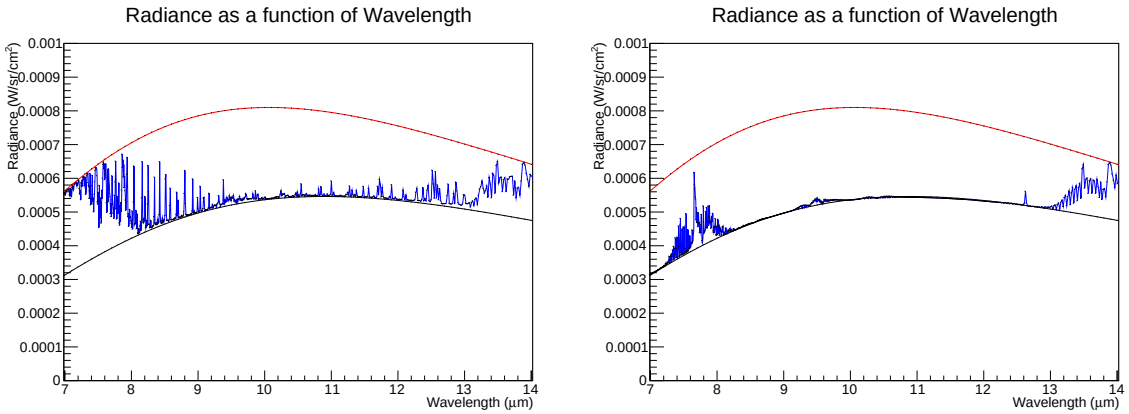

Figure 8.26: Radiance curves (in blue) for an atmosphere with (left) and without (right) water vapour.

expected from the atmosphere. There are two effects to consider that can increase the known sky temperature.

Thick cloud in the atmosphere emits blackbody radiation over a continuous wavelength range which is dependent only on its temperature. As discussed in the previous section, the Earth similarly radiates as a blackbody with a temperature of $\sim 288\,\mathrm{K}$. This radiation coming from the Earth can be absorbed and/or reflected by the cloud, which heats it up, causing the cloud to be warmer than expected. This radiation coming from Earth can also be absorbed by molecules in the atmosphere, and re-emitted at characteristic wavelengths. These two sources of extra radiation that come from the Earth are detected by our infrared cloud cameras, causing the systematic in cloud temperature that causes it to be overestimated. Water vapour is one of the largest absorbers of infrared radiation in the atmospheric window, thus knowledge of the amount of water vapour present in the atmosphere, and its effect on the temperature bias, must be better understood (as the bias effect caused by water vapour varies, dependent upon the total amount of water vapour present) [71]. The latent heat released within the cloud during the condensation process also raises the temperature of the cloud, another effect that must be accounted for [67].

The effect of water vapour emission (compared to all other atmospheric gases) is clear when looking at the radiance curves with and without water vapour present. This is illustrated in Figure 8.26.

The results of Figure 8.26 show that the radiance increases substantially due to the presence of water vapour. The largest effect of water vapour occurs between 7 and 9 μm, where the radiance is saturated by absorption of radiation close to the ground [51]. This leads to the largest overestimate in radiance being in this wavelength range, as illustrated by Figure 8.27.

The other effect mentioned previously (due to the cloud absorbing and re-emitting infrared radiation from the Earth) is also present when looking in a sub-region of the atmospheric window [51]. Figure 8.28 shows the radiation (between 9 and 13 μm) of an atmosphere with a cloud present.

Figure 8.28 shows that the radiation between 9 and 13 μm approximates a blackbody radiator, at a slightly warmer temperature than expected. This is due to the absorption of infrared radiation from the Earth (by the cloud) which causes it to be warmer than

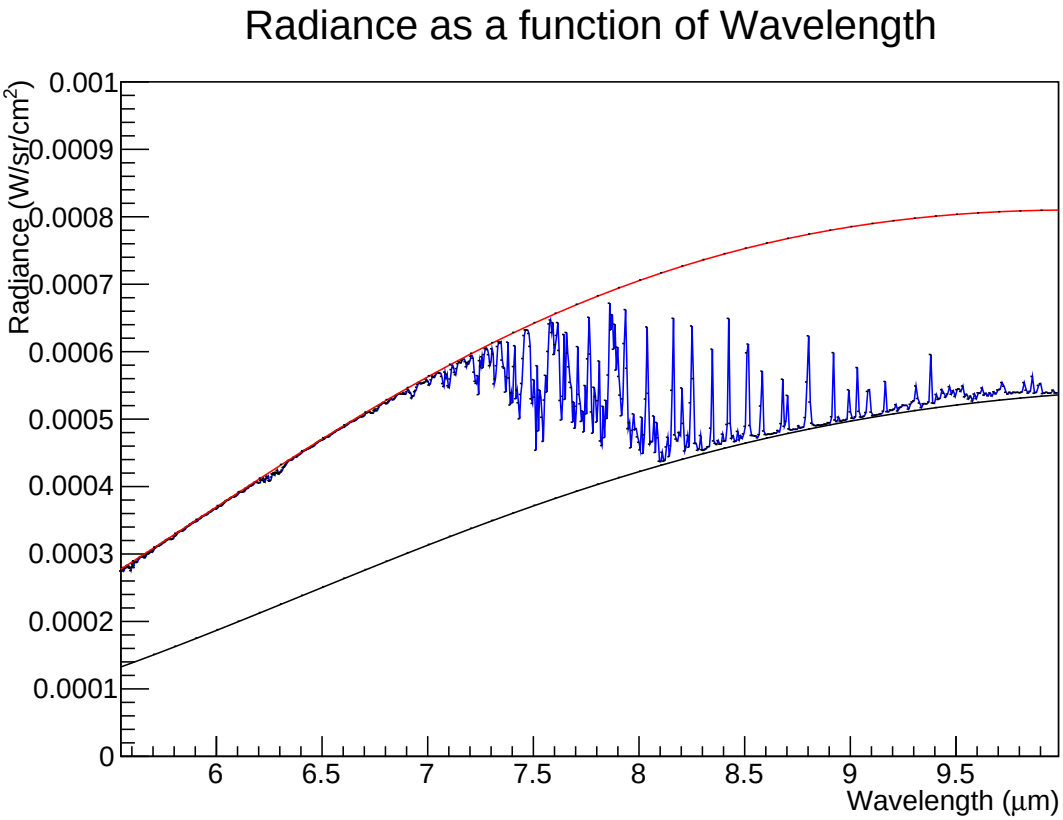

Figure 8.27: The same radiance curve as in Figure 8.25 between 7-9 μm. Saturation occurs due to the presence of water vapour in the atmosphere, meaning that water vapour is radiating close to the ground temperature (red curve).

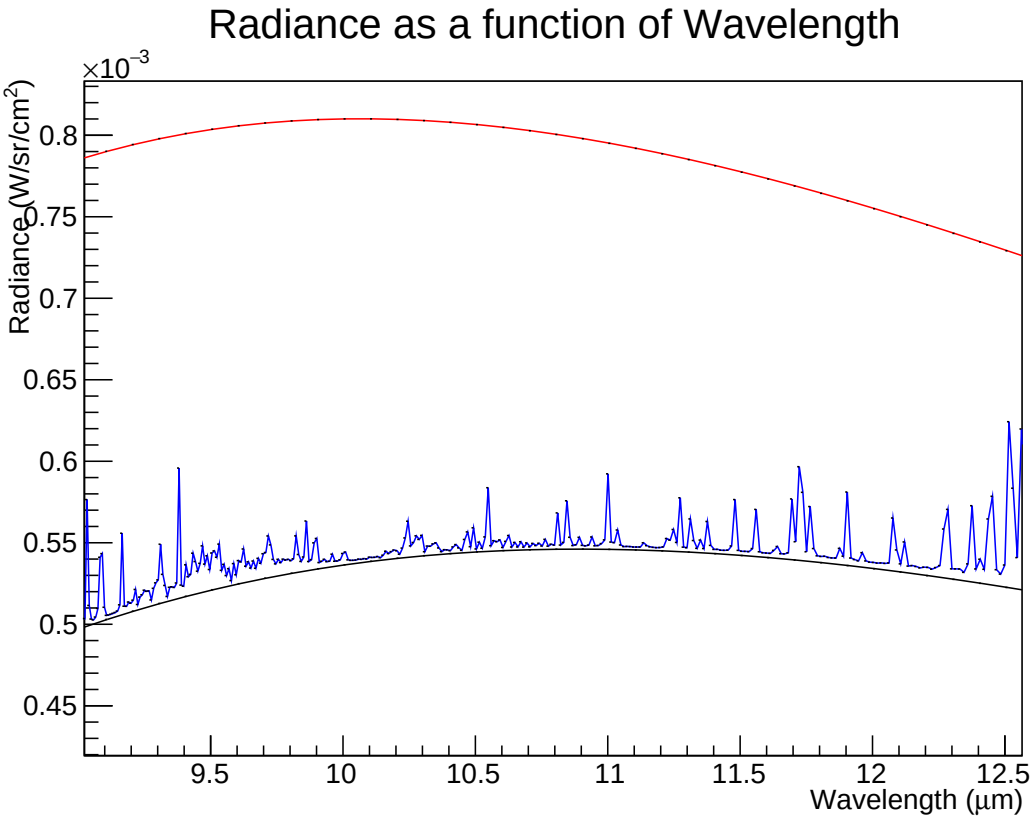

Figure 8.28: Same radiance curve as in Figure 8.25 in the wavelength range 9-13 μm. The radiance is slightly higher than expectation (black curve) in this range.

what is expected (the black curve) given its altitude [51]. The latent heat released by the cloud also contributes to this overestimate as well. This effect cannot be separated from the other component in the bias due to the cloud (absorption, re-emission and reflection of terrestrial radiation). As a result, these two cloud effects will be treated as a single contribution to the bias.

These two effects (latent heat/absorption by the cloud, and atmospheric gases present between the ground and the cloud) combine together to create a cloud temperature that is warmer than expected for a blackbody at a given height. As a result, they need to be taken into account when estimating a cloud temperature from the cloud camera. The effect caused by water vapour will be separated from the the effect caused by the presence of all other atmospheric gases (which will be combined with the latent heat/absorption effects of the cloud). As discussed in Chapter 5, the cloud cameras at Auger are calibrated to a blackbody object at a known temperature [48]. Thus, the signal received by the cloud camera is approximated as an equivalent blackbody temperature. As a result, the blue radiance curves in Figures 8.26 and 8.28 need to be transformed into a blackbody curve with the same integrated radiance (thus the same temperature), in order to best simulate what would be the cloud camera response.

In order to transform the blue radiance curve (from Figures 8.26 and 8.28) into an equivalent blackbody curve, the total radiance I needs to be calculated [in the atmospheric window] for the blue curve via equation 8.16

$$I = \int_{7}^{14} R \, d\lambda \tag{8.16}$$

where R is the radiance of the blue curve at a given wavelength, and I is taken between 7 and 14 μm.

One can similarly calculate the total radiance of the black curve (which has an already known temperature), which will be smaller. Thus, one must adjust the temperature of the black curve until its total radiance I matches the total radiance of the blue curve. One then has a new blackbody curve (shown in green in Figure 8.29) which now has a known temperature and a total radiance I that matches the total radiance of the blue curve (also shown in 8.29).

Given that both the green and black curves in Figure 8.29 have known temperatures, one then knows the exact value of the overestimation of cloud temperature, given as the difference in temperature between these two curves in Figure 8.29. One can label this difference as ΔT, which is made up of two components as shown by equation 8.17

$$\Delta T = \Delta T_{cloud} + \Delta T_{vapour} \tag{8.17}$$

where ΔT_{cloud} is the effect of various sources (absorption of Earth's radiation by the cloud, latent heat released by the cloud and the absorption/re-emission of Earth's radiation by other atmospheric trace gases), and ΔT_{vapour} is the effect of water vapour re-emitting Earth's radiation between the ground and the cloud.

It is important to understand how ΔT changes as a function of both cloud height and amount of water vapour present. The goal is to hopefully be able to split ΔT into its two components and be able to correct for each of them, independent of the other. Figure 8.30 shows how ΔT changes as a function of cloud height, with each series of points corresponding to a different amount of water vapour (labelled as precipitable water vapour [or PWV]).

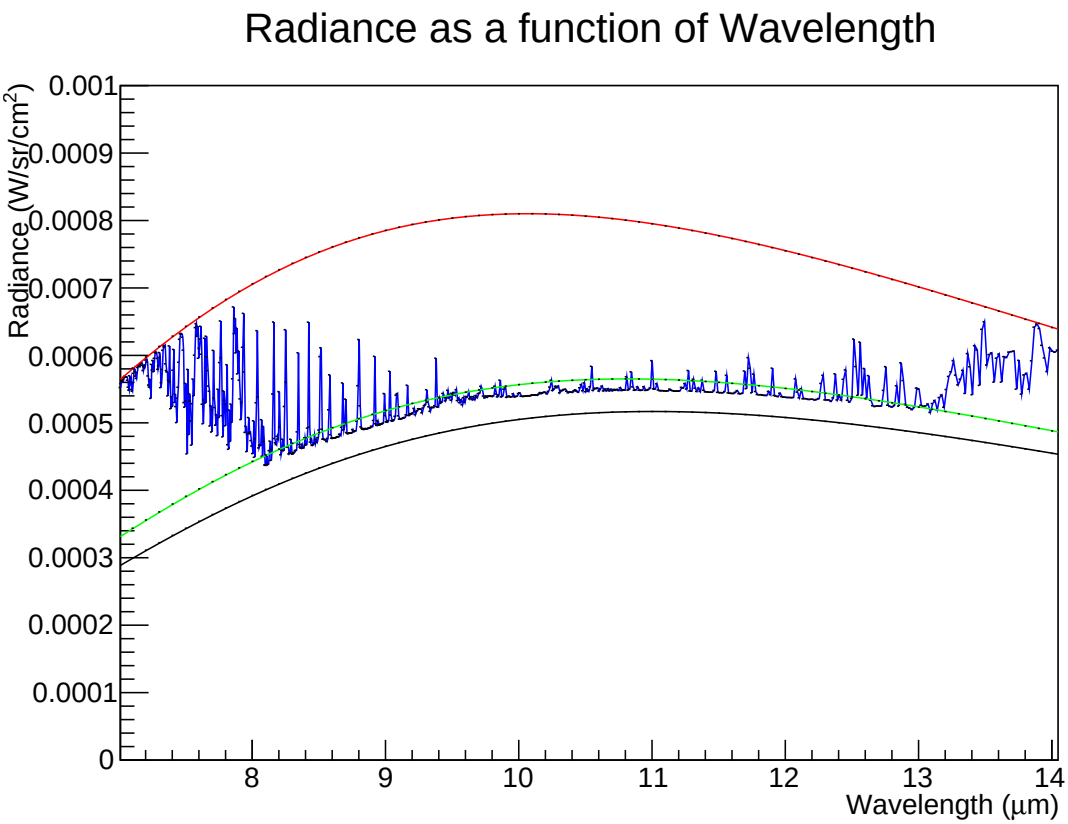

Figure 8.29: The radiance curve in green has the same radiance as the blue radiance curve (the radiative signature of the atmosphere with a cloud at 4 km) when integrating in the atmospheric window. This information can be used to calculate the systematic uncertainty on sky temperature.

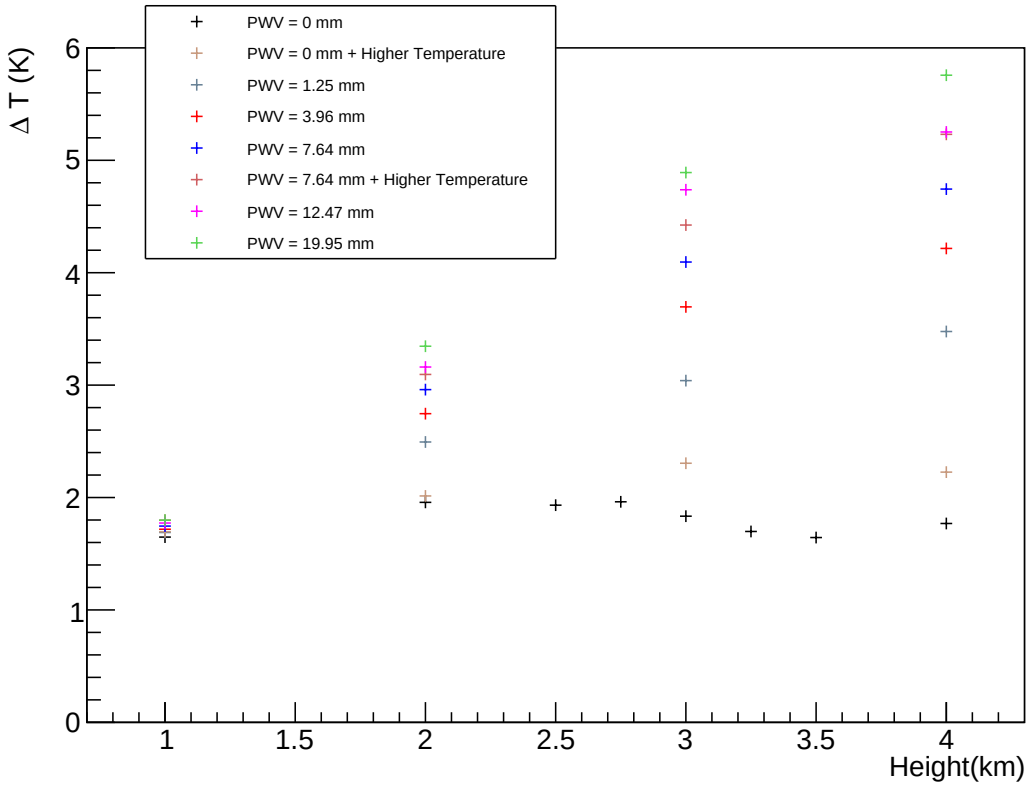

Figure 8.30: ΔT [see text] vs cloud height for various amounts of water vapour and for two different ground temperatures as indicated. Cloud used was a cumulus cloud and the two ground temperatures used were 287.8 and 297.8 K.

The results of Figure 8.30 shows that the ΔT correction depends upon the PWV. It was shown earlier that the lapse rate is important for the sky temperature estimation due to the absorbing properties of water vapour. The PWV (typically measured in mm) is how much liquid water one would have on the ground if all the water vapour in a unit area column condensed into liquid and was placed into a unit area box. This is the best measure of the total amount of water vapour in the atmosphere. To measure it, one must first calculate the density of water vapour at a given temperature and pressure, given by equation 8.18 [103].

$$\rho_w = \frac{0.0022 p_w}{T} \tag{8.18}$$

In equation 8.18, ρ_w is the density of water vapour (in kg/m^3), p_w is the partial pressure due to water vapour (in pascals) and T is the atmospheric temperature (in K). The partial pressure $p_w(z)$ is taken from GDAS for a given height and temperature. Equation 8.18 is derived from the ideal gas law, as shown by equation 8.19 for pressure p, density ρ, specific gas constant $R_{specific}$ and temperature T.

$$p = \rho \, R_{specific} \, T \tag{8.19}$$

The specific gas constant is given by the ratio of the molar gas constant (8.314 J/mol/K) to the molecular weight of the gas M_w [0.018 kg/mol for water vapour]. Assuming water vapour is an ideal gas and rearranging for density ρ_w yields equation 8.20

$$\rho_w = \frac{M_w}{R}\left(\frac{p_w}{T}\right) = \frac{0.018}{8.314}\left(\frac{p_w}{T}\right) = \frac{0.0022 p_w}{T} \tag{8.20}$$

as shown previously.

PWV is then calculated by equation 8.21 up to a specified height range [84]. For the purposes of this study, PWV is only calculated up to 4 km above ground level, as the atmosphere is too cold above this height to retain water vapour.

$$PWV = \sum_i (z_{i+1} - z_i)\rho_i \tag{8.21}$$

In equation 8.21, the z_i refer to a particular height in the atmosphere (e.g. $z_i = 0$ for ground level).

The units of the PWV are [kg/m^2], which is not the mm units required. In order to get the PWV into length units, the PWV is divided by the density of liquid water (997 kg/m^3), to then convert to mm. This value is what is used to measure the amount of water vapour in the atmosphere, in order to understand the systematic effect on the temperature. There are a number of other parameters that depend upon the amount of water vapour in the atmosphere, most notably the fit parameters of the clear sky from Chapter 5. This is shown in section 8.6.

8.6 Cloud Camera Fit parameter Correlations

It is worth testing to see if the clear sky fit parameters A and B (as shown in section 5.4) also depend on the amount of water vapour present in the atmosphere, since these parameters are found after determining temperature T as a function of zenith angle θ in the cloud camera field of view. If one can determine the correlation between the fit parameters and the water vapour, then the overestimation of cloud temperature will be better understood. As a reminder, from equation 5.10, the A parameter is the vertical sky temperature from the cloud camera (for zenith angle $\theta = 0°$) and the B parameter controls how quickly sky temperature increases as a function of zenith angle.

The correlation between these parameters and water vapour is found by plotting A and B against the PWV (Precipitable Water Vapour - in units of g/cm^2).

One can plot the PWV value as a function of the fit parameters from the cloud cameras. Starting off with the A parameter, the result of which is shown in Figure 8.31.

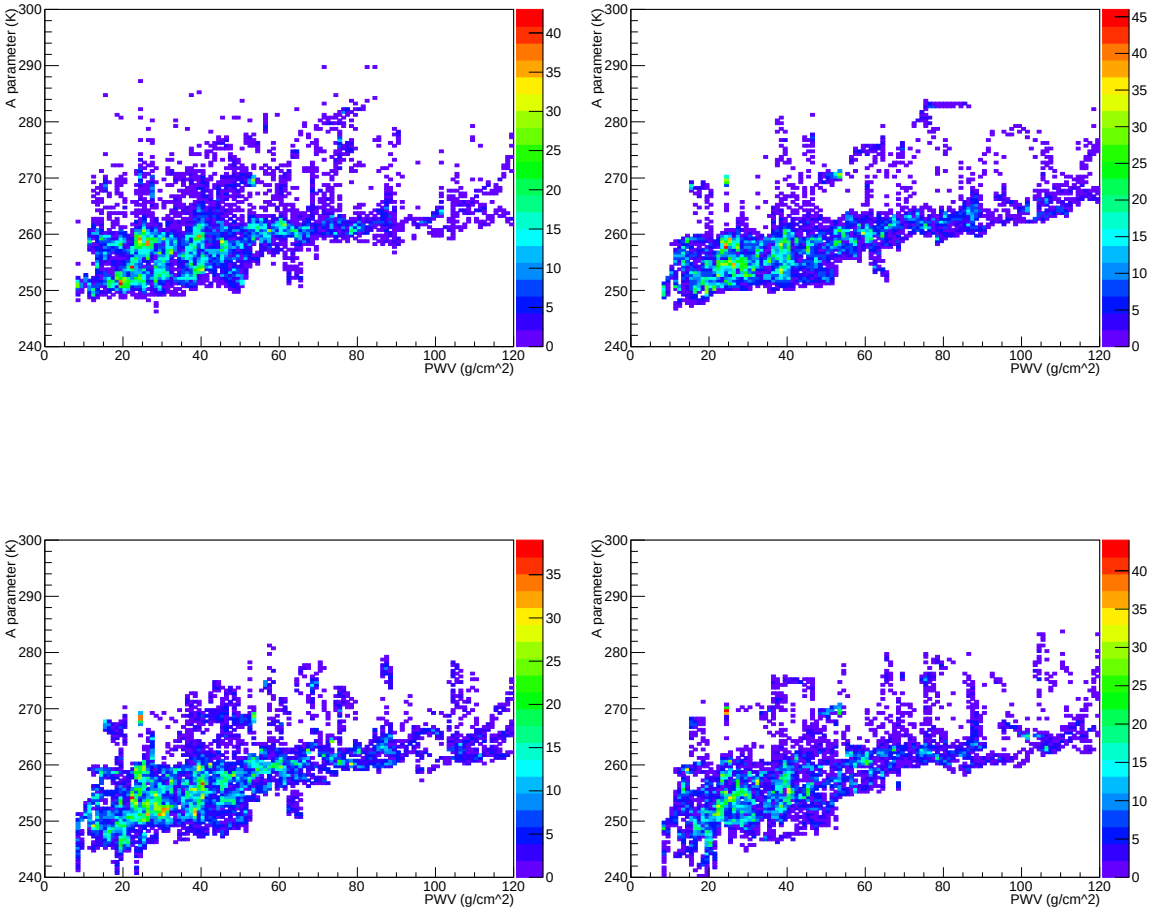

Figure 8.31: A vs. PWV for each cloud camera in 2018, for all cloud camera scans (clear and cloudy). Top left is Coihueco, top right is Loma Amarilla, bottom left is Los Leones and bottom right is Los Morados.

Additionally, the correlation of the B parameter with the PWV is shown in Figure 8.32.

Figures 8.31 and 8.32 show that there is a level of correlation between the PWV and the two fit parameters. Two lines of correlation are present in the B parameter, due to the distinction between clear sky and cloudy sky. As discussed in Chapter 5, the $\ln \sec \theta$ curve is flat for overcast skies, thus the slope (B) is ~ 0. In the case of overcast skies, no correlation between B and water vapour is expected. However, for clear and partially cloudy skies, the correlation between B and water vapour is evident. The A parameter also depends upon the amount of water vapour present in the atmosphere, as shown in Figure 8.31, thus the effect of water vapour on cloud temperature is an important factor in the estimation of cloud heights.

Another interesting comparison to make is between the fit parameters from the cloud camera and the ground temperature given by the weather stations present at each site. Figures 8.33 and 8.34 show the relationship between A and B respectively with the ground temperature.

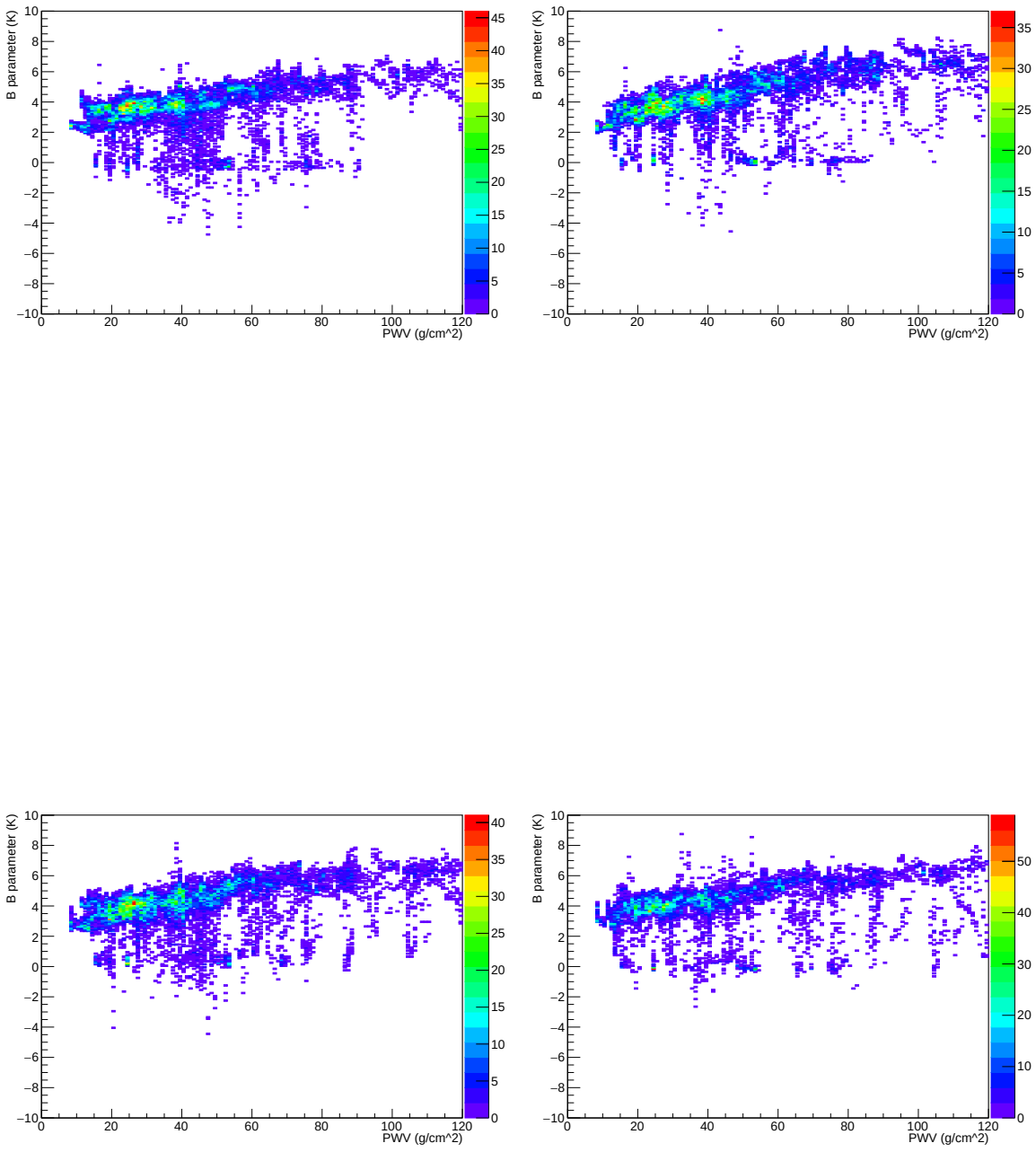

Figure 8.32: B vs. PWV for each cloud camera in 2018 for all cloud camera scans (clear and cloudy). Top left is Coihueco, top right is Loma Amarilla, bottom left is Los Leones and bottom right is Los Morados.

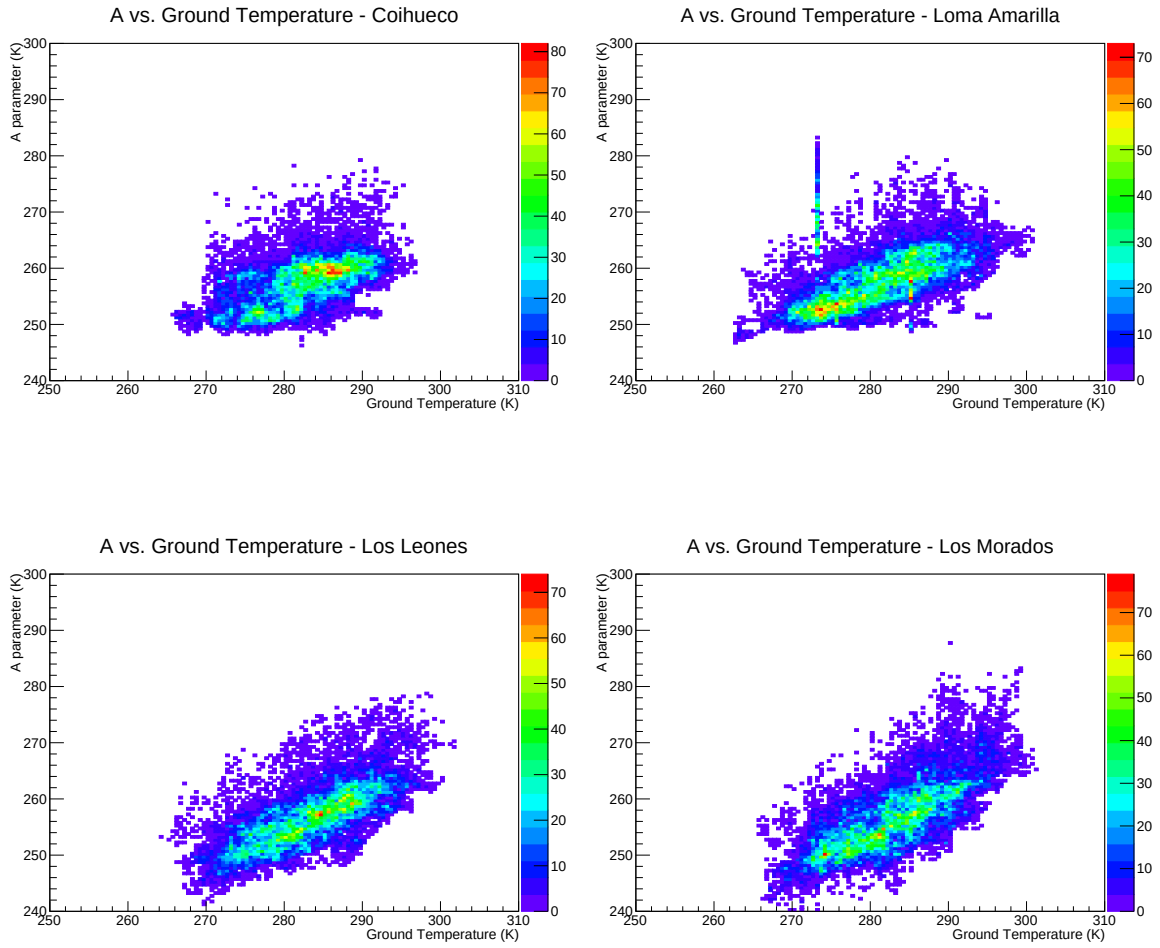

Figure 8.33: A parameter vs. Ground Temperature for each cloud camera in 2017/2018 for both clear and cloudy skies. Top left is Coihueco, top right is Loma Amarilla, bottom left is Los Leones and bottom right is Los Morados.

The A parameter is the zenith temperature, which depends upon the ground temperature (as shown by Figure 8.33), which is expected since the radiation from Earth (at the ground temperature) is responsible for the amount of absorption of radiation by the atmosphere (hence its temperature).

It is not as obvious when looking at Figure 8.34 if there is a strong correlation between the B parameter and the ground temperature, as the B parameter is more highly variable between clear and totally overcast conditions (which are not differentiated here).

The correlations of A and B with these atmospheric parameters is the reason why it is important to understand the overestimation of sky temperature ΔT, and why it depends on the amount of water vapour present in the atmosphere, as well as the ground temperature.

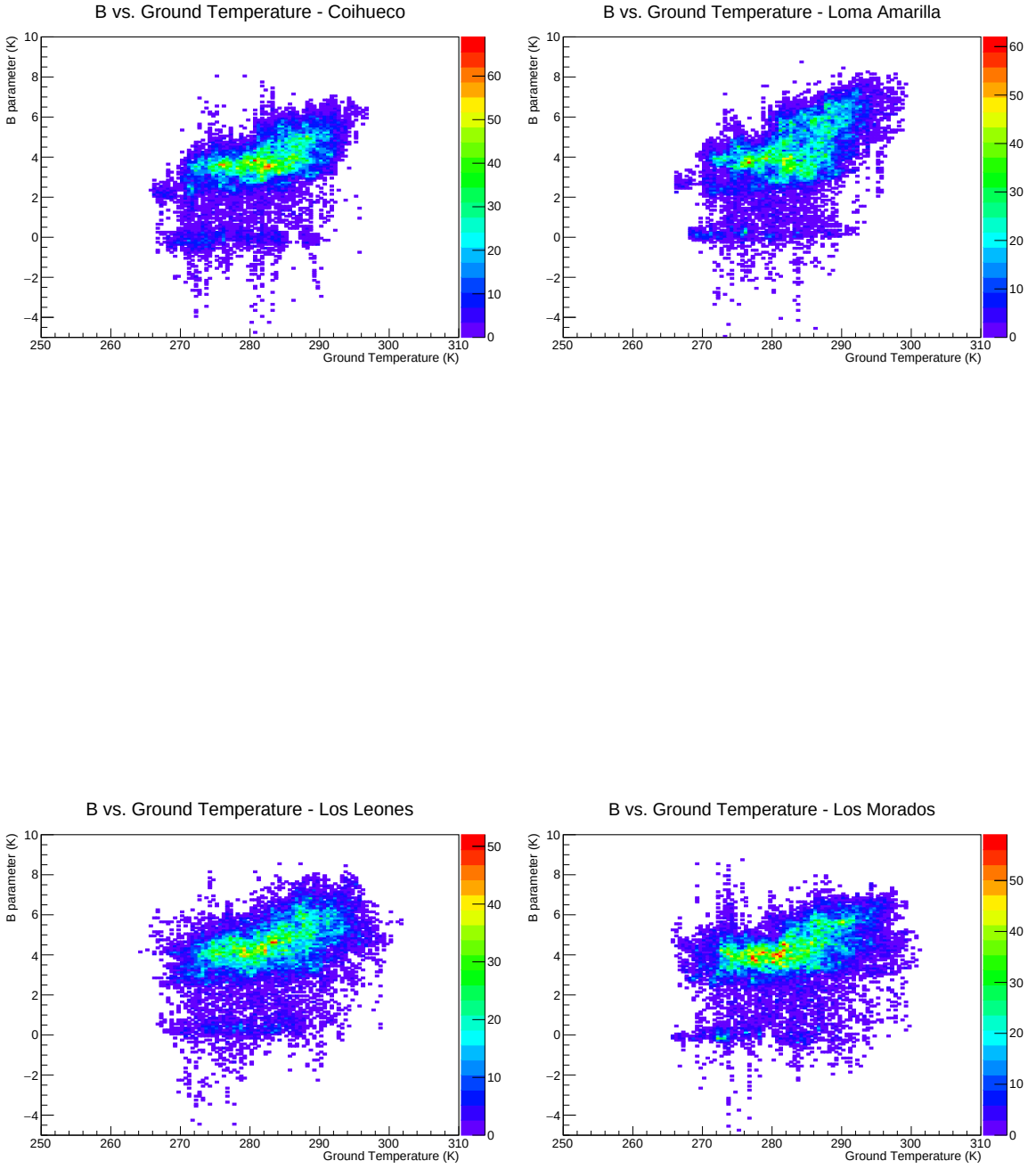

Figure 8.34: B parameter vs. Ground Temperature for each cloud camera in 2017/2018 for both clear and cloudy skies. Top left is Coihueco, top right is Loma Amarilla, bottom left is Los Leones and bottom right is Los Morados.

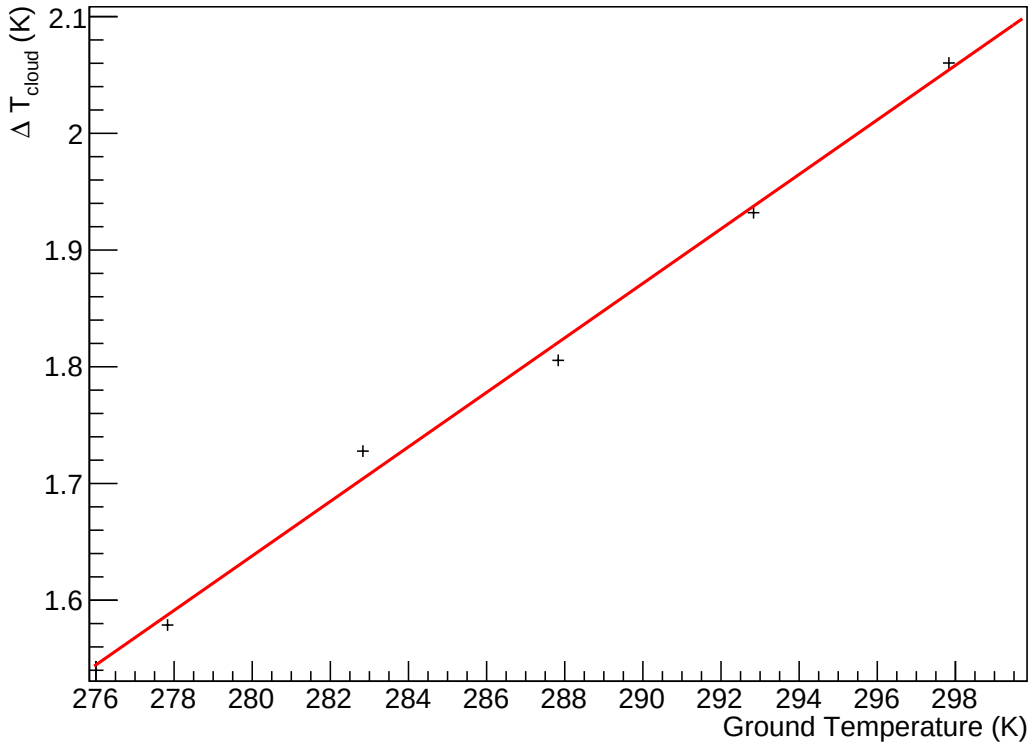

Figure 8.35: ΔT_{cloud} fitted as a function of the Ground temperature.

Delta T Correlations

As well as depending on the amount of water vapour present in the atmosphere, ΔT also depends upon the cloud itself, since even in the absence of water vapour there is still a measurable temperature difference. This temperature difference is approximately constant as a function of cloud height (as seen in Figure 8.30), implying a flat baseline value due to the cloud which can be subtracted from all subsequent temperature measurements. This means that ΔT can be split into two components, labelled ΔT_{cloud} and ΔT_{vapour}. The only dependency of ΔT_{cloud} is due to the ground temperature, which can be corrected for. This is done by fitting a horizontal line to the ΔT curve as a function of cloud height (refer to Figure 8.30) with no water vapour present, and calculating the intercept as a function of ground temperature (the slope is assumed to be 0). The result of this is shown in Figure 8.35.

As can be seen by Figure 8.35, the dependency of ground temperature on ΔT_{Cloud} seems to be linear (from MODTRAN). This is due to the infrared radiation emitted by the warm ground (which is assumed to be a blackbody radiator). As it happens, the range of ΔT_{cloud} is only $\sim 0.5\,\mathrm{K}$, which corresponds to a $\sim 70\,\mathrm{m}$ height difference This comes from assuming a lapse rate of $6.5\,\mathrm{K/km}$ and dividing the temperature difference by the lapse rate, as done in equation 8.2 for calculating cloud height via method B. It is only a minor effect, even over a large range of ground temperatures ($\sim 25\,\mathrm{K}$) as shown by Figure 8.30. As a result, an average offset of $\sim 1.8\,\mathrm{K}$ can be applied. In any case, the result of a linear fit to ΔT_{cloud} as a function of ground temperature is shown in equation 8.22 (in units of K). This is used to apply a more accurate correction due to

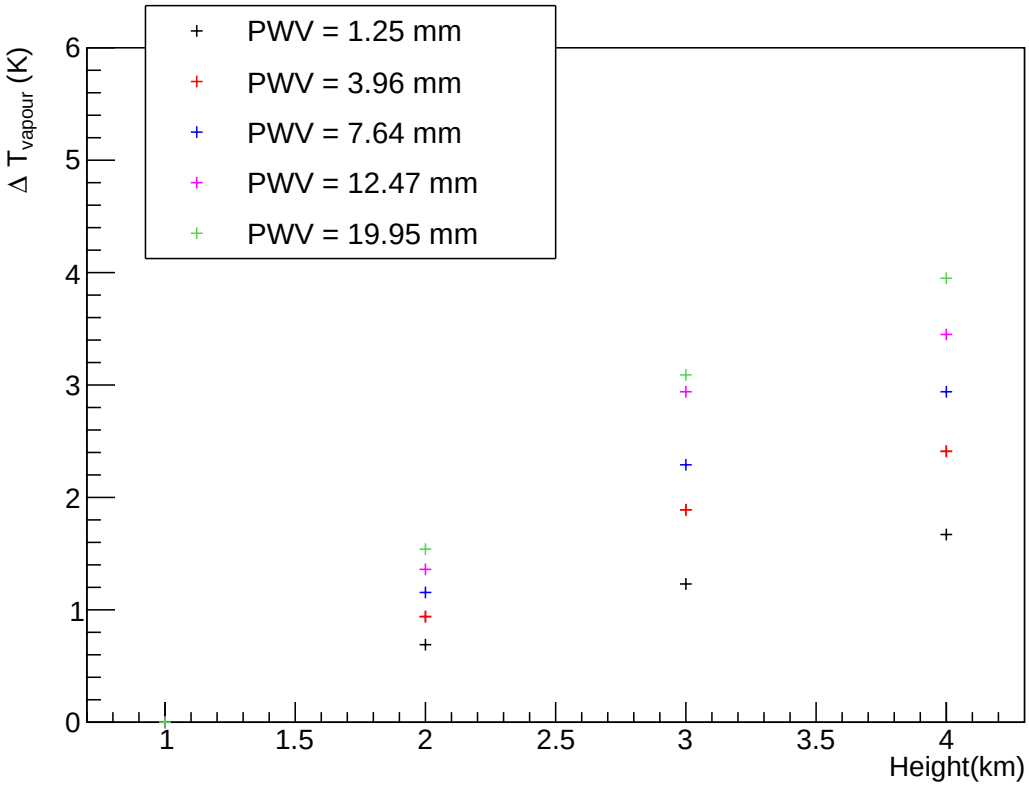

Figure 8.36: ΔT_{vapour} as a function of cloud height, for various values of PWV.

the absorption of Earth's radiation by both the cloud and the other atmospheric trace gases (not including water vapour), plus the effect of latent heat as discussed earlier.

$$\Delta T_{cloud} = 0.02(T_{ground}) - 4.90 \tag{8.22}$$

Now, looking again at Figure 8.30, one can notice that the difference in ΔT, with and without water vapour, does not change for different ground temperatures. This means that ΔT_{vapour} does not depend on the ground temperature. The parameterisation of ΔT_{vapour} is done by subtracting the baseline ΔT_{cloud} from Figure 8.30. The result of this is shown in Figure 8.36.

Once the baseline ΔT_{cloud} has been subtracted, a straight line can be fit to ΔT_{vapour} as a function of cloud height (for different values of PWV). From this, the slope and intercept can be extracted. The slope and intercept values will then be parametrised in terms of PWV. This is shown in Figure 8.37 for the slope (left) and intercept (right).

As is shown by Figure 8.37, a quadratic fit to each fit parameter is the best parameterisation for ΔT_{vapour} as a function of cloud height (for different values of PWV). The results of these quadratic fits are given by equations 8.23 and 8.24 for the slope (M) and the intercept (C).

$$M = -0.002(PWV^2) + 0.084(PWV) + 0.48 \tag{8.23}$$

$$C = 0.002(PWV^2) - 0.073(PWV) - 0.431 \tag{8.24}$$

Now that the sky temperature systematics are known as a function of PWV and ground temperature, the next step is to apply these ΔT corrections to the cloud height estimation techniques shown in section 8.1.

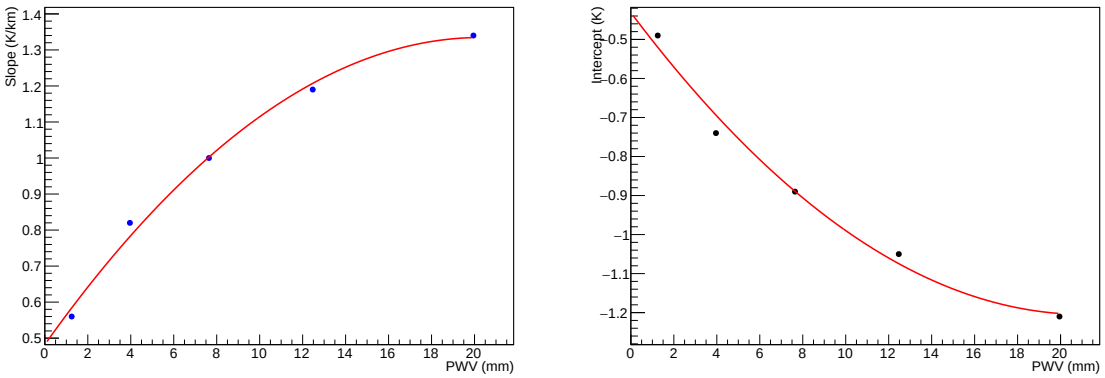

Figure 8.37: Slope (left) and intercept (right) of the ΔT_{vapour} fits vs. cloud height as a function of PWV. A quadratic fit is included as it best parameterises the data.

This is done iteratively using the following steps;

- Estimate the cloud base height via both methods (uncorrected)
- Calculate ΔT from the measured ground temperature and PWV
- Subtract ΔT from the sky temperature first measured by each method to calculate the corrected sky temperature
- Re-calculate cloud base height given this new sky temperature

In order to get the best estimation of cloud base height, one can compare the cloud base heights between the CLF and the direct GDAS profile method (method A), as well as the CLF with the lapse rate method (method B). This was done at the beginning of this chapter, without the ΔT corrections applied. The results of this, with the ΔT, lapse rate, sky temperature and ground temperature corrections applied, is shown in Figure 8.38.

Figure 8.38 illustrates that the ΔT correction has improved the estimation of cloud base heights (compared to Figure 8.3), however there are still inaccuracies when comparing the cloud base height to that of the CLF. There is a bias in the CLF recorded cloud base heights due to cloud base heights below 1 km not being recorded, as the laser cannot be seen at that low elevation by the FDs. As a result, estimated cloud heights below 1 km will not agree, creating a biased underestimate in the cloud base height.

To best check how well these techniques have worked (from Figure 8.38) when applying all corrections, one can again look at the residual difference between the calculated and measured cloud base heights. This is shown in Figure 8.39.

Figure 8.39 shows good consistency in the estimated cloud base height to that given by the CLF, with method A (matching the sky temperature of the cloud camera to the GDAS profile) performing slightly better overall. The residuals here show an improvement from the uncorrected cloud base height comparisons (see Figure 8.4), indicating that the ΔT corrections have indeed improved the estimation of cloud base height.

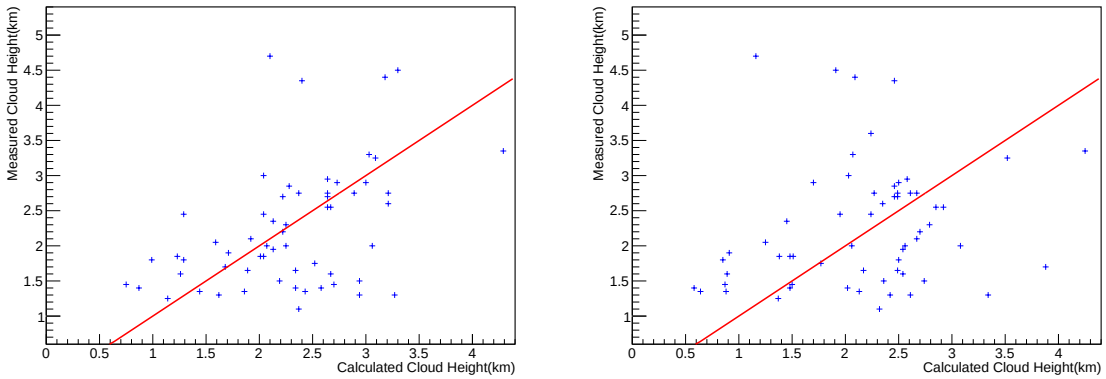

Figure 8.38: Comparison of calculated cloud height to that given by the CLF with all corrections applied. Left: Cloud height calculated by comparing the sky temperature to the GDAS profile, matching the ground temperature between GDAS and the local weather station, and applying ΔT corrections (Method A). Right: Cloud height calculated by the lapse rate method (equation 8.2), using the warmest peak value for sky temperature, lapse rate fitted from ground level to 5 km A.G.L, and applying the ΔT corrections (Method B).

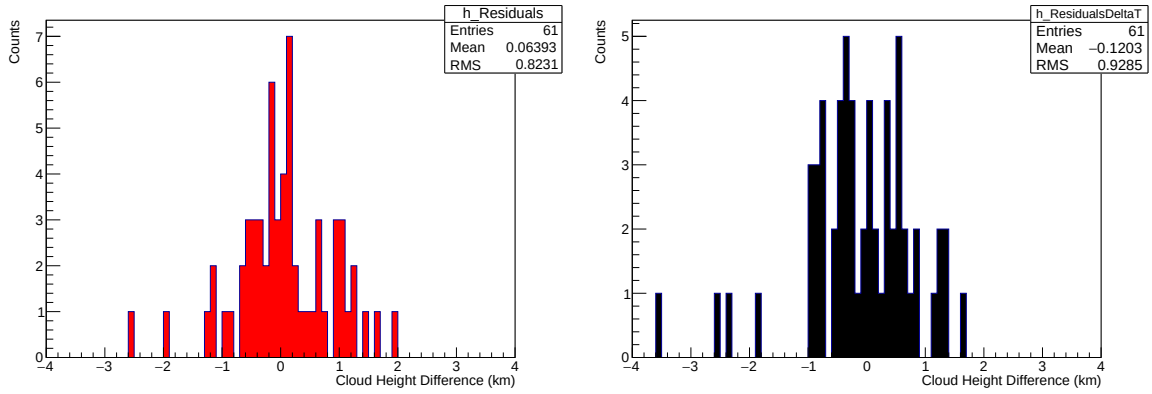

Figure 8.39: Residuals between the cloud base height estimated via method A (left) and method B (right) with the cloud base height measured by the CLF with all corrections applied.

It is worth looking at a couple of examples of the application of these cloud base height estimation methods, by illustrating what information is gathered from each instrument (cloud camera, GDAS, and the CLF) to estimate the cloud base height. Firstly, Figure 8.40 shows the cloud camera image, temperature profile (from GDAS), vertical sky temperature distribution for each pixel of the vertical cloud camera image, and the light profile used to estimate the cloud height with the CLF, before the ΔT corrections are applied. The light profile (top right in Figure 8.40) shows the return signal of the CLF laser at different heights in the atmosphere (above ground level). If the laser passes through a cloud, the cloud will increase the scattering of light in the direction of the FD, which shows up as a peak in the light profile, as shown in Figure 8.40. The lowest peak from each of the 4 data windows (each 15 minute period, as discussed in section 8.1) is the chosen cloud height from the CLF for that particular hour. The blue curve in the light profile is the laser signal received on a reference night, which is simply defined as a clear night with no aerosols present in the atmosphere. As can be seen from the light profile, the presence of cloud creates a much stronger signal (yellow curve) than what is seen on an ideal night, allowing for cloud to be easily identified. Using the GDAS temperature and cloud camera image gives a first estimate to the cloud base height, before the ΔT corrections are applied which will recalculate the cloud temperature/base height. In this example, both height estimation methods agree with the base height given by the CLF (2.7 km) to within 100 m after the ΔT corrections are applied.

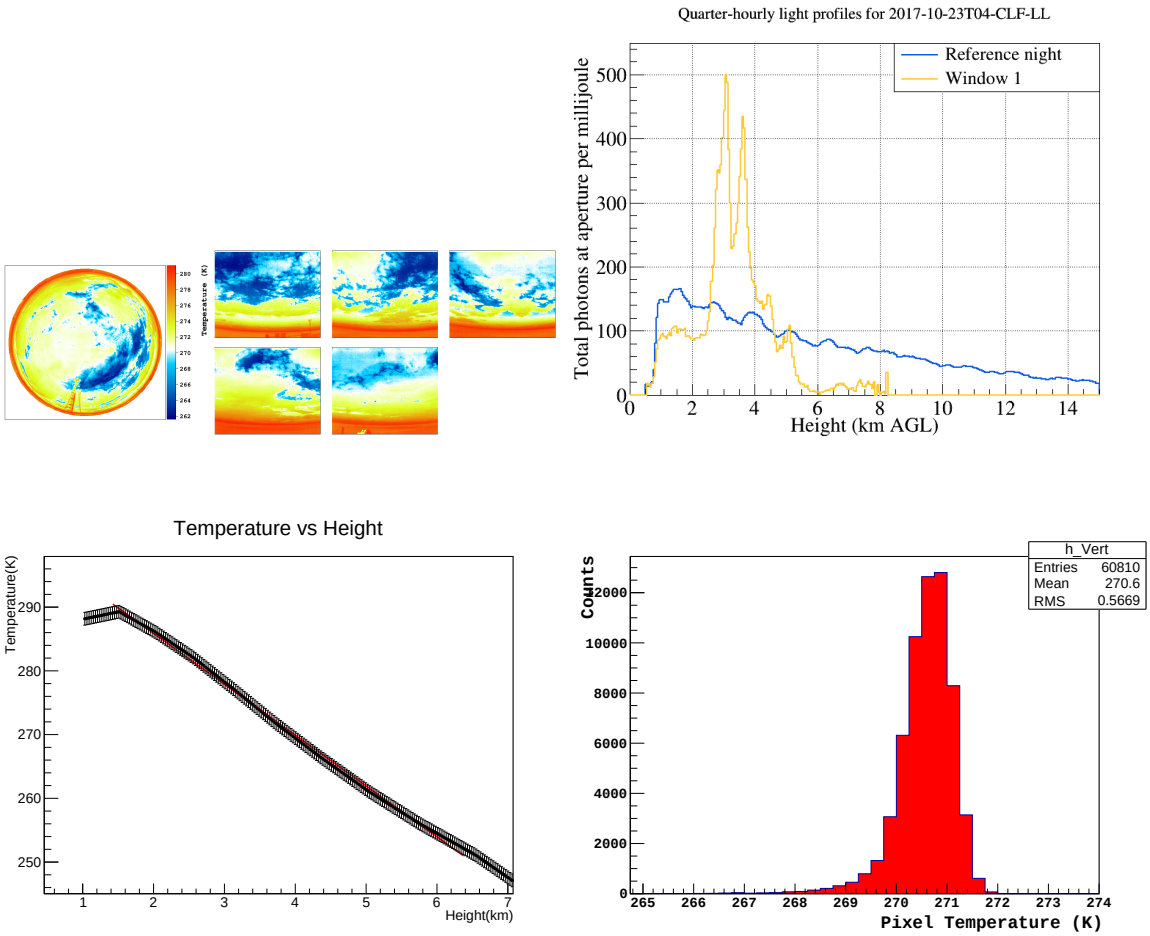

Figure 8.40: Top left: Cloud camera scan clearly showing a cloud in the vertical image. Top right: The light profile from the CLF for this hour. Bottom left: Temperature profile from GDAS at this time, showing that a linear fit to obtain the lapse rate is accurate. Height is given as above sea level (A.S.L). Bottom right: Pixel sky temperature distribution from the vertical cloud camera image.

Figure 8.40 illustrates that the two cloud height estimation techniques are accurate when the data obtained is accurate. There is a clear peak in the light profile (top right) at around 2.7 km, which is much brighter than expected for a clear night (blue curve) due to an increase in scattering of photons out of the laser beam, caused by the presence of a cloud. The temperature profile from GDAS (bottom left) shows that the assumption of a linear decrease in temperature with height is reasonable. The low spread in the pixel temperature distribution (bottom right) means that the average pixel temperature is an accurate estimate of the cloud temperature. However, this high degree of agreement between the cloud height estimation technique and the CLF does not always occur, as shown by Figure 8.41.

Figure 8.41 shows the estimation of cloud height given by the CLF is inaccurate. Peaks in the CLF light profile below 2 km are clearly visible, which would better correspond to the cloud height estimated from equation 8.2 (a height of $\sim$ 1.3 km above ground level, or 2.7 km above sea level).

The reference night light profile is also shown, which is the laser light seen for ideal conditions (clear sky + no aerosols), which is much brighter than the data for

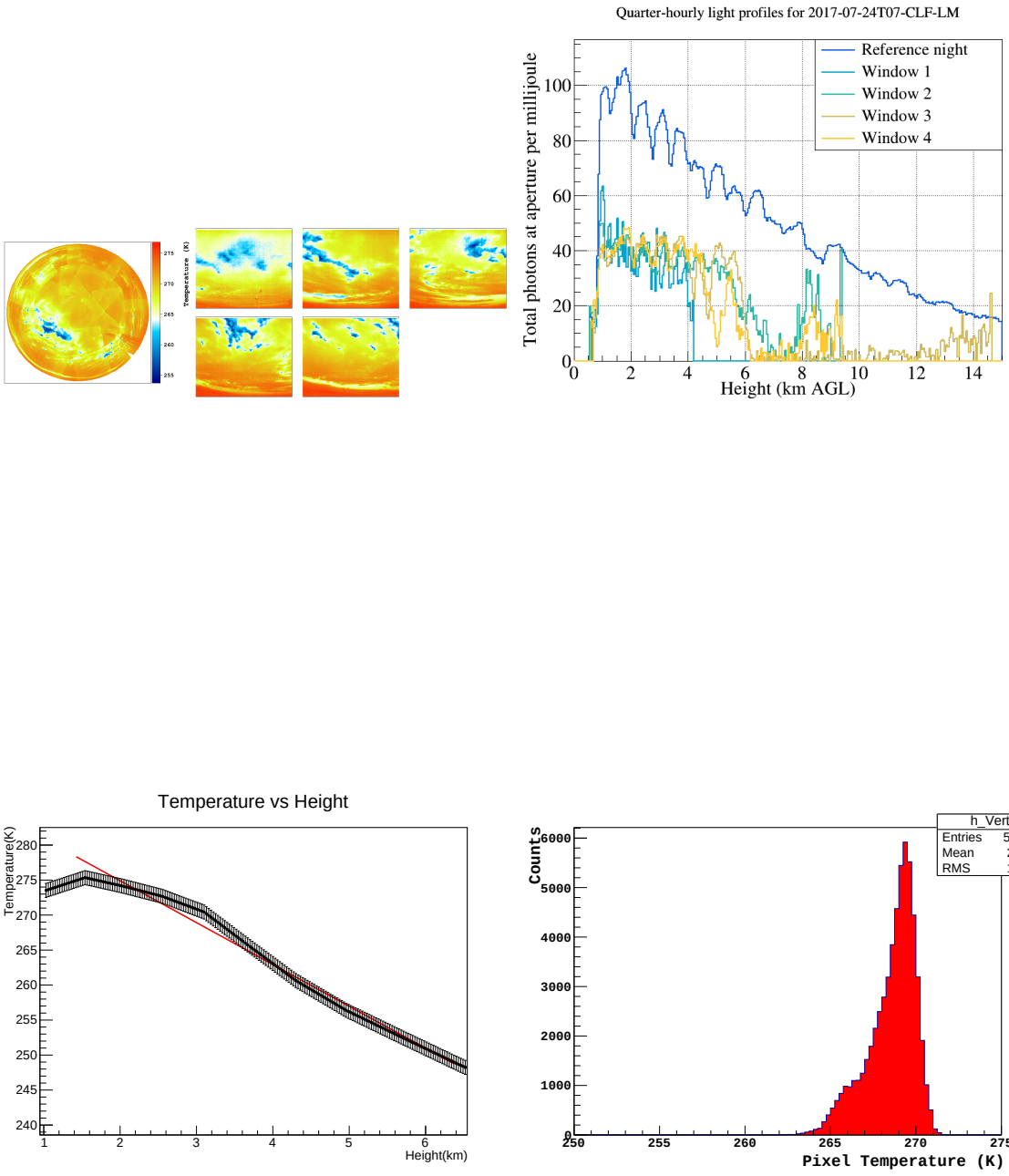

Figure 8.41: Top left: Cloud camera scan clearly showing a cloud in the vertical image. Top right: The light profile from the CLF for this hour. Bottom left: Temperature profile from GDAS at this time. Height is given as above sea level (A.S.L). Bottom right: Pixel sky temperature distribution from the vertical cloud camera image.

this particular hour. Cloud and aerosols can scatter a large amount of photons out of the CLF laser beam. The relative difference between light profiles in each window is quite large around 4.7 km, which can be falsely identified as a cloud in rare cases. From observation of all other images in Figure 8.41, it is clear that the measured cloud height of 1.3 km is a reasonable estimate for the true cloud height. There is a small spread in the pixel temperature distribution, thus the temperature at the peak of the distribution is used for the sky temperature. However, this and the temperature profile appear to be able to discern the cloud temperature and lapse rate quite accurately, suggesting that the only inaccuracy is the false interpretation of the cloud in this case. This type of circumstance is rare, but can explain the outliers seen in Figure 8.38.

Further examples of these light profile comparisons can be found in Appendix E.

Overall, applying the ΔT corrections has improved the cloud height estimation technique, and has given us a much greater understanding of the effects of water vapour and cloud in the infrared.

8.7 Summary

Overall, the application of the two cloud base height estimation techniques using the cloud camera compare favourably with the measured height from the CLF. Once all the proper corrections for lapse rate, ground temperature and sky temperature have been properly applied, the resulting cloud base height agreed well with expectation and was shown to be a much better estimate in some cases. It has been shown from various outlier points, that in the case of conditions not being totally overcast, that the cloud camera and CLF may not be viewing the same cloud. However, in this case the cloud camera was shown to provide a reasonable cloud height estimate, showing that in the case of the CLF missing the cloud seen by the camera, the cloud height estimation technique can provide a good estimate of the true cloud base height in the atmosphere. For conditions that are not completely overcast, the sky temperature can be estimated by the peak in the pixel temperature distribution for the cloudy pixels, allowing for an accurate as possible measurement of the cloud temperature (hence its height). Limitations in the CLF (such as missing cloud, and not being able to estimate cloud heights below 1 km) and not knowing the emissivity of a cloud in the cloud camera FOV are still uncertainties to be taken into account; however it has been shown that the cloud height estimation technique can accurately calculate the base height of clouds in the atmosphere. This technique can be employed as a supplementary check on cloud base heights in order to determine if cloud has influenced the longitudinal profile of an extensive air shower event, allowing for more events to be used in subsequent analysis, which will improve the overall statistics of the Pierre Auger Observatory.

Chapter 9

Conclusions

The major aim of this thesis was to better understand the effects that the atmosphere (specifically clouds) above the Pierre Auger Observatory has on Extensive Air Shower (EAS) measurements and reconstruction.

Currently there are two instruments used to extract the height of clouds in the atmosphere, namely the LIDAR, and the CLF. An important goal of my studies has been to determine if there are other methods that can be employed to measure the height of clouds in the atmosphere, in the event of these monitoring instruments not operating. I presented two methods, one using a known atmospheric parameter known as the lifting condensation level (LCL), and the other using the infrared cloud cameras installed at Auger and housed on top of each fluorescence detector building. I also utilised data from the local weather stations at each FD site, and the data assimilation model known as GDAS.

Firstly, the lifting condensation level method (discussed in Chapter 4), which is the height in the atmosphere at which condensation of water vapour occurs and clouds begin to form. It is defined as the height at which the temperature and dew points match, such that there is 100% relative humidity. This parameter was found to overestimate the height of clouds in the atmosphere (compared to the LIDAR), due to a number of uncertainties in the analysis.

Firstly, GDAS ground temperatures do not match the temperature given by the local weather stations (located at ground level) and as such, this correction must be applied in order to best match the temperature profile with the local weather conditions. As well as this, fitting a lapse rate to the GDAS temperature profile was shown to be inconsistent, as the approximation of a linear decrease in temperature as a function of height is not always applicable. The release of latent heat due to the condensation of water vapour (as discussed in section 4.2.1) causes the atmosphere around the water droplet (i.e. within the cloud) to heat up, slowing the steady rate of decrease of temperature with height [67]. This can be seen as a characteristic bump in radiosonde profiles (see Figure 4.6), which will decrease the lapse rate, hence increase the calculated LCL. This is thought to be the main cause of the overestimate of cloud height via the LCL technique.

However, the dew point depression (or DPD - defined as the difference between the temperature and dew point at a given height) from the LCL could be used to predict whether a particular time is likely to be cloudy. This was done by defining threshold values on the DPD (dependent upon atmospheric temperature) that would define cloud layers in the atmosphere and their parameters (such as cloud base height, cloud top

height and thickness) [79]. Optimisation of the thresholds was then performed to find the best values possible to achieve the highest overall accuracy of the prediction of cloud. This was shown to be accurate $\sim$ 70% of the time, making it a reasonable supplementary check of atmospheric conditions in the absence of other instruments. It was not shown to improve the estimation of cloud base heights however, meaning another technique needed to be employed for better accuracy of cloud base height estimation.

The other cloud base height estimation method, using the infrared cloud cameras, was shown to perform better than the LCL method for estimating cloud base heights. It is well established from previous studies that the brightness temperature of a cloud in the infrared is generally overestimated by thermal imaging cameras, due to the release of latent heat via the condensation process [67] [97]. There are also atmospheric molecules (most notably water vapour) present below the cloud that absorb and re-emit infrared radiation at similar wavelengths to the cloud, which also contribute to the overestimation in temperature. Once these effects were well understood and simulated using MODTRAN, they were able to be corrected for, allowing for a better measurement of cloud temperature (thus cloud base height). Whilst there are still uncertainties on the temperature due to the emissivity of a cloud (which is not a perfect blackbody absorber), the technique was able to show a high degree of accuracy in estimating cloud base height compared to the CLF. This technique can be employed in the absence of other instruments in order to estimate cloud base height in the atmosphere (for thick clouds), which will improve the overall statistics and quality of extensive air shower reconstruction data.

The prevalence of cirrus cloud in the cloud camera field of view is also important, and has not been as well understood. Due to the fact that cirrus cloud is both high altitude, and optically thin, it has a different spectral response in the atmospheric window to all other liquid water clouds. The best tool for understanding the effect of cirrus cloud on our cloud camera images is to simulate their spectral response with MODTRAN. Using this program has allowed us to simulate the response of cirrus cloud in the atmosphere, which has been conclusively shown to have a very similar spectral response to that of a clear sky. Cirrus cloud is typically high up in the atmosphere, beyond the point at which most extensive air showers develop and deposit their maximum energy in the atmosphere. This fact, coupled with the clear sky similarity, means that one can ignore cirrus cloud in the field of view of the FDs since this will not impact air shower reconstruction. Previous work has included the effects of cirrus cloud in the cloud quality cuts (see section 3.5.9) since the cloud camera thresholds were adjusted to image cirrus cloud in their field of view. This results in air shower events being removed from the air shower analysis since they did not pass the quality cuts. However, with this new information, we are able to keep a larger fraction of air shower events by ignoring cirrus cloud, which improves the overall statistics of the reconstruction.

The previously completed analysis on the processing of cloud camera images has allowed us to extract greater detail on the cloud cover within each FD pixel for the purposes of air shower reconstruction [48]. A key goal from my work on this process has been to implement the HEAT telescopes in the analysis, since this had not been done previously. Testing of the modifications proved fruitful, showing that information from HEAT could also be used and sorted into the cloud database for further analysis. As a result, we were able to modify the scanning sequence of the cloud camera located at the Coihueco FD site, in order to include HEAT data as frequently as the other

FDs. The inclusion of HEAT has allowed for cloud quality cuts to be extended to lower energy showers than was done previously, thus improving the statistics at these lower energies.

Another key addition to this process was to implement a GUI (Graphical User Interface) in order to provide the collaboration with a much faster and more user-friendly method of processing cloud camera images. The use of the GUI has greatly increased the productivity of the analysis.

Finally, the importance of the characteristics of cloud on our analysis has always been high, and as a result, a study was performed to understand the seasonal variations in the cloud base height and cloud coverage in the atmosphere. These results showed that cloud is much more prevalent in the winter months, an important finding due to the nights (and as a result, FD observing times) being much longer compared to summer.

Appendix A

Error Estimates of the LCL

Given there is built in assumptions to the calculation, uncertainties in the quantities are present. Radiosonde flights from the GDAS/Auger comparison paper (done at the Balloon Launching Station on site) quoted a temperature difference of 1 K [37]. This is the value of temperature uncertainty that is used henceforth. For the other derived quantities, this uncertainty is unknown. However, given the relation of vapour pressure with height (shown by equation A.1 for constants A and B) we can write;

$$\Delta e(h) = A\exp(Bh) \tag{A.1}$$

Via the same analysis of radiosonde comparison, the quoted errors at various heights (1 hPa at sea level and 0.37 hPa at 2 km a.s.l - corresponding to the scale height of water vapour [37]) can be substituted to find the constant values A and B. This is shown by equation A.2.

$$\Delta e(0) = A\exp(B(0)) = 1 \tag{A.2}$$

$$\therefore A = 1 \tag{A.3}$$

$$\Delta e(2) = \exp(B(2)) = 0.37 \tag{A.4}$$

$$2B = \ln(0.37) \tag{A.5}$$

$$\therefore B = -0.5 \tag{A.6}$$

In the same way, errors can be estimated for the Dew Point Temperature. When the air temperature is above 0° C, the Dew Point Temperature is given by;

$$T_d = \frac{237.3\ln(\frac{e}{6.108})}{17.2 - \ln(\frac{e}{6.108})} \tag{A.7}$$

To estimate the error in the Dew Point, start by taking the derivative, which is given by;

$$\frac{dT_d}{de} = \frac{237.3(\frac{1}{e})(17.2 - \ln(\frac{e}{6.108}) - 237.3\ln(\frac{e}{6.108})(-\frac{1}{e})}{(17.2 - \ln(\frac{e}{6.108}))^2} = \frac{4081.56}{e(17.2 - \ln(\frac{e}{6.108}))^2} \tag{A.8}$$

Using error estimates given by the variance σ^2, the error in the Dew Point can be written;

$$\sigma^2_{T_d} = (\frac{dT_d}{de})^2\sigma^2_e \tag{A.9}$$

Now, we can write the uncertainty in Dew Point as;

$$(\sigma_{T_d})^2 = (\frac{4081.56}{e(17.2 - \ln(\frac{e}{6.108}))^2})^2(\Delta e)^2 \tag{A.10}$$

$$\therefore \sigma_{T_d} = \pm(\frac{4081.56}{e(17.2 - \ln(\frac{e}{6.108}))^2})(\Delta e) \tag{A.11}$$

This gives an approximate error in the dew point of $\pm 1.8\,\mathrm{K}$ for air temperatures above the freezing point. When going below the freezing point of water, complications arise. The errors in dew point temperature become so large that the values calculated are now meaningless. Thus, another equation is required to deal with the step in temperature. As it happens, the saturated vapour pressure is divided piece-wise above and below the freezing point. Above zero degrees, the saturated vapour pressure is given by

$$e_s = 6.108 \times \exp(\frac{17.2T}{237.3 + T}) \tag{A.12}$$

which leads to the expression for the Dew Point in the same regime (taking $e_s(T_d) = e$), whereas for temperatures below zero the saturated vapour pressure is given by;

$$e_s = 6.107 \times \exp(\frac{21.88T}{265.5 + T}) \tag{A.13}$$

Applying the same definition as above and rearranging to solve for T_d we find;

$$T_d = \frac{265.5\ln(\frac{e}{6.107})}{21.88 - \ln(\frac{e}{6.107})} \tag{A.14}$$

Similarly, performing the same technique as previously the error can be estimated as;

$$\therefore \Delta T_d = \pm(\frac{5809.14}{e(21.88 - \ln(\frac{e}{6.107}))^2})(\Delta e) \tag{A.15}$$

Once the errors in both the air temperature and dew point temperature have been determined, the error in DPD (Dew Point Depression) can be determined from error propagation techniques. The estimate of the variance in DPD is done via the *error propagation equation* [104].

$$\sigma_x^2 = \sigma_u^2(\frac{\partial x}{\partial u})^2 + \sigma_v^2(\frac{\partial x}{\partial v})^2 + 2\sigma_{uv}^2(\frac{\partial x}{\partial u})(\frac{\partial x}{\partial v}) + ... \tag{A.16}$$

The cross terms are weighted differently and if the fluctuations in the measured quantities u and v are uncorrelated, then on the average, one should expect the term to vanish in the limit of large random selection of observations [104]. These terms can then be ignored and the final result is as follows;

$$\sigma_x^2 = \sigma_u^2(\frac{\partial x}{\partial u})^2 + \sigma_v^2(\frac{\partial x}{\partial v})^2 \tag{A.17}$$

Substituting the variables for temperature and dew point gives the error in the depression (σ_{Dep}) as;

$$\sigma_{Dep} = \sqrt{\sigma_T^2 + \sigma_{T_d}^2} \tag{A.18}$$

Appendix B

Optimum Value for Dew Point Thresholds

This distance l is given as;

$$l = \sqrt{s_1^2 + s_2^2} \tag{B.1}$$

with s_1 and s_2 given by;

$$s_1 = x_1 - (x_1 a_1 + x_2 a_2) a_1 \tag{B.2}$$

$$s_2 = x_2 - (x_1 a_1 + x_2 a_2) a_2 \tag{B.3}$$

Here, x_1 is the distance along the x-axis of the maximum point (i.e. the FPR), and x_2 is the distance along the y-axis (TPR). Direction vectors a_1 and a_2 are given by;

$$(a_1, a_2) = (\cos\theta, \sin\theta) \tag{B.4}$$

where θ is the angle of the random guess line (the black line in Figure 4.19) - in this case 45°. As a result;

$$(a_1, a_2) = (\cos 45, \sin 45) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \tag{B.5}$$

The maximum distance l can now be simplified by finding the form of s_1 and s_2.

$$s_1 = x_1 - (\frac{x_1}{\sqrt{2}} + \frac{x_2}{\sqrt{2}})\frac{1}{\sqrt{2}} = x_1 - (\frac{x_1}{2} + \frac{x_2}{2}) \tag{B.6}$$

$$s_2 = x_2 - (\frac{x_1}{\sqrt{2}} + \frac{x_2}{\sqrt{2}})\frac{1}{\sqrt{2}} = x_2 - (\frac{x_1}{2} + \frac{x_2}{2}) \tag{B.7}$$

Hence

$$s_1 = x_1 - \frac{x_1}{2} - \frac{x_2}{2} = \frac{x_1 - x_2}{2} \tag{B.8}$$

$$s_2 = x_2 - \frac{x_1}{2} - \frac{x_2}{2} = \frac{x_2 - x_1}{2} = -s_1 \tag{B.9}$$

As a result;

$$l = \sqrt{s_1^2 + s_2^2} = \sqrt{s_1^2 + (-s_1)^2} = \sqrt{2}(s_1) \tag{B.10}$$

Appendix C

Derivation of Lifting Condensation Level

Using the assumption that temperature decreases linearly with height at a constant rate Γ, then the temperature T at a given height h is given as

$$T(h) = T_0 - \Gamma h \tag{C.1}$$

where T_0 represents the temperature at a freely chosen reference level. By applying the same assumption to the dew point temperature (decreases at a constant rate Γ_w), we can write

$$T_d(h) = T_1 - \Gamma_w h \tag{C.2}$$

where T_1 represents the dew point temperature at the same reference level as the temperature. Now, since T and T_d are equal at the LCL, equations C.1 and C.2 can be combined giving:

$$T_0 - \Gamma(LCL) = T_1 - \Gamma_w(LCL) \tag{C.3}$$

$$T_0 - T_1 = \Gamma(LCL) - \Gamma_w(LCL) = (\Gamma - \Gamma_w)(LCL) \tag{C.4}$$

and rearranging:

$$LCL = \frac{T_0 - T_1}{\Gamma - \Gamma_w} \tag{C.5}$$

In equation C.5 above, the LCL is calculated at some reference level, which can be freely chosen. In other words, the LCL can be calculated at any level (height) in the atmosphere. The value calculated from this, is the height above the given reference level where clouds form. For the purposes of this study, the LCL was calculated at the ground level of the array ($\sim 1.42\,\text{km}$ above sea level).

Appendix D

Instructions for using the Cloud Mask GUI

Here are the instructions in order to run the Cloud Mask Graphical User Interface (or GUI for short) for processing of cloud camera images (which is what fills the cloud database). This appendix will go through the file structure of all the files required in order to run the program, as well as outputs and errors with the program. Everything you need for the program to run is outlined in the following sections. The cloud mask directory which contains everything required is as follows: "/remote/riposte/jday/CloudMasks-master". This directory is able to be accessed for the ease of cloud mask analysis.

File structure

Files are zipped by site and by day, which is done for the purposes of file size compression. File sizes can get quite large due to the high number of images from each camera, thus it is advised to only have a few months worth of data at any one time (or if looking at an entire year only have one site's data at a time). All data is readily accessible on the "comptona" local machine, and the extraction of images is straightforward. The directory containing all raw data on comptona is as follows: "/remote/comptona/CloudCamera/Data/Current_System_Data/

Raw_Data__Post_2017/Auger/SS" where SS is the two digits corresponding to each site (LL/LM/LA/Co). Each site is then grouped by year, month, and day.

The main directory needed contains the cloud mask program (as well as all other files required to compile and run the program), as well as several subdirectories. There are two header files required in the main cloud mask directory, labelled "CameraInfo.h" and "DatesTimes.h". The camera info header file contains the ImageMagick libraries required for the processing and display of cloud camera images, whilst the date/time header file contains the boost libraries required to manipulate dates and times (i.e. converting from GPS time to UTC time). A "Make-depend" and "Makefile" are also required in this main directory, in order to compile and run the program.

Now, to discuss all the subdirectories and files required. The first of these subdirectories is labelled "extract_camera_data" which contains a bash script ("extract.sh"). Inside of the extract_camera_data folder, there are two other subdirectories known as "Raw" and "Processed". The data from comptona must be copied into the Raw directory first. As an example, in order to copy all data for 2020 (must be done one site

Figure D.1: Directory structure for the cloud camera image files that will be processed. In this example, "CloudMasks_Violet" is the name of the main cloud masks directory which contains the cloud mask program.

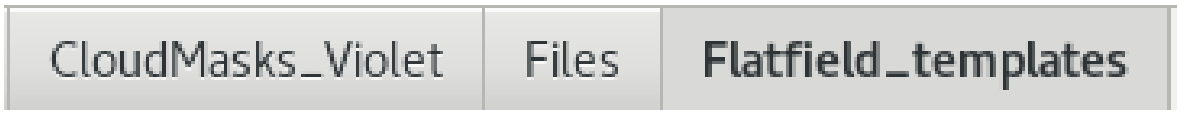

Figure D.2: Directory structure for the flat-field correction text files that will be applied by the cloud mask program. In this example, "CloudMasks_Violet" is the name of the main cloud masks directory which contains the cloud mask program.

at a time) you must make a 2020 directory in Raw, then once you are in that directory put the following statement into the command line:

"cp -r /remote/comptona/CloudCamera/Data/Current_System_Data/ Raw_Data__Post_2017/Auger/SS/2020 .". The "." at the end is a required statement for copying zip files.

Once the files are copied from comptona into the "Raw" directory, return to the extract_camera_data directory and run the bash script. This will extract all individual images into the "Processed" directory (sorted by month and day) and delete the data from the "Raw" directory. After this, copy/move all files from "Processed" into "Files/CameraData/SS" where SS is the two letter acronym for each FD site (LL/LM/LA/Co/HE) ["Files" is a subdirectory of the main directory]. Note that within these site directories, the files will again need to be sorted by year, month, and then day. HEAT uses the same data as Coihueco and thus does not, require its own subdirectory. Once this is done, all cloud camera data is ready to be run through the program to generate cloud masks. Figure (D.1) shows the structure of these data directories.

The second important subdirectory in the main cloud masks directory is known as "Flatfield_Templates". This directory contains the text files for the flat-fielding image correction outlined in Chapter 5. Again, these are sorted by each FD site (LL/LM/LA/Co/HE) and there is one text file for each FD. HEAT also uses the same flat-field correction as Coihueco. Make sure that these templates are in the correct directory as shown by Figures (D.2 and D.3).

There are other important files contained within the "Files" directory which must also be present in order for the cloud mask program to run. Firstly, there is a header file "CameraInfo.h" which contains the camera data structures for reading the files. There is also a "DatesTimes.h" header file which for a given input can provide dates and times (using boost libraries) to make the input and output of images simple. These will be placed in the directory where you will execute the program.

As well as the flat-fielding templates, there is another essential directory inside the Files directory labelled "FDpixels". This directory also sorts the FDs by site, and contains the pixel pointing directions for each telescope of the FDs. An example for the Coihueco site is illustrated in Figure D.4.

The last required file in this File directory is an "ontimes.dat" file which contains the on-times of the FDs for times being processed. This file speeds up the analysis

Figure D.3: This is the flat-field correction template for each FD site, contained in the directory shown in Figure D.2. These files must exist in order to apply the flat-field correction to all processed cloud camera images.

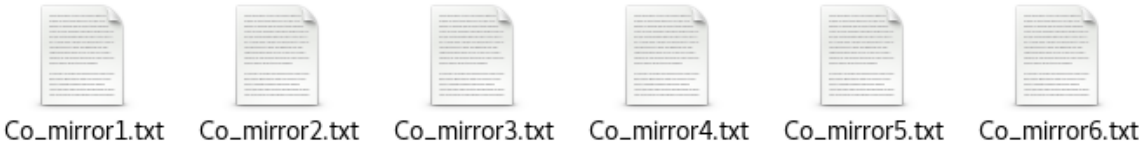

Figure D.4: Example pixel pointing direction text files for each telescope of Coihueco. There are copies of these text files for each site.

process by skipping over times when the FDs were not observing. As an example, "ontimes_2005_2019.dat" contains all the FD ontimes from the start of 2005 to the end of 2019.

```
riposte:[jday/CloudMasks-master]$ ./masks -t 3.5 0 "2019-01-01" "2020-01-01"
```

Figure D.5: Example execution of the cloud mask program. The first option "-t" sets the threshold to run the images over (3.5 for LL/LM/LA and 4.5 for Co/HE). The second number (0 in this example) corresponds to the site that you wish to process (0 = LL, 1 = LM, 2 = LA, 3 = CO and 4 = HE). The dates in "" are the day(s) [start and end] which you want to process, ordered by year, month and then day. The dates defined will perform the analysis over each entire day, from start to end. Dates are defined here by the time of the last image (e.g. if the last image was 2019-01-10 09:54:22 then the day of all images is 2019-01-10). The hour, minute and second can also be defined within "" (for example "2019-01-01 01:20:00") if only a select few images are required.

GUI Execution

Once all the required files are present, the next step is to run the cloud mask program to generate the cloud camera images. This is done by executing the program as illustrated by Figure D.5 in the CloudMasks-master directory. This is the same program that runs the GUI (the program itself is "Backend.cc" in the src directory).

From Figure D.5, the -t option and number (3.5 for LL,LM and LA, 4.5 for Co and HE) are required in order to set the default thresholds for each camera in order to produce the images that will be displayed by the GUI. Additionally, the site ID number and dates are required to be input, else the program will not execute. The site ID is 0,1,2,3 or 4 for LL/LM/LA/Co/HE respectively. The dates input are whichever day(s) you wish to process. It is done inclusively, meaning that the second date input is the final day you wish to process. As an example, if you wished to run over all of September 2018 for Los Leones, the execution would be " ./masks -t 3.5 0 "2018-09-01" "2018-09-30" ".

Once all of the site data has been processed, these images will be placed into the output directory within CloudMasks-master. Inside of the CloudMasks-master directory you will find the following files and directories:

- camera_data: A symbolic link that points to where the camera image data is stored [as illustrated by Figure D.1].
- extract_camera_data: Contains the Raw and Processed directories in order to download the raw cloud camera data.
- resources: Contains the pixel pointing direction files, flat-field template files and the ontimes data file.
- src: Contains all source code for the GUI (i.e. header files and executable files) and the original cloud mask program written by Patrick [48].
- output: This is where the processed cloud camera images and cpd files from running the cloud mask program will be placed. Re-running the program in the GUI will then place all updated images and cpd files in this directory after changes are incorporated.
- masks: This is the executable program to use the GUI, and process/produce cloud camera images.

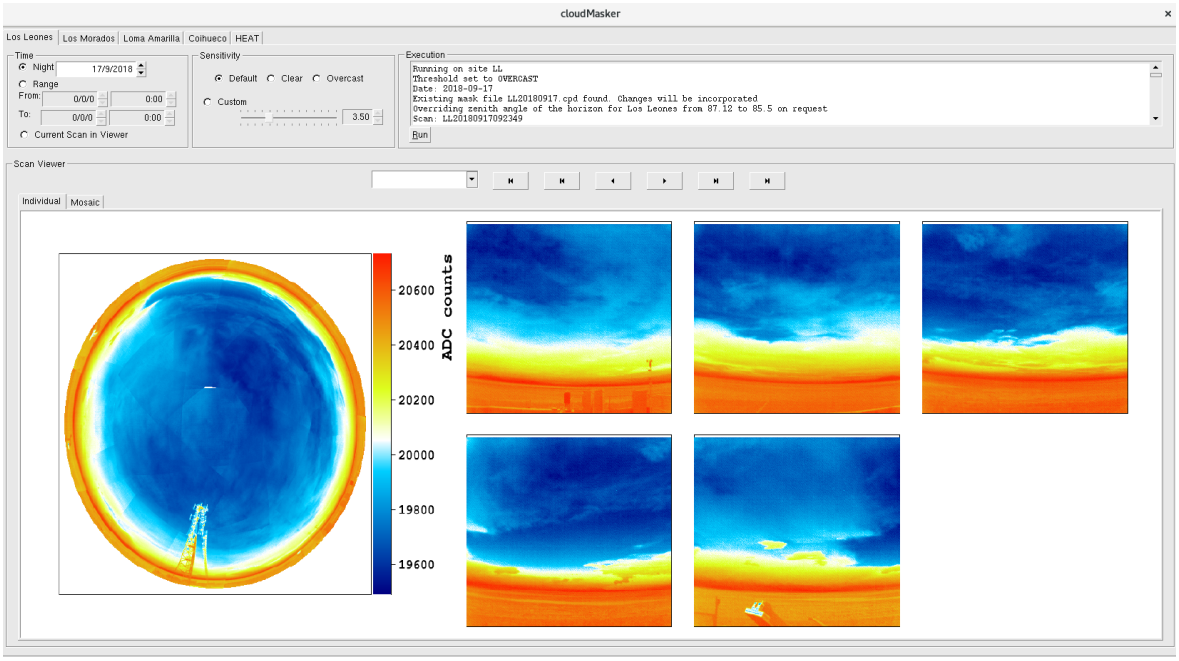

Figure D.6: Screenshot of the Graphical User Interface for the Cloud Mask program.

- README.md: Instructions on the structure of the GUI in case one runs into any errors.

- AUTHORS: If you are stuck on any one particular step in the implementation, the relevant person(s) to contact is listed.

Executing "./masks" without any options will bring the GUI up, displaying every image (that has been processed as shown by Figure D.5) for each site that is in the output directory from earliest to latest. This GUI display is illustrated by Figure D.6.

This is one way of displaying the images, by showing the field of view images of that particular scan. In order to properly make the necessary adjustments to the images, the mosaic view is required by clicking the mosaic button on the left hand side. The mosaic view is shown in Figure D.7.

Once in this mosaic view, the relevant changes to each individual image can be made. This is done via the threshold buttons at the top of the page. There is a slider that can be used to adjust the cloud threshold, or buttons to set a particular image to overcast or clear. If an image has not captured all of the cloud in the field of view, the threshold will need to be lowered (moving the slider to the left). In the case where too much of the image has been marked as cloudy, the threshold will need to be increased. This adjustment will never be required to be more than 1 K either way (as discussed in Chapter 5). The changes are then implemented by clicking the "Run" button, which will instantly display the results of the change ("Run" is the equivalent of typing the statement from Figure D.5 into the terminal for a single time). In order for the Run button to work the camera data needs to be available in the relevant directory. The recommended way to make these changes is to focus on one site at a time (and delete the camera data from the site once all changes are made), since all camera data for all sites for an entire year can completely fill up the riposte machine.

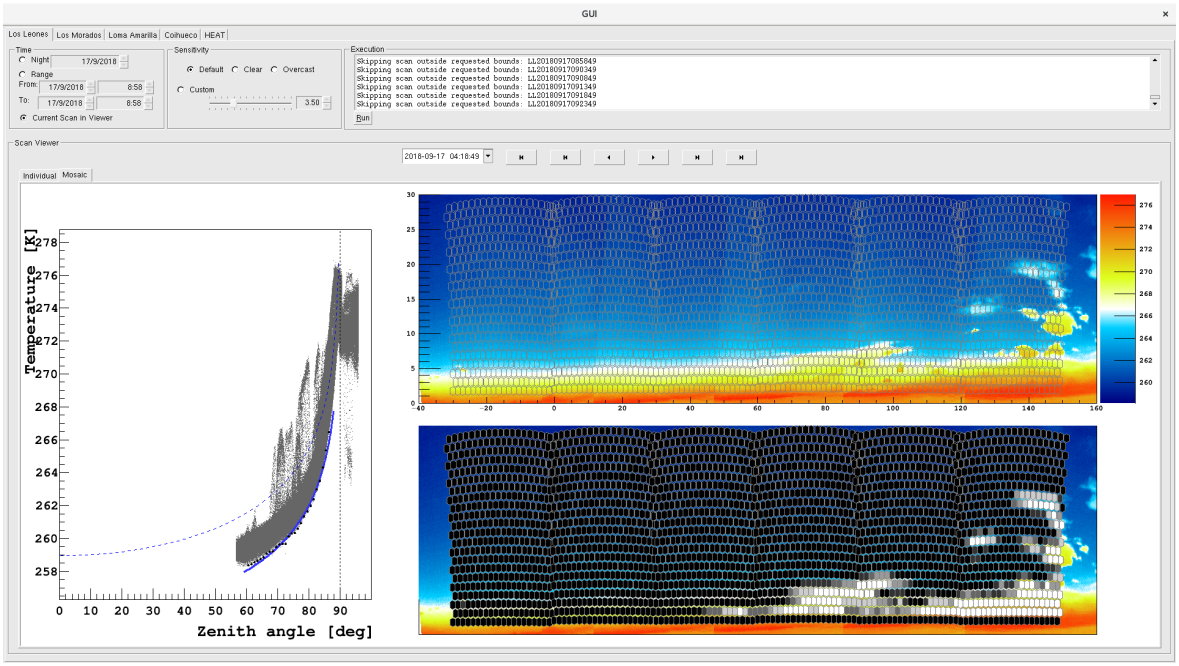

Figure D.7: Screenshot of the Graphical User Interface for the Cloud Mask program.

For each night, all images' cloud fraction values will be put into a cpd file, a text file that contains the site ID, GPS time and the cloud fraction value for each pixel. These cpd files are what is processed into the database, so once a particular day/month/time period is completed, copy all of the cpd files (from output) into a separate directory. It is recommended that this directory is also organised by site and by year, as this makes the entry into the database much simpler. Once all cpd files are ready, contact Bruce Dawson (bruce.dawson@adelaide.edu.au) with the folder that contains all of these cpd files (as an example, it might be contained in "/remote/riposte/jday/cpdfiles/SS" where SS is the two letter site acronym [LL/LM/LA/Co/HE]), so that the cloud masks can be stored into the database for the collaboration to use.

Appendix E

Cloud Height Estimation Technique - example data

This appendix contains the raw data used in the cloud height estimation technique from Chapter 8. There are a few different examples present here, some where the estimated cloud height agrees with the information given by the CLF, and some where it does not. By looking at these individually, one can obtain a better understanding of the reason(s) why the technique can work, and what can be done in cases where it doesn't work.

The first example shown here (Figure E.1) is from the Los Leones cloud camera, where the two cloud height estimation techniques both agree on a height of $\sim 0.8\,\mathrm{km}$, but the CLF height is given as 1.8 km.

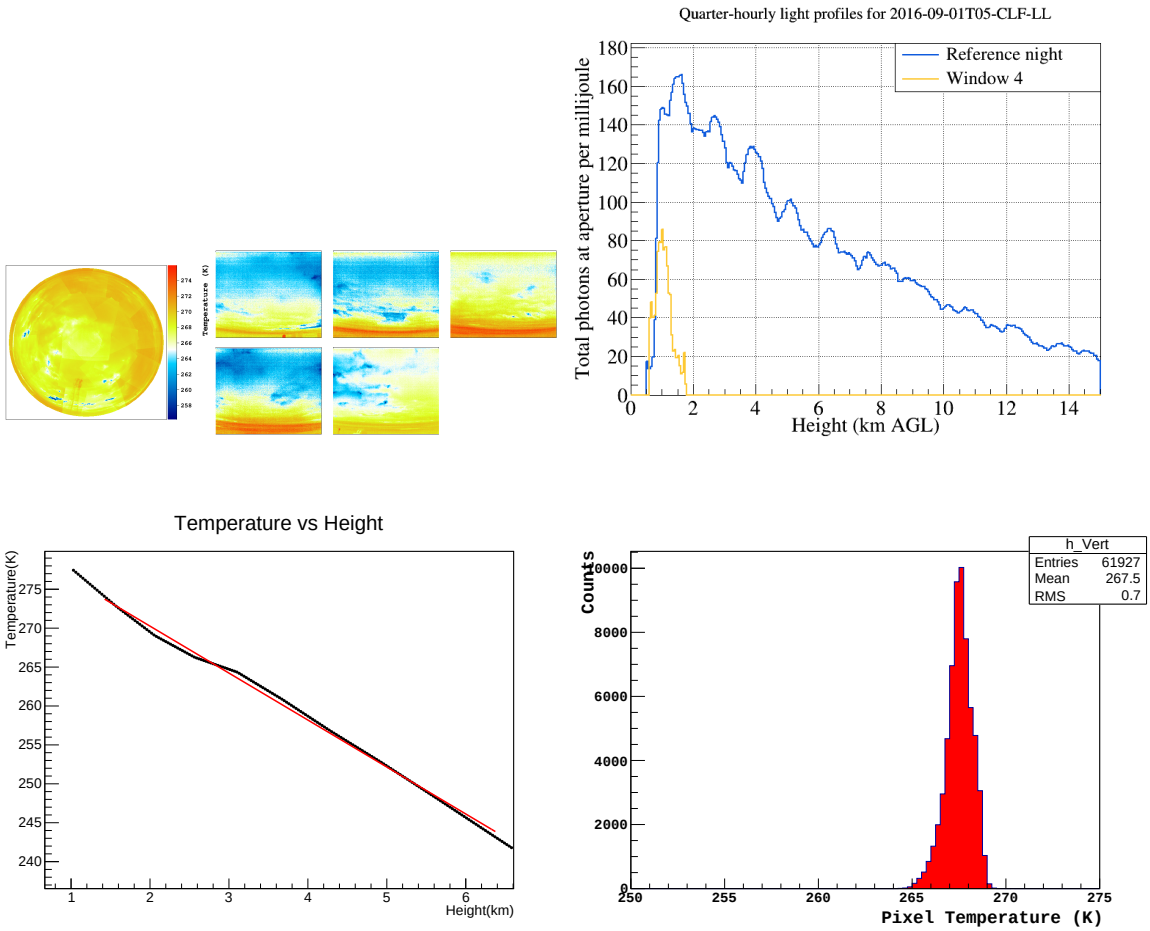

Figure E.1: Top left: Cloud camera image from Los Leones. Top right: Light profile from the CLF for this particular hour. Bottom left: Vertical image pixel temperature distribution. Bottom right: Temperature profile from GDAS for this hour.

Next, Figure E.2 shows an example where the lapse rate fit to the GDAS profile is not very accurate. However, the first peak in the XLF light profile agrees with the heights calculated via both cloud height estimation methods ($\sim 0.5\,\mathrm{km}$), but the second peak (at $\sim 1.4\,\mathrm{km}$) is the height that is compared to. The reason for the first peak not being used is due to the FD not seeing light from the XLF laser at such a low altitude, due to the slight tilt (1.6°) in elevation of the FD pixels. High concentrations of aerosols close to the ground can result in this spike in the brightness of laser light, which is unreliable as an estimate of cloud height.

The next example (Figure E.3 shows a case where the estimated cloud height from the two techniques is low ($< 1500\,\mathrm{m}$) but the height given by the CLF is very high ($> 4\,\mathrm{km}$). The reason for such a high height from the CLF is as discussed in Chapter 8, that the relative difference in brightness of each separate window is large, resulting in a false detection of a cloud layer.

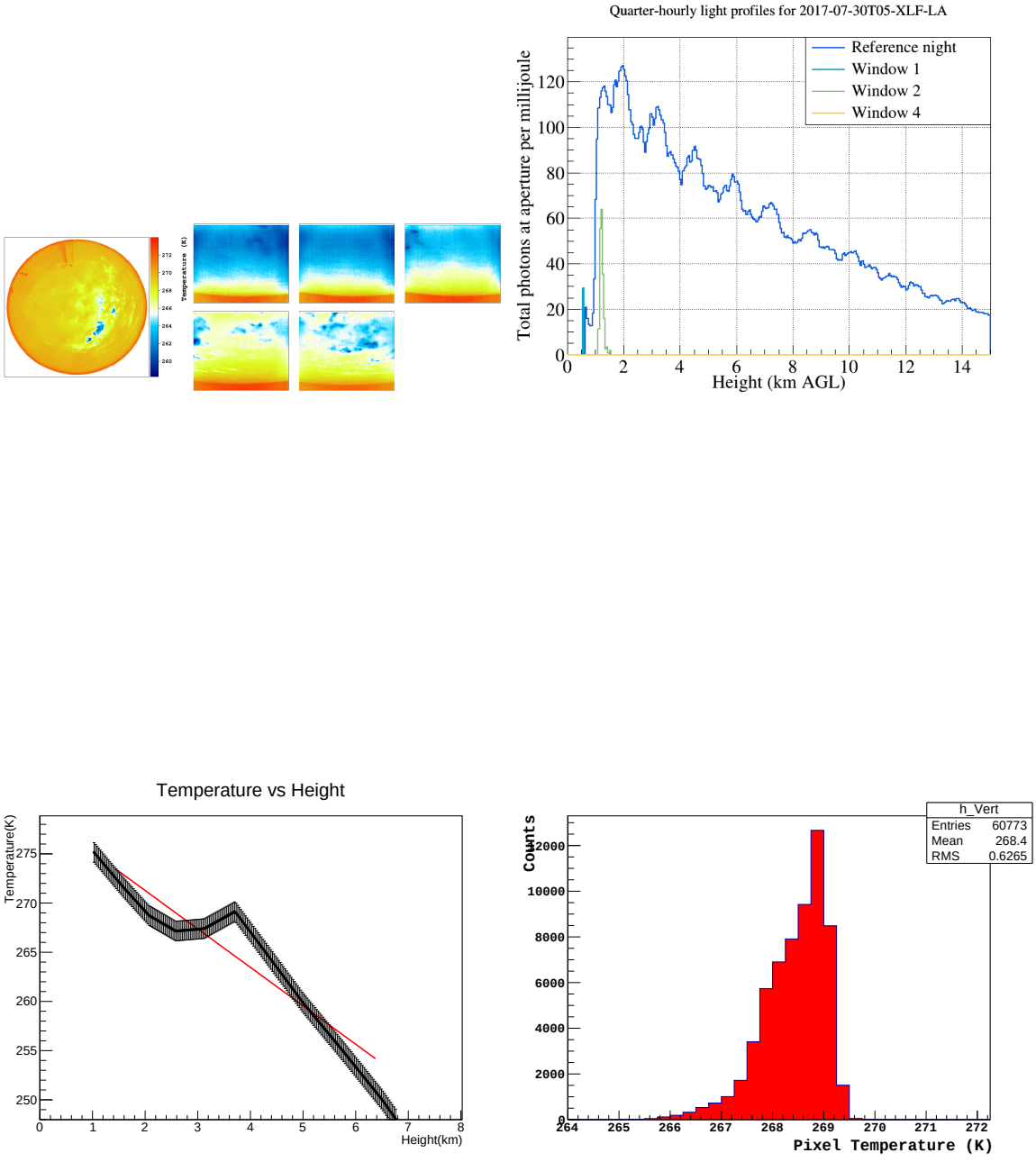

Figure E.2: Top left: Cloud camera image from Loma Amarilla. Top right: Light profile from the CLF for this particular hour. Bottom left: Vertical image pixel temperature distribution. Bottom right: Temperature profile from GDAS for this hour. Height is given as above sea level (A.S.L).

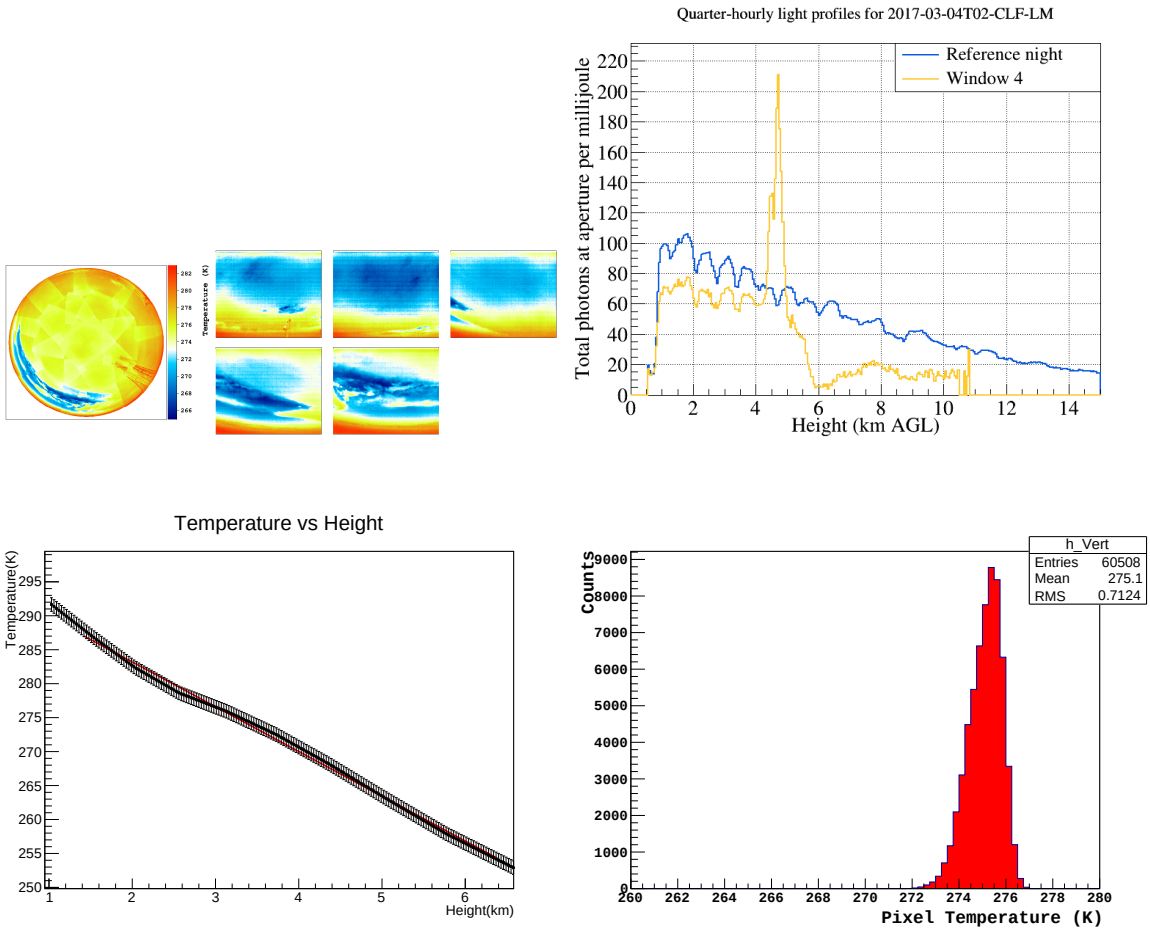

Figure E.3: Top left: Cloud camera image from Los Morados. Top right: Light profile from the CLF for this particular hour. Bottom left: Vertical Image pixel temperature distribution. Bottom right: Temperature profile from GDAS for this hour. Height is given as above sea level (A.S.L).

The following examples are cases when multiple sites view the same height from the CLF. Each of the heights calculated for each site separately agree reasonably well with each other. Firstly, Figure E.4 where three sites (LL,LM and LA) all view a cloud height of $\sim 2.7\,\mathrm{km}$.

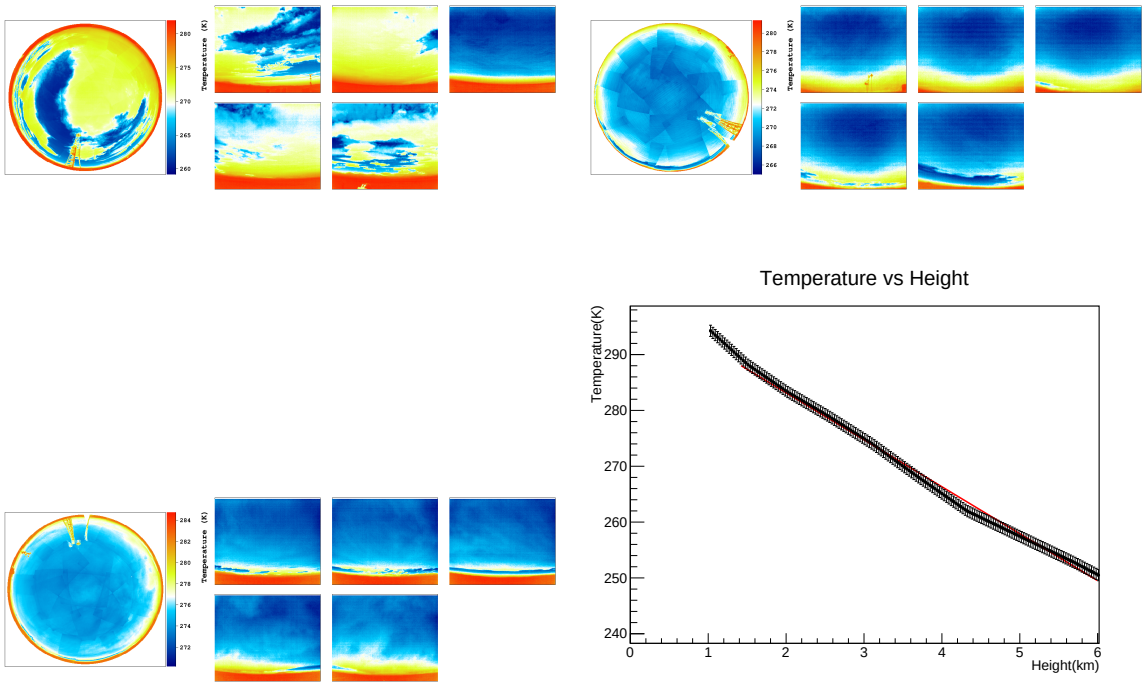

Figure E.4: Cloud camera images from Los Leones (top left), Los Morados (top right) and Loma Amarilla (bottom left). Bottom right: Temperature profile from GDAS for this hour. Height is given as above sea level (A.S.L).

Figure E.5 shows the pixel temperature distributions for each cloud camera image from Figure E.4. Figure E.6 shows the CLF light profiles from each of these sites for the same hour.

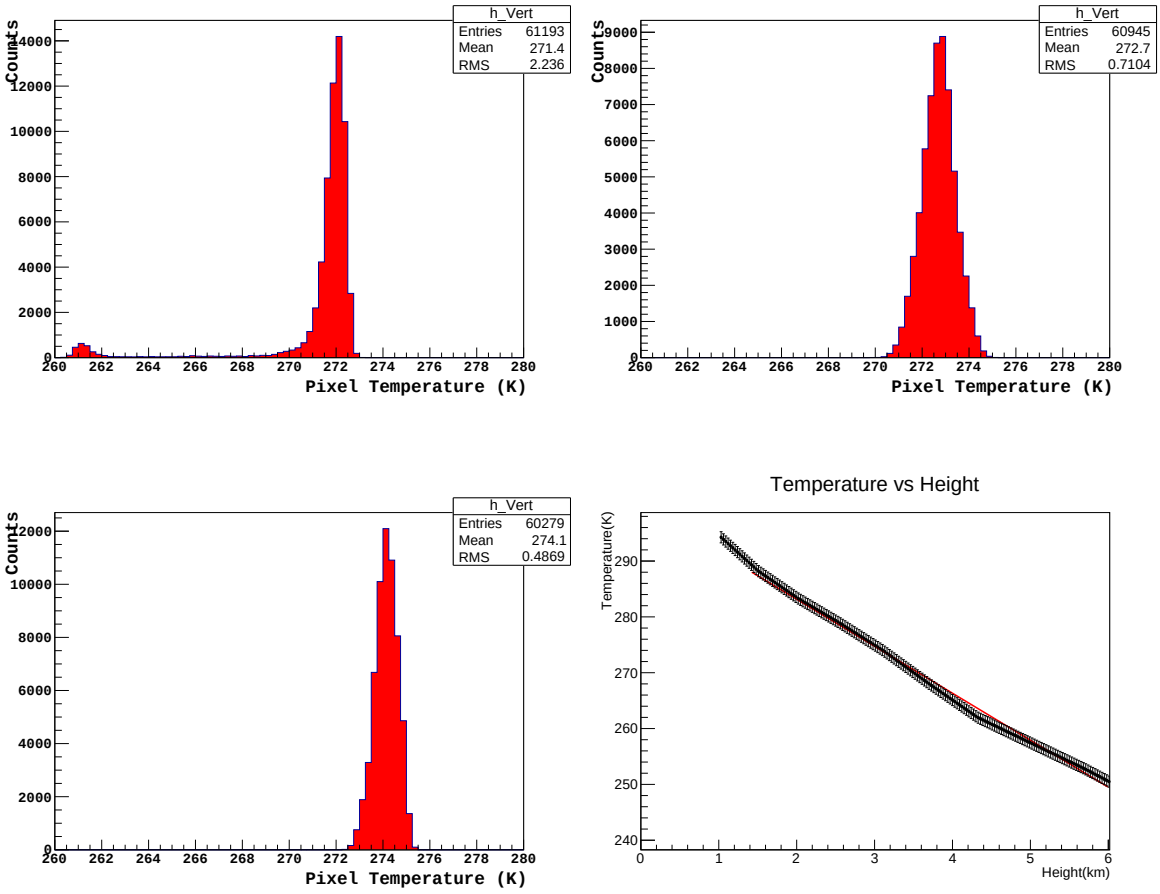

Figure E.5: Vertical image pixel temperature distributions from Los Leones (top left), Los Morados (top right) and Loma Amarilla (bottom left). Bottom right: Temperature profile from GDAS for this hour. Height is given as above sea level (A.S.L).

Upon examination of the light profiles in Figure E.6 it is not clear that the heights given by the CLF/XLF should agree, since each ones lowest peak does not line up [the height from the light profiles that was used, was 2.7 km].

The final example shown here is another case of the CLF and estimated cloud heights agreeing quite well, as shown in Figure E.7.

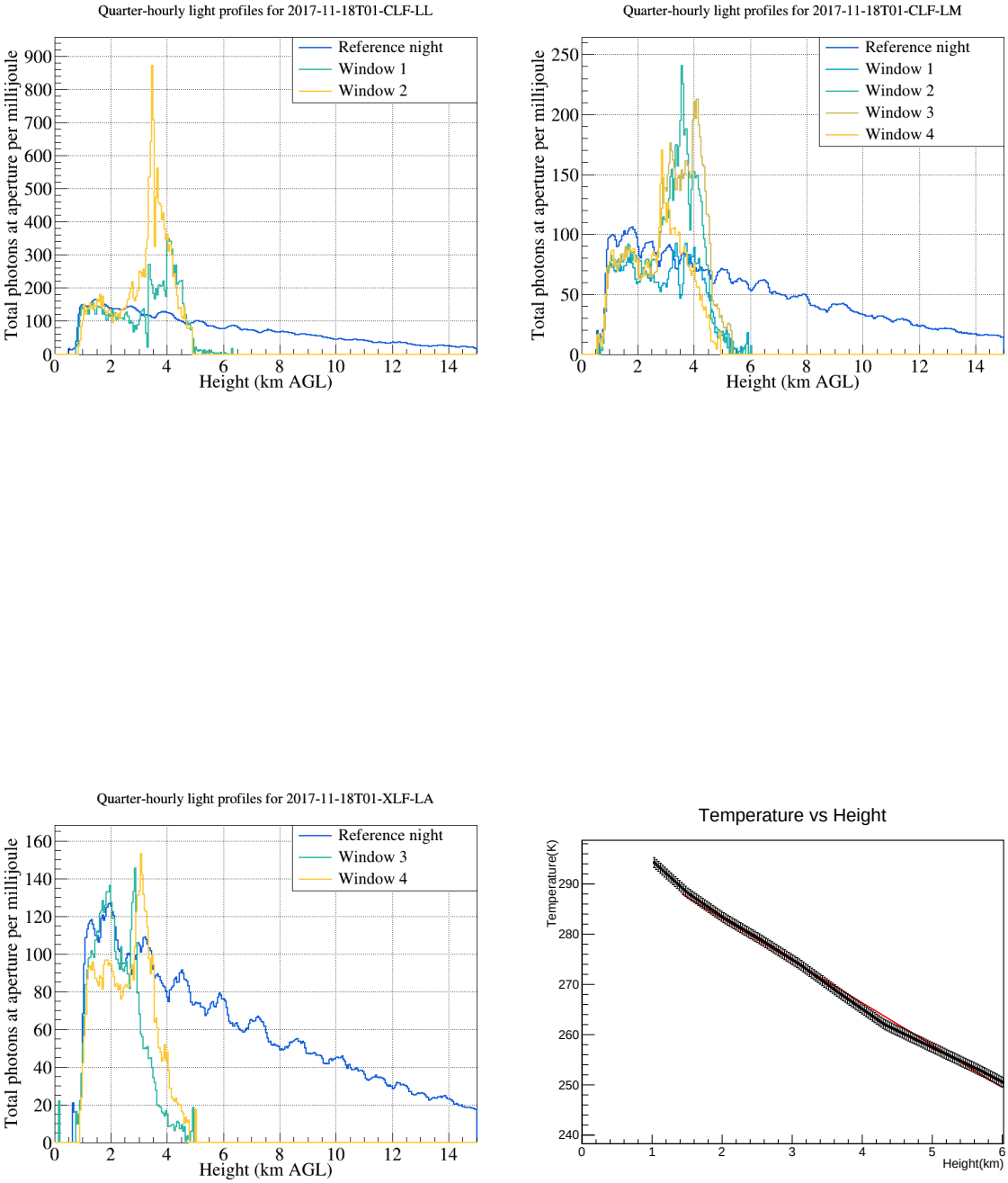

Figure E.6: Light profiles from the CLF for Los Leones (top left), Los Morados (top right) and Loma Amarilla (bottom left) [XLF in this case]. Bottom right: Temperature profile from GDAS for this hour. Height is given as above sea level (A.S.L).

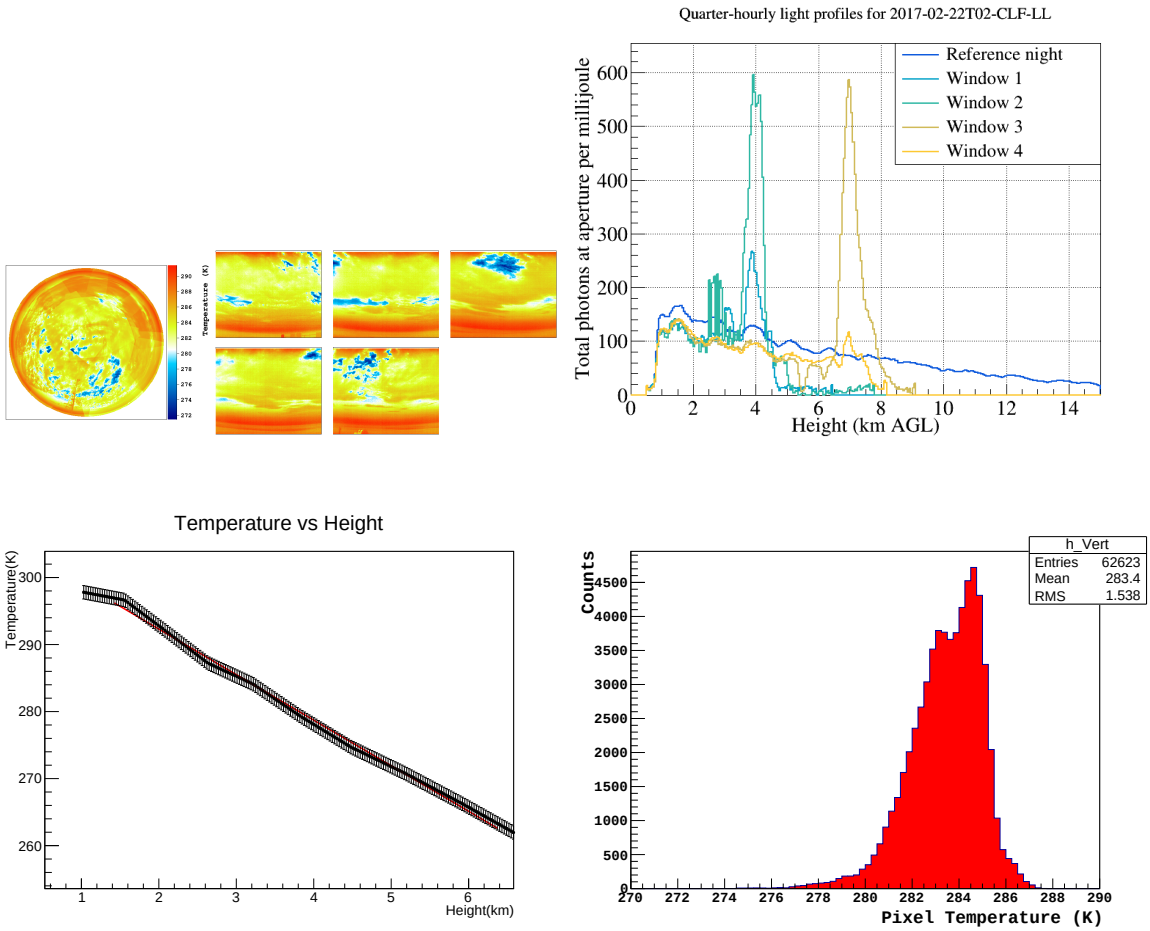

Figure E.7: Top left: Cloud camera image from Los Leones. Top right: Light profile from the CLF for this particular hour. Bottom left: Vertical image pixel temperature distribution. Bottom right: Temperature profile from GDAS for this hour. Height is given as above sea level (A.S.L).

The results of Figure E.7 show that the measured cloud height via the lapse rate method has matched well with the height given by the CLF (the bright peak at around 2.5 km), showing that this technique and its comparison to the CLF can be very accurate.

To summarise all of these case studies, the height given by the CLF/XLF is not always accurate in overcast conditions, as it quite often overestimates the true height of the cloud calculated via the estimation techniques outlined in Chapter 8. The results from those cloud height estimation techniques was shown to be quite accurate in calculating the temperature, hence height, of a cloud imaged by the infrared cloud cameras at Auger. As a result, in the case of overcast conditions, it is more reliable a measure of cloud height, information which is used for much of the extensive air shower reconstruction and cloud quality cuts as described in Chapter 3. Uncertainties on the sky temperature due to the uncertainty on the pixel temperature distributions and the cloud itself not being a perfect blackbody still exist, however, for thick, low, cloud, this technique has been shown to estimate cloud height to within 500 m quite well, an improvement from previous attempts [97] [48].

Bibliography

[1] K. H. Kampert and A. A. Watson. Extensive Air Showers and Ultra High-Energy Cosmic Rays: A Historical Review. *The European Physical Journal H*, 37:359–412, 2012.

[2] C. Grupen. The history of Cosmic Ray Studies after Hess. *Nuclear Physics B*, 239:19–25, 2013.

[3] C. T. R. Wilson. On an Expansion Apparatus for making Visible the Tracks of Ionising particles in Gases and some results obtained by its use. *Proc. R. Soc. Lond. A*, 87:277–292, 1912.

[4] J. Matthews. A Heitler model of Extensive Air Showers. *Astroparticle Physics*, 22:387–397, 2004.

[5] R. Engel. Theory and Phenomenology of Extensive Air Showers. http://moriond.in2p3.fr/J05/trans/sunday/engel1.pdf. Last accessed August 7th, 2020.

[6] The Pierre Auger Collaboration. Muons in air showers at the Pierre Auger Observatory: Measurement of atmospheric production depth. *Phys. Rev.*, 90:5–10, 2014.

[7] G. Matthiae. The cosmic ray energy spectrum as measured using the Pierre Auger Observatory. *New Journal of Physics*, 12:075009, 2010.

[8] C. Amsler et al. Review of Particle Physics. *Physics Review Letters*, 1:254–260, 2008.

[9] F. G. Schroder. Cosmic Ray Indirect Rapporteur. *Proceedings of Science (ICRC2019) 030*, 1:1–20, 2019.

[10] T. Beau E. Armengaud, G. Sigl and F. Miniati. CRPropa: A numerical tool for the Propagation of UHE Cosmic Rays, Gamma Rays and Neutrinos. *Astroparticle Physics*, 28:463–471, 2006.

[11] A. Letessier-Selvon. Theoretical and Experimental Topics on Ultra High Energy Cosmic Rays. *International Journal of Advanced Engineering Research and Science*, 6:1–46, 2000.

[12] M. Settimo for the Pierre Auger Collaboration. Measurement of the mass composition of ultra-high energy cosmic rays with the Pierre Auger Observatory. *Journal of Physics*, 718:37–52, 2016.

[13] G. Morlino. High Energy Cosmic Rays from Supernovae. *Handbook of Supernovae*, pages 1711–1736, 2017.

[14] The Pierre Auger Collaboration. Evidence for a mixed mass composition at the 'ankle' in the cosmic-ray spectrum. *Physics Letters*, 762:288–295, 2016.

[15] The Pierre Auger Collaboration. Limits on point-like sources of ultra-high-energy neutrinos with the Pierre Auger Observatory. *Journal of Cosmology and Astroparticle Physics*, 11:4–28, 2019.

[16] The Pierre Auger Collaboration. Indication of anisotropy in arrival directions of ultra-high-energy cosmic rays through comparison to the flux pattern of extragalactic gamma-ray sources. *The Astrophysical Journal Letters*, 853:L29, 2018.

[17] The Pierre Auger Collaboration. Observation of a Large-scale Anisotropy in the Arrival Directions of Cosmic Rays above 8×10^{18} eV. *Science*, 357:1266–1270, 2017.

[18] DESY. Multi messenger astrophysics. https://astro.desy.de/theory/multi_messenger_astrophysics/index_eng.html. Last Accessed December 1st, 2019.

[19] The HESS Collaboration. Acceleration of petaelectronvolt protons in the Galactic Centre. *Nature*, 531:476–479, 2016.

[20] M. G. Baring. Diffusive Shock Acceleration: the Fermi Mechanism. *Very High Energy Phenomena in the Universe*, 1:1–10, 1997.

[21] The Pierre Auger Collaboration. The Pierre Auger Cosmic Ray Observatory. *Nuclear Instruments and Methods in Physics Research A*, 798:172–213, 2015.

[22] AIRFLY Collaboration. Measurement of the pressure dependence of air fluorescence emission induced by electrons. *Astroparticle Physics*, 28:41–57, 2007.

[23] B. Keilhauer et. al. Nitrogen fluorescence in air for observing extensive air showers. *EPJ Web of Conferences*, 53:01010, 2013.

[24] M.G. Winnick. Cloud Cameras at the Pierre Auger Observatory. *PhD Thesis, University of Adelaide*, 2010.

[25] J. Abraham for the Pierre Auger Collaboration. The Fluroescence Detector of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics A*, 620:227–251, 2010.

[26] L. Perrone for the Pierre Auger Collaboration. Auger FD: Detector Response to Simulated Showers and Real Event Topologies. *Nuclear Physics B.*, 136:407–414, 2004.

[27] N. Scharf and C. Meurer for the Pierre Auger Collaboration. HEAT - a low energy enhancement of the Pierre Auger Observatory. *Instrumentation and Methods for Astrophysics*, 7:183–186, 2011.

[28] J. Bellido for the Pierre Auger Collaboration. Depth of maximum of air-shower profiles at the Pierre Auger Observatory: Measurements above $10^{17.2}$ eV and Composition Implications. *Proceedings of Science*, 1:1–8, 2017.

[29] J Vicha and J. Chudoba. Data Processing at the Pierre Auger Observatory. *J. Phys. Conf. Ser.*, 608:2, 2015.

[30] A. Castellina for the Pierre Auger Collaboration. AugerPrime: the Pierre Auger Observatory Upgrade. *EPJ Web of Conferences*, 210:06002, 2019.

[31] M. Platino et al. AMIGA at the Auger Observatory: the telecommunications system. *Journal of Instrumentation*, 8:1–12, 2013.

[32] C. Glaser. Results and Perspectives of the Auger Engineering Radio Array. *EPJ Web Conf.*, 135:1–6, 2016.

[33] F. G. Schroder. Radio detection of high-energy cosmic rays with the Auger Engineering Radio Array. *Nuclear Instruments and Methods in Physics Research Section A*, 824:648–651, 2016.

[34] T. Fodran. Cosmic Ray Indirect Rapporteur. *Proceedings of Science (ICRC2021) 037*, 1:1–20, 2021.

[35] The Pierre Auger Collaboration. A study of the effect of molecular and aerosol conditions in the atmosphere on air fluorescence measurements at the Pierre Auger Observatory. *Astroparticle Physics*, 33:108–129, 2011.

[36] NOAA (National Oceanographic and Atmospheric Administration). GDAS Data Archive. https://www.ready.noaa.gov/gdas1.php. Last Accessed May 12th, 2018.

[37] The Pierre Auger Collaboration. Description of Atmospheric Conditions at the Pierre Auger Observatory using the Global Data Assimilation System (GDAS). *Astroparticle Physics*, 35:591–607, 2011.

[38] B. Fick et al. The Central Laser Facility at the Pierre Auger Observatory. *Journal of Instrumentation*, 1:11–31, 2006.

[39] Violet M. Harvey for the Pierre Auger Collaboration. Real Time Measurements with Atmospheric Instruments at the Pierre Auger Observatory. *Proceedings for the 36th International Cosmic Ray Conference*, 2019.

[40] T. Storelvmo. Aerosol Effects on Climate via Mixed-Phase and Ice Clouds. *Annual Review of Earth and Planetary Sciences*, 45:199–222, 2017.

[41] J. Blazek for the Pierre Auger Collaboration. The FRAM Telescope at the Pierre Auger Observatory. *Proceedings for the 35th International Cosmic Ray Conference*, 1:201–208, 2017.

[42] E. Hog et. al. The Tycho2 catalogue of the 2.5 million brightest stars. *Astronomy and Astrophysics*, 355:27–30, 2000.

[43] M. Iarlori V. Rizi, A. Tonachini and G. Visconti for the Pierre Auger Collaboration. Atmospheric monitoring with LIDARs at the Pierre Auger Observatory. *European Physical Journal*, 127:92–103, 2007.

[44] S.Y. BenZvi et al for the Pierre Auger Collaboration. The LIDAR system of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research*, 574:171–184, 2007.

[45] D. Hauf. Geostationary Operational Environmental Satellite R-Series. https://www.lockheedmartin.com/en-us/products/goes-r-series.html. Last accessed December 12th, 2019.

[46] The Pierre Auger Collaboration. Identifying Clouds over the Pierre Auger Observatory using Infrared Satellite Data. *Astropart. Phys.*, 50:92–101, 2013.

[47] Xenics Infrared Solutions. Smart and affordable: Xenics gobi-384. https://www.xenics.com/press-release/smart-and-affordable-xenics-gobi-384/. Last Accessed December 19th, 2020.

[48] Patrick van Bodegom. Remote Sensing of Clouds with Longwave Infrared Cameras at the Pierre Auger Observatory. *PhD Thesis, University of Adelaide*, 2018.

[49] National Oceanographic and Atmospheric Administration. Layers of the Atmosphere. https://www.weather.gov/jetstream/layers. Last Accessed July 21st, 2020.

[50] T. Xian and C. R. Homeyer. Global tropopause altitudes in radiosondes and reanalyses. *Atmospheric Chemistry and Physics*, 19:5661–5678, 2019.

[51] D. H. McIntosh and A. S. Thom. Essentials of Meteorology. *Wykeham Publications London*, 1969.

[52] Met Office. What is a Global circulation pattern? <https://www.internetgeography.net/topics/what-is-global-atmospheric-circulation/>. Last accessed November 22nd, 2020.

[53] S. Fueglistaler et al. Tropoical Tropopause layer. *Reviews of Geophysics*, 47:1–31, 2009.

[54] B. D. Santer et al. Contributions of Anthropogenic and natural forcing to recent tropopause height changes. *Science*, 301:479–483, 2003.

[55] G. I. Golodets. Chapter VII: The Dissociation of Molecular Oxygen, The Recombination of Oxygen Atoms and Isotopic Exchange with Molecular Oxygen. *Heterogeneous Catalytic Reactions Involving Molecular Oxygen*, 15:180–199, 1983.

[56] NASA Earth Observatory. Ozone balance in the Stratosphere. https://earthobservatory.nasa.gov/features/Ozone. Last Accessed September 3rd, 2020.

[57] J. S. Daniel R. W. Portmann and A. R. Ravishankaa. Stratospheric ozone depletion due to nitrous oxide: influences of other gases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367:1256–1264, 2012.

[58] S. Chapman. On ozone and atomic oxygen in the upper atmosphere. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10:369–383, 1930.

[59] M. Hamilton. Cloud Microphysics. *Advanced Atmospheric Physics - Lecture Notes*, 2016.

[60] D. Barahona. Common Cloud Names, Shapes and Altitudes. http://nenes.eas.gatech.edu/Cloud/Clouds.pdf. Last Accessed February 8th, 2019.

[61] Stack Overflow. Why is the Cumulonimbus cloud formation so dangerous? https://aviation.stackexchange.com/questions/16700/why-is-the-cumulonimbus-cloud-formation-so-dangerous. Last Accessed March 5th, 2020.

[62] U. Lohmann and B. Gasparini. A cirrus cloud climate dial? *Science*, 357:248–249, 2017.

[63] A. J. Baran. On the scattering and absorption properties of cirrus cloud. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 89:17–36, 2004.

[64] D. Cziczo et al. Clarifying the Dominant Sources and Mechanisms of Cirrus Cloud Formation. *Science*, 340:1320–1324, 2013.

[65] M. B. Baker. Cloud Microphysics and Climate. *Science*, 276:1072–1078, 1997.

[66] J. Houghton. *The Physics of Atmospheres*. Cambridge University Press, 1986.

[67] R. R. Rogers and M. K. Yau. *A short course in Cloud Physics*. Butterworth Heinemann, 1989.

[68] J. H. Shaw R. Sloan and D. Williams. Infrared Emission Spectrum of the Atmosphere. *Journal of the Optical Society of America*, 45:455–460, 1955.

[69] R. West. Modelling Cloud Droplet Formation. *University of Oxford*, 1:3–10, 2008.

[70] J. R. Allen. Measurements of Cloud Emissivity in the 8-13 μm waveband. *Journal of Applied Meteorology*, 10:260–265, 1971.

[71] J. T. Kiehl and K. E. Trenberth. Earth's Annual Global Mean Energy Budget. *National Center for Atmospheric Research, Boulder, Colorado*, 78:197–208, 1997.

[72] Michael D. King. Radiative Properties of Clouds. *International Geophysics*, 54:123–149, 1993.

[73] KP Shine. Radiative Forcing of Climate Change. *Space Science Reviews*, 94:363–373, 2000.

[74] National Snow and Ice Data Centre. Thermodynamics: Albedo. /cryosphere/seaice/index.html. Last accessed June 14th, 2019.

[75] A. A. Sinkevich and R. P. Lawson. A Survey of Temperature Measurements in Convective Clouds. *Journal of Applied Meteorology*, 44:1133–1145, 2005.

[76] J. Su et. al. Cloud temperature measurement using rotational Raman lidar. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 125:45–50, 2013.

[77] J. Abraham et. al for the Pierre Auger Collaboration. A Study of the Effect of Molecular and Aerosol Conditions in the Atmosphere on Air Fluorescence Measurements at the Pierre Auger Observatory. *Astroparticle Physics*, 33:108–129, 2010.

[78] R. Stull G. West and R. Howard. Temperature Inversions. https://www.eoas.ubc.ca/courses/atsc113/snow/met_concepts/06-met_concepts/06c-inversions/index.html. Last Accessed January 7th, 2021.

[79] J. Zhang et al. Analysis of cloud layer structure in Shouxian, China using RS92 radiosonde aided by 95 GHz cloud radar. *Journal of Geophysical Research*, 115:1–13, 2010.

[80] L. Lusted. Signal Detectability and Medical Decision-Making. *Science*, 171:1217–1219, 1971.

[81] Air Resources Laboratory. Ready archived meteorology. http://ready.arl.noaa.gov/READYamet.php. Last Accessed April 20th, 2020.

[82] J.L. Marshall and D.L. DePoy. Flattening scientific CCD imaging data with a dome flat field system. *ArXiv Astrophysics e-prints*, 2005.

[83] C. Hinnah P. Kask, K. Palo and T. Pommerencke. Flat field correction for high-throughput imaging of fluorescent samples. *Journal of Microscopy*, 263:328–340, 2016.

[84] A. Maghrabi. Ground based measurements of atmospheric infrared radiation from clear and cloudy skies. *PhD Thesis, University of Adelaide*, 2007.

[85] B. A. Myers. User Interface Software Tools. *ACM Transactions on Computer-Human Interaction*, 2:64–103, 1995.

[86] P.K. Acharya A. Berk, G.P. Anderson and E.P. Shettle. MODTRAN 5.2.0.0 user's manual. July 2008.

[87] A. Berk et. al. Band Model Method for Modelling Atmospheric Propagation at Arbitrarily Fine Spectral Resolution. *U.S. Patent*, 2008.

[88] P. Conforti A. Berk and F. Hawes. An accelerated line-by-line option for MODTRAN combining on-the-fly generation of line centre absorption with 0.1 cm-1 bins and pre-computed line tails. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI*, 9472:17, 2015.

[89] A. Berk et. al. MODTRAN6: a major upgrade of the MODTRAN radiative transfer code. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, 9088:7, 2014.

[90] D.K. Lynch and S. Mazuk. New Cirrus Models for MODTRAN4. 2002.

[91] S. Ellingson. *Electromagnetics*. Virginia Tech Libraries, 2018.

[92] T. Posch M. Kocifaj and H.A Solano Lamphar. On the relation between zenith sky brightness and horizontal illuminance. *Monthly Notices of the Royal Astronomical Society*, 446:2895–2901, 2015.

[93] University of Arizona Department of Physics. Lecture 9 - Langley method, Airmass and solar ephemeris. https://uweb.engr.arizona.edu/~ece583/lecture_9_10_quad.pdf. Last Accessed March 22nd, 2021.

[94] K. P. Mollman and M. Vollmer. Infrared Thermal Imaging: Fundamentals, Research and Applications. *Wiley*, 2010.

[95] Tydex. Germanium data sheet. http://www.tydexoptics.com/materials1/for_transmission_optics/germanium/. Last Accessed October 23rd, 2020.

[96] T. Aoki M. Hori and T. Tanikawa et. al. In-situ measured spectral directional emissivity of snow and ice in the 8-14 μm atmospheric window. *Remote Sensing of Environment*, 100:486–502, 2006.

[97] D. Riordan et. al. Cloud Base temperature measurements using a simple longwave infrared cloud detection system. *Journal of Geophysical Research*, 110:1–9, 2005.

[98] R. S. Vose I. Durre and D. B. Wuertz. Overview of the Integrated Global Radiosonde Archive. *Journal of Climate*, 19:53–68, 2006.

[99] P. Green et. al. Recent advances in measurement of the water vapour continuum in the far-infrared spectral region. *Philosophical Transactions of the Royal Society*, 370:2637–2655, 2012.

[100] S. Zhang Q. Zhou, H. Liu and X. Deng. Sensitivity analyses of precipitable water vapour retrieval from the ground-based infrared measurements in clear sky conditions. *Journal of Applied Remote Sensing*, 13:4–19, 2019.

[101] F. Kneizys et. al. Atmospheric Transmittance/Radiance: C*Z Computer Code LOWTRAN 6. 1:64–67, 1983.

[102] E. F. Lowry. The Infrared Absorption Spectrum of Carbon Monoxide. *Journal of the Optical Society of America*, 8:647–658, 1924.

[103] O. A. Alduchov and R. E. Eskridge. Improved Magnus' form approximation of saturation vapour pressure. *Journal of Applied Meteorology*, 35:601–609, 1996.

[104] Philip R Bevington and D Keith Robinson. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, New York, 2003.