VRIJE
UNIVERSITEIT
BRUSSEL

Graduation thesis submitted in partial fulfilment of the requirements for the master's degree of Physics and Astronomy

THE PERFORMANCE OF A NOVEL CALIBRATION METHOD FOR THE AUGERPRIME SURFACE DETECTOR SIGNALS AND THE ULTRA-HIGH-ENERGETIC COSMIC-RAY SPECTRA ON & OFF THE GALACTIC PLANE

Katarína Simkova

August 2022

Promotors: Prof. Dr. Ioana Codrina Mariș (ULB) Prof. Dr. Stijn Buitink (VUB)
Sciences and Bioengineering Sciences

Abstract

Pierre Auger Observatory is the world's largest experiment for measuring ultra-high-energy cosmic rays (UHECRs) and is currently expected to take data at least until 2030. An upgrade of the Observatory, the so-called AugerPrime, started in 2016 and is still ongoing. The Surface Detector (SD) array was enhanced by mounting Surface Scintillator Detector (SSD) on almost every Water Cherenkov Detector (WCD), to measure the particles of the extensive air showers reaching the ground. The extensive air showers are produced in the atmosphere by UHECR primaries.

Besides the enhancement on the sensitivity to the nature of the UHECRs, the new detector part enables an improvement in the calibration for WCD: requiring coincidence between WCD and SSD suppresses the electromagnetic background in the calibration histograms and records primarily atmospheric muons. This suppression, combined with the different detector geometry when requiring coincidence triggers, results in relative difference Δ between muon peak in currently-used calibration of WCD, and the coincidence calibration of WCD. We have analysed data of the tests on 154 SD stations and determined the expectation values of these differences: $\langle\Delta\rangle_{1,3} = 2.51 \pm 0.08$, $\langle\Delta\rangle_2 = 6.62 \pm 0.13$ for PMT 1 & 3, and for PMT 2, respectively. The correlation between value of Δ and the measure of a station aging was observed. The larger the electromagnetic background, the lower the charge of the apparent muon hump. Tools for event signals analysis using two different calibrations were developed and it was shown that the coincidence calibration performs at least as well as the usual calibration. The study has been partly limited by the available statistics.

The last part of this work focuses on checking possible sky differences on the flux of UHECRs in different sky regions. The latest results of the Pierre Auger Collaboration using only Fluorescence Detector data hint at heavier composition of UHECRs coming from on the Galactic plane compared to off the Galactic plane with post-penalisation significance of $3.3\,\sigma$. We have used the events measured with the SD, more than $300,000$ above $2.51\,\mathrm{EeV}$ to check if this sky distribution, combined with the evolution of the mass composition with energy can produce visible effects on the SD data. The SD data are split to on- & off-plane regions and the ratio of the corresponding fluxes is determined. The composition of UHECR gets heavier at higher energies. To confirm the heavier on-plane composition, the increasing trend in the ratio (on-/off-plane) is expected. The trend is not observed in this work, however the statistics presented allows for the trend of the ratio to be at $2.5\,\%$ at the energies of $32\,\mathrm{EeV}$.

Contents

Chapter 1

Introduction

Ultra-high-energy cosmic rays (UHECRs) are the most energetic particles in the Universe, reaching energies of more than 10^{20} eV. Such energy, corresponding to a tennis ball flying at 100 km/h, is concentrated into a single proton or a heavier nucleus. Comparing with the center-of-mass energy in the largest human-made accelerator - the Large Hadron Collider, the energy is three orders of magnitude higher. As such, UHECRs are an intriguing probe of the Standard Model of Particle Physics at these extreme momenta.

From astrophysical point of view, UHECRs must get accelerated in the most violent places in the Universe. What are their sources and acceleration mechanisms is still the subject of very active research. The complications of answering the question are two-fold: firstly, UHECRs suffer from propagation effects and secondly, their flux at Earth is very low.

To resolve the issue of the flux, one needs an observatory with very large aperture. Pierre Auger Observatory is the world's largest UHECR observatory covering the surface of 3000 km^2. The second part of this chapter will describe its layout, observation techniques and its upgrade.

1.1 Challenges and Basics of UHECR Observations

Starting with the challenges of UHECR observations, let us first look at the propagation of UHECRs towards us. UHECRs are charged particles. Hence, their trajectories bend in the galactic and extragalactic magnetic fields. These fields are another subject of active astrophysical research and their models are still under development [1]. Tracking of the UHECR trajectories starting from their arrival directions on Earth back to their sources is thus not possible. However, at the UHECR energies of more than 10^{18} eV, the trajectories bend roughly by 3-10° only, ranging from proton to iron. Thanks to this property, they offer an obvious advantage if looking for their sources, compared to cosmic rays at lower energies [2].

In addition to trajectory distortion, another of propagation effects is when UHECRs interact with other particles in their way, for instance cosmic microwave background. In such case, some of the UHECR energy is deposited into new particle creation resulting in UHECR energy losses.

The second major difficulty arising in UHECR observations is their low flux. Above 10^{10} eV, the spectrum of cosmic rays, i.e. their number as the function of their energy $\mathrm{d}N/\mathrm{d}E$ decreases as a power law E^{-b} with $b = 2.7$ until the knee at 10^{15} eV, $b = 3.1$ up to the ankle at 10^{18} eV, and then $b = 2.6$ until the cutoff at 10^{20} eV [2]. The spectrum and its features can be seen in Fig. 1.1. Therefore, the higher the energy of cosmic rays, the more rare they become. At the energies of 10^{19} eV, their flux is approximately 1 particle/km^2/century. It is obvious that one requires a very large observatory to obtain enough statistics in their measurements.

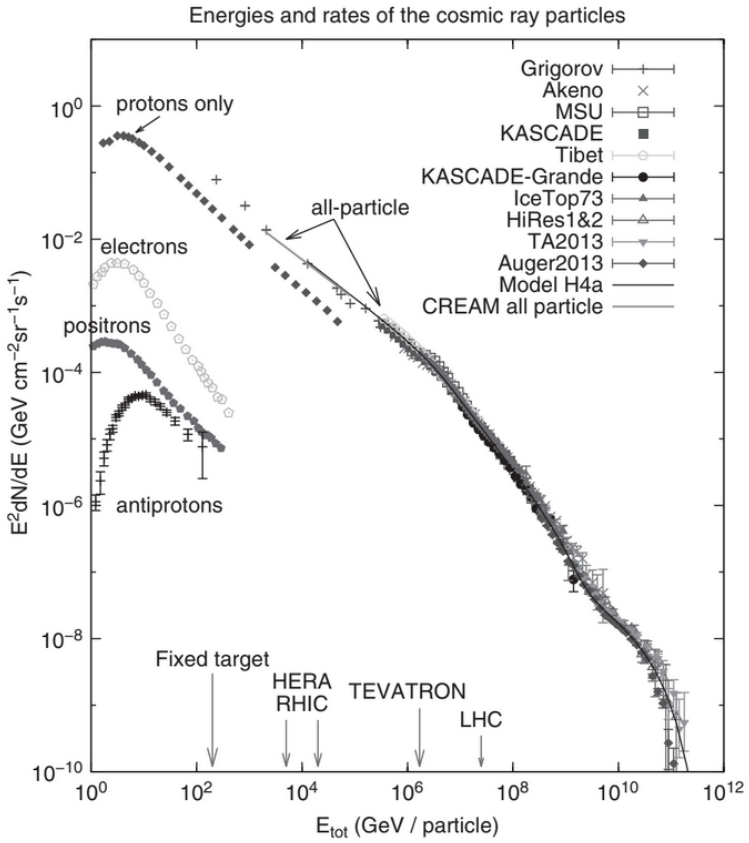


Figure 1.1: Figure of the cosmic-ray spectrum taken from [2]. The differential flux $\mathrm{d}N/\mathrm{d}E$ is multiplied by the square of energy to better distinguish the features of the spectrum. Above 10 GeV, the spectrum is well-described by a power law with exponents 2.7 (up to 10^{15} eV), 3.1 (up to 10^{18} eV), and 2.6 up to the cutoff at 10^{20} eV. The center-of-mass energy comparison with the human-made accelerators is also shown.

The knee of the spectrum can be explained by the upper limit of possible acceleration mechanism in the supernova remnants of our Galaxy. The spectrum then falls off more rapidly due to fewer sources in the Galaxy, while at the ankle, extragalactic sources are believed to come into play due to lack of small-scale anisotropy in these UHECRs, and due to the dipole measurement with the excess not on the plane of Milky Way [3]. Going above 10^{18} eV therefore offers us observations of more interesting, extragalactic messengers.

Chapter 3 of this work focuses on the analysis of the highest-energetic UHECRs comparing flux from two different parts of the sky above $10^{18.4}$ eV. The lower-energetic cosmic rays in the range of $10^9 - 10^{14}$ eV are also worth a notice however, thanks to their stable flux of roughly a particle per m^2 per second. As such, they are useful tool for calibration of particle detectors, which is the topic of Chapter 2 of this work.

How can we measure these cosmic bullets, namely their arrival direction, energy and mass? Fortunately, the Earth's atmosphere is our shield from them, while it also enables multiple means

Figure 1.2: Figure to illustrate an extensive air shower (EAS) developing in the atmosphere. EAS cannot be seen by bare eye, the colours are only illustrative. Red represents the electromagnetic component of the EAS, while dark particles around the center are hadrons. Resource of the figure: [4].

of measurements. When an UHECR hits the atmosphere, a hadronic cascade of secondary particles called an extensive air shower (EAS) is initiated, as illustrated by Fig. 1.2. The characteristics of the EAS can tell us about the energy and the mass of the primary UHECR. Firstly, the total number of secondary particles in the shower as well as the extent of its footprint is proportional to the primary energy. Secondly, the lighter the UHECR, the deeper in the atmosphere it will interact. One of the crucial variables when observing UHECRs, X_{max}, is the depth of the atmosphere traversed first by the primary UHECR and continued by the secondary particles in the EAS, where the number of particles in the EAS is at maximum. Lastly, EAS are composed of hadrons close to the core, of muons, and of electromagnetic (EM) component encapsulating photons, electrons and positrons. This composition, especially the number of muons in the EAS is another variable that informs us about the mass of the primary UHECR.

Naturally, having large amounts of particles in the EAS makes it much easier to measure UHECRs than trying to catch a single particle. It is safe to create sparse arrays sampling particles of the EAS on the ground level, as footprint radii of these showers extend to tens of kilometers. With sparse arrays, large areas can be covered and therefore a large exposure accumulated to make up for the low flux. This was the idea behind the Pierre Auger Observatory, the world's leading facility to measure UHECRs.

1.2 Pierre Auger Observatory

The Pierre Auger Observatory is the world's largest observatory intended to measure EAS both during their development in the atmosphere and when they reach the ground. Its extent is $3000\,\mathrm{km}^2$, with roughly 60 km across, as seen in the Fig. 1.3 showing its layout. The Observatory is at the altitude of 1340-1610 m above sea level [5], located on Pampa Amarilla near the city Malargüe in the province of Mendoza, Argentina. The location was chosen carefully so that the SDs are on a plane with little vegetation and in the altitude, where the EAS started by

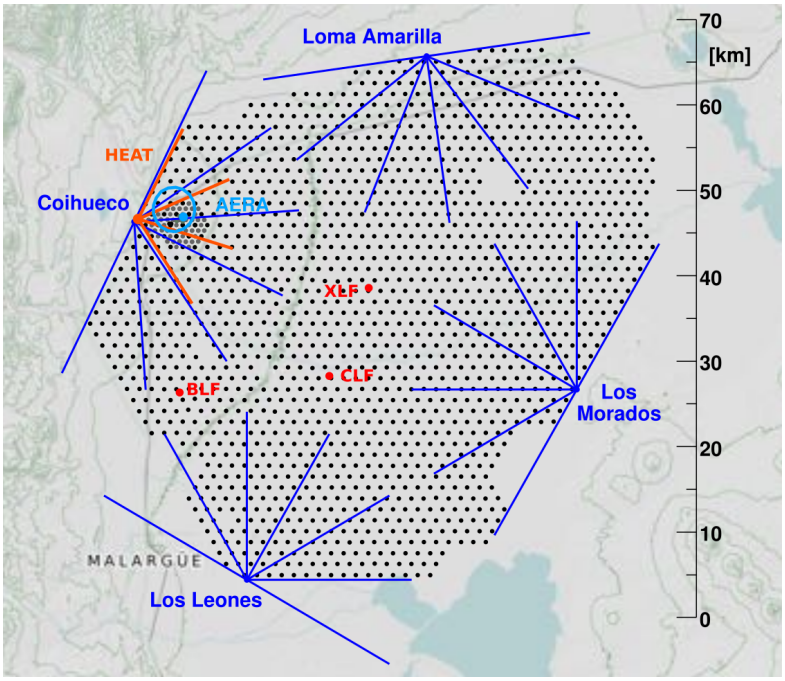


Figure 1.3: Figure of the Pierre Auger Observatory layout taken from [6]. Black dots represent one SD station each. Blue dots show FD sites with FD eye field of view drawn as blue lines. Red dot and lines show HEAT telescopes. Other red dots denote locations of laser facilities used for atmosphere monitoring. AERA is the site of engineering array for radio detection of EAS.

UHECRs of energies above 10^{18} eV are in the stage of their development most suitable for their observation, having passed the vertical depth of $\approx 890\,\mathrm{g\,cm^2}$.

The Observatory has two main components: the surface detector (SD) array and the fluorescence detectors (FDs). SD consists of 1664 SD stations creating the array having triangular cells of the 1500 m side-length. In Fig. 1.3, each black dot represents one SD station. FDs are stationed in groups of 6 at four locations denoted by blue dots in Fig. 1.3 around the SD array. Their field of view (blue lines in Fig. 1.3) is so that they overlook the whole SD array, namely the atmosphere just above the array. Each location sees 180° in azimuth and looks 1.5° in elevation [6]. Three additional FDs at Coihueco site, denoted by red in Fig. 1.3, are oriented to see higher in the atmosphere than the other FDs to observe lower-energetic EAS, which develop higher in the atmosphere. They are complementary to the infill SD arrays in lowering the energy threshold above which the Pierre Auger Observatory is capable to measure.

SD array takes a snapshot of the EAS on the plane of the ground, as shown in Fig. 1.4. Based on the timing of the signal, the shower geometry, thus the direction of the primary UHECR is determined. Signals collected in multiple SD stations are fitted to infer the energy of the primary. FDs work mostly individually, seeing the longitudinal development of the EAS just before it hits the ground, as illustrated in Fig. 1.4. This naturally enables to infer more characteristics of the EAS (and the primary UHECR) than pure SD measurements. However, FD eyes are telescopes working only on dark, moonless nights. The duty cycle of FD is 15 %, while the SD array has almost 100 % uptime.

FD makes use of the fluorescence of air, when an EAS passes through it. Secondary particles

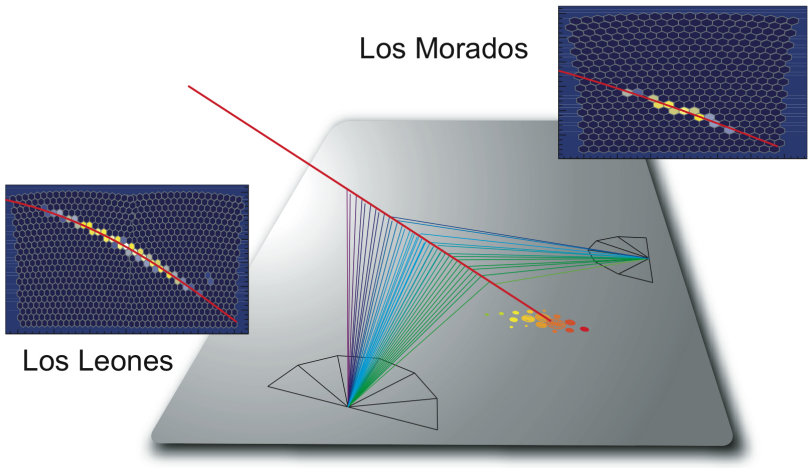


Figure 1.4: Figure of a hybrid event seen by 3 FD eyes while the EAS develops in the atmosphere and 17 SD stations on the ground. Figure is from [4]. Colours represent the timing of the signal, while the size of SD stations represents the signal strength in individual SDs and shows the measured lateral distribution of the EAS.

of the shower excite the molecules of nitrogen, which then take approximately 10 ns to emit the light mostly at 300-400 nm to de-excite [2]. Using this technique, it is possible to reconstruct the longitudinal profile of the EAS, deriving $X_{\max}$ and the primary energy from it, as the light intensity corresponds to energy deposit of the EAS in the atmosphere.

FDs are composed of hexagonal mirrors with overall curvature diameter of the spherical mirror being 3.4 m. Focused fluorescence light is then collected by 440 PMTs, i.e. pixels on the spherical focal surface of 1.7 m. This allows the reconstruction of the projected direction and timing of the EAS as it developed in the atmosphere. A schematics of an FD can be seen in Fig. 1.5. One such eye of the FD has a field of view of $30^\circ \times 30^\circ$.

Part of the events, corresponding to the duty cycle of FDs, are measured by both the FDs and the SD array. These are so-called hybrid events. An event can be even golden-hybrid, if it is measured by at least 3 SDs and an FD. In this case, the primary energy and direction can be inferred both from FD and SD observations. The energy from FD measurement is determined by integrating the longitudinal development of the shower. At the start of the operation of the Observatory, these FD energy measurements were used for cross-calibration with energy estimators deduced from the lateral distribution of particles based on calibrated SD signals. Thanks to that, the energy measurement of the primary UHECR from SD array is very accurate even in the cases when FDs have to be switched off. Lastly, the mass of the primary can be estimated with good precision using FD measurements of $X_{\max}$.

The composition from pure SD measurements is currently available though deep-neural-network analysis of these measurements. However, this method has lower resolution compared to FD measurements. Remember, the heavier the nucleus, the more deflected it gets in magnetic fields on its way towards Earth. It is therefore highly-desirable to obtain the mass of the primary

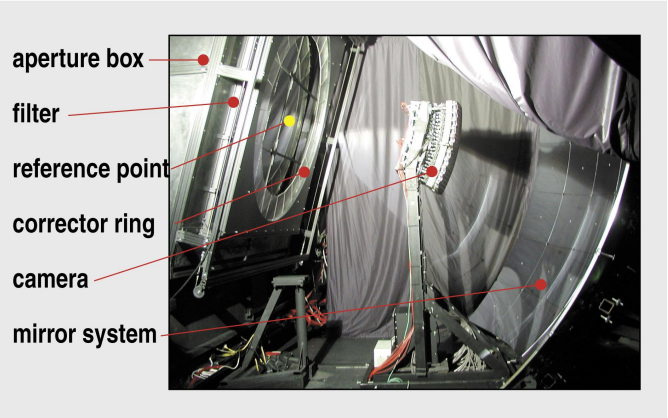


Figure 1.5: A photo of one eye of the fluorescence detector with description of its main parts. Camera is composed of 440 PMTs (pixels) that collect the fluorescence light, after the light passed through the aperture system and UV filter, and was reflected from the segmented mirrors. Resource: [7].

with high accuracy in 100% of the measurements, not only 15% of them. This could be remedied by gifting additional ability of measuring the muon component of the EAS to the SDs, not only through deep learning. The need for muon component measurements is one of the reasons behind AugerPrime, the upgrade of the Pierre Auger Observatory, taking place at this very moment.

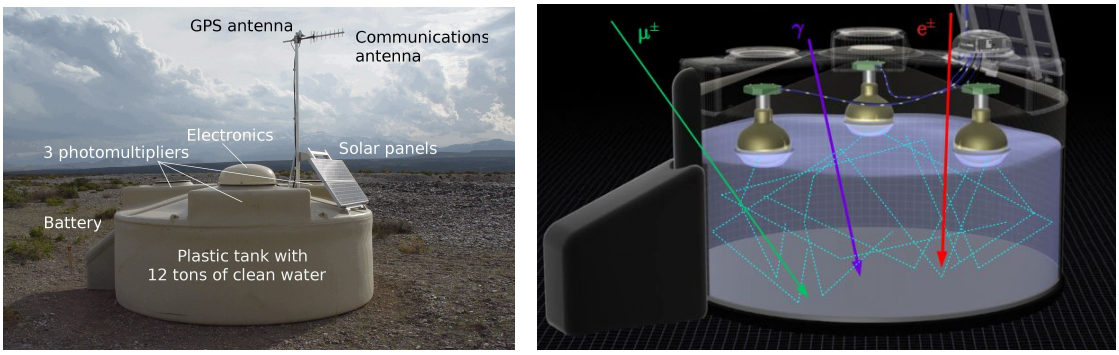
1.2.1 Surface Detector of the Pierre Auger Observatory

Let us now look at an SD station in more detail, since the SD and its data is the focus of the work in this master thesis. With the distance of 1.5 km between two SD stations, an SD station is a self-standing unit. Each SD is energetically independent, possessing a solar panel and a battery to provide on average 10 W of power to the electronics and communication & data transmission (GPS, radio) components. A look on an SD station from outside is shown in Fig. 1.6a. SD stations communicate with the Central Data Acquisition System (CDAS) only via radio and they are not interconnected otherwise. Thus, they require reliable calibration method ensuring uniform response of the whole array.

The central part of each SD station is a Water Cherenkov Detector (WCD) sketched in Fig. 1.6b. It contains 12 t of the hyper-pure water stored in a reflective Tyvek liner of 1.2 m height and 3.6 m diameter. The water tanks, as designed at the beginning of their deployment, contain 3 large photomultiplier tubes (PMTs) called PMT 1, PMT 2 and PMT 3. The PMTs look downwards from the top of the liner to collect the photons produced by Cherenkov effect, spaced symmetrically 1.2 m away from the tank's central axis [5].

Cherenkov effect occurs when relativistic particles traversing a medium faster than the speed of light in the given medium produce Cherenkov radiation, which can be measured by PMTs. Specifically, this happens when $\beta n > 1$, with n the refractive index of the medium, and β the ratio of particle speed and the speed of light in vacuum c. The light is emitted in the so-called Cherenkov cone in the direction of particle movement. The opening angle of the cone θ, depends on the energy of the particle: $\cos\theta = (\beta n)^{-1}$, as $\beta = [1 - (m_0 c^2/E_{\text{tot}})^2]^{1/2}$ with E_{tot} the total relativistic energy of a particle and m_0 its rest mass [9].

The voltage produced after collecting the photoelectrons is then converted by flash analogue-



(a) (b)

Figure 1.6: a) Close-up of a station of the surface detector before the upgrade with the description of its outer parts from [4]. b) A look inside the Water Cherenkov Detector. Charged secondary particles of an EAS generating Cherenkov radiation are illustrated in colour: muons/antimuons in green, and electrons/positrons in red. Photons in dark blue get detected as well. Resource: [8].

to-digital converter (FADC). The details of this process as well as the characteristics of the FADCs will be detailed in the next sections and chapters, as they are being upgraded and the calibration of WCD belongs to the central part of this work.

If the signal passes level-two trigger (T2) at an individual station level, the CDAS of the Observatory is notified. If such signal is approved as T3 (passing the third-level trigger) by CDAS, which looks for time coincidences between the T2 signals, all the data are requested from all the stations that took part in this event, and these data are stored for offline reconstruction. The offline event reconstruction is then based on signals from multiple SDs and has to pass level-4, and level-5 triggers (denoted T4 and T5). In that case, using the timing of the start of the signal and its magnitude in each SD, one can infer the direction and the energy of the primary UHECR [5]. In purely SD events, mass inference of the primary was possible with not so good resolution before the upgrade of SD stations, which started in 2016 [6] and is still ongoing.

1.3 AugerPrime, the Upgrade of the Observatory

The upgrade consists of multiple parts and detectors to enhance the Observatory's resolution and its abilities in determining the mass of the primary UHECRs in 100 % of the measurements.

Two components of the upgrade are important for this work. The first one is the Surface Scintillator Detector (SSD), which has been mounted on top of almost every SD station. In scintillators, passing secondary particles excite molecules or electrons of the material (depending on the scintillator type), which then emits light when de-exciting. The key fact in this detector is that it rarely interacts with photons as opposed to WCD.

The second important component that is being upgraded is the SD station electronics. The upgraded unified board (UUB) increases the resolution when digitasing the signal. Only the signals of the SD stations where both the components had been upgraded were analysed.

The upgrade started in 2016 and the deployment was on schedule despite the world pandemic. The upgrade continues at a very satisfying pace. As of 17$^{\text{th}}$ July 2022, 1437 SSDs were installed in the field, accompanied by 405 UUBs [10]. The state of the upgrade is shown on the map of SD array in Fig. 1.7. For comparison, on 13$^{\text{th}}$ March 2022 the upgraded numbers were 1410 SSDs and 256 UUBs.

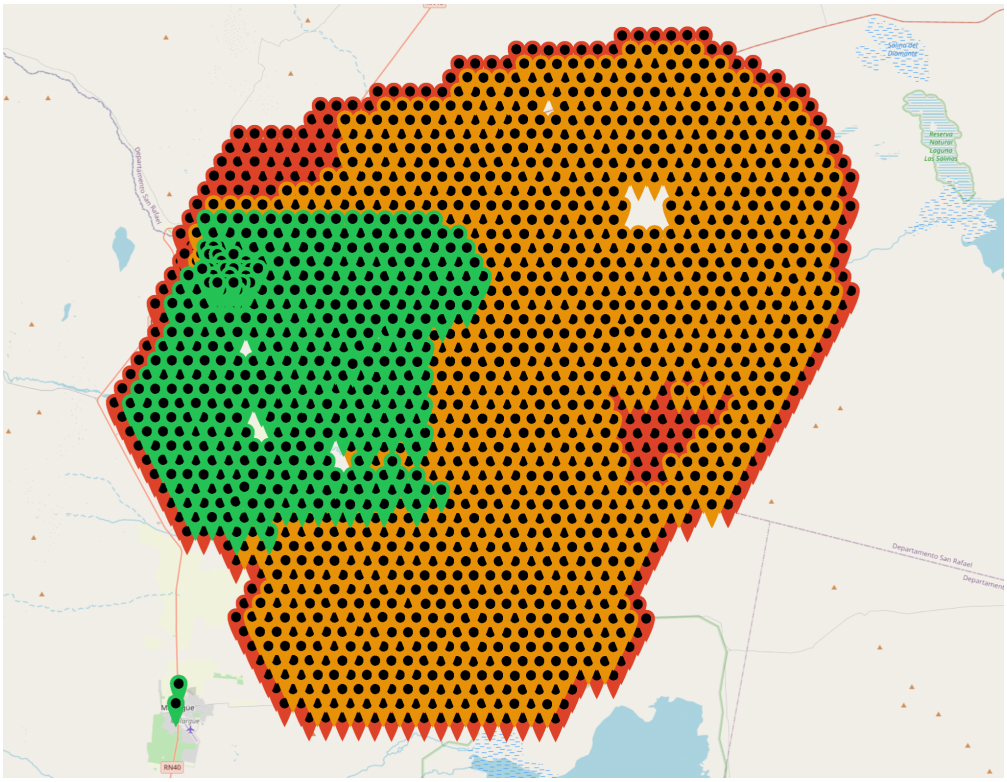


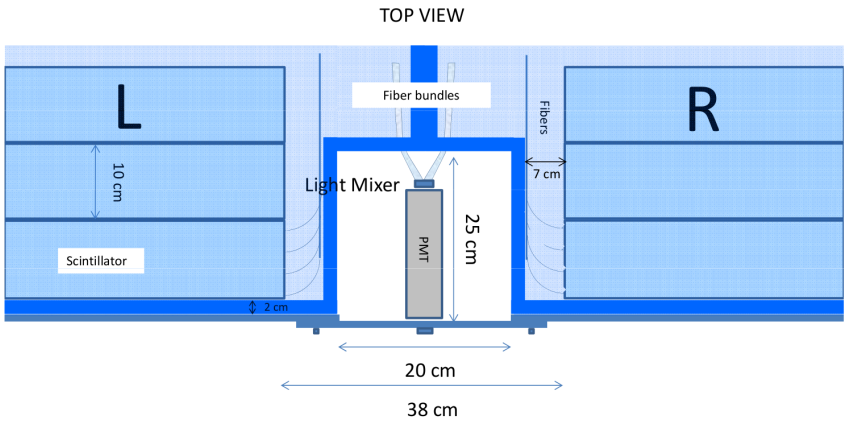
Figure 1.7: State of the upgrade of the SD array as of 17$^{\text{th}}$ July 2022. Green stations have both SSD and upgraded electronics. Orange stations have SSD, but old electronics. Red stations do not have a SSD nor new electronics. Figure from [10].

1.3.1 Surface Scintillator Detector

The SSD consists of two plastic scintillator sub-modules 1.9 m^2 each, which are composed of smaller scintillator bars. The two modules are denoted by "L" and "R" in Fig. 1.8b. The scintillation light is collected by Kuraray Y11(300)M S-type wavelength-shifting fibres, whose one end is located inside holes in the scintillator bars. The light is transmitted through the U-turns of the fibres, which were designed to maximise the light yield. The attenuation length of light in the fibres is (312 ± 3) cm. The other end of the fibers is optically-coupled to bi-alkali Hamamatsu R9420 PMT with quantum efficiency of 18 %. Everything is in a light-tight and waterproof box. SSD also have double roof of aluminium to minimise the temperature variations due to sunshine. The measured temperatures inside SSD usually go up to 40°C.

After the assembly, each SSD module was checked for uniformity of response. The average gain of photoelectrons per one vertical minimum-ionising particle (MIP) was 30 ± 2 based on the measurements. The dynamic range required in SSD is 1 to 2×10^4 MIP, ranging from calibration to event signals. This is acquired by splitting the anode signals into two: one attenuated by factor 4 and the second amplified by factor 32 [11].

(a)



(b)

Figure 1.8: The new Surface Scintillator Detector (SSD) have been mounted on top of almost every SD station. (a) Real-world photo of the upgraded SD station. (b) SSD model. Figures taken from [4] and [6].

1.3.2 Electronics Upgrade

The local electronics handles the high voltage of PMTs, up to second-level triggers, communications with the CDAS and of course, the signal recording [11]. The FADC AD9628 of UUB has sampling rate of 120 MHz, three-times better compared to 40 MHz of pre-upgrade unified board (UB). The new dynamic range is 12 bit, a 2-bit improvement from UB [6]. Upgrading GPS receivers for signal timing to I-Lotus M12M, SD reaches 2 ns accuracy compared to the previous one of 10 ns [5]. This assures smaller uncertainty in estimating the arrival direction of an event.

The implications of the parameters of the SD electronics on calibration and signal recording will be discussed in more detail in Chapter 2, which is dedicated to a new calibration method of the SD array.

Chapter 2

Calibration of the Surface Detectors

The SD stations are self-standing units 1.5 km apart communicating with the CDAS via radio with restricted bandwidth and not inter-connected otherwise. They need a robust calibration method allowing the comparison and synthesis of the signals of EAS particles. This means the signals have to be measured consistently at all positions no matter where in the SD array a WCD is located.

The method for calibration is based on measuring the atmospheric muons. These are secondary particles coming from lower-energetic EAS, which developed high in the atmosphere. More specifically, they come from the decay of charged pions and kaons of the hadronic cascades. Their energy spectrum varies between 1 GeV and 10 TeV with the most probable value of 100 GeV [2]. The signal of a muon passing through a WCD is predominantly influenced by its direction, i.e. path length in water of a WCD. The second quantity influencing signal less is its energy, due to Cherenkov effect [12]. The unit of signal in SD is defined as the signal produced by the muon passing vertically through the center of a WCD - a vertical equivalent muon (VEM). The signals of atmospheric muons having other directions are then distributed around this value as discussed and shown in the following section about calibration histograms.

However, the secondary particles entering the WCD are not only vertical muons, a large number of photons, electrons and positrons is also present, as shown in Fig. 2.1. This is the electromagnetic (EM) component of secondaries not related to events measured by the Observatory. The signals of the EM component are generally lower than VEM due to their lower energy deposit in water compared to muons. This results in the contamination of the calibration histograms by the EM background and thus finding the charge collected by PMTs corresponding to 1 VEM is not always straightforward.

As it turns out, one of the effects of WCDs aging is merging of the muon contribution and the EM component in calibration histograms. In these cases, the conventional offline calibration of WCDs[1] might fail. To overcome this obstacle by eliminating the EM component, a new offline calibration method based on coincidence between SSD and WCD signals was proposed by Auger group of the IIHE in 2021 [14]. The first short test to see if the parameters were set correctly was performed on 4 tanks. The results were encouraging and the next step towards the new calibration method was collecting and analysing more data to compare the new and the

[1] Offline calibration of WCDs, the subject of this chapter, is used when reconstructing the events later after data acquisition. WCDs also have an online calibration based on signal rates, to adjust the high voltage of PMTs in real time [13].

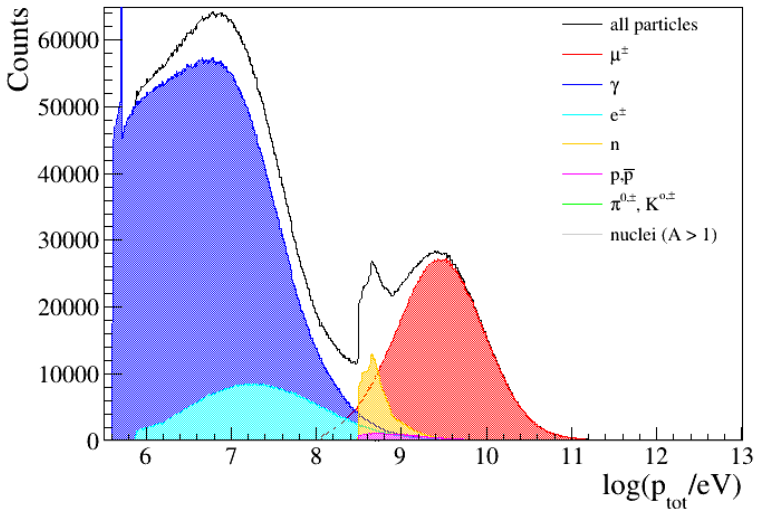


Figure 2.1: Histogram of secondary particle momenta in the atmosphere simulated at the ground level from [8]. Electromagnetic background is shown in blue, while the atmospheric muons crucial for SD calibration are in red.

currently-used calibration methods. This is where the author of this thesis took over the work.

Two more and considerably larger tests were performed on 154 SDs in total. In next sections, the analysis methods and the results of the tests will be detailed.

2.1 Pulse Height and Charge Calibration Histograms

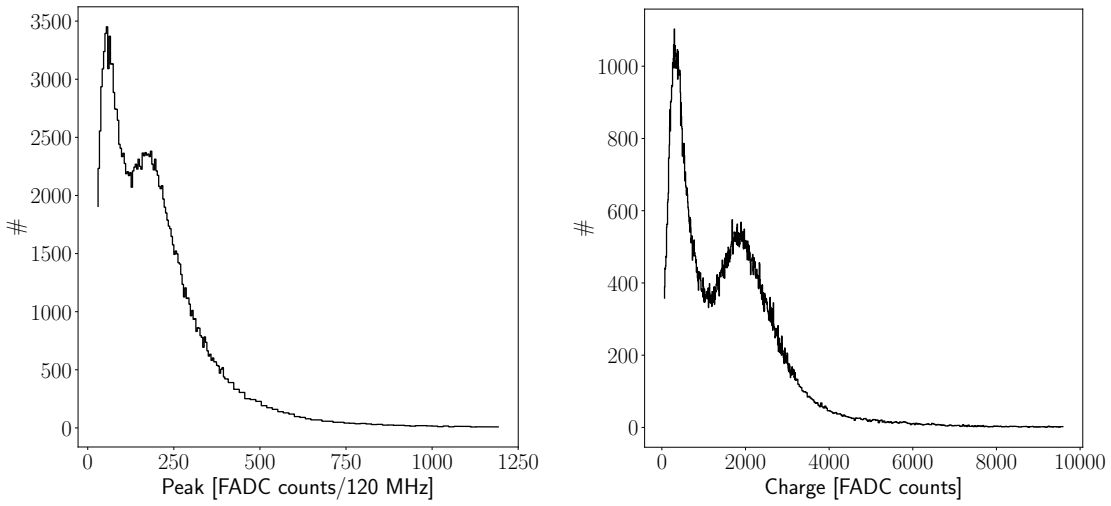
To understand the calibration of WCDs and the meaning of the calibration histograms, one needs to comprehend how the photons collected by the PMTs of a WCD are digitised. The voltage at the anode of a PMT is proportional to the number of collected photoelectrons, which naturally depends on the number of photons in the WCD. The voltage is then digitised by an FADC. The FADC takes the value of 0-2 V and with its 12 bits ($2^{12} = 4096$), it assigns an integer in the range 0-4095 FADC counts/sampling frequency to the measured voltage. With the FADC's sampling frequency of 120 MHz, this happens every 8.3 ns. An amplitude histogram of these FADC count values/120 MHz collected for 60 s is the first of the two calibration histograms. An example can be seen in Fig. 2.2a.

To obtain the second of the calibration histograms - the charge histogram, one needs to integrate voltage values V_i (in FADC counts/sampling frequency) over time. In this discrete case, the integration is a sum and the integration window is 70 time bins of the FADC [15]. Using the definition of electrical current $I = \mathrm{d}Q/\mathrm{d}t$:

$$V(t) = R\frac{\mathrm{d}Q}{\mathrm{d}t} \qquad \Longrightarrow \qquad Q = \frac{1}{R}\sum_{i=1}^{70} V_i \Delta t, \tag{2.1}$$

with R the resistance across which the voltages are measured and $\Delta t = 8.3$ ns. An example of the resulting charge calibration histogram can be seen in Fig. 2.2b.

In the histograms, two peaks can be observed. The first at lower charge and amplitude is the one due to signals from the EM component of the background. The second peak is due to the



(a) (b)

Figure 2.2: Figure of the example pulse height (a) and charge (b) calibration histograms of the SD station 1819, PMT 3. In UUB, PMT voltage is digitised every 8.3 ns giving it the value between 0-4095 FADC counts/120 MHz. Collecting the values over 60 s results in the amplitude height histogram of the background signals on the right. Integrating (summing) the pulse heights in the sliding time window of 70 time bins gives charge histogram on the left. In both cases the left, higher peak is due to EM component of background particles. The second peak is the result of atmospheric muons passing through the tank, on which the WCD is calibrated.

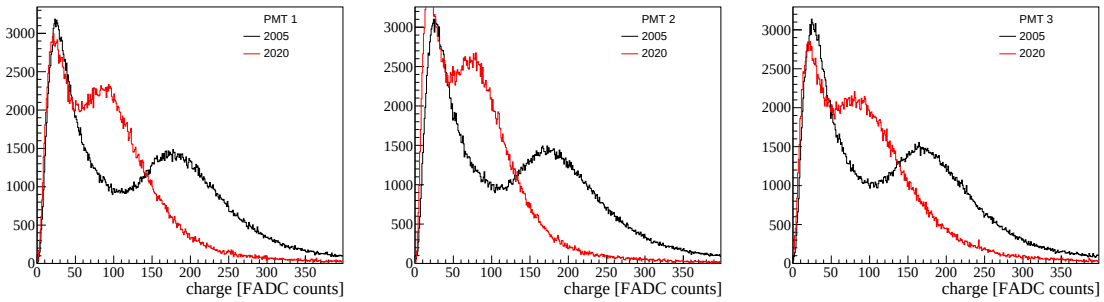
through-going muon signals. Using additional external muon scintillators to trigger only on the muons with the suitable directions [12], the constant conversion factor between the FADC count of the maximum of the second peak $Q^{\text{peak}}_{\text{VEM}}$, and the actual maximum of the VEM peak, VEM was established as $Q^{\text{peak}}_{\text{VEM}} = (1.03 \pm 0.02)\text{VEM}$ [13] for individual PMT signals.

What is done further in the offline analysis of event signals is, that the second peak of the charge calibration histogram (CH) is fitted by the polynomial of the second degree to find the FADC count of the second peak, $Q^{\text{peak}}_{\text{VEM}}$. This is converted to VEM and finally, the calibrated signals in units of VEM in various SDs can be consistently calculated to reconstruct the events.

2.2 The Effect of Aging of the Surface Detectors on Their Calibration

The construction of the Observatory took place in years 2002-2008 [6], with the start of data acquisition in 2004. This means some of the WCDs have been already almost 20 years in the field, and time is taking its toll on them. The effects are multiple, but the crucial one for this work is the merging of the VEM peak with the EM peak in the calibration histogram, as shown in Fig. 2.3. The reason for the merging is that the total signal per particle becomes smaller due to higher light losses [16]. The merging makes the estimation of the $Q^{\text{peak}}_{\text{VEM}}$ produced by fitting more complicated. Another concern is the possibility of the bias of $Q^{\text{peak}}_{\text{VEM}}$ by the EM component leftwards more than accounted for by the VEM conversion factor.

To characterise this merging quantitatively, one may look at the ratio of valley to hump v/h.



(a) (b) (c)

Figure 2.3: Figure of the charge calibration histograms of SD Maya Evelyn (ID number 601) in 2005 (black) and 2020 (red) taken from [14]. As the station ages, the light signal per particle in WCD is lower, which results in merging of the muon peak into the EM background.

The valley is the value of the charge calibration histogram at the minimum between the EM and the muon peaks. Hump is the value of the maximum of the muon peak. This ratio is at most 1, with good PMTs defined as having $v/h < 0.8$.

2.3 Coincidence Histogram - a New Calibration Method

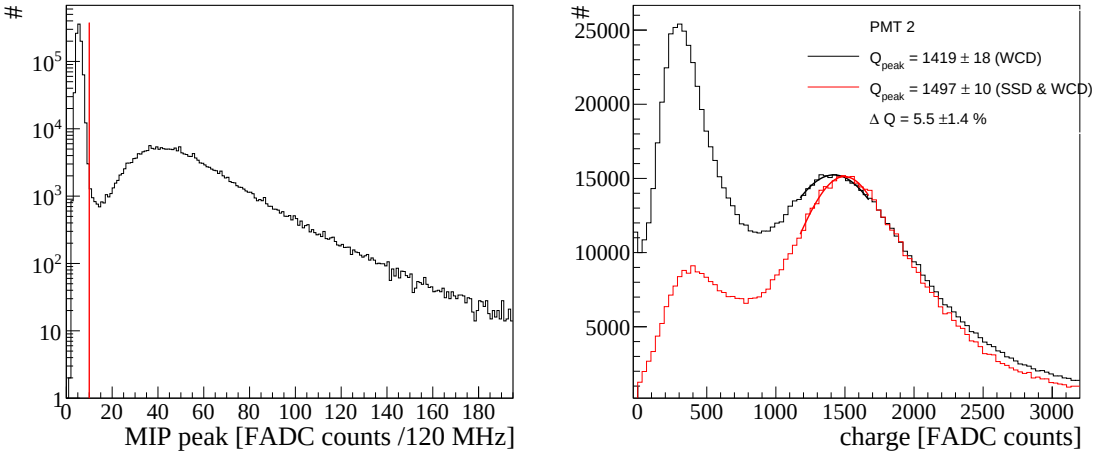
Fortunately, almost all SD array obtained an additional detector - the SSD as described in 1.3.1, with the border stations being excluded. Compared to WCD, SSD responds to few photons. Therefore, it is possible to suppress the bulk of EM background in WCD calibration signal if requiring coincidence between WCD and SSD signals, i.e. recording mostly signals of muons, electrons and positrons [14].

The coincidence procedure is as follows: The signals from the SSD are buffered. If a signal in WCD exceeds 30 FADC counts/120 MHz above baseline, a signal of more than 10 FADC counts/120 MHz is searched in SSD in the time window of ± 50 time bins (417 ns). Signals with the peaks satisfying these conditions are recorded for the coincidence calibration histograms [17]. The rest of the procedure of the construction of the coincidence calibration histograms is the same as for the common calibration histograms.

Previous studies included collecting 1 minute of data from the WCD and SSD of station Trak Jr. onto a USB stick. The pulse heights from SSD were investigated by the means of histogram shown in Fig. 2.4a to determine the threshold of 10 FADC counts/120 MHz in SSD to have the coincident muon signal.

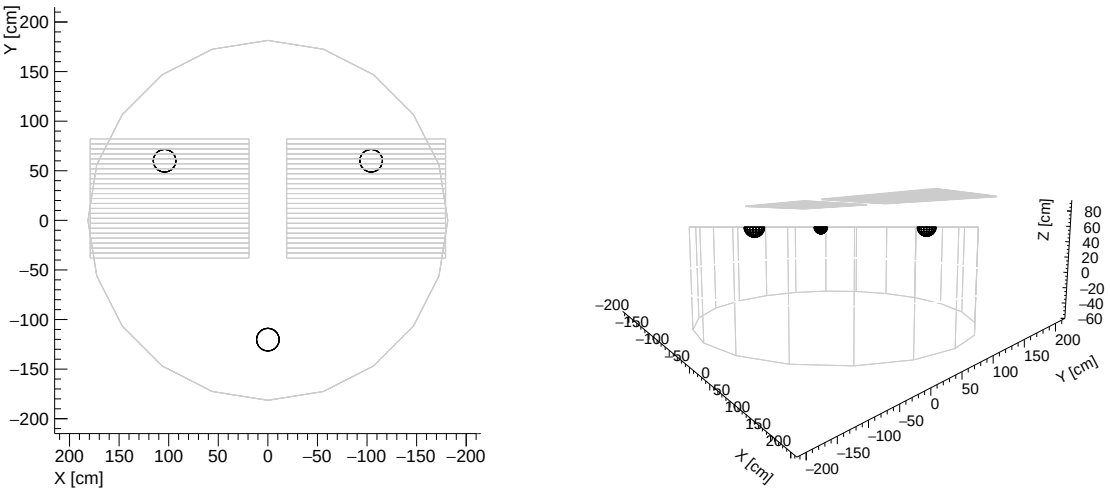
To verify the method, data from 3 more SDs (Zapata, Bac Dau, and Van Mieu) were collected using again a USB stick. The data confirmed the correctness and effectiveness of the coincidence calibration method. Previous studies included simulations as well. This was to see the differences between $Q^{\text{peak}}_{\text{VEM}}$ in the usual charge calibration histogram (CH) and the coincidence charge calibration histogram (CCH), denoted as $Q^{\text{peak}}_{\text{CH}}$ and $Q^{\text{peak}}_{\text{CCH}}$ respectively. This difference, as well as the first example of CCH can be seen in Fig. 2.4b.

There are two causes of these differences. The first is the geometry of the detector as a whole. The SSD covers only a part of the top of the SD. Therefore, not all the directions of the through-going muon signals normally present in a CH are recorded in a CCH. The second reason is the absence of the EM peak in the coincidence case, which convolved with the muon peak shifts the estimated peak in CH leftwards.



(a) (b)

Figure 2.4: Figures taken from [14]. a) Pulse height histogram of SSD of station Trak Jr. The red vertical line shows the threshold condition of 10 FADC/120 MHz for coincidence between SSD and WCD signals. b) Charge histogram is shown in black. To compare, normalised coincidence charge histogram (in red) was constructed from the test data of Van Mieu imposing the condition for the coincidence of calibration signals.



(a) (b)

Figure 2.5: SSD mounting schematics from [15]. Look at a station of the SD from above (a) and from the side (b). Rectangular SSD is shown as grey stripes, the three large PMTs are shown as circles/half-spheres. PMT 2 is the one not under the SSD. b) Looking at the SD station from the side, the distance between water volume (the top of cylinder) and the SSD is shown.

Moreover, there is asymmetry in the mounting of SSD from the viewpoint of the PMT 2 and PMTs 1 & 3, as shown in Fig. 2.5. For PMT 1 & 3 a mirror symmetry around a vertical plane through the SD's middle can be found, thus the difference between $Q_{\mathrm{CH}}^{\mathrm{peak}}$ and $Q_{\mathrm{CCH}}^{\mathrm{peak}}$ should be the same for these two PMTs. For PMT 2, this difference will be non-identical with PMTs 1&3 due to the asymmetry of the viewpoint of PMT 2.

To quantify these differences between $Q_{\mathrm{CH}}^{\mathrm{peak}}$ and $Q_{\mathrm{CCH}}^{\mathrm{peak}}$, a variable Δ was defined as:

$$\Delta = \frac{Q_{\mathrm{CCH}}^{\mathrm{peak}} - Q_{\mathrm{CH}}^{\mathrm{peak}}}{Q_{\mathrm{CH}}^{\mathrm{peak}}}. \tag{2.2}$$

This is a relative difference with respect to the charge histogram and will be quoted in %.

The simulations showed expected Δ to be $0.6 \pm 1.7\,\%$, $4.2 \pm 1.5\,\%$, and $0.6 \pm 1.6\,\%$ for PMT 1, 2 and 3, respectively [14]. The data of the four previously-investigated SDs were consistent with these predictions.

2.4 New Tests of Coincidence Calibration

Next steps towards the verification of the coincidence calibration required much broader testing on multiple WCDs in February 2022. The analysis of these new data is the subject of this chapter. The first test on 02/02/2022 was 26 min long and included 61 SDs (excluding SDs with ID$\leq$100 as engineering array). The second test of 18.75 h on 09-10/02/2022 was performed on all of the array with upgraded electronics and data were collected from 150 SDs in total. The data from all but 4 SDs (545, 1743, 1745, 1747) of the test of Feb. 2^{nd} were present in the data set of Feb. $9^{\mathrm{th}} - 10^{\mathrm{th}}$ as well. In turn, the test of Feb. $9^{\mathrm{th}} - 10^{\mathrm{th}}$ contained 93 additional SDs compared to Feb. 2^{nd} dataset.

The coincidence histograms were transmitted via radio this time, in a file containing events which passed the third-level trigger. To perform the test, the coincidence histograms were not part of the standard SD file and they got split from the event data when reaching the CDAS. The coincidence histograms were then stored in a binary file with the time stamp of the event. Thus, one had to pair them with their events at first. That brings us to the next section - the analysis of the new data.

2.5 Analysis of the New Tests

The analysis of the new calibration method consisted of two main parts. The first was to find $Q_{\mathrm{CH}}^{\mathrm{peak}}$, $Q_{\mathrm{CCH}}^{\mathrm{peak}}$ and v/h of each PMT of each SD and to determine their individual Δ. Then, the expectation value of $\langle\Delta\rangle_{1,3}$, and $\langle\Delta\rangle_2$ for all PMTs 1&3, and PMTs 2 across all the SD stations was calculated to have a conversion factor between the calibrations.

The second analysis part made use of the found Q^{peak} values to calibrate the event signals from FADC units to approximately VEM[2]. The stability of the calibration was then assessed using the signal differences between PMTs of the same WCD comparing signals calibrated by the two calibration methods. The first was using the muon peak found from CH, and the second used muon peak found from CCH with $\langle\Delta\rangle$ applied to account for the muon peak differences due to the detector geometry.

[2]The conversion factor $Q_{\mathrm{VEM}}^{\mathrm{peak}} = (1.03 \pm 0.02)$VEM was omitted as a constant unimportant for this kind of analysis, where the method is to compare relative differences between PMTs and this conversion factor gets cancelled.

2.5.1 Differences between peaks in conventional and coincidence calibration (Δ studies)

The binary files containing coincidence histograms were read by the C++ script provided along with the first set of test data [17]. The output of the script was ASCII format, where coincidence histograms were stacked and were separable using the header. The time stamp of the event was also provided. What followed after each header were two tables, one with coincidence pulse height histogram and the other with CCH. The first column contained the bin number (0-599) for CCH, and the three other columns contained the FADC count for each bin.

For this analysis, only the CCH were used. The reason was to focus on the charge histograms as only they are used in the offline reconstruction of the events. In pulse height histograms, the merging of the muon hump into the EM background is generally more severe among WCDs, as shown in Fig. 2.6a. The dominant effect on the difference between muon peak of the conventional and the coincidence pulse height histogram was therefore expected to be that of aging, not of the geometry of the detector. In fact, pulse height histograms are not used for calibration anymore. Investigating the coincidence pulse height histograms is a prospective future study.

First, the CCHs were read from the text output. The bin counts were divided by bin width, to work consistently with $\mathrm{d}N/\mathrm{d}Q$. The bin numbers were converted to bin centers based on the bin width. Specifically for CCH, the bin width up to and including bin 402 (starting from 0) has width of 8 FADC counts and the bins 403-599 are of higher width due to small statistics of the high-energy muons. CCH of SDs with ID$\leq$100 were automatically rejected, because they are part of the engineering array. Another CCH that were excluded straightaway were those considered as faulty PMT and were found by having the maximum count within the first 36 bins of CCH. The second type of faulty PMTs was cut automatically if they were empty, i.e. the average $\mathrm{d}N/\mathrm{d}Q < 0.05$. The examples of these two cases are shown in Appendix A.

The CCH then needed to be paired with CH. CH were extracted from unreconstructed SD event files in the form of ASCII files as one file per event per SD per PMT[3]. The CCH and CH were then paired based on the time stamp of the CCH and the event. It was noted, that the difference in the event time stamp was +2 s compared to the CCH timestamp. In the case of the test of Feb. 2, all the CCH were paired, however in the test of Feb. 9-10, 98 out of 3285 timestamps were not found. If a timestamp was missing among the SD events, the CCHs of all the PMTs of the given SD with this timestamp were discarded. The CH were loaded after pairing and the consistent scaling of the histogram to have $\mathrm{d}N/\mathrm{d}Q$ was assured as well. In the case of CH, it is the first 400 bins that have width of 8 FADC counts and the last 200 are 32 FADC counts wide.

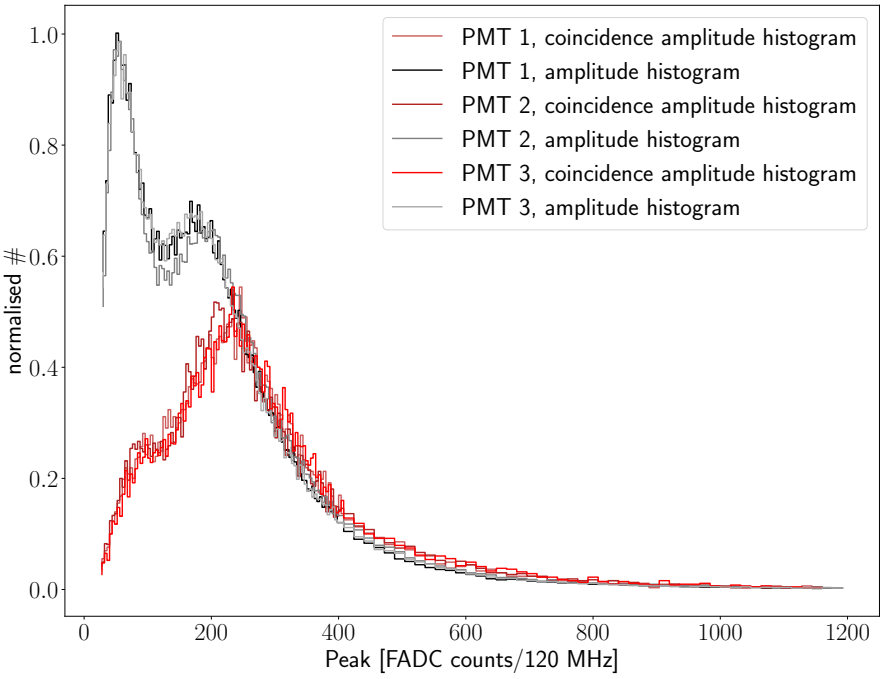
Histogram statistics and faulty PMTs

A raw CCH is shown in Fig. 2.6b. The (Poisson) statistics of coincidence CCH after dividing by the bin width B is:

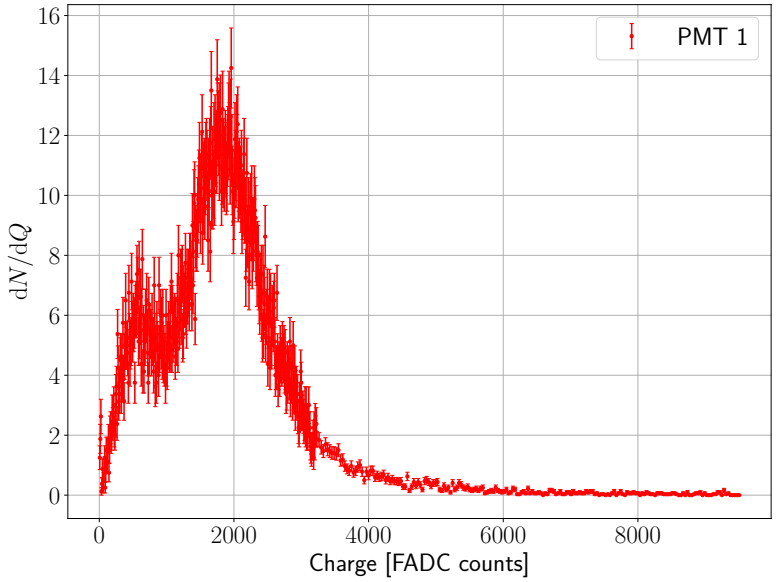
$$\frac{\mathrm{d}N}{\mathrm{d}Q} = \frac{N}{B} \pm \frac{\sqrt{N}}{B}. \tag{2.3}$$

By inspection, some faulty PMTs were not rejected automatically in the loading step. Mostly it was because an unexplained peak in the valley of the CH, as seen in Fig. 2.7a (SD 816 PMT 3, SD 1840 PMT 2, and SD 1849 PMT 2. Additional case was SD 851 PMT 3 with faulty behaviour in both CH (Fig. 2.7b) and CCH. All the calibration histograms of these PMTs were rejected for further analysis.

[3]by Dr. Mauricio Suárez Durán

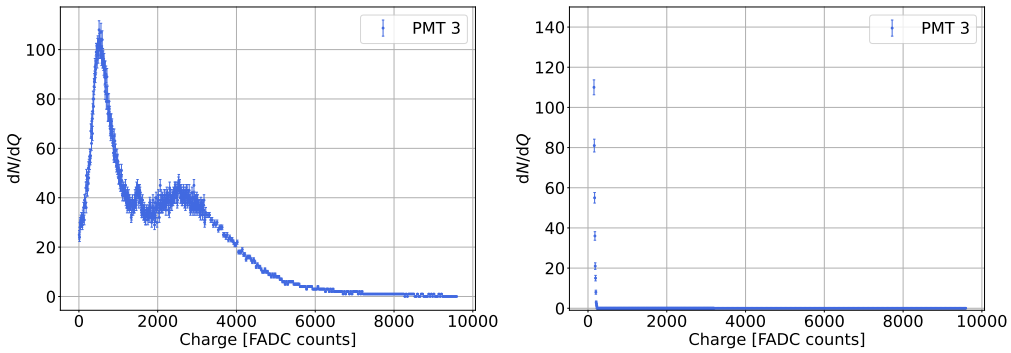


(a) SD 1819, normalised pulse height histogram (in black/gray) and coincidence pulse height histogram (in red) rescaled to match the tail of usual pulse height histogram. It is clear the muon hump of the conventional amplitude histogram is shifted by the EM background much to the left.



(b) SD 1819, raw CCH at GPS time 1328472272.

Figure 2.6

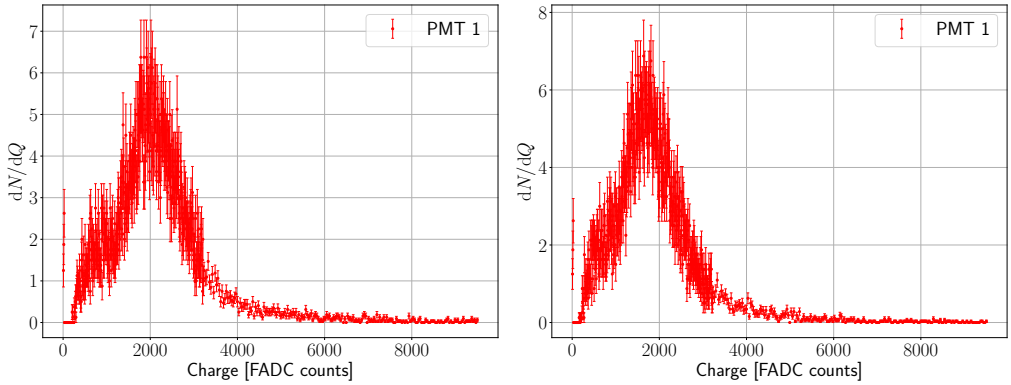


(a) SD 816, charge histogram at GPS time 1328503785. There was an unexplained extra peak in the CH valley of this PMT.

(b) SD 851, charge histogram at GPS time 1328502388.

Figure 2.7: Two examples of PMTs with faulty behaviour.

Some of the CCH and CH have a peak at 0, illustrated in Fig. 2.8. This is caused by a large baseline for a few PMTs (CCH: 666, 1; 851, 1. CH: 909, 1; 909, 2; 847, 1; 1227, 1; 1227, 2; 1798, 1). The peak at zero was discarded in the analysis.



(a) SD 666, coincidence charge histogram at GPS time 1328516896.

(b) SD 851, coincidence charge histogram at GPS time 1328502388.

Figure 2.8: Example of CCH wih high-baseline, which can be seen as a few $\mathrm{d}N/\mathrm{d}Q$ at the lowest charge followed by $\mathrm{d}N/\mathrm{d}Q = 0$ and only then typical CCH shape. Furthermore, compared to other CCHs such as that in Fig. 2.6b, the peak of the CCHs shown here reaches only the half-values of $\mathrm{d}N/\mathrm{d}Q$, with around 14 being a usual value.

Fitting the peaks and the valley

To find Q^{peak} and v/h, the muon peak in both CCH and CH, and the valley in CH were fitted by the second degree polynomial of the form:

$$f(Q) = aQ^2 + bQ + c, \tag{2.4}$$

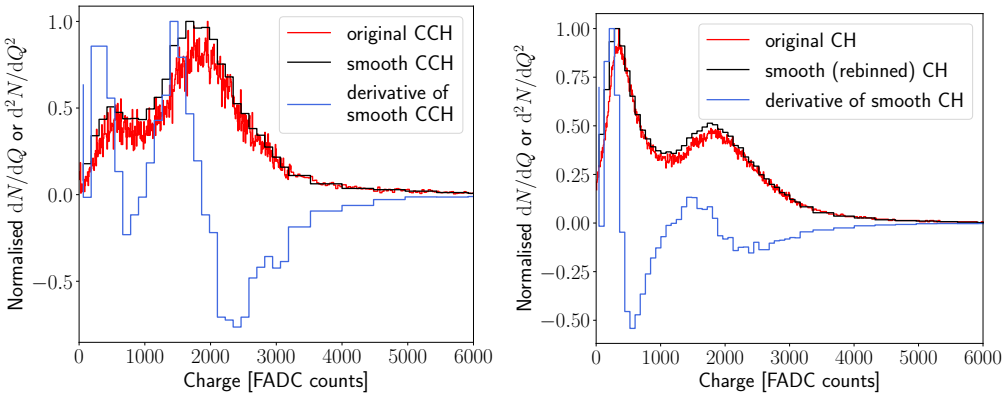
with a, b, c the free parameters.

The least-squares fitting method of python's scipy.optimize.curve_fit [18] was used.

The fitting range was found automatically based on specific shapes of the first derivations of re-binned (smoother) CH and CCH. The fitting was done on the original calibration histograms however (not on the smoothed ones). The example of a CCH, its smoothed version and its derivative can be seen in Fig. 2.9a, the same for CH in Fig. 2.9b. Starting with CCH, the bins were merged by 15. The derivative of the i^{th} channel of the smoothed histogram was found as:

$$\frac{\mathrm{d}^2N}{\mathrm{d}Q^2} = \frac{\left(\frac{\mathrm{d}N}{\mathrm{d}Q}\right)_{i+1} - \left(\frac{\mathrm{d}N}{\mathrm{d}Q}\right)_{i-1}}{Q_{i+1} - Q_{i-1}}, \tag{2.5}$$

with Q in FADC counts.



(a) SD 1819, CCH at GPS time 1328472272.

(b) SD 1819, CH at GPS time 1328472272.

Figure 2.9: Examples of the CCH and CH and their smooth versions, re-binned by 15 and 10 bins respectively. Derivative of the smooth version used for finding the fitting range automatically is shown in blue. All the histograms are normalised by their maximum for the display purposes.

The algorithm then looked for the zeros of the derivative, i.e. the peak and potentially the smaller peak on the left slope of CCH. The minimum of the derivative specifying the inflection point on the tail of CCH (to the right of the muon peak) was also found. The muon peak was then found as the first zero to the left of the minimum of the derivative. The boundaries were expanded around the peak until the count in the smoothed histogram at the boundary was 75 % of the peak.

In case of CH, the re-binning merged 10 bins. The derivative was found in the same way as for CCH. The algorithm then looked for the minimum of the derivative, which was the inflection point before the valley, and for the maximum of the derivative after the minimum, which was the inflection point after the valley but before the muon hump in CH. Zero of the derivative just before and just after this maximum were then the valley and the hump of CH. The left boundary of the valley was set as the place of the derivative minimum plus 10 % of the distance between the minimum and the second maximum of the derivative. The right boundary of valley and left boundary of the hump fit were set to be identical, initially at half the distance between the valley and the hump. The right boundary of the hump was set as the steepest point of the tail of the muon hump minus 40 % of the distance between this steepest point and the muon hump. The right hump border is further shifted rightwards if the value at it is not below 75 % of the value at

the hump. The left valley border is adjusted to be at 130 % of the valley count of the smoothed histogram. Lastly, if valley-to-hump ratio of the smoothed histogram is more than 0.85, the left hump border is the point of the valley and the right valley border is the point of the hump. Otherwise, this valley-hump boundary is shifted leftwards until it reaches 50 % of hump count (hump overall value minus the valley value)[4].

Subsequently, Q^{peak} and Q^{valley} values were determined using the parameters returned by the fit function, i.e. finding its local extremum:

$$Q^{\text{peak}} = \frac{-b}{2a} \pm \sqrt{\frac{\varepsilon_{\text{b}}^2}{4a^2} + \frac{b^2\varepsilon_{\text{a}}^2}{4a^4} - \frac{b}{2a^3}M_{1,2}}. \tag{2.6}$$

Here M is the 3×3 covariance matrix returned by the fit function in use. The uncertainties of parameters are determined also from M as $\varepsilon_{\text{a}} = \sqrt{M_{1,1}}$ and $\varepsilon_{\text{b}} = \sqrt{M_{2,2}}$.

The uncertainty on Q^{peak} in Eq. (2.6) was derived from the general formula for uncertainty propagation, where for a function $f(x)$ with a, b free parameters [19]:

$$\varepsilon_{\text{f}} \approx \sqrt{\left(\frac{\partial f}{\partial a}\right)^2 \varepsilon_{\text{a}}^2 + \left(\frac{\partial f}{\partial b}\right)^2 \varepsilon_{\text{b}}^2 + 2\frac{\partial f}{\partial a}\frac{\partial f}{\partial b}M_{1,2}}. \tag{2.7}$$

The values of the peak and valley in CH at the point of local extremum of the fit needed to be determined for the calculation of v/h as well. For the values, let us call them H and V, Eq. (2.4) was used of course, substituting Q^{peak} and Q^{valley} for Q respectively. The uncertainty on their values $\varepsilon_{\text{HorV}}$ was found based on Eq. (2.7) as:

$$\varepsilon_{\text{HorV}} = \left[Q^4\varepsilon_{\text{a}}^2 + Q^2\varepsilon_{\text{b}}^2 + \varepsilon_{\text{c}}^2 + 2Q^3M_{1,2} + 2Q^2M_{1,3} + 2QM_{2,3}\right]^{\frac{1}{2}}, \tag{2.8}$$

where the variables have the same meaning as in Eq. (2.6).

The example of a fit can be seen in Fig. 2.10. Goodness of fit was judged by its reduced χ^2 (χ^2/ndf) and by its p-value. χ^2/ndf was calculated as:

$$\chi^2/\text{ndf} = \frac{1}{m-3}\sum_{i=1}^{m}\left[\frac{(\text{d}N/\text{d}Q)_i - f(Q_i)}{\varepsilon_{\text{d}N/\text{d}Q,i}}\right]^2, \tag{2.9}$$

with m the number of data points in the fitting range, and $\varepsilon_{\text{d}N/\text{d}Q}$ the error on the data points as defined in Eq. (2.3). For a good fit, this value should be close to 1. If much lower than 1, the uncertainties on data points are overestimated or one is overfitting. If it is much higher than one, the uncertainties are underestimated, or the fit is bad due to wrong theoretical model for example.

The p-values were found using the python's function scipy.stats.chi2.sf [20] as:

$$\text{pval} = 1 - \int_0^{\sum\chi^2} f(x, m-3)\,dx, \tag{2.10}$$

with $\sum\chi^2$ the same sum as in the numerator of Eq. (2.9) and $f(x, m-3)$ the χ probability distribution with $m-3$ degrees of freedom [19]. The p-value of a fit varies from 0 to 1, and gives the probability of finding a worse value of χ^2 than the one obtained in the fitting procedure. It is expected to follow uniform distribution between 0 and 1.

[4]The method of automatic range assignment failed at 9 valleys, 5 CH humps and 4 CCH humps and was corrected by hand in those cases.

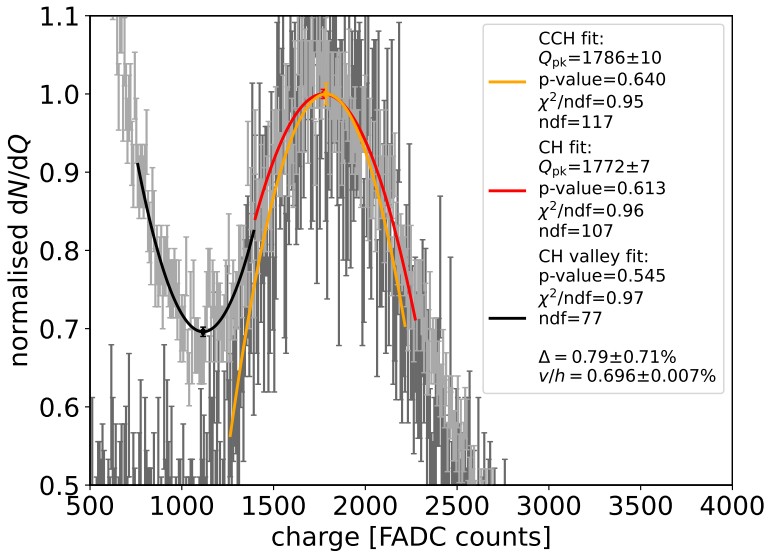


Figure 2.10: Example fit of SD 1819, PMT 1 at time 1328479404 in a zoomed region of muon hump. CH is shown in light grey while CCH is in dark grey. The histograms are normalised by the value of their respective muon peaks found from the fit. CH muon hump and valley fit are shown in red and black respectively, CCH fit is shown in orange. Fit curves span the automatically-found fitting range. The uncertainties of the fit are drawn at the point of the local extremum of the fit curve.

The p-values of all the fits in the data set of Feb. 2^{nd}, and Feb. 9^{th}-10^{th} altogether are shown in Fig. 2.11. A slight tilt observed towards lower p-values is observed, which might indicate the peak is not properly described by a second-degree polynomial (subject for a further study). However, this is not expected to influence the results as the peak is well reconstructed, while the statistical uncertainties on the results might be slightly underestimated.

The fitting range was tuned in the best way possible, requiring certain accuracy of $Q_{\text{CH}}^{\text{peak}}$, as with lowering the fitting range the uncertainty on $Q_{\text{CH}}^{\text{peak}}$ grows. Inspecting the lowest-p-value fits, the fit range appeared acceptable even in those cases and the lower p-values might have resulted from statistically unexpected fluctuations in counting.

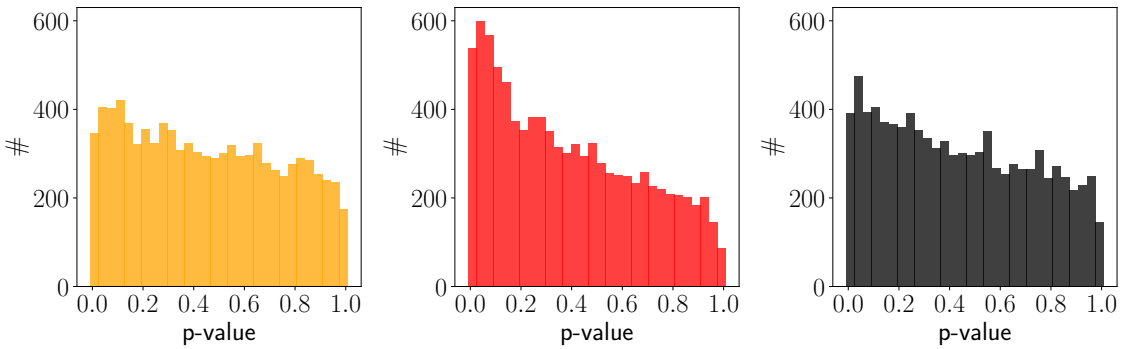
Calculation of Δ and v/h

Δ was then calculated for each PMT and each event of each SD separately according to Eq. (2.2). The uncertainty on Δ, ε_Δ was derived using Eq. (2.7):

$$\varepsilon_\Delta = \sqrt{\frac{\varepsilon^2_{Q^{\text{peak}}_{\text{CCH}}}}{\left(Q^{\text{peak}}_{\text{CH}}\right)^2} + \frac{\left(Q^{\text{peak}}_{\text{CCH}}\right)^2 \varepsilon^2_{Q^{\text{peak}}_{\text{CH}}}}{\left(Q^{\text{peak}}_{\text{CH}}\right)^4}}. \tag{2.11}$$

The valley to hump ratio in CH was calculated as:

$$v/h = \frac{V}{H} \pm \sqrt{\frac{\varepsilon^2_{\text{V}}}{H^2} + \frac{V^2 \varepsilon^2_{\text{H}}}{H^4}}. \tag{2.12}$$



(a) CCH fits. (b) CH hump fits. (c) CH valley fits.

Figure 2.11: p-values of all the fits of the data sets of both tests.

Merging the results from different events

The data of each event of an SD were kept and fitted separate up to this point. Practically, this meant that each event was one instance of the same python class in the code, and the next step was to average the results of Δ and v/h of the SDs present in multiple events, to obtain one Δ and v/h per PMT.

For this, another class which merges the results was created. It uses the method of weighted average, assuming the uncertainty on Δ and v/h is of Gaussian nature. Then, the weighted average values were calculated as [19]:

$$\overline{\Delta} = \frac{\sum_{i=1}^{n} \frac{\Delta_i}{\varepsilon^2_{\Delta,i}}}{\sum_{i=1}^{n} \frac{1}{\varepsilon^2_{\Delta,i}}} \pm \left(\frac{1}{\sum_{i=1}^{n} \frac{1}{\varepsilon^2_{\Delta,i}}} \right)^{\frac{1}{2}} \tag{2.13}$$

with n the number of events when the PMT was present. The same weighted average was used for averaging v/h.

Results of Δ and v/h

Histograms of Δ shown in Fig. 2.12 contain Δ values of PMTs with $v/h < 0.8$ to exclude the SDs where the aging is expected to dominate the Δ results. The data of PMTs 1 & 3 were merged to increase the statistics, as due to geometry these PMTs are expected to have the same Δ. The data of these histograms are predicted to follow a ratio distribution of two correlated non-central normal distributions. The correlations are difficult to predict therefore the shape and moments of such distribution are difficult to estimate. Normal distribution is a reasonable approximation [21] for the purposes of this analysis.

Thus, the expectation values of Δ (to have the conversion factor between the new and the usual offline calibration method) were found by fitting the histograms in the specified range using scipy.stats.norm.fit [22] to fit by a Gaussian. In those ranges, there were 234 Δ values to find $\langle\Delta\rangle_{1,3} = 2.51 \pm 0.08$ and 127 Δ values to infer $\langle\Delta\rangle_2 = 6.62 \pm 0.13$. The uncertainties on μ, σ (the parameters of the fit) quoted in Fig. 2.12 were calculated as $\sigma/\sqrt{n}$ and $\sigma/\sqrt{2n}$ respectively, with n the number of Δ values in the fitted set.

In Fig. 2.13, all the resulting Δ from fit and their weighed averages are shown along with the previously-found expectation values (as red bars). Many of them are inconsistent with the expectation value, hinting at systematic effects of SDs aging distorting their distributions.

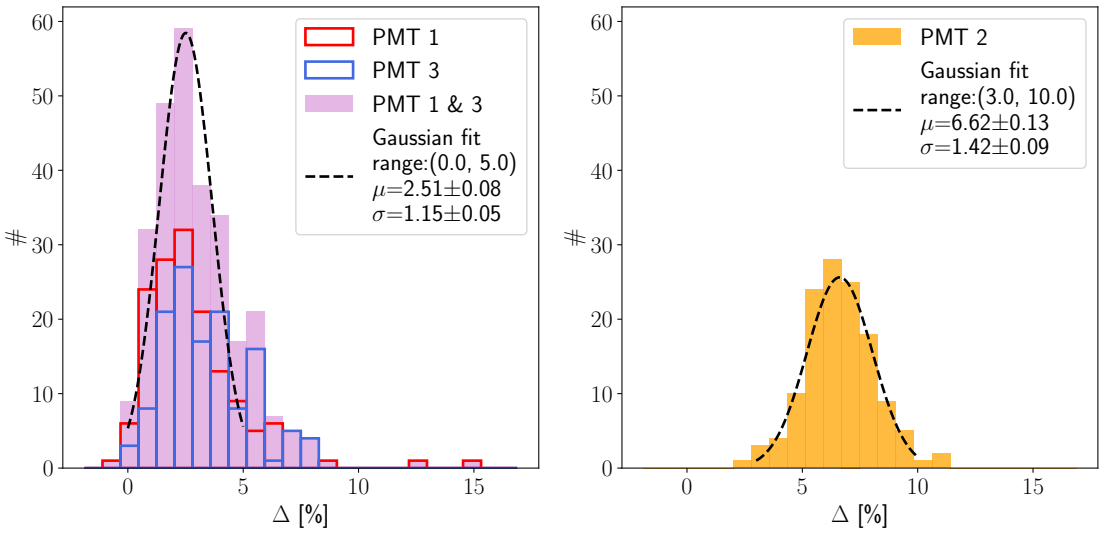
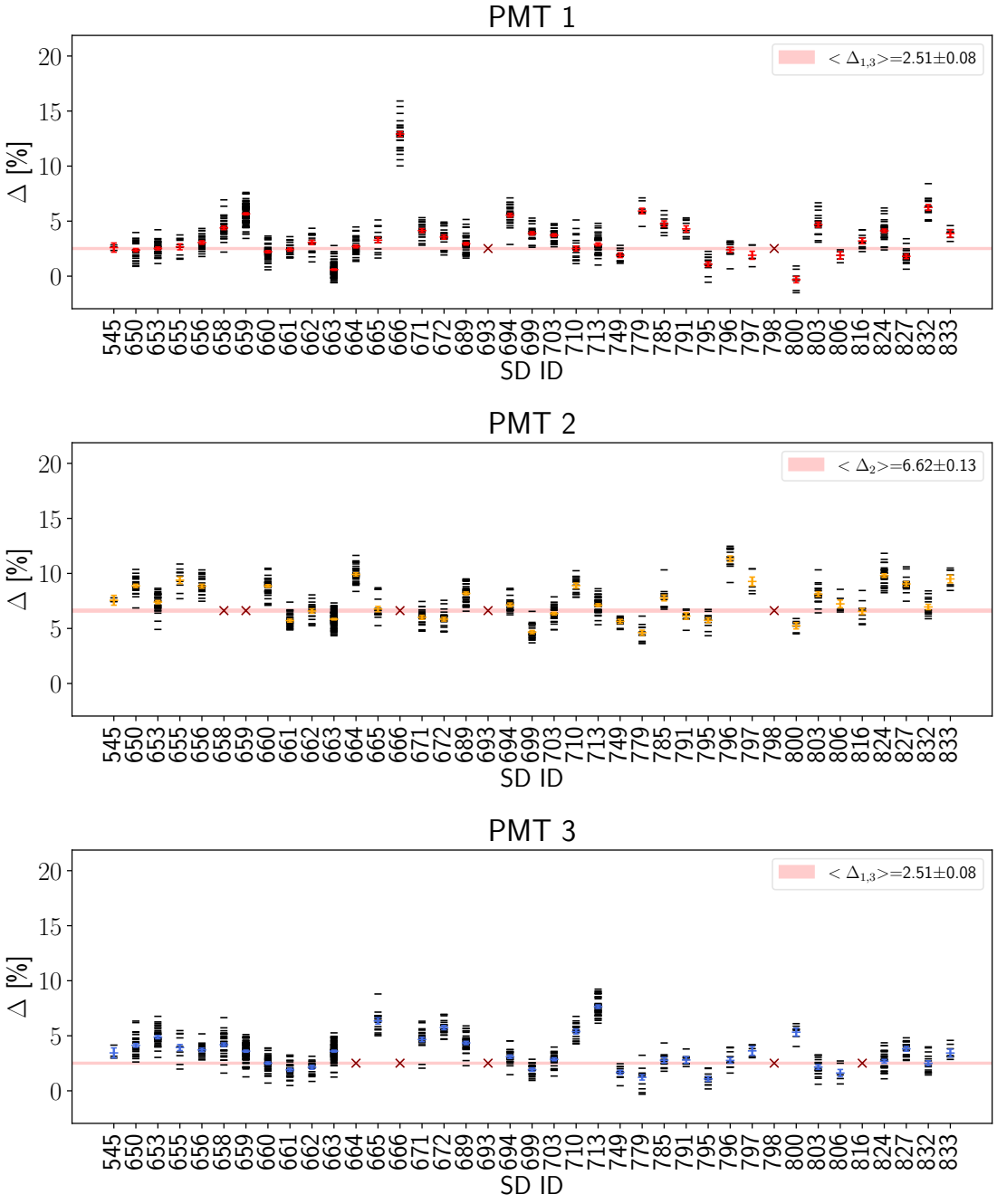
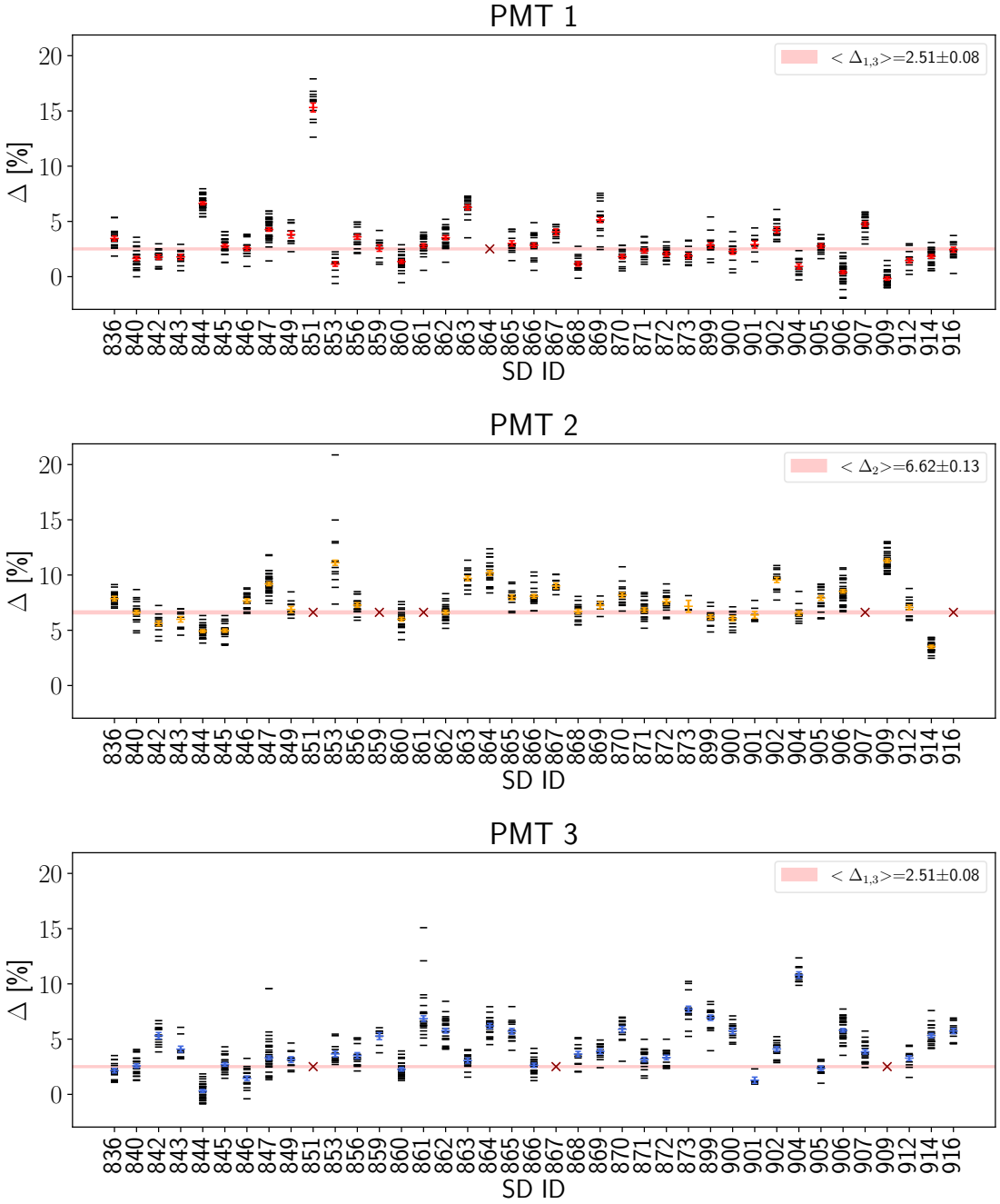


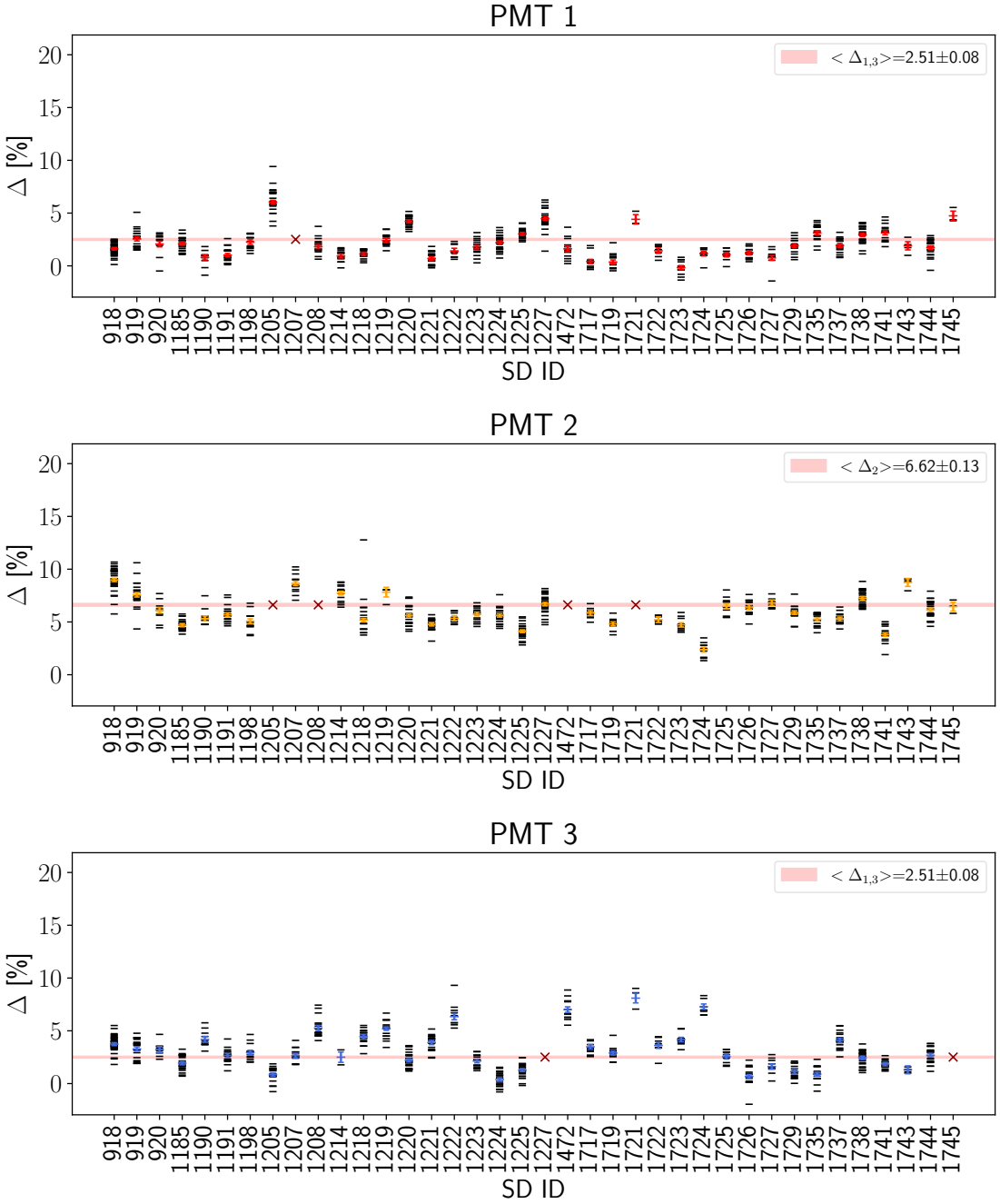
Figure 2.12: Histograms of Δ of PMTs with $v/h < 0.8$. Fit of normal distribution used to find the expectation value of $\langle\Delta\rangle_{1,3}$ over 234 individual Δs and $\langle\Delta\rangle_2$ over 127 Δs is shown by black dashed line only in the fitting range. Parameter μ represents the resulting Δ expectation values.


PMT 1
< Δ1,3>=2.51±0.08
Δ [%]
SD ID
PMT 2
< Δ2>=6.62±0.13
Δ [%]
SD ID
PMT 3
< Δ1,3>=2.51±0.08
Δ [%]
SD ID

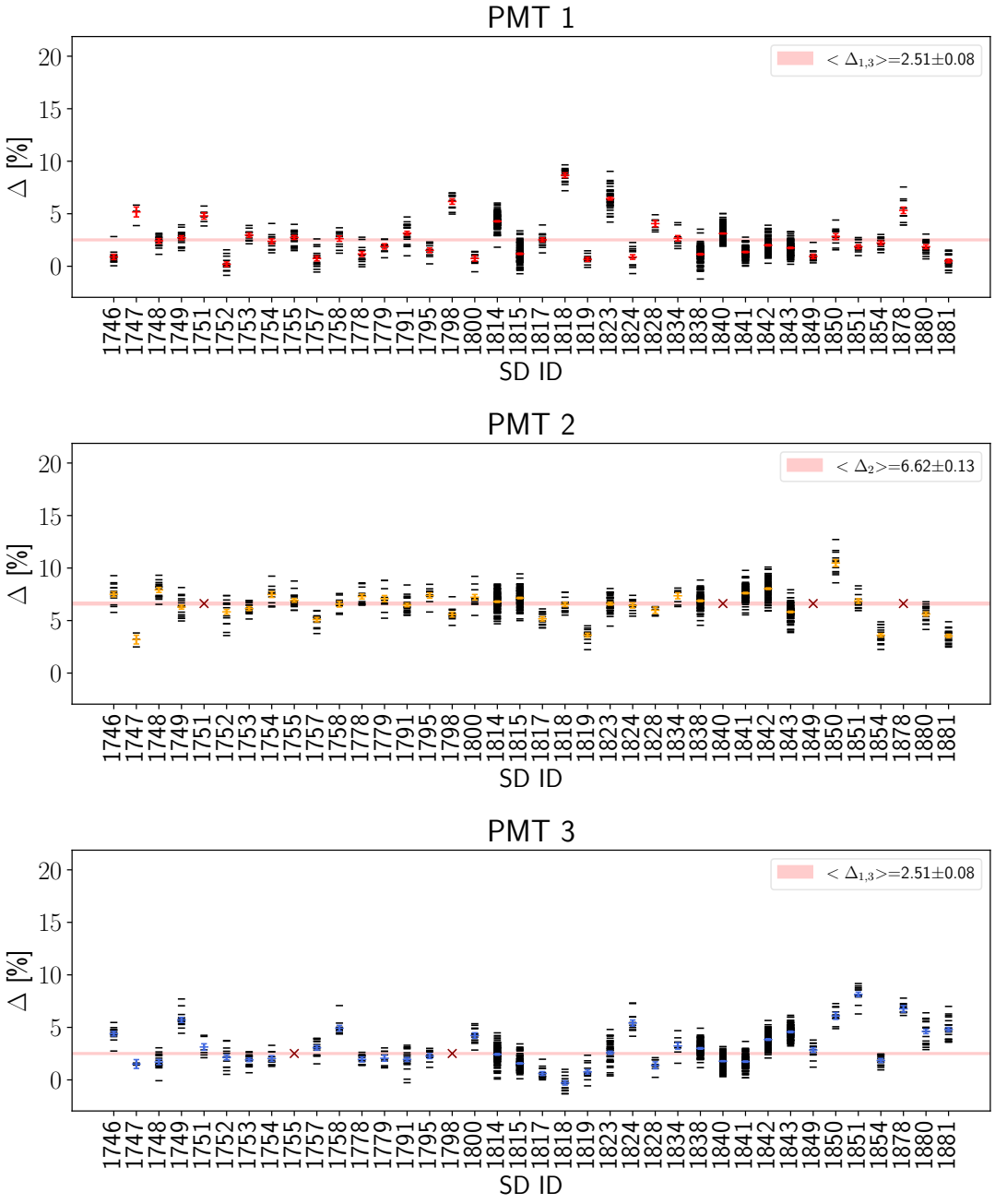
(a)


PMT 1
< Δ1,3>=2.51±0.08
Δ [%]
SD ID
PMT 2
< Δ2>=6.62±0.13
Δ [%]
SD ID
PMT 3
< Δ1,3>=2.51±0.08
Δ [%]
SD ID

(b)


PMT 1
< Δ1,3>=2.51±0.08
Δ [%]
SD ID
PMT 2
< Δ2>=6.62±0.13
Δ [%]
SD ID
PMT 3
< Δ1,3>=2.51±0.08
Δ [%]
SD ID

(c)



(d)

Figure 2.13: Resulting Δ, differences between peaks of conventional charge calibration histogram and coincidence calibration histogram as defined in Eq. (2.2). Individual fits are shown as black bars, with their weighted average as errorbar in colour. Dark red × denotes PMTs without data or fit. The expectation value of Δ is shown as red bar.

The most obvious outliers were inspected. With $\Delta = 20.9 \pm 1.9\,\%$, the highest-$\Delta$ outlier was PMT 2 in SD 853 at time 1327848949, shown also in Fig. 2.14. The reason for this behaviour was very clear: With $v/h = 0.95$, its muon peak has considerably merged with the EM background. This distorted the peak much leftwards, resulting in higher difference between coincidence muon peak and the distorted muon peak of the usual calibration, thus higher Δ. The position of the coincidence peak is expected to be very stable. This stresses how important it is to remember the aging effect when calibrating the event signals offline by usual (non-coincidence) method. Using the same conversion factor between $Q_{\mathrm{CH}}^{\mathrm{peak}}$ and VEM for all the SD stations might not yield accurate calibrated signals anymore.

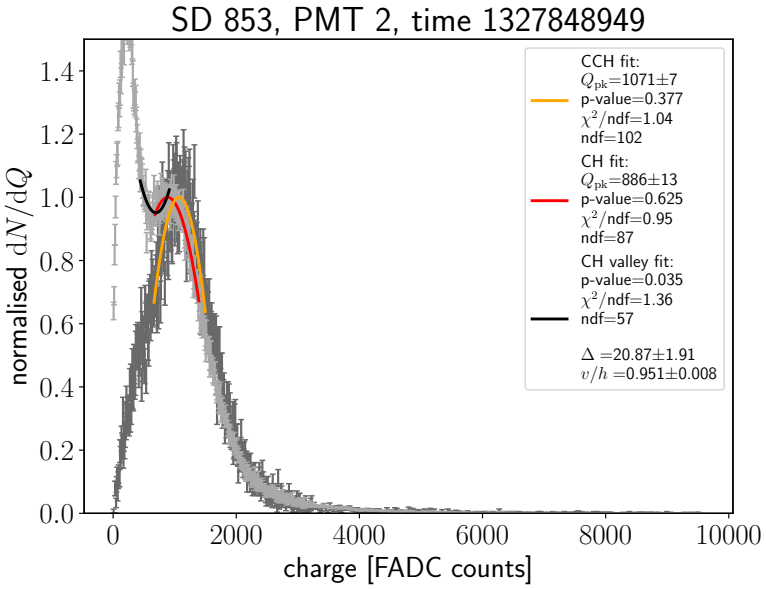


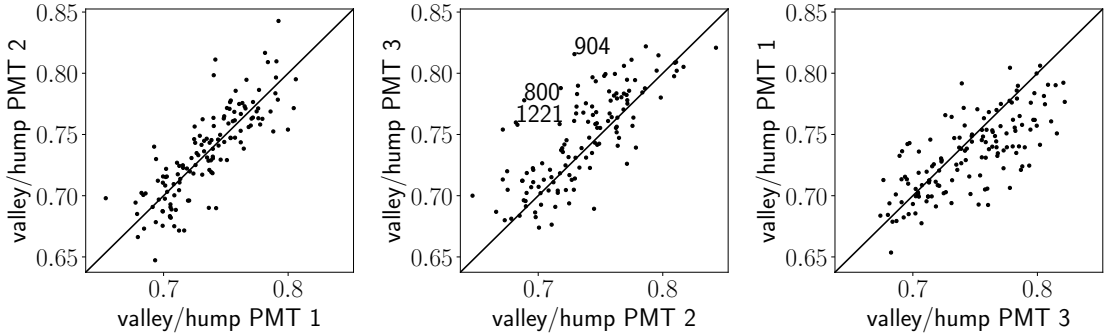
Figure 2.14: The most obvious fit outlier in the plot of all Δ (Fig. 2.13b). Clearly, the high $\Delta = 20.9 \pm 1.9\,\%$ is due to merging of the muon peak into the EM part, which distorts the peak leftwards, while CCH should remain fixed in its position on the x-axis.

As mentioned, the aging results primarily from higher losses of Cherenkov light in WCD. Higher v/h is therefore expected in all three (two) working PMTs of the same tank. To verify this, one may look at v/h of one PMT plotted against v/h of another PMT of the same SD station, as shown in Fig. 2.15a. As expected, the data points lie reasonably close to $y = x$ line. The outliers might be explained by the worse optical coupling of a PMT, which results in different v/h of one PMT compared to the other two (one) PMTs of the same SD station.

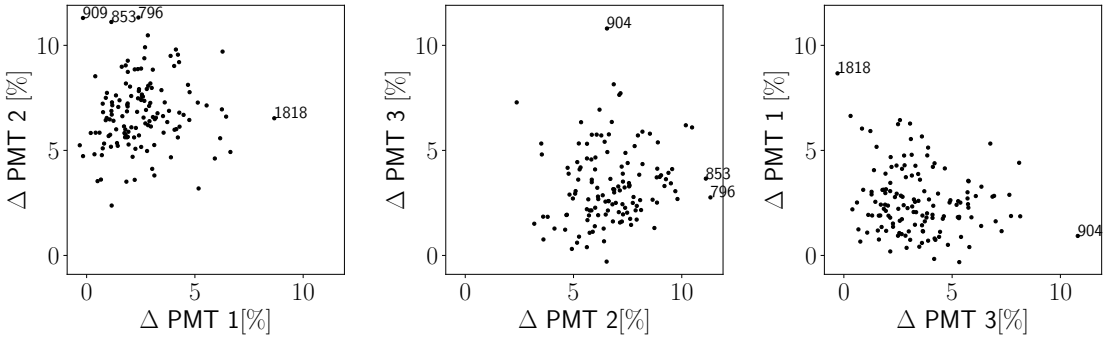
The same approach can be used to compare Δ of the two different PMTs of the same tank, shown in Fig. 2.15b. In this case, only the outlier SD 904 is observed in both $(v/h)_i$ vs. $(v/h)_j$ and Δ_i vs. Δ_j plot. In the other outliers, no explanation was available. The fits were inspected and it was confirmed that the outlier values were not caused by bad fits.

Summary of data cuts

The summary of the amount of data used in the CCH (Δ) analysis after specified cuts, and to infer the presented results is provided in Table 2.1.



(a) v/h vs. v/h of two PMTs of the same SD station.



(b) Δ vs. Δ of the PMTs of the same SD station.

Figure 2.15

PMT	1	2	3
All CCH	3496	3496	3496
Paired with CH	3398	3398	3398
SD ID > 100	3378	3378	3378
$(\mathrm{d}N/\mathrm{d}Q)_{\max}$ at bin num. >36 or $\langle \mathrm{d}N/\mathrm{d}Q \rangle > 0.05$	3314	3024	3159
Non-faulty PMTs by hand	3314	2827	3148
Successful fits	3305	2826	3145
	Averaged fits		
	150	136	142
$v/h < 0.8$	148	130	131
used to find $\langle \Delta \rangle$	130	127	104

Table 2.1: Summary of the numbers of analysed CCH and CH data. The table shows also the number of Δ results included in the previous figures.

Correlation between Δ and v/h

As mentioned in the previous section, correlation between v/h and Δ is expected. To explore this more quantitatively in the data, the Pearson correlation coefficient r was used. It is defined

as [19]:

$$r = \frac{\sum_{i=1}^{n} \left[(v/h)_i - \overline{v/h}\right] \left[\Delta_i - \bar{\Delta}\right]}{\left[\sum_{i=1}^{n}((v/h)_i - \overline{v/h})^2)\right]^{1/2} \left[\sum_{i=1}^{n}(\Delta_i - \bar{\Delta})^2\right]^{1/2}} \tag{2.14}$$

with bar denoting the mean values and n the total number of PMTs with calculated Δ and v/h. The uncertainty on r was calculated as:

$$\varepsilon_{\mathrm{r}} = \frac{1 - r^2}{\sqrt{n}}. \tag{2.15}$$

Two PMTs, namely PMT 1 in both SD 666 and SD 851 had very high Δ and low v/h and were obvious outliers in the correlation calculations. It is worth noting, that these were the only two PMTs with high baseline in CCH. They were therefore rejected in this correlation study and will require further investigation.

In Figs. 2.16a to 2.16c, individual Δ and v/h of each PMT can be seen. In Figs. 2.16d and 2.16e, the profile plots of the data, binned based on v/h are illustrated. The Pearson correlation coefficient is quoted in the legends. To increase the statistics, the results of PMT 1 & 3 were merged in the profile plot. With $r = 0.39 \pm 0.05$ in case of PMT 1 & 3, and $r = 0.75 \pm 0.04$ in PMT 2 results, a high correlation was shown (where higher r means higher positive correlation). The higher the v/h, the higher the difference between the conventional calibration muon peak and coincidence calibration muon peak, with the effect more prominent in PMT 2. Furthermore, in profile plot of PMT 1 & 3 a plateau can be seen at low v/h. This was also expected due to EM peak influencing Δ more if the EM and muon peak are more-severely merged.

2.5.2 Event signal analysis

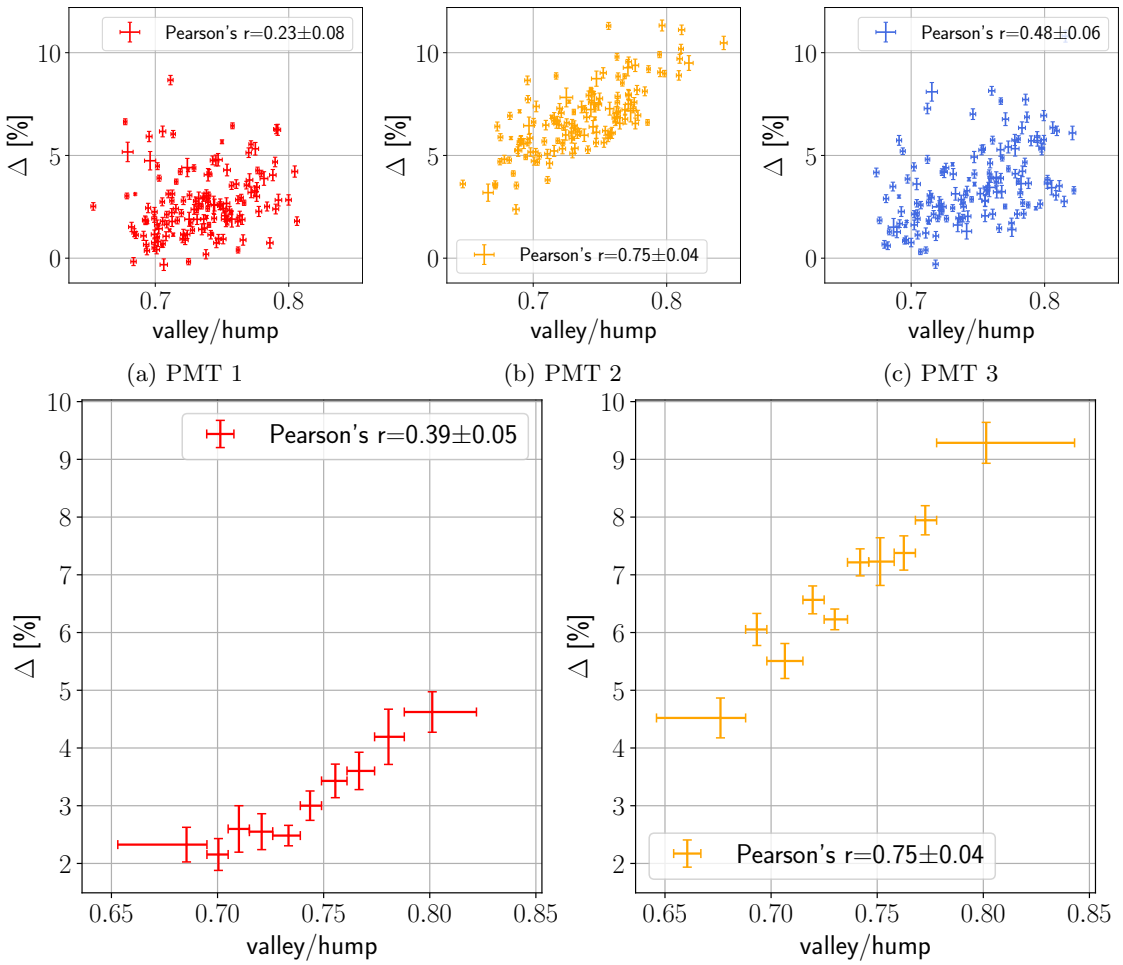
Having estimated CCH and CH muon peaks, the next step to verify the stability of CCH calibration was to analyse the event signals coming from the time period of the second test, more exactly between 1644430994 and 1644497773 in UTC time. These are all the events that passed the third level trigger of the SD array. The traces were extracted from unreconstructed SD event files and provided in the form of one ASCII file per SD per PMT[5]. Each line of the file contained event ID and the UTC time of the event separated by space. The rest of the line were FADC counts/sampling frequency of each event of length 2048 time bins (17.067 μs as shown in Fig. 2.17.

Data processing and exclusion

First, all the events, which do not have a successful CH or CCH fit were discarded. Then, all the saturated traces were discarded. A saturated event has an FADC count of 4095 in at least one time bin. There were 455 saturated traces.

After then, an algorithm for the subtraction of the baseline in all the traces was developed. The baseline can be seen in the raw trace of Fig. 2.17 as a constant value of roughly 250 FADC counts. Among all the analysed PMTs, the baseline values were between 210 and 290 FADC counts as shown in Fig. 2.18. There were cases however, when after the event, the baseline returned to a different value. Or in some traces, an accidental muon at the start or the end of the trace occurred, which then distorted the mean of the first (or last) bins of the trace. Those

[5]by Dr. Mauricio Suárez Durán



(a) PMT 1

(b) PMT 2

(c) PMT 3

(d) PMT 1 & 3. The data are binned on v/h, the averages of the bins are shown.

(e) PMT 2. The data are binned on v/h, the averages of the bins are shown.

Figure 2.16: Δ as function of v/h. (a)-(c) show individual PMT results. (d),(e) show binned results. Δ of PMT 1 of SDs 666 and 851 were excluded from these plots as outliers requiring further investigation.

are supposed to be empty. To simplify the method of the baseline subtraction, the decision was made to exclude all the traces where either of this happens.

The method to subtract the baseline first checked if the baseline was the same in the beginning and at the end of the trace. It looked at the first and the last 100 bins of the trace separately. It found the mean and the standard deviation of each of these segments. Then the method chose the smaller of the two standard deviations σ_{min}. If the difference between the mean of the first and the mean of the last 100 bins was more than $2\sigma_{\text{min}}$, the trace was discarded.

Otherwise the baseline of one trace was found as the mean of the first and the last 100 bins, so 200 bins in total, and its value was subtracted from all the bins of the trace. The value of the standard deviation of the baseline (of the 200 bins in total) σ_{B} got stored to be used in the

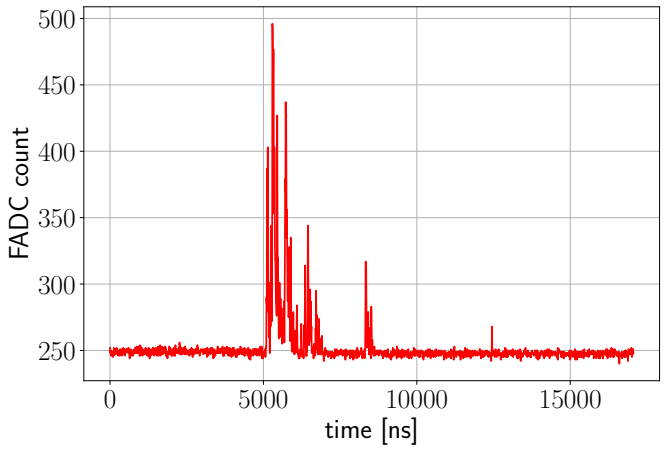


Figure 2.17: A raw trace of SD 1819, PMT 1, UTC time 1644468290.

following step - the summation of the trace counts.

At this step, one mean baseline value (the mean of the baseline of all traces of given PMT) and its standard error per PMT were plotted in the histograms of Fig. 2.18, and for individual SD IDs in Fig. 2.19.

The summation of traces

To compare the calibrated signals between PMTs of the same tank, it was not of much concern if spurious signals such as accidental muons not belonging to an event ended up in the sum. What was important was to sum the same time interval in all three (two) PMTs of one SD. To obtain the total signal, the methods for identification of the time range of interest in the traces were adapted from [23] and applied.

At first, the traces of PMTs of each station were analysed individually. For each trace, candidate signal regions were found where in the trace, bin counts were more than $5\sigma_{\mathrm{B}}$. Then the method started from the candidate segment, which contained a bin with the most counts (thus was automatically a signal segment). The segments were merged consecutively with the neighbours to the left and to the right as long as the break between them was less than 60 bins (500 ns). For each PMT, the left and the right boundary of the final signal segment got stored.

Finally, the minimum time to the left and the maximum time to the right of the trace was found as the boundaries of all the PMTs of the SD. Five bins were added to the left to ensure that the beginning of the signal was captured. If the resulting interval was less than 30 bins wide, it was made to be at least 30 bins. This same interval was then applied to all the three traces of the SD. An example of the resulting interval to be summed is shown in Fig. 2.20.

Let us look at the outlier baselines $\overline{\sigma_{\mathrm{B}}} > 3$ FADC counts/120 MHz in Fig. 2.19. For each of these PMTs, the highest-σ_{B} trace was inspected to verify the signal-search method. This is shown in Fig. 2.21.

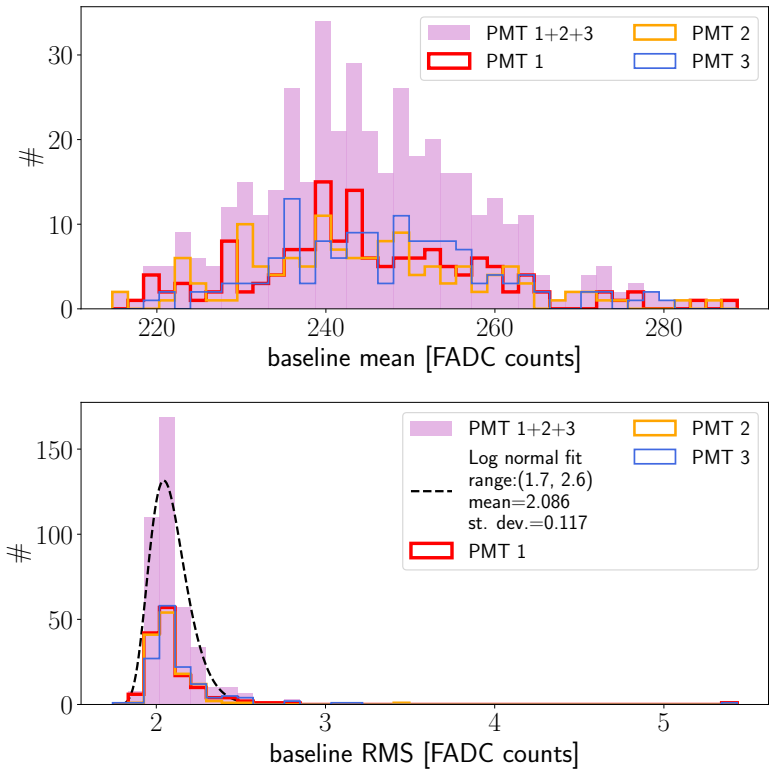


Figure 2.18: Histograms of baseline means averaged among events and their averaged standard deviations (RMS) in the traces data set. There is one baseline per PMT. The baselines, which made the cut, had successful CH and CCH fits, they were not saturated, and they were consistent before and after the event. These histograms are one of the first insights into the upgraded electronics of the SD stations. Baseline RMS histogram was fitted by log normal to find the expectation value of 2.086 FADC counts/120 MHz with the standard deviation of 0.117 in the new electronics.

Converting to three signal types

Finally, three signal types were defined to be able to compare the calibrations. They were named S_{CH}, S_{CCH}, and S_{CO}. S_{CH} was calculated as:

$$S_{\mathrm{CH}} = \frac{\sum FADC}{Q^{\mathrm{peak}}_{\mathrm{CH}}}, \tag{2.16}$$

with $\sum FADC$ the sum of trace over interval containing signal identified as described in previous sub-subsection. $Q^{\mathrm{peak}}_{\mathrm{CH}}$ was the peak of the muon hump of the conventional calibration histogram found by fitting. S_{CCH} was found similarly, just dividing the sum of trace by $Q^{\mathrm{peak}}_{\mathrm{CCH}}$ instead of $Q^{\mathrm{peak}}_{\mathrm{CH}}$.

For the S_{CO}, the idea was to recover the S_{CH} from S_{CCH} approximately, by using the constant conversion factor (expectation value of Δ) across all the PMTs of the same number i ($i = 1, 2, 3$):

$$S_{\mathrm{CO}} = \frac{1 + \langle \Delta \rangle_i}{Q^{\mathrm{peak}}_{\mathrm{CCH}}} \sum FADC. \tag{2.17}$$

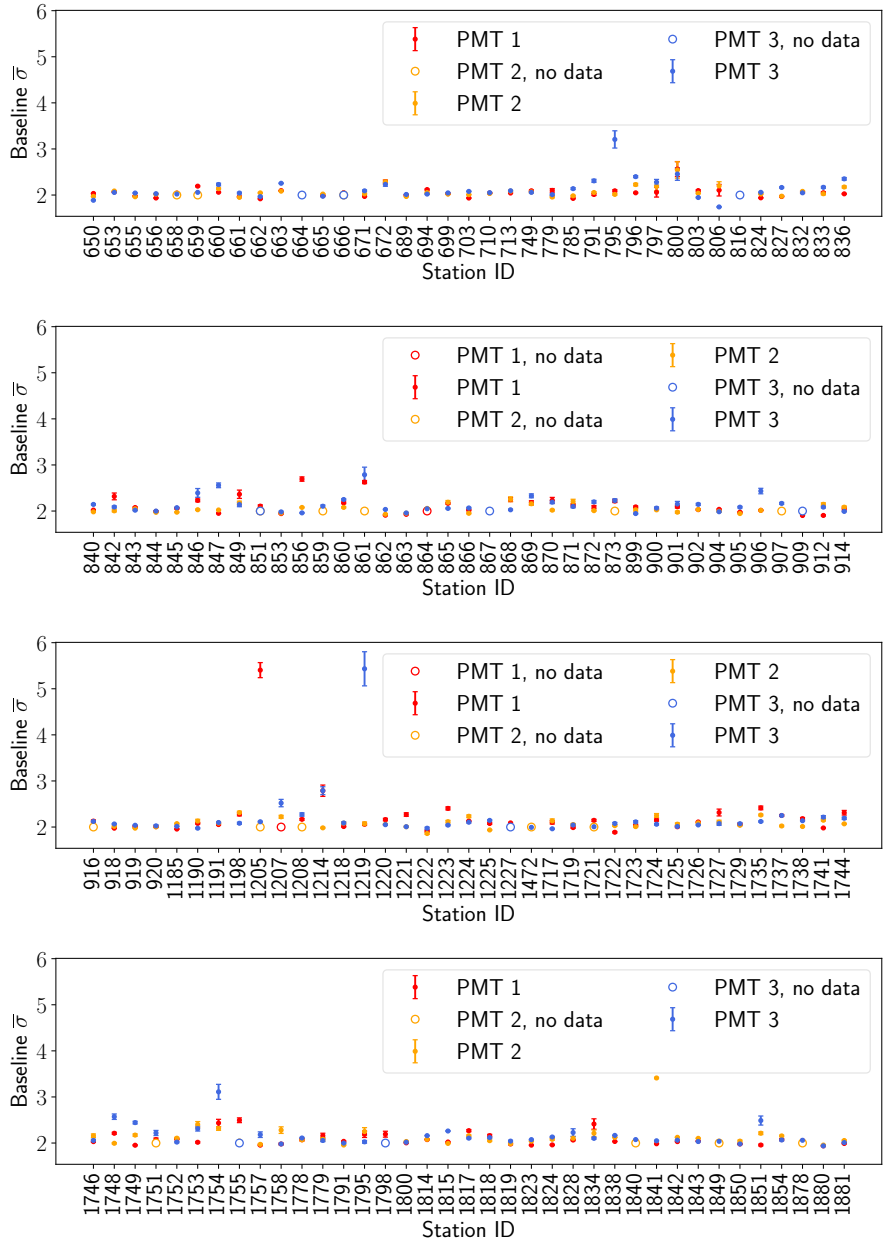
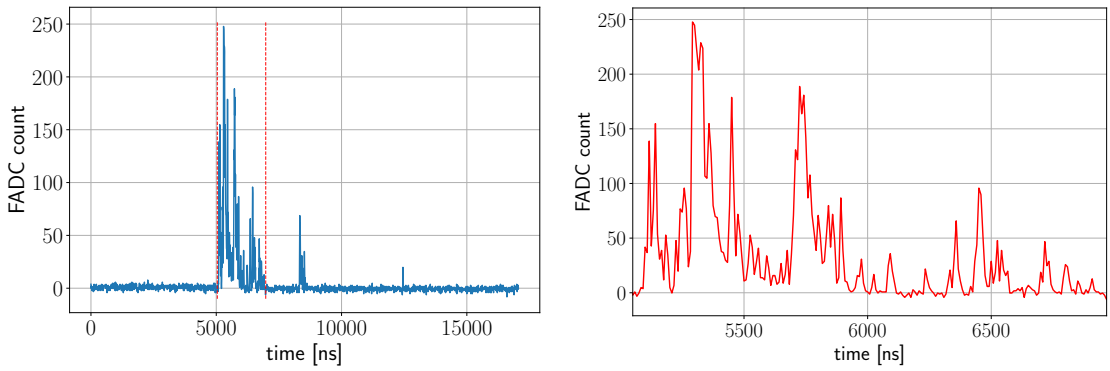


Figure 2.19: The averages of baseline RMS with errorbars denoting their standard error on the mean in each analysed SD. The PMTs that have not participated in any event are shown as open symbols.

S_{CH}, S_{CCH}, and S_{CO} were approximately in units of VEM, only a constant conversion factor was omitted.[6]

[6]The uncertainty on S could not be determined for the following reason: The uncertainty is generally dominated by the arrival direction of the event [23] and with events not having been reconstructed, this information was not



(a) (b)

Figure 2.20: A trace of SD 1819, PMT 1, UTC time 1644468290 with subtracted baseline. a) Summation boundaries are shown as red vertical lines. b) Zoom into the summation region of the trace.

The behaviour of the differences between PMT signals $(S_i - S_j)/ < S_{i,j} >$ with $i, j = 1, 2, 3$ and $i \neq j$ was studied. Especially it was their distributions, which show if the coincidence calibration is more or a least as stable as the conventional calibration. For the studies looking at standard deviation (RMS) specifically, the difference was divided by $\sqrt{2}$ to account for the two Gaussian distributions in this equation.

Results

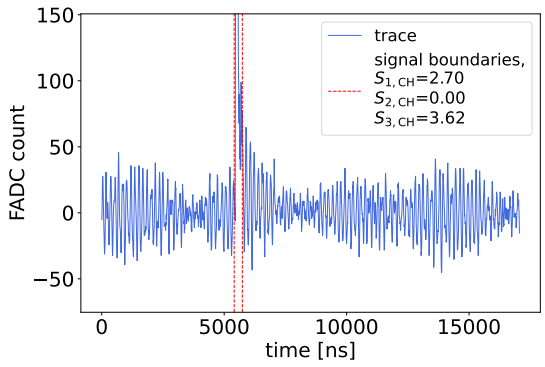
Only the events with all 3 PMT signals present in an SD were included in these results. The lower threshold on the average $< S_{\mathrm{CH}} >$ of the three PMTs was set to 3. This gave 1475 events in different SDs in total. In events with particular direction, i.e. the cone of Cherenkov radiation being directed at one of the PMTs, the signal in this PMT is much larger than the signals in the other two PMTs, as shown in Fig. 2.22c. These cases can bias the analysis significantly and the cut of $S_i^{\mathrm{CH}}/(S_j^{\mathrm{CH}} + S_k^{\mathrm{CH}}) < 1.5$ (with i, j, k the permutations of $1, 2, 3$) was applied. The events before and after the cut are shown in Fig. 2.22. The final number of analysed events was then 1474.

Final cut was done on the outliers of $-1 < (S_1^{\mathrm{CH}} - S_2^{\mathrm{CH}})/\langle S_{1,2}^{\mathrm{CH}} \rangle < 1$ which biased the analysis significantly in all three signal types. The cut is shown in 2.23, in this case only 2 events were discarded.

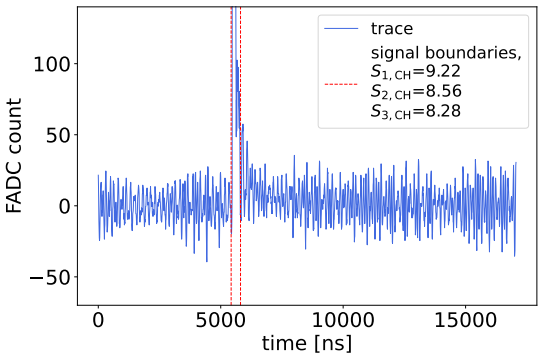
To verify the signal selection, we may look at the relative difference between each PMT and the average of the PMT signals, $S_i/\langle S_{1,2,3} \rangle - 1$ in %, as a function of $\log_{10}\langle S_{1,2,3}^{\mathrm{CH}} \rangle$, i.e. the average signal charge in Fig. 2.24. What we expect is the means of signals from the usual calibration to be roughly zero. The means of S_{CCH} signals should be smaller compared to the usual calibration by approximately $\langle \Delta_i \rangle$, as the muon hump from coincidence calibration is at higher charge due to detector geometry by approximately the value of $\langle \Delta_i \rangle$. The rescaled signals of the coincidence calibration, S_{CO}, should then return to the values of approximately the usual calibration. Fig. 2.24 confirms this, remembering that $\langle \Delta \rangle_{1,3} = 2.51 \pm 0.08$ and $\langle \Delta \rangle_2 = 6.62 \pm 0.13$.

The histograms of the relative signal differences between PMTs in the same SD and the same event are shown in Fig. 2.25. Means and standard deviations of the distributions of both methods

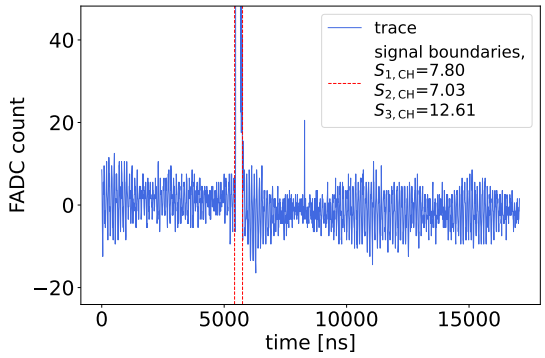
available.



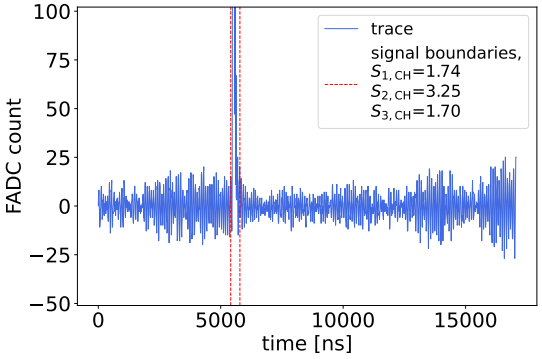
(a) SD 1205, PMT 1, UTC 1644488155, $\sigma_{\mathrm{baseline}} = 15.07$



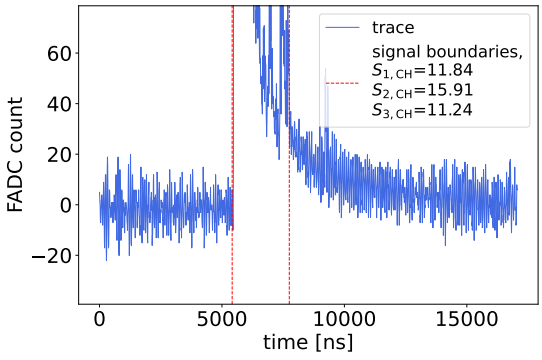
(b) SD 1219, PMT 3, UTC 1644487982, $\sigma_{\mathrm{baseline}} = 14.00$



(c) SD 1754, PMT 3, UTC 1644489195, $\sigma_{\mathrm{baseline}} = 4.82$



(d) SD 795, PMT 3, UTC 1644454761, $\sigma_{\mathrm{baseline}} = 10.21$



(e) SD 1841, PMT 2, UTC 1644489700, $\sigma_{\mathrm{baseline}} = 7.87$

Figure 2.21: Traces with the highest baseline RMS found amongst the PMTs with $\overline{\sigma_{\mathrm{B}}} > 3$. The method to find signal boundaries worked reliably in these cases, with the signal boundaries denoted by red bars. Signals in each PMT are quoted in the legends to verify that the high baseline fluctuations had only a small influence on the resulting signals.

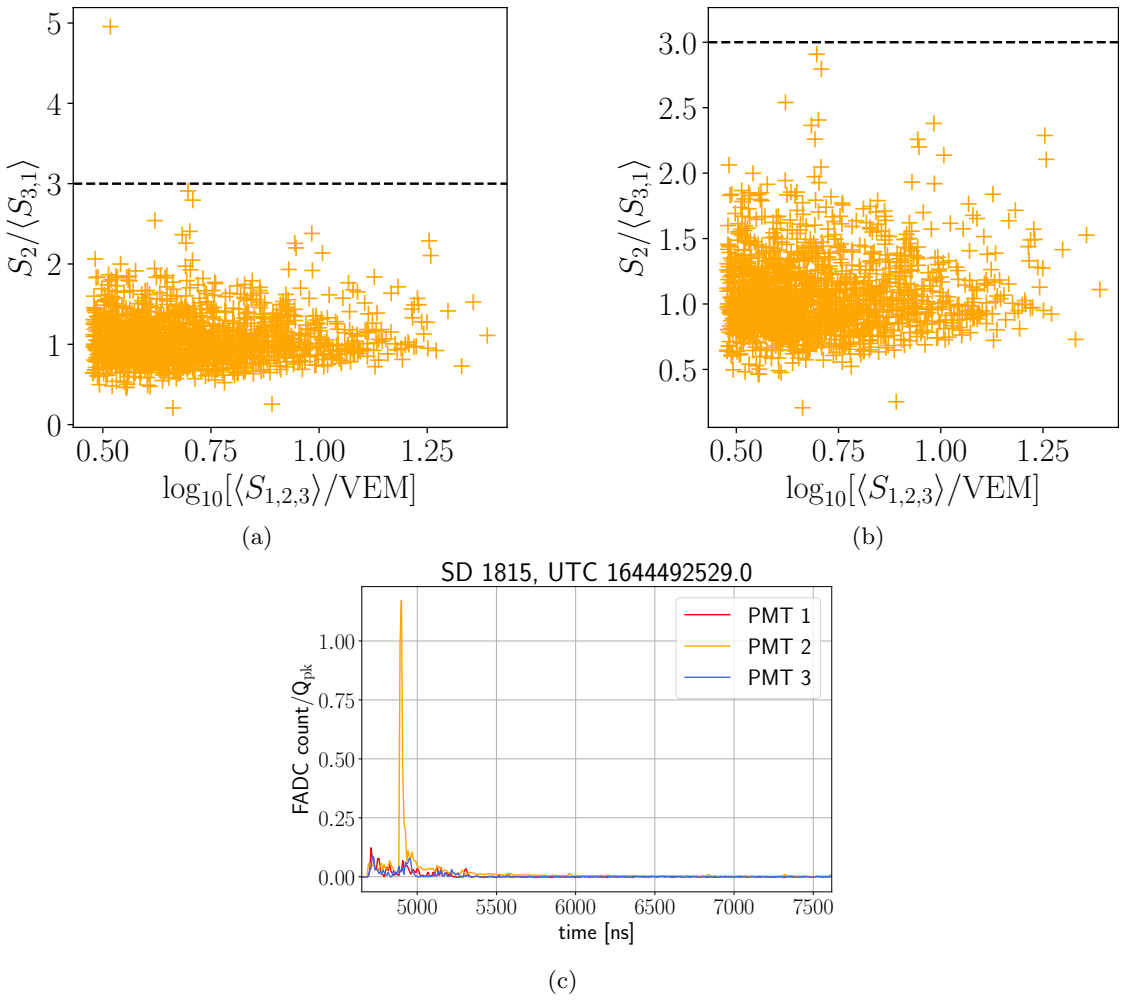


Figure 2.22: a) CH signals before the $S_i^{\mathrm{CH}}/\langle S_{j,k}^{\mathrm{CH}}\rangle < 3$ data cut. b) shows the CH signals after the cut was made. In c), there is the event trace that was eliminated by the cut. A stricter cut can be applied in future to eliminate also the clear outliers above 2.

are approximately the same, which tells us that the new calibration method performs at least as good as the usual offline calibration method.

Thanks to the mirror symmetry between PMT 1 & 3, it is possible to compare the relative signal differences of these two PMTs using S_{CCH}, thus the signals not rescaled via $\langle\Delta_i\rangle$. It can be interesting to look at the stability of the new and the usual calibration using different lower-charge thresholds, as shown in Fig. 2.26. The means and standard deviations of the distributions are again very similar in all cases. With the threshold of 10 VEM the mean of coincidence calibration is hinting towards a closer value to the zero as opposed to the usual calibration. However with this amount of statistics, it is not possible to make any certain conclusions about the improvements of the new calibration for the highest-magnitude signals.

Instead of applying the lower-magnitude threshold on the signals, we may look at the RMS (standard deviation), i.e. the stability of the signals binned as the function of average CH signal,

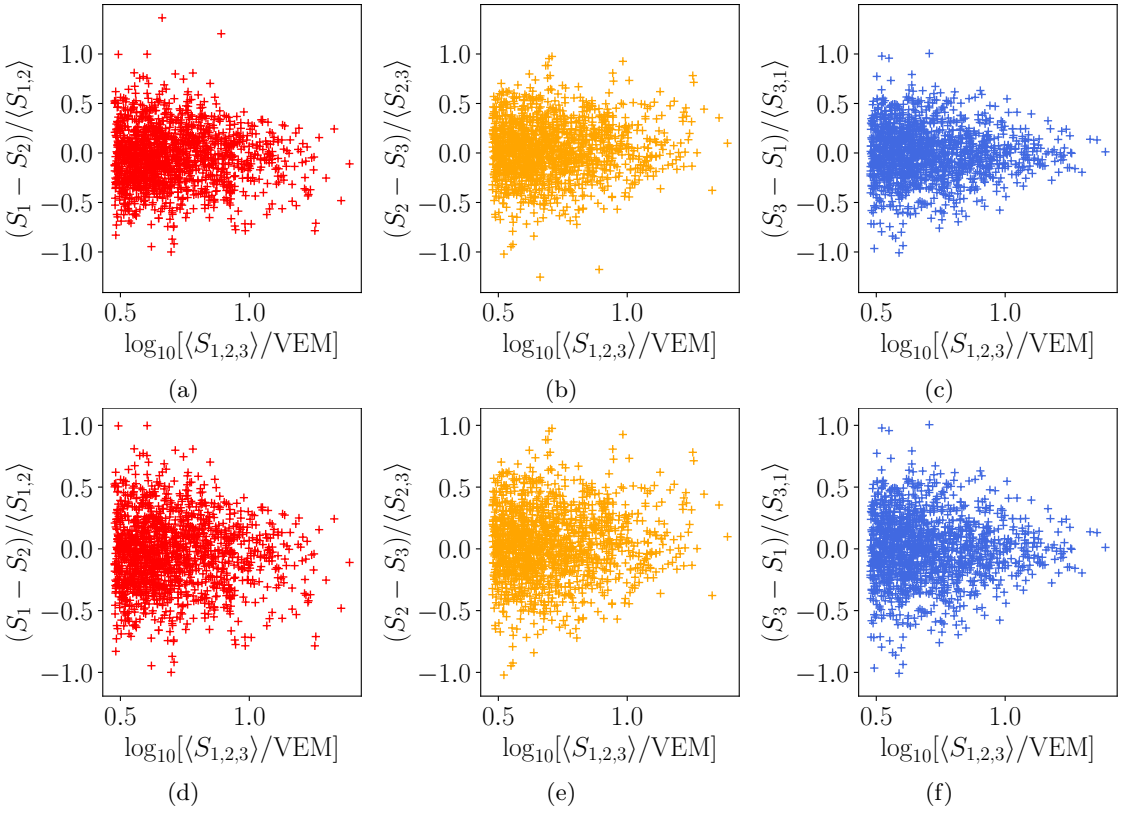
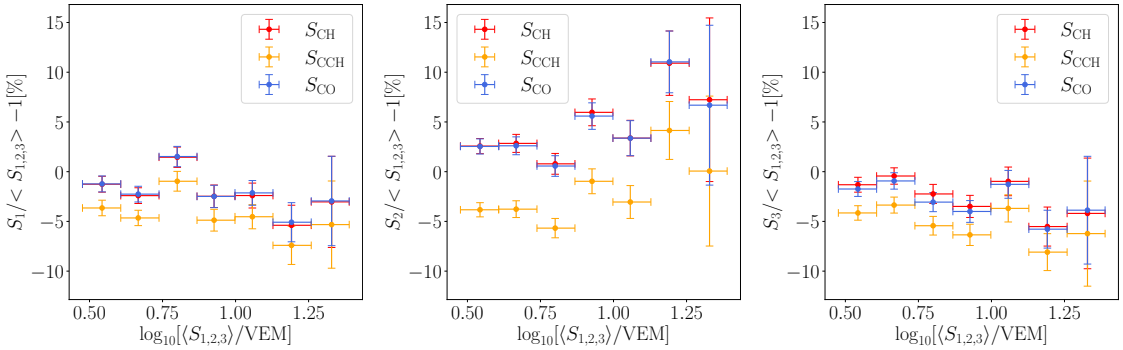


Figure 2.23: a-c) CH signals before the cut on $-1 < (S_1^{\mathrm{CH}} - S_2^{\mathrm{CH}})/\langle S_{1,2}^{\mathrm{CH}}\rangle < 1$. d-f) show the CH signals after the cut was made.

or its decadic logarithm again. This is shown in Fig. 2.27. Comparing primarily S_{CH} and S_{CO} signals, we see slight improvement in PMT 1 & 2, and PMT 2 & 3 towards higher signal charges. The differences in PMT 3 & 1 are negligible. It is obvious however, that the differences between the resolution of the usual and the new offline calibration are still consistent. Thus to prove the improvement in resolution, it will be necessary to collect much more events containing the coincidence histograms to reduce the uncertainties in Fig. 2.27.

Lastly, let us look at the results of PMTs with $v/h > 0.8$, i.e. those that aged significantly and for which the new calibration brings the hope of more reliable muon hump estimation. The events were split on those where both PMTs have $v/h < 0.8$ and on those with at least one PMT having $v/h > 0.8$. This is shown in Fig. 2.28. The plots present the comparison of relative differences between PMTs using the conventional (CH) and the coincidence calibration (CO). The improvement of the new calibration would manifest as the black triangles being closer to the horizontal line, which would mean smaller differences between PMTs in the CO signals. This can be seen in the first quadrant of Fig. 2.28c, however the third quadrant of this figure contains fewer, but still some events which contradict this result. Again, higher number of events containing CCH will be required to fully explore if the new calibration brings significant improvement to aging stations.



(a) PMT 1 (b) PMT 2 (c) PMT 3

Figure 2.24: Signal at one PMT divided by the average signal of all the three PMTs as the function of decadic logarithm of the average CH signal to verify the signals analysis. The difference between S_{CH} and S_{CCH} corresponds to the expectation values of Δ found previously, as it should. Rescaled S_{CO} signals then return to roughly the S_{CH} values in agreement with the purpose of this signal type.

Summary of data cuts

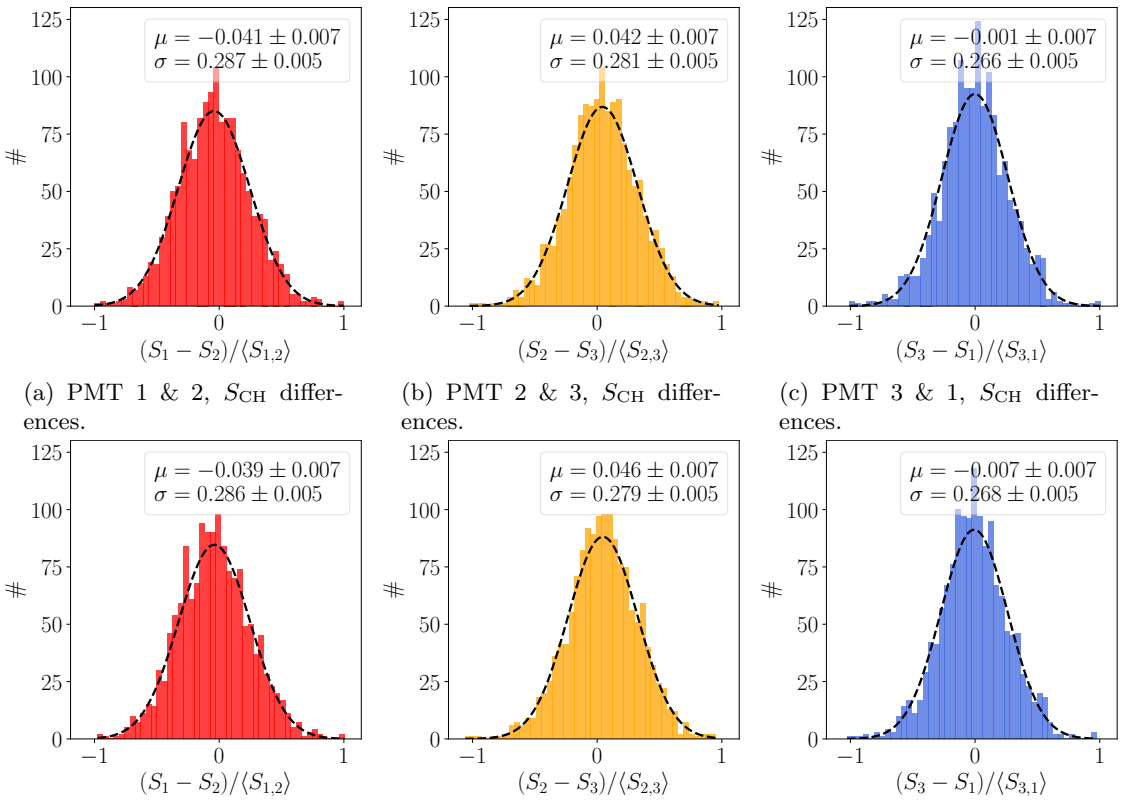
The summary of the above-mentioned data cuts and the amount of data in each of the plots is shown in Table 2.2.

PMT	1	2	3
Successful CH & CCH fit	3107	2655	2955
Non-saturated	2945	2518	2799
Baseline consistent at both ends	2850	2434	2707
All 3 PMTs present & $\langle S_{CH}\rangle_{1,2,3} > 0.05$		2214	
$\langle S_{CH}\rangle_{1,2,3} > 3$		1475	
$S_{CH,i}/(S_{CH,j} + S_{CH,k}) < 1.5$		1474	
$-1 < (S_{CH,1} - S_{CH,2})/\langle S^{CH}_{1,2}\rangle < 1$		1472	
$\langle S_{CH}\rangle_{1,2,3} > 6$		439	
$\langle S_{CH}\rangle_{1,2,3} > 10$		112	

Table 2.2: Summary of the number of events used in each step of the signal analysis.

2.6 Discussion

All the test data of coincidence calibration available until mid-July have been analysed. Overall, the relative difference between the charge of the muon peak between the coincidence and the usual calibration was found for 428 PMTs of the SD-1500 array. Strong positive correlation was shown between these relative differences and the PMT's v/h (valley/hump) ratio characterising the severity of the merging of the muon and EM peak in the usual calibration histogram: The higher v/h, the more is the muon peak of the usual calibration shifted leftwards and the higher the relative difference is. After a careful selection, the common expectation values of the relative differences were found to be $\langle\Delta\rangle_{1,3} = 2.51 \pm 0.08$ for PMT 1 & 3, and $\langle\Delta\rangle_2 = 6.62 \pm 0.13$ for PMT 2. These are borderline-consistent with the results from simulations of the previous



(a) PMT 1 & 2, S_{CH} differences.

(b) PMT 2 & 3, S_{CH} differences.

(c) PMT 3 & 1, S_{CH} differences.

(d) PMT 1 & 2, S_{CO} differences.

(e) PMT 2 & 3, S_{CO} differences.

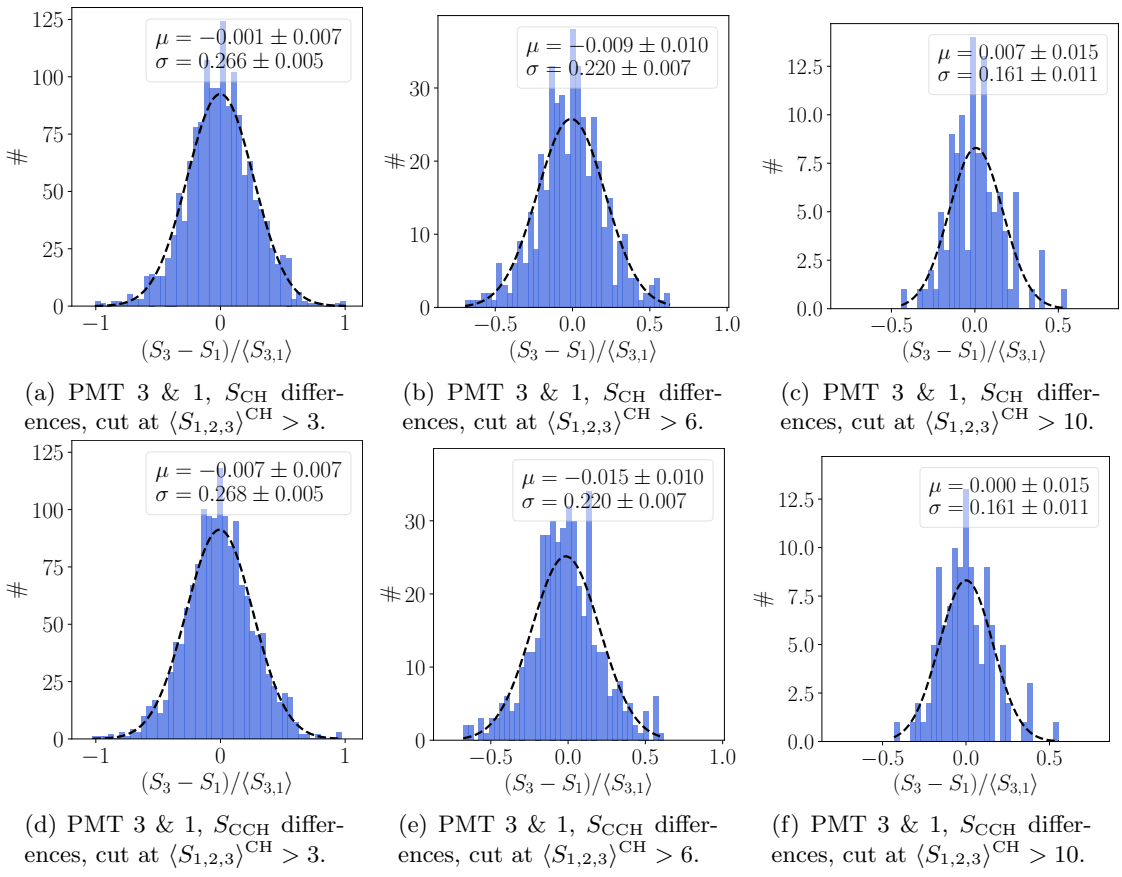
(f) PMT 3 & 1, S_{CO} differences.

Figure 2.25: Histograms of signal differences between the PMTs of the same tank fitted by Gaussian with resulting parameters quoted in the legends. The signal types are defined in Eq. (2.16)&2.17. The first row uses the usual calibration method to find the muon hump. The second row uses the muon hump found in the new coincidence charge calibration histogram, and rescales the signal via $\langle \Delta_i \rangle$ to account for geometric effects of the coincidence condition. The signal calibrated by the new method is then comparable with the one obtained by usual calibration method.

study [14] of $0.6 \pm 1.7\,\%$, $4.2 \pm 1.5\,\%$, and $0.6 \pm 1.6\,\%$ for PMT 1, 2 and 3, respectively. It is the first hint of the difference between the data and simulations, and opens a discussion to further investigation.

Thanks to this study, using 1472 events containing coincidence histograms and all three PMTs working, it was shown that this calibration method is at least as stable as the usual offline calibration method of charge histograms. Looking at a potential improvement in resolution of the new method, namely RMS of the differences in PMT signals of the same WCD, only small improvement within uncertainty of the usual method was observed. Indeed, with RMS of 0.2, if the improvement is for example 3 %, the RMS of the new method would be 0.197. The uncertainties on the data will have to be improved by analysing much more events to make any conclusions on these improvements.

Looking at the more-aged stations, for which the new method will bring a more reliable offline



(a) PMT 3 & 1, S_{CH} differences, cut at $\langle S_{1,2,3}\rangle^{\mathrm{CH}} > 3$.

(b) PMT 3 & 1, S_{CH} differences, cut at $\langle S_{1,2,3}\rangle^{\mathrm{CH}} > 6$.

(c) PMT 3 & 1, S_{CH} differences, cut at $\langle S_{1,2,3}\rangle^{\mathrm{CH}} > 10$.

(d) PMT 3 & 1, S_{CCH} differences, cut at $\langle S_{1,2,3}\rangle^{\mathrm{CH}} > 3$.

(e) PMT 3 & 1, S_{CCH} differences, cut at $\langle S_{1,2,3}\rangle^{\mathrm{CH}} > 6$.

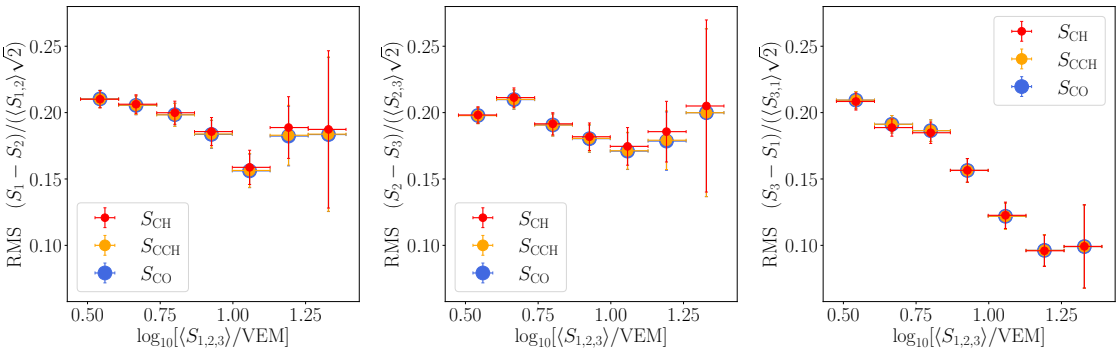
(f) PMT 3 & 1, S_{CCH} differences, cut at $\langle S_{1,2,3}\rangle^{\mathrm{CH}} > 10$.

Figure 2.26: Histograms of the relative signal differences between PMT 1 & 3 imposing different lower-charge threshold fitted by the Gaussian with the parameters quoted in the legends. The first row (a-c) shows the signals calibrated by the muon peak obtained from the usual offline calibration method. The second row (d-f) shows the signal differences of calibrating by the muon hump found from the coincidence charge histogram only. The comparison in the case of PMT 1&3 pair is possible because of their mirror symmetry.

calibration, statistics were lacking for quantifying the improvement. A small improvement was seen, but more data will be needed to further investigate it.

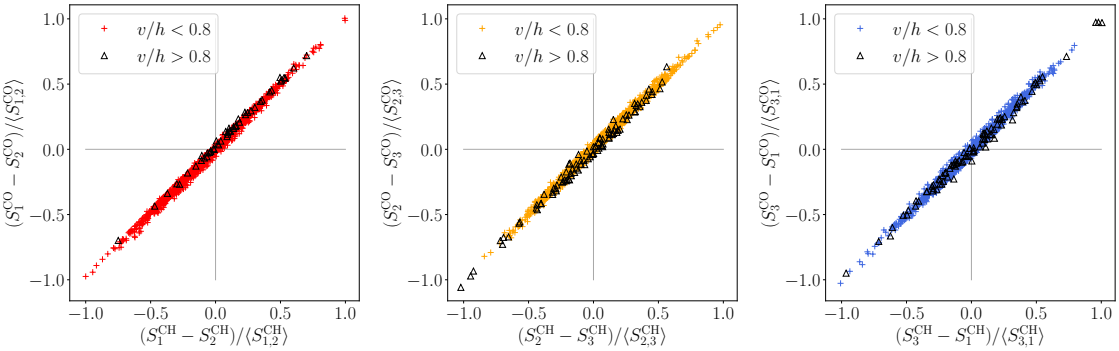
Potential extensions to this analysis include investigating correlation between Q^{peak} found by new calibration method and the muon peak estimate from the online calibration of the WCDs. If they agreed, it would hint at the online and the new calibration working more reliably as opposed to the usual offline calibration method. Another extension of the analysis and the coincidence calibration overall would be looking at seasonal variations of MIP peak in the SSDs to verify the threshold of 10 FADC counts/120 MHz used for the coincidence condition. Lastly, having more data on the relative differences and v/h relationship from not only more tests, but also more SDs as the upgrade of the Observatory continues, one may start to consider correction of individual Δ based on v/h values.

The analysis developed in this study is the first kick on the further wide testing and implementation of the coincidence calibration in the upgraded SD array of the Pierre Auger Observatory.



(a) PMT 1 - PMT 2 (b) PMT 2 - PMT 3 (c) PMT 3 - PMT 1

Figure 2.27: Resolution of the usual and the new calibration as a function of CH signal magnitude investigated using RMS of the relative differences in signals between PMTs of the same tank.



(a) PMT 1 - PMT 2 (b) PMT 2 - PMT 3 (c) PMT 3 - PMT 1

Figure 2.28: Comparison of relative differences between PMT signals using the new (S_{CO}) and the usual (S_{CH}) calibration methods, with the focus on aging stations. Black triangles denote event signals where at least one of the two PMTs whose differences are investigated has $v/h > 0.8$. The improvement brought by the new calibration would be shown by black triangles being closer to the horizontal line, i.e. as smaller spread in coincidence calibration signals.

Chapter 3

UHECR Energy Spectra in the Extra-galactic Plane

3.1 Motivation

Taking the high-quality FD hybrid data between 2004 and 2018, a hint for the change of the mass composition over the sky has been observed. The mass-composition variable, $X_{\max}$ showed a $3.3\,\sigma$ difference between the events coming from the region on the Galactic plane (Galactic latitude $|b| \leq 30°$) and off the Galactic plane at the energies higher than $10^{18.7}$ eV (5 EeV) [24]. The primaries coming from on the Galactic plane are generally heavier according to the study as shown in Fig. 3.1.

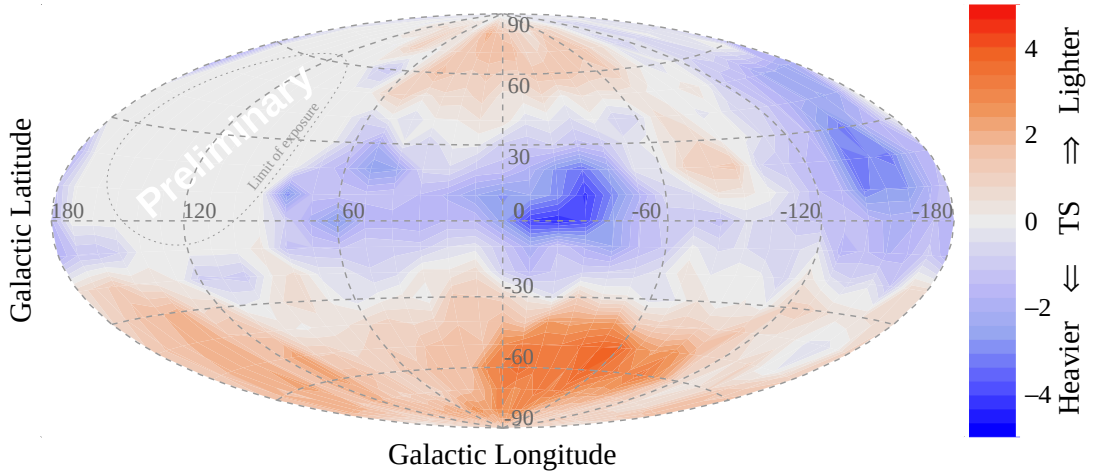
This can be explained by propagation of UHECR in Galactic Magnetic Field. According to the latest models, UHECR can propagate diffusively up to the energy threshold of $6 * Z$ EeV (with Z the atomic number of the nucleus) [25]. Thus heavier nuclei at UHECR energies will propagate diffusively, while the propagation of lighter components at these energies can become ballistic. Moreover, at the energies higher than 5 EeV (the ankle of the cosmic-ray spectrum), the origin of UHECR is extragalactic, supported experimentally by the dipole anisotropy in the UHECR flux at more than 8 EeV [24].

The mass-dependent anisotropy comparing on & off plane regions may then be the experimental result of the following hypothesis: The UHECR observations along the Galactic plane will contain mostly diffusively-propagating heavy component coming from extragalactic sources with Galactic coordinates of the sources both close to and further away from the Galactic plane.

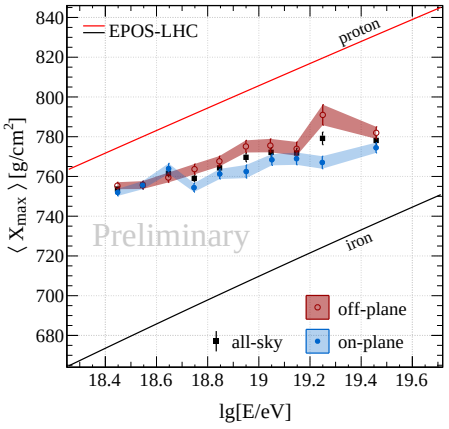
This study was done based on FD measurements. Compared to SD array, FDs have considerably lower duty cycle, thus lower statistics of events. However, SD events are missing accurate $X_{\max}$ measurement. Checking the mass-dependent anisotropy in the SD-array data is therefore not so straightforward.

Making use of one of the other latest key results of the Pierre Auger Observatory: The depth of shower maximum $X_{\max}$, thus the composition of the UHECR primaries tends towards heavier component at higher energies. Fits of $X_{\max}$ using hadronic (EAS) models to find the fractional composition of UHECR were performed to show that at the cut-off of the spectrum, the UHECR are primarily in the mass range of helium or nitrogen, with iron having non-zero fraction at the highest UHECR energies [26]. This is shown in Fig. 3.2.

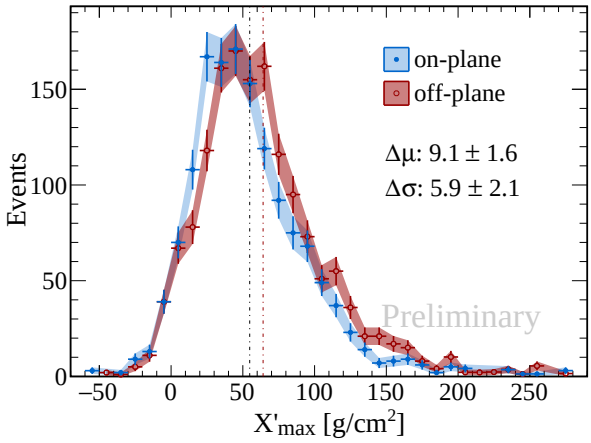
Combining this information with the claim that the on-plane flux contains heavier UHECR, one expects to see more flux coming from on-plane compared to off-plane region at the highest



(a) Sky map of UHECR composition with lower energy threshold of 5 EeV.



(b) First X_{max} moment of the on- & off-plane regions in different energy bins.



(c) X_{max} distributions of all the hybrid data.

Figure 3.1: Figures from [24] showing the key results of the proceeding.

UHECR energies. Taking the ratio of the flux on-plane and off-plane, this ratio would increase from the value of 1 when going higher in energy in the SD data. A short analysis which uses this reasoning was performed on all 5T5 SD data until January 2021 and it will be described in the following sections.

3.2 Analysis

The data set for this analysis included all the 5T5 events (signal in the central SD station and 5 stations in the hexagon around it) between 2004 and 2021, the reconstruction version v17r0 downloaded from [27]. The events come with the zenith angle of less than 60 °. The lower

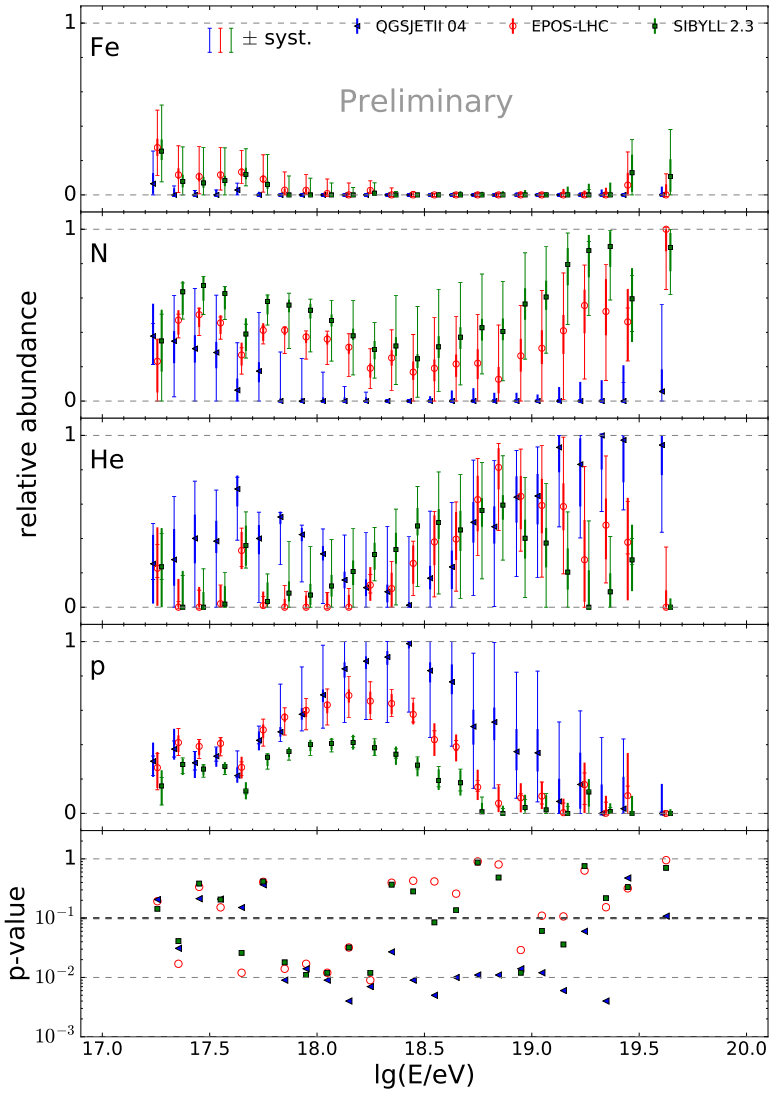


Figure 3.2: Figure from [26] with updated data. Fractional composition of UHECRs is estimated using $X_{\max}$ data and three different hadronic interaction models.

primary energy threshold was set to $E_{\text{lower}} = 10^{18.4}$ eV. The cuts ensure the full efficiency of SD. The total number of events was 301 581, approximately 40 % increase in statistics compared to 215 030 6T5 events (signal in the central SD station and all 6 stations in the hexagon around it), which are more typically used for UHECR spectrum analysis. The reason behind using 5T5 events for this study was the obvious increase in statistics.

Part of the analysis scripts was adapted from a Jupyter Notebook of the Open Data website of the Pierre Auger Observatory. This notebook contained functions to correct the UHECR flux for the exposure of the Observatory [28], which was needed for this study.

3.2.1 Raw spectrum

The highest-energy event in the data set was at $E_{\text{highest}} \approx 168\,\text{EeV}$. Two binnings were investigated, one of 10 and one of 6 bins, with 0.1 and 0.2 width in $log_{10}[E/\text{eV}]$ respectively. The last bin was the integrated bin to ensure enough statistics.

The data were split using their Galactic latitude b to those on-plane $|b| \leq 30°$ (as in [24]) and the rest off-plane. This meant 149 977 and 151 604 events respectively. The event counts N were binned appropriately and with the sufficient amount of statistics which was ensured in each bin, Poisson statistics with symmetric uncertainty of $\sqrt{\text{N}}$ was used.

The flux in a bin not corrected for exposure is then $\Phi = N/\Delta E$ with ΔE the width of the bin. When taking the ratio R, it is simply:

$$R = \frac{\Phi_{\text{on}}}{\Phi_{\text{off}}} = \frac{N_{\text{on}}}{N_{\text{off}}} \left(1 \pm \sqrt{\frac{1}{N_{\text{on}}} + \frac{1}{N_{\text{off}}}}\right) \tag{3.1}$$

with the error propagation derived using the approximation of Gaussian limit for sufficient number of counts. The bin centers show the simple mean energy of the events in the bin.

3.2.2 Exposure of the Observatory

Using the functions from [28], the data were loaded into the binned sky map of HEALPix library. The script contained the function for plotting the count map shown in Fig. 3.3. Binning of the map nside=30 was chosen.

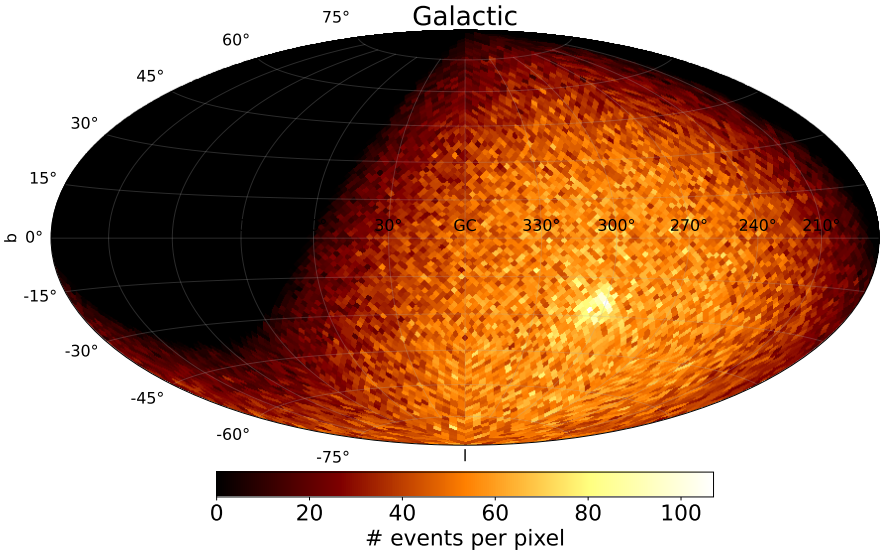


Figure 3.3: Count map of all the 5T5 events with the energy above $10^{18.4}$ eV and zenith angle below 60° of the Pierre Auger Observatory between 01/01/2004 and 31/01/2021, using the binned HEALPix map as in [28], with nside=30.

Part of the sky on upper left side of the plot is missing as it is the part of the sky not seen by the Pierre Auger Observatory. The bright spot of events at $(l, b) = (330°, -30°)$ is the result of the high exposure of the Observatory at those Galactic coordinates. Obviously, if comparing the flux from different parts of the sky, one has to correct for the effect of exposure. The exposure

of an observatory at a fixed location (latitude) on Earth is purely declination-dependent [29]. Converting the exposure from right ascension-declination (RA-dec) to Galactic coordinates then gives the exposure of the similar pattern as seen in the countmap in Fig. 3.3.

Let us call the function defining exposure at one part of the sky $\varepsilon(\delta)$ when it depends on declination of the RA-dec coordinate system, or $\varepsilon(l, b)$ if it is a function of Galactic coordinates. E_{tot} is the total exposure of an observatory. Then the exposure of a part of the sky we are interested in, E_{part} can be calculated as:

$$E_{\text{part}} = \frac{\int_{\text{part}} \varepsilon(\delta) E_{\text{tot}} \, \mathrm{d}\delta}{\int_{\text{all sky}} \varepsilon(\delta) \, \mathrm{d}\delta} = \frac{\int_{\text{part}} \varepsilon(l,b) E_{\text{tot}} \, \mathrm{d}l \, \mathrm{d}b}{\int_{\text{all sky}} \varepsilon(l,b) \, \mathrm{d}l \, \mathrm{d}b}. \tag{3.2}$$

The exposure-corrected flux is $\Phi_{\text{corr}} = \Phi / E_{\text{part}}$, and the exposure-corrected flux ratio is then:

$$R = \frac{\int_{\text{off}} \varepsilon(l,b) \, \mathrm{d}l \, \mathrm{d}b}{\int_{\text{on}} \varepsilon(l,b) \, \mathrm{d}l \, \mathrm{d}b} \left[\frac{N_{\text{on}}}{N_{\text{off}}} \left(1 \pm \sqrt{\frac{1}{N_{\text{on}}} + \frac{1}{N_{\text{off}}}} \right) \right]. \tag{3.3}$$

The functions included in [28] calculate exposure given coordinates of a pixel in HEALPix binned sky map using the formula given in [29]. Therefore in the analysis script, the exposure map of the whole sky was produced using these previously-written functions, using nside=64.

The new part to the script split the exposure maps in two, one for on-plane and one for the off-plane region. These maps are seen in Fig. 3.4. The total exposure of the SD array of the Pierre Auger Observatory between 2004 and 2021 used in these plots was the one given by the official exposure calculator [30], with the choice of "5T5 POS1 Bad period DAQ excluded" exposure type. The total exposure $E_{\text{tot}} = 84\,697.65\,\text{km}^2\,\text{sr year}$.

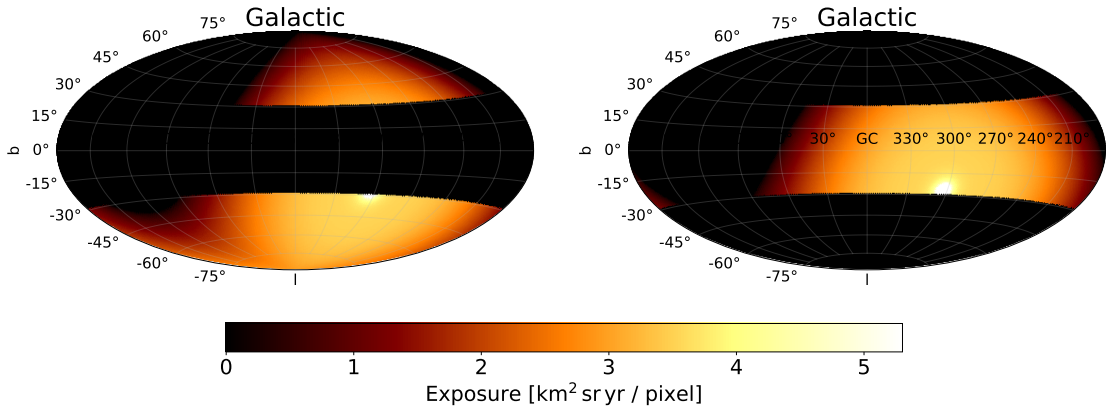


Figure 3.4: Exposure of the Pierre Auger Observatory between 2004 and 2021, using the binned HEALPix map as in [28], with nside=64. The exposure is split at $|b| = 30°$ for the on- & off-Galactic-plane region.

Finally, the ratio of the exposure as described in Eq. (3.3) was calculated using the sums over on-/off- plane pixels of the exposure map:

$$\frac{\int_{\text{off}} \varepsilon(l,b) \, \mathrm{d}l \, \mathrm{d}b}{\int_{\text{on}} \varepsilon(l,b) \, \mathrm{d}l \, \mathrm{d}b} \approx \frac{\sum_{\text{off}} \varepsilon(l,b)}{\sum_{\text{on}} \varepsilon(l,b)} = 1.013. \tag{3.4}$$

3.3 Results & Discussion

The resulting ratio of the flux on & off the Galactic plane is shown in Fig. 3.5. The ratio does not have a clear trend, it is roughly consistent with $R = 1$ (where $R = 1$ means no difference between on & off the plane flux). It is possible the difference is so small (around 2.5 %) that the trend lies within the statistic error. Next steps would therefore include predicting the ratio using the data containing the mass composition information to see if the prediction exceeds the statistical uncertainty found in this study.

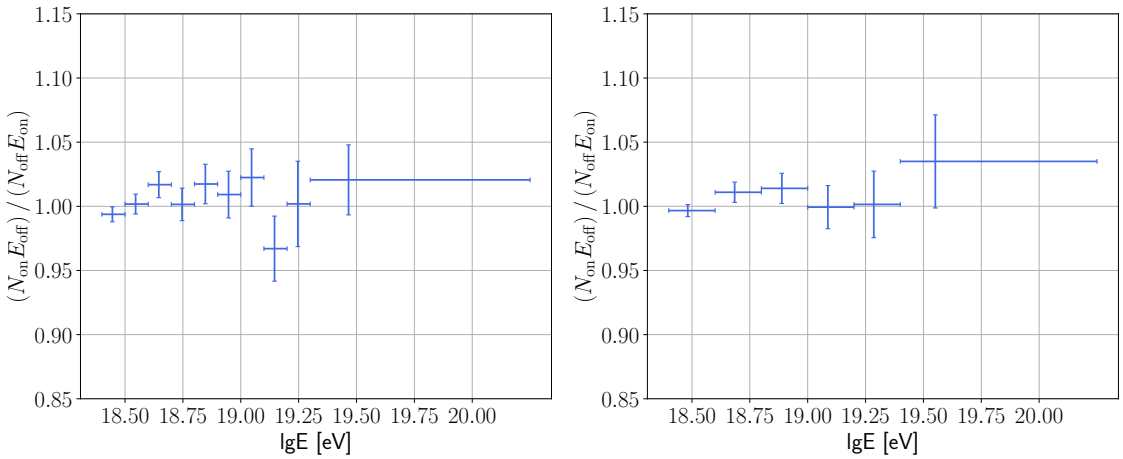


Figure 3.5: Exposure-corrected ratio of on- & off-Galactic-plane SD-event flux between 01/01/2004 and 31/01/2021 using two different binnings.

Considering the energy resolution of the Observatory, some events migrate between the energy bins. This would be a systematic effect. Taking the ratio of the flux is expected to cancel these systematics however.

Conclusion and Outlook

The last years in Pierre Auger Observatory have been very exciting - apart from groundbreaking results at the forefront of science, the new data from fully-upgraded SD array stations started to accumulate. The oldest stations in the array have operated since 2004. In rough conditions of Pampa Amarilla, it is not a surprise that the detectors' aging became apparent. The Observatory is planned to be the largest of its kind and therefore crucial for the UHECR measurements at least until 2030. It is therefore a pressing matter to resolve the issue of aging in SD stations. One of the effects of aging is the light losses in WCDs, which indirectly impacts the precision of the offline calibration. Namely, the atmospheric muon hump merges into electromagnetic background, making it more difficult to accurately determine the charge corresponding to 1 VEM, the quantity used for signal calibration. Previously and also in this work, it was shown the merging typically happens in all three PMTs of the same SD station, making it probably the effect inherent to the water tank and not individual PMTs.

The problem of muon peak and electromagnetic peak merging can be resolved either by deconvolving the two peaks, or by making use of the new detector in the upgraded array - the SSD at the top of almost each SD station. This novel technique, requiring coincidence between the SSD and WCD, reduces the photon contribution in the calibration histograms and allows to better estimate the charge corresponding to 1 VEM. To use this new calibration histograms as well as for backwards compatibility, the task that was achieved in this work is to determine the relation between VEM and the muon peak estimate from coincidence calibration. The difference in the relation compared to usual calibration comes from different geometry of the detector as a whole, and from the absence of electromagnetic peak in the new histograms. In this study, the relation was determined experimentally using comparison between the usual calibration and the new calibration, the relative difference in muon peak estimates Δ. For PMT 1 & PMT 3 $\langle\Delta\rangle_{1,3} = 2.51 \pm 0.08$, and for PMT 2 $\langle\Delta\rangle_2 = 6.62 \pm 0.13$.

Strong correlation between Δ and valley to hump ratio v/h (the measure of how severely the muon peak merges into electromagnetic background in usual calibration) was observed with Pearson correlation coefficient $r = 0.39 \pm 0.05$ for PMT 1 & 3 combined, and $r = 0.75 \pm 0.04$ for PMT 2. The current method uses a constant factor to convert between muon peak estimate and VEM, which is the same for all SD stations. This study showed the importance of accounting for aging in the calibration. For the most-aged stations the muon hump in standard calibration appears to be distorted to lower charge by more than what is accounted for by the conversion factor.

Using all the events containing coincidence calibration histograms until mid-July, it was shown the new calibration behaves at least as good as the usual one. More statistics will be needed to properly assess potential improvements. This work developed the basis for the analysis pipeline for future studies. At this moment, the coincidence calibration is being implemented into the new data format of the SD event files and into the official Auger analysis tools. Thus in near future, new results including increased statistics can be expected, with the hope of improvements

brought by the new calibration becoming more apparent. The coincidence calibration is now a calibration complementary to the usual one. Eventually, in the most-aged stations in the next $\sim$ 10 years of the Observatory's operation, the new calibration might become the only method that provides reliable calibrations. In this study we have set the foundations of this new method.

The second part of this work looked at higher-level analysis of SD data. In the last few years, FD data were used for investigating mass anisotropy in different regions of the sky. At ICRC 2021, the Pierre Auger Collaboration presented a $3.3\,\sigma$ mass-composition change over the sky, data showing heavier composition in the Galactic-plane region. After then, the call remains open to cross-check this result with the SD data, which are more in amount, however they are missing accurate primary-mass information.

Using the results of the interpretation of the maximum of the air-shower development, we know that the higher the energy, the higher the fraction of heavy elements in UHECR flux. This study therefore explored flux differences on & off the Galactic plane by the means of their ratio. The increasing trend toward higher energies if the composition on plane is heavier would hint toward the conclusion about mass anisotropy inferred from FD data. The trend was not observed, which tells us it is at most 2.5 % at the energies of 32 EeV, limit given by the statistical power of SD data.

The next steps expanding this analysis would include predicting the expected ratio using the mass composition information, to see if the prediction is consistent with SD statistics. This analysis also shows how important AugerPrime is for the future of UHECR discoveries, since it will enhance the sensitivity of the surface detector to perform similar studies, a separation by the primary composition for different sky regions. In the end, this mass-enhanced anisotropy study can constrain the models of UHECRs propagation & magnetic fields they traversed, bringing us a step further in revealing the sources of UHECRs, a mystery as old as the field of astroparticle physics.

Bibliography

[1] Coleman A, Eser J, Mayotte E *et al.* 2022 Ultra-high-energy cosmic rays: The intersection of the cosmic and energy frontiers URL https://arxiv.org/abs/2205.05845

[2] Gaisser T K, Engel R and Resconi E 2016 *Cosmic Rays and Particle Physics* 2nd ed (Cambridge University Press)

[3] null null, Aab A, Abreu P *et al.* 2017 *Science* **357** 1266–1270 (*Preprint* https://www.science.org/doi/pdf/10.1126/science.aan4338) URL https://www.science.org/doi/abs/10.1126/science.aan4338

[4] Fluorescence Detector album on flickr https://www.flickr.com/photos/134252569@N07/albums [accessed 06/09/2022]

[5] 2015 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **798** 172 – 213 ISSN 0168-9002 URL http://www.sciencedirect.com/science/article/pii/S0168900215008086

[6] Aab A *et al.* (Pierre Auger) 2016 (*Preprint* 1604.03637)

[7] Fluorescence Detector album on flickr https://www.flickr.com/photos/134252569@N07/22414306455/in/album-72157654097333143/ [accessed 05/08/2022]

[8] Zapparrata O 2022 Private communications

[9] Knoll G F 2000 *Radiation Detection and Measurement, 3rd ed.* 3rd ed (New York: John Wiley and Sons) ISBN 978-0-471-07338-3, 978-0-471-07338-3

[10] staff P A O 2022 Private communications

[11] Castellina A and Pierre Auger Collaboration 2019 *European Physical Journal Web of Conferences* (*European Physical Journal Web of Conferences* vol 210) p 06002 (*Preprint* 1905.04472)

[12] Aglietta M, Allison P, Andres E C *et al.* 2005 *29th International Cosmic Ray Conference (ICRC29), Volume 7* (*International Cosmic Ray Conference* vol 7) p 83

[13] Bertou X, Allison P, Bonifazi C *et al.* 2006 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **568** 839–846 ISSN 0168-9002 URL https://www.sciencedirect.com/science/article/pii/S0168900206013593

[14] Mariș I C, Sato R and Zapparrata O 2021 Using the augerprime scintillators for improving the calibration of the water cherenkov detectors Auger internal note 2021-046

[15] Durán M S 2022 Private communications

[16] Orazio Zapparrata K C and Mariș I C 2019 The time evolution of the number of stations triggered by air-showers Auger internal note 2019-066

[17] Sato R 2022 Private communications

[18] scipy.optimize.curve_fit https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html [accessed 10/02/2022]

[19] Cowan G 1998 *Statistical Data Analysis* Oxford science publications (Clarendon Press) ISBN 9780198501558 URL https://books.google.be/books?id=ff8ZyW0nlJAC

[20] scipy.stats.chi2 https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.norm.html [accessed 21/02/2022]

[21] Jack Hayya Donald Armstrong N G 1975 *Management Science* **21** 1338–1341

[22] scipy.stats.norm https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.norm.html [accessed 05/04/2022]

[23] Aab A, Abreu P, Aglietta M *et al.* 2020 *Journal of Instrumentation* **15** P10021–P10021 URL https://doi.org/10.1088/1748-0221/15/10/p10021

[24] Mayotte E, Abreu P, Aglietta M *et al.* 2021 *PoS* **ICRC2021** 321

[25] Erdmann M, Müller G, Urban M and Wirtz M 2016 *Astroparticle Physics* **85** 54–64 ISSN 0927-6505 URL https://www.sciencedirect.com/science/article/pii/S0927650516301451

[26] Aab A, Abreu P, Aglietta M *et al.* 2014 *Physical Review D* **90** URL https://doi.org/10.1103%2Fphysrevd.90.122006

[27] Observer - the data production for the Pierre Auger Observatory at KIT using the Offline framework https://web.ikp.kit.edu/observer/data.html [accessed 08/12/2021]

[28] Large-scale Anisotropy https://www.kaggle.com/code/augeropendata/large-scale-anisotropy/notebook [accessed 01/12/2021]

[29] Sommers P 2001 *Astroparticle Physics* **14** 271–286 ISSN 0927-6505 URL https://www.sciencedirect.com/science/article/pii/S0927650500001304

[30] SD Exposure Calculator http://ipnwww.in2p3.fr/~augers/AugerProtected/ExpoCalc.html [accessed 13/12/2021]

Additional Activities

Apart from the work on analyses described in the previous chapters, the author of this thesis traveled to Malargüe and operated FDs and LIDAR as the part of the FD Night Shift of 26^{th} January to 10^{th} February 2022.

She showed the progress of the analysis of Chapter 2 at two Operations & Long-term Performance online calls. She also made the presentation of the work from Chapter 2 at the Auger Analysis Meeting in Wuppertal on 18^{th} July 2022.

Within her university (Vrije Universiteit Brussel) and the research institute (IIHE Brussels), she took part in the outreach to high school students in form of a Masterclass: Particles from the cosmos. The masterclass took place on 27^{th} April 2022 and contained introductory lecture about astroparticle physics by professors of the institute. One half of the exercises was demonstrated by the Auger group of IIHE, where the author also helped out. In the afternoon, the remote FD control room was shown to the high-school students and the author spoke about her FD night shift experience.

Acknowledgements

My first and foremost thanks goes to the first of my supervisors, Prof. Dr. Ioana Mariș, who always did everything in her power to make this a successful master's thesis. This included guiding and checking my steps in this project, creating the best environment possible for me to work, and enabling me to travel to the Observatory in Malargüe, and to the collaboration meeting in Wuppertal.

I would like to offer my special thanks to the Pierre Auger Observatory staff and the Pierre Auger Collaboration, especially the Auger group of IIHE. The group was always happy to help me with any questions or issues, whether regarding the Observatory or anything else. Namely it means Orazio Zapparrata and Dr. Mauricio Suárez Durán, who even provided some of the data and materials as cited in the text. It was also Dr. Nicolás González, who was always happy to help me with anything regarding Offline and more. At the start of my work, Dr. Daniela Mockler was part of this group and available to help as well. Out of the people at the Observatory, I am particularly grateful to Dr. Ricardo Sato, who was kind enough to dedicate some time of his busy schedule to meeting with me in person in Malargüe and to collecting the test data of coincidence calibration analysed in Chapter 2.

Next, I would like to express my sincere gratitude to my second supervisor Prof. Dr. Stijn Buitink, to the head of my master's examination committee Prof. Dr. Alberto Mariotti, and to the Inter-University Institute For High Energies (IIHE), who made it possible for me to do this thesis in Pierre Auger Observatory, even though I am not at the university directly participating in the collaboration. I am grateful also to the IIHE Secretariat, Audrey Terrier and Sofie Van Den Bussche, and the Management Assistant of the Physics Department at the Vrije Universiteit Brussel (VUB) Nina Hindrikx, who always pointed me in the right direction despite my more-complicated bureaucratic situation.

I would like to thank VUB as the home university of my master's and the provider of my international students' scholarship, and the physics professors there, who prepared me to undertake such an extensive project at the scientific level. My thanks go also to Université Libre de Bruxelles (ULB), the host university of my master's project, and Universiteit Gent as a host university of some of my master's courses.

I feel grateful also to the two readers of my master's thesis, Prof. Dr. Krijn De Vries, and Prof. Dr. Steven Lowette, who kindly agreed to fulfill this necessary task to read through and grade my work.

Moving on to my personal acknowledgments, I would like to express how deeply grateful I am to my parents, Martin and Mária, and my brother Martin Jr., who are always there to support me, no matter how difficult or demanding a challenge I decide to undertake. I offer my gratitude also to the family I chose, that is my friends, who are always there to listen to my issues and worries and are able to make them lighter. So these last thanks go to Mitja, Yasmine, Quique, Simon, Rose, Dieder, Yarno, Xander, Jethro, Pablo, Paul, Tia, Samo, Martin, Marka, Soňa, Ráchel, Slavo, Lucka, Adka, and Philipp.

Appendix

A Faulty CCH

Automatically-excluded CCH that were considered faulty based on either of following two criteria. The first was the maximum count within the first 36 bins of CCH as presented in Fig. A.1a. The second type was empty CCH, i.e. the average $\mathrm{d}N/\mathrm{d}Q < 0.05$ with the example shown in Fig. A.1c.

The complete list of these automatically-rejected CCH is provided below:

SD 658, PMT 1, GPS time: 1328530029.

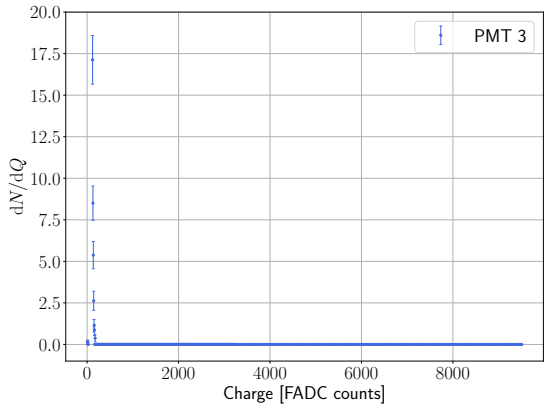
SD 658, PMT 2, GPS times: 1328490101, 1328492596, 1328473979, 1328490294, 1328487516, 1328480999, 1328527058, 1328531330, 1328504254, 1328504167, 1328532085, 1328509278, 1328492738, 1328528052, 1328503749, 1328482978, 1328493027, 1328483062, 1328475739, 1328528087, 1328488468, 1328504562, 1328519767, 1328523857, 1328519351, 1328530029, 1328495148, 1328527745, 1328527431.

SD 659, PMT 2, GPS times: 1328530818, 1328496139, 1328510703, 1328481635, 1328479413, 1328508913, 1328491091, 1328492537, 1328510656, 1328494931, 1328520543, 1328493618, 1328468783, 1328498700, 1328530672, 1328481389, 1328523139, 1328494781, 1328520970, 1328473459, 1328531636, 1328467479, 1328493716, 1328515121, 1328522402, 1328525255, 1328471261, 1328473047, 1328524866, 1328532685, 1328478182, 1328493728, 1328479751, 1328499652, 1328508886, 1328503965, 1328499261, 1328524019, 1328466562, 1328482872, 1328525441, 1328522109, 1328494776, 1328478617, 1328510132, 1328506578, 1328477814, 1328513790, 1328487543, 1328511090, 1328504128, 1328475080, 1328524174, 1328472366, 1328504308, 1328472631, 1328500324, 1328468250, 1328484524, 1328471617, 1328487925, 1328469188, 1328500287, 1328480955, 1328496924, 1328514628, 1328525100, 1328485633, 1328531356, 1328518935, 1328491027, 1328531503, 1328489258, 1328499623, 1328489462, 1328504424, 1328515367, 1328524261, 1328529838, 1328509350, 1328511159, 1328509969, 1328509890, 1328473971, 1328492901, 1328487857, 1328494536.

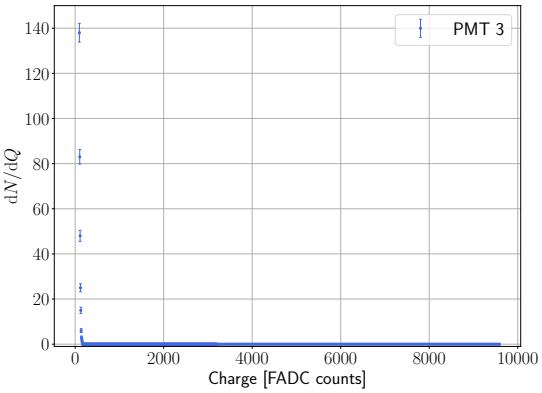
SD 664, PMT 3, GPS times: 1328473350, 1328526256, 1328501814, 1328492180, 1328499162, 1328509795, 1328529143, 1328473888, 1328515385, 1328497919, 1328495082, 1328506056, 1328472952, 1328525795, 1328471419, 1328485446, 1328505449, 1328486910, 1328512486, 1328493371, 1328476319, 1328531093, 1328490123, 1328502169, 1328521296, 1328530659, 1328525209, 1328497587, 1328507038, 1328507404, 1328474765, 1328497644, 1328480955, 1328477746, 1328510338, 1328475937, 1328486709.

SD 666, PMT 2, GPS times: 1328516896, 1328512233, 1328481024, 1328475054, 1328508270, 1328520291, 1328530265, 1328467903, 1328482872, 1328496687, 1328527047, 1328508008, 1328506571, 1328500298, 1328493340, 1328466646, 1328524866, 1328530477, 1328520213, 1328500872, 1328469665, 1328473469, 1328489073, 1328502551, 1328495348, 1328495254, 1328500290, 1328513481, 1328505550, 1328527713, 1328470137, 1328484970, 1328487629, 1328530952, 1328485662, 1328528087.

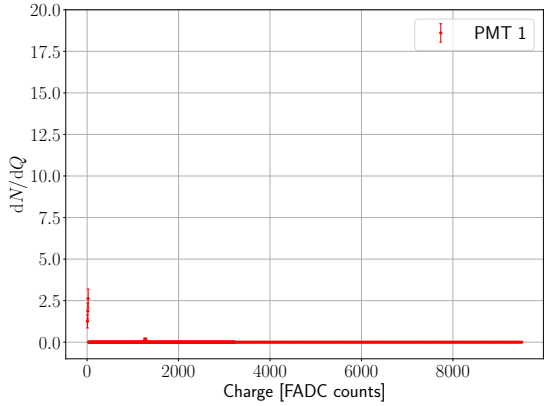
SD 666, PMT 3, GPS times: 1328516896, 1328512233, 1328481024, 1328475054, 1328508270, 1328520291, 1328530265, 1328467903, 1328482872, 1328496687, 1328527047, 1328508008, 1328506571, 1328500298, 1328493340, 1328466646, 1328524866, 1328530477, 1328520213, 1328500872, 1328469665, 1328473469, 1328489073, 1328502551, 1328495348, 1328495254, 1328500290, 1328513481, 1328505550,



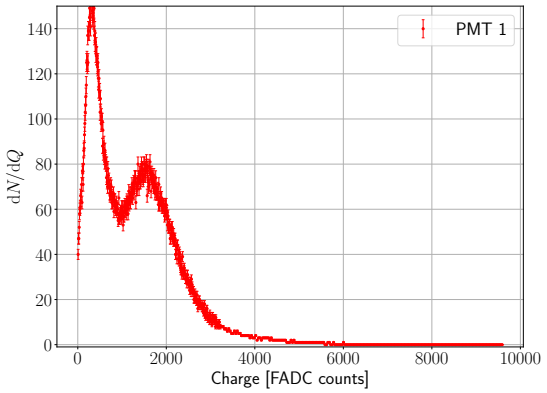
(a) CCH of SD 909, PMT 3 at GPS time 1328510540 as an example of maximum count within the first 36 bins.



(b) CH of SD 909, PMT 3 at GPS time 1328510540.



(c) CCH of SD 693, PMT 1 at GPS time 1328480763 as an example of empty CCH.



(d) CH of SD 693, PMT 1 at GPS time 1328480763.

Figure A.1: Examples of automatically-excluded CCH in a), c). Their corresponding usual charge histograms, CH, are shown for reference in b), d).

1328527713, 1328470137, 1328484970, 1328487629, 1328530952, 1328485662, 1328528087.

SD 693, PMT 1, GPS times: 1328480763, 1328525127, 1328525209, 1328526838, 1328526600, 1328505781, 1328517981, 1328496213.

SD 693, PMT 2, GPS times: 1328480763, 1328525127, 1328525209, 1328526838, 1328526600, 1328505781, 1328517981, 1328496213.

SD 693, PMT 3, GPS times: 1328480763, 1328525127, 1328525209, 1328526838, 1328526600, 1328505781, 1328517981, 1328496213.

SD 798, PMT 1, GPS times: 1328516405, 1328509826, 1328522768, 1328498320, 1328511159, 1328521122, 1328513382, 1328468281, 1328499841, 1328522110, 1328522924, 1328510110.

SD 798, PMT 2, GPS times: 1328516405, 1328509826, 1328522768, 1328498320, 1328511159, 1328521122, 1328513382, 1328468281, 1328499841, 1328522110, 1328522924, 1328510110.

SD 798, PMT 3, GPS times: 1328516405, 1328509826, 1328522768, 1328498320, 1328511159, 1328521122, 1328513382, 1328468281, 1328499841, 1328522110, 1328522924, 1328510110.

SD 851, PMT 1, GPS times: 1328522966, 1328523198, 1328520240.

SD 851, PMT 2, GPS times: 1328502388, 1328473027, 1328477023, 1328522966, 1328476497, 1328516860, 1328479380, 1328510653, 1328523198, 1328490292, 1328512369, 1327848946, 1327848584, 1327849531.

SD 851, PMT 3, GPS times: 1328502388, 1328473027, 1328477023, 1328522966, 1328476497, 1328516860, 1328479380, 1328510653, 1328523198, 1328490292, 1328512369, 1328520240, 1327848946, 1327848584, 1327849531.

SD 859, PMT 2, GPS times: 1328499523, 1328475004, 1328531660, 1328473346, 1328522749, 1328506429, 1328528087, 1328480071, 1328477280.

SD 861, PMT 2, GPS times: 1328504826, 1328491992, 1328524187, 1328517901, 1328480482, 1328471581, 1328469081, 1328520456, 1328519886, 1328498398, 1328481235, 1328520549, 1328532348, 1328521075, 1328520433, 1328491388, 1328494247, 1328482665, 1327849539, 1327848592, 1327848955.

SD 864, PMT 1, GPS times: 1328505627, 1328522749, 1328514201, 1328528748, 1328515807, 1328511592, 1328506429, 1328493178, 1328471997, 1328478992, 1328531660, 1328498795, 1328473858, 1328488283, 1328499523, 1328507681, 1328498937, 1328532384, 1327848602, 1327848964, 1327849549.

SD 867, PMT 3, GPS times: 1328504062, 1328513687, 1328523371, 1328477330, 1328488820, 1328473384, 1328485737, 1328468644, 1327848970, 1327848608, 1327849554.

SD 869, PMT 2, GPS times: 1328524635, 1328516785, 1328493050, 1328525173, 1328520530, 1328527179, 1327848976, 1327848614, 1327849562.

SD 873, PMT 2, GPS times: 1328488820, 1328532405, 1328529038, 1328514960, 1328494778, 1328519886, 1328473858, 1328479050, 1328514201, 1328501753.

SD 907, PMT 2, GPS times: 1328519499, 1328520433, 1328507333, 1328512783, 1328502355, 1328516785, 1328507551, 1328527179, 1328503988, 1328467989, 1328492255, 1328524635, 1328507523, 1328504454, 1328472929, 1328479019, 1327848635, 1327849583, 1327848998.

SD 909, PMT 3, GPS times: 1328510540, 1328469726, 1328493333, 1328488195, 1328530659, 1328467362, 1328503067, 1328508081, 1328522213, 1328523327, 1328485766, 1328492596, 1328502140, 1328505279, 1328503794, 1328528684, 1328496182, 1328497916, 1328480956, 1328505085, 1328474720, 1328476377, 1328530401, 1328512380, 1328500130.

SD 916, PMT 2, GPS times: 1328504424, 1328529038, 1328507503, 1328479128, 1328509051, 1328470441, 1328531500, 1328528711, 1328525227, 1328511606, 1328482899, 1328479745, 1328506487, 1328468645.

SD 1205, PMT 2, GPS times: 1328520433, 1328498398, 1328504826, 1328498247, 1328480482, 1328470690, 1328528058, 1328516785, 1328503397, 1328473858, 1328478992, 1328481235, 1328491992, 1328515379, 1328492255, 1328523371, 1328521075, 1327849004, 1327848641, 1327849588.

SD 1207, PMT 1, GPS times: 1328490292, 1328500492, 1328479975, 1328502388, 1328499978, 1328524746, 1328527157, 1328493916, 1328488490, 1328506930, 1327849592, 1327849007, 1327848644.

SD 1208, PMT 2, GPS times: 1328522749, 1328528058, 1328476497, 1328479021, 1328505412, 1328477528, 1328470690, 1328529384, 1328485447, 1328501219, 1328509477, 1327848648, 1327848919, 1327849009, 1327849595.

SD 1214, PMT 3, GPS times: 1328524187, 1328473470, 1328502388, 1328510253, 1328502355, 1328478332, 1328486800, 1328509372, 1328530474, 1328476426, 1328527601.

SD 1219, PMT 2, GPS times: 1328509533, 1328470690, 1328503785, 1328479059, 1328510592, 1328488195, 1328492492, 1328472958, 1327848668, 1327849616, 1327848919, 1327849030.

SD 1227, PMT 3, GPS times: 1328479021, 1328492017, 1328521075, 1328507457, 1328491388, 1328509016, 1328525173, 1328470690, 1328527157, 1328480015, 1328467989, 1328516785, 1328492255, 1328494110, 1328486800, 1328505475, 1328505412, 1328510592, 1328523371, 1328503397, 1328515379, 1328507523, 1328514694, 1328478992, 1328528058, 1327849052, 1327849637, 1327848689.

SD 1472, PMT 2, GPS times: 1328522286, 1328494778, 1328476131, 1328521859, 1328503785, 1328488283, 1328512368, 1328472970, 1328469588, 1328515563, 1327849055, 1327848693, 1327849639.

SD 1721, PMT 2, GPS times: 1328515616, 1328512017, 1328470528.

SD 1745, PMT 3, GPS times: 1327848714, 1327849076, 1327849661.

SD 1751, PMT 2, GPS times: 1328515879, 1328526850, 1328487202, 1327849084, 1327849669, 1327848723.

SD 1755, PMT 3, GPS times: 1328531466, 1328528634, 1328499845, 1328497043, 1328468131, 1328477280, 1328476180, 1328524722, 1328519360, 1328491115, 1328530575, 1328485031, 1328523132, 1328513882, 1328503128, 1328500298, 1328517899, 1328506143.

SD 1798, PMT 1, GPS times: 1328527469, 1328521802, 1328524411, 1328527526, 1328528748, 1328488407.

SD 1798, PMT 2, GPS times: 1328527469, 1328521802, 1328524411, 1328527526, 1328528748, 1328488407.

SD 1798, PMT 3, GPS times: 1328507474, 1328527469, 1328521802, 1328511592, 1328524411, 1328514365, 1328479912, 1328527526, 1328470690, 1328528748, 1328488407, 1328501568, 1327848744, 1327849106, 1327849692.

SD 1878, PMT 2, GPS times: 1328528899, 1328509077, 1328514893, 1328522492, 1328473600, 1328526336, 1328511766, 1328516161, 1328528429, 1328513134, 1328515411.