The energy spectrum and the effective area based on AERA measurements

Katharina Holland

Master Thesis in Physics and Astronomy

Supervisor: Dr. Charles Timmermans

Radboud University Nijmegen

July 2012

Preface

This thesis contains the results of my master research. I did my project at the Experimental High Energy Physics department at the Institute for Mathematics, Astrophysics and Particle Physics (IMAPP) at the Radboud University Nijmegen. My supervisor for this time was Charles Timmermans.

Section 1 gives an overview about cosmic rays, a description of the energy spectrum and the processes in the atmosphere. The Pierre Auger Observatory, where the data that is used comes from, and the reconstruction packages for air showers are explained in the second chapter. The data sets are described in section 3.
In section 4, parameterizations for the radio induced signal strength are introduced. Furthermore, the effective area for different energies of an optimal working array is calculated. The parameterizations in combination with the AERA and surface detector data of April-June 2011 are analyzed in section 5. The calculations lead to changes in the parameterizations. The changes and improvements of the calculation with the April-June data, are used for a time dependent analysis of data taken between November 2011 and February 2012. With this data an AERA based energy spectrum is calculated. The results are shown in section 6. In the seventh section, a new parameterization and a corresponding energy spectrum are derived. The thesis ends with a conclusion in chapter 9. A comparison of the effective area, for all parameterizations used in this thesis, dependent on the energy and threshold is shown in the appendix.

Katharina Holland
July 2012

Acknowledgments

First I would like to thank my supervisor Charles Timmermans for his advice, support and encouragement. My thanks goes also to Stefan, Stefan and Harm for answering questions and fixing Offline. Furthermore I would like to thank Guus, Harm, Kimmely, Marie, Stefan, Stefan and the rest of the high energy department for the fun we had in our office, the secretary and the breaks.

I like to thank the department and NIKHEF, for the financial support for the six weeks stay at CERN as a summer student and giving me the opportunity to see science in a different environment.

I want to thank my parents for the financial support during my study. Also a thanks to them, Barbara and the whole family for the motivational support and believing in me. Danke!

Contents

Glossary and acronyms **9**

1 Cosmic Rays **11**

1.1 Discovery of cosmic rays 11

1.2 Energy spectrum 12

1.3 Extensive air showers - the Heitler model 14

1.3.1 Light emission 16

1.3.2 Radio emission 17

1.4 Shower parameters 17

2 Pierre Auger Observatory **19**

2.1 Fluorescence detector - FD 19

2.2 Surface detector - SD 21

2.3 Auger Engineering Radio Array - AERA 21

2.4 Reconstruction of the air shower 22

2.4.1 CDAS 22

2.4.2 Offline 23

3 Data set **27**

4 Parameterization of cosmic ray induced radio signals **33**

4.1 Simulation of the effective area 36

5 How to count SD events that should be measured by AERA **41**

5.1 AERA random trigger data 41

5.2 Missing high probability events 43

5.3 Modification of the parameterizations 47

6 Second data set and a time dependent threshold **51**

6.1 Comparison April-June data to November-February data 54

6.2 Energy spectrum 57

7 A data based parameterization **63**

7.1 Division into distance bins 63

7.2 Energy spectrum 64

7.3 Comparison of the two formulae 68

8 Conclusion **71**

A Energy dependent effective area - a comparison of the different formulae **73**

B Additional plots for the data based parameterization **79**

Glossary and acronyms

α	Angle between shower axis and Earth magnetic field
AERA	Auger Engineering Radio Array, array of radio stations
angular direction	The zenith and azimuth angle, that characterize the arrival direction of the shower
CDAS	Central Data Acquisition System, reconstruction program
Coincident events	Events that are measured with SD and AERA
eV	Electron Volt, eV $= 1.602\,10^{19}\,C$
Infill array	Area with a smaller spacing of water tanks, $750\,m$ instead of $1500\,m$
FD	Fluorescence detector
Kathy Turner	SD station in the middle of AERA
Offline	Reconstruction program
Openings angle	Angle between two shower axis
Phi, ϕ	Azimuth angle, measured with respect to the east
PMT	Photo multiplier tube
RD	Radio detector
SD	Surface detector
Shower parameters	Parameters that describe the shower (θ, ϕ ...)
Shower core	Position where the shower axis hits the ground
Theta, θ	Zenith angle, measured with respect to the vertical

1 Cosmic Rays

The Earth is continuously bombarded with high energetic particles, called cosmic rays. Cosmic rays come from outer space, and most charged rays are protons. But also nuclei have been measured. These particles have energies ranging from $10^9 - 10^{20}$ eV, the energy spectrum can be seen in figure 1 [1] and is explained in this chapter. A large flux of low energetic particles arrive on Earth. Therefore it is possible to measure these cosmic rays directly, using satellites located above the atmosphere. Cosmic rays with higher energies occur less often. For example cosmic rays with an energy of at least 10^{17} eV occur about once per km^2 and day. The typical size of a detector on a satellite is a few m^2, hence such a detector would seldom measure a high energetic cosmic ray.

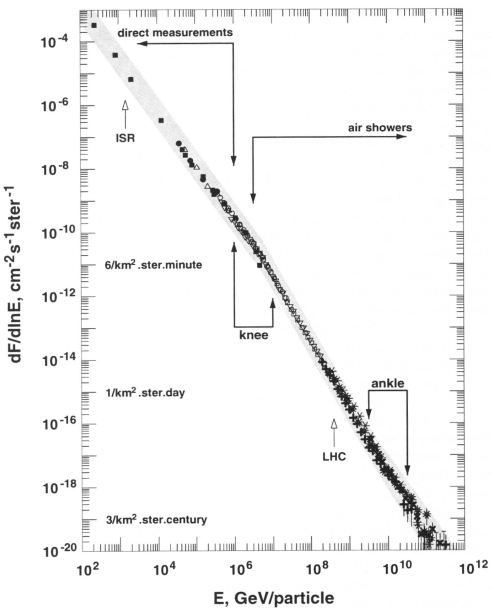


Figure 1: The energy spectrum of cosmic rays, the flux is given per cm^2 s ster. High energetic cosmic rays are very rare.

Therefore large detectors are built on Earth, which measure extensive air showers. An air shower starts, when a comic ray enters the atmosphere and interacts with a particle from the atmosphere. An air shower is described later on in this chapter.
The energy of cosmic rays is much larger than the energies reached by man made colliders. The center of mass energy of the Large Hadron Collider is in the range of 10^{13} eV [2].

1.1 Discovery of cosmic rays

The story of the discovery of cosmic rays starts in the beginning of the twentieth century. Theodor Wulf made the observation that the amount of ionizing radiation does not decrease according to the expectations with altitude. These measurements where done on the Eiffel tower. The idea was that the ionizing radiation originates from radioactive decays from the Earth and should therefore decrease with altitude [3].
Between 1911 and 1913 Victor Hess and Werner Kohlhörster made several balloon flights and measured that that amount of ionizing radiation increases above an al-

titude of 1.5 km. The conclusion was, that the original radiation comes from above. During the solar eclipse of 17th of April 1912 Hess made another balloon flight and concluded, that the radiation comes from outer space and excluded the Sun as possible source [4]. Hess named the radiation 'Höhenstrahlung' and received the Nobel prize for his measurements in 1936.
The phrase cosmic rays for the 'rays of cosmic origin' comes from Millikan and Cameron [5]. They concluded that not (only) the cosmic rays but also secondary particles are measured (beta radiation), which are stimulated by the primary particle.
In 1939 Pierre Auger measured coincidences of secondary particles at sea level and in the Alps [6]. For this phenomena he used the phrase extensive air showers. From the number of particles that he has been measured, he assumed that the energy of the primary particles was in the order of 10^{15} eV.

1.2 Energy spectrum

The following section is based on the description by J. Blümer et al. [9]. In figure 1 the energy spectrum of cosmic rays is shown. The flux is described by a power law, $\frac{\mathrm{dN}}{\mathrm{dE}} \propto \mathrm{E}^{\gamma}$. Up to an energy of $\sim 10^{15}$ eV, the spectral index γ is about -2.7. At higher energies γ changes to -3.0. This behavior is known as the knee in the cosmic ray spectrum. There are different scenarios, explaining the change of the spectral index.

1. The maximum energy attained during the acceleration in galactic sources is reached. Therefore, this is a rigidity dependent cut off. Due to the theory of acceleration of cosmic rays, the maximum energy reached by Supernova remnants is in order of $\mathrm{Z}\,10^{15}$ eV with Z the proton number.

2. There is a leakage from the galaxy, the magnetic fields are not strong enough to bend the cosmic rays. They escape and do not reach the Earth. Above the knee the escape of particles commences.

3. There are interactions with other particles in the galaxy, which lead to an energy loss.

4. New / unknown physics is happening.

To show the change of the spectral index, the flux is often multiplied with $\mathrm{E}^{2.5}$ or E^{3} to show the characteristics of the spectrum. Such an energy spectrum is shown

in figure 2. It is clearly visible, that the spectral index changes round 10^{15} eV. Around $4 \cdot 10^{17}$ eV, the spectrum becomes a little bit steeper, this is called the second knee. Its existence however is not so clear and still under debate. At an energy of about 10^{18} eV, the spectrum seems to become flatter again. This feature is called the ankle. Again, the origin of this feature is unclear. It might be due to the influx of extra galactic particles in our galaxy, or an energy reduction of even higher energy cosmic rays. At highest energies, the flux drops extremely fast, this is may caused by the GZK effect [7, 8]. Greisen, Zatsepin and Kuzmin predicted a cut off at $\sim 6 \cdot 10^{19}$ eV, due to the photon and pion production in the interaction between ultra high energetic cosmic rays and the cosmic microwave background. An other reason could be, that the maximum energy, that can be gained by acceleration is achieved.

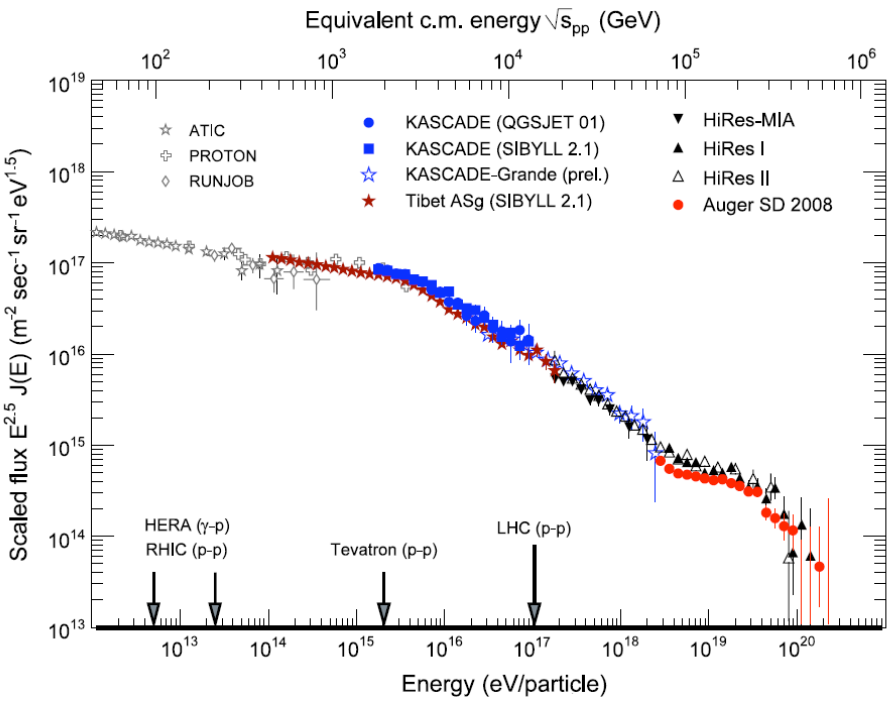


Figure 2: All-particle cosmic-ray energy spectrum as obtained by direct measurements above the atmosphere by the ATIC, PROTON, and RUNJOB as well as results from air shower experiments. Shown are Tibet AS results , KASCADE data (interpreted with two hadronic interaction models), preliminary KASCADE-Grande results, and Akeno data. The measurements at high energy are represented by HiRes-MIA, HiRes I and II, and Auger. [9]

As indicated by the red dots, in figure 2, the Pierre Auger Observatory measures cosmic rays with energies above 10^{18} eV. The energy spectrum obtained from Auger air shower measurements is shown in figure 3. However, the energy range in which radio technique is used in Auger is between $10^{17} - 10^{18}$ eV, which is the energy region in which a spectrum is obtained in this thesis.

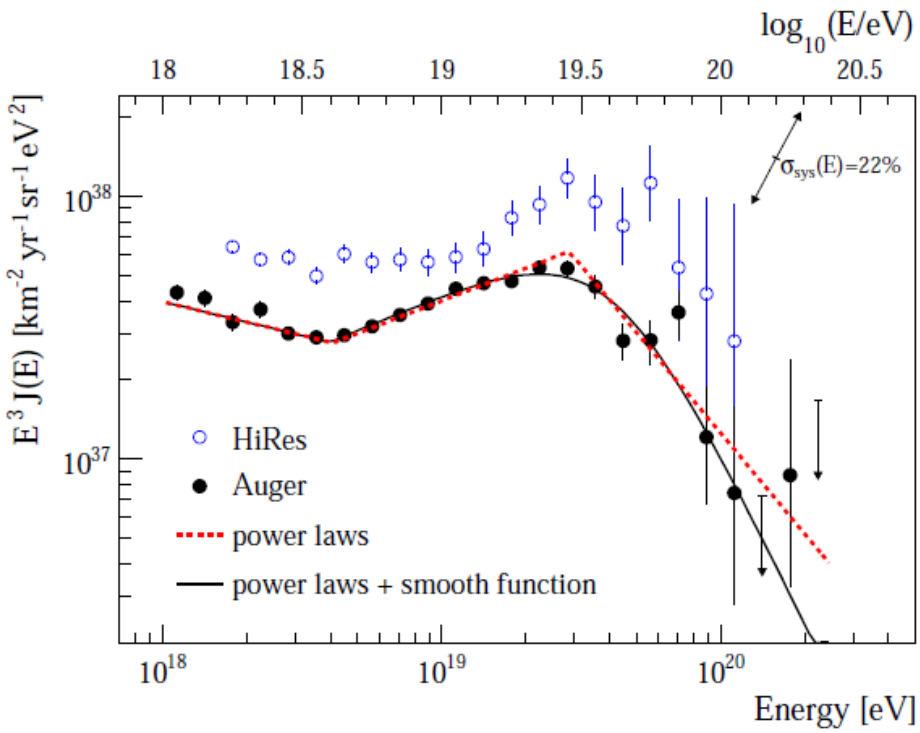


Figure 3: The energy spectrum derived from data of the Pierre Auger Observatory. [10]

1.3 Extensive air showers - the Heitler model

When a cosmic ray enters the atmosphere of the Earth, an extensive air shower starts. A cosmic ray and a molecule in the atmosphere interact and secondary particles are produced, which interact again with molecules in the atmosphere, thereby creating even more particles. In this manner a cascade is formed. A 10^{15} eV proton creates an air shower that contains about 10^6 particles. Such an air shower can be divided into different components. In the first collisions, mostly pions and protons are produced. The neutral pion (π^0) decays immediately into two photons, while the charged pions ($\pi^{\pm}$) decay to muons or have further collisions.

Electromagnetic cascade

98% of the particles in an extensive air shower are electrons, positrons and photons. They are produced in pair production ($\gamma \rightarrow e^+e^-$) and Bremsstrahlung ($e \rightarrow e + \gamma$), once a photon is produced. The electromagnetic shower is fed from the hadronic interaction by the creation and subsequent decay of neutral pions.

Muonic cascade
In an air shower, 1.7% of the particles are muons. They are produced during the decay of $\pi^{\pm}$. Due to their long live time, some muons reach the Earth and do not decay further.

Hadronic cascade
In the first collision, hadrons are produced. These particles undergo further collisions, where further hadrons are produced. The number of hadrons that reach Earth is energy dependent and usually small (0.5%).

The development of an air shower can be approximated by the Heitler model [11, 12] which is schematically shown in figure 4. The Heitler model is a simple model to describe the development of an electromagnetic cascade and is expanded for hadronic showers. According to this model, every particle undergoes a splitting into two particles after encountering a fixed amount of matter, which is called a length measured in g/cm^2. This means, that there are 2^n particles with equal energy after n interactions. The number of particles is maximal when the critical energy $\mathrm{E_c^e}$ of the electron is reached. At this energy (85 MeV), it is not possible to produce more particles. The total number of particles is therefore $\mathrm{N_{max}} = 2^{\mathrm{n_c}} = \mathrm{E_p}/\mathrm{E_c^e}$ with $\mathrm{E_p}$ the primary energy of the cosmic ray and $\mathrm{n_c}$ the number of interactions. Here it is assumed, the cascade is purely electromagnetic.
One splitting length is $\mathrm{d} = \lambda_\mathrm{r} \ln(2)$ in which λ_r is a radiation length, which depends on the medium. The total shower length of the shower maximum is $\mathrm{n_c} \times \lambda_\mathrm{r} \ln(2)$. This distance $\mathrm{X_{max}} = \lambda_\mathrm{r} \ln(\mathrm{E_p}/\mathrm{E_c^e})$ is energy dependent and measured from the top of the atmosphere. The more energetic the primary cosmic ray is, the larger is the penetration depth of the shower. Also the depth of the first interaction in the atmosphere depends on the energy of the primary particle. Massive particles interact earlier with the atmosphere.
The hadronic cascade is described in a similar fashion. After a depth of $\mathrm{d} = \lambda_\mathrm{i} \ln(2)$, with λ_i the interaction length, neutral and charged pions are produced. Roughly 2/3 of the particles are charged and form a hadronic cascade. The neutral pions decay immediately to photons and are then part of the electromagnetic cascade (the 'feeding' of the electromagnetic shower component is not taken into account). This process continues, until the critical energy of the $\pi^{\pm}$ is reached. This energy is estimated as the energy at which the decay length becomes less than the interaction length. When the critical energy is reached, the pions decay to muons.
Assuming that all pions decay to muons, the total energy is divided between the muons and the particles of the electromagnetic cascade. $\mathrm{E_p} = \mathrm{E_c^\pi}\,\mathrm{N}_\mu + \mathrm{E_c^e}\,\mathrm{N_{max}}$.

Furthermore, when knowing the number of particles, it is possible to estimate the energy of the primary particle.
For the description above, it is assumed that the primary particle is a proton. For nuclei, the superposition model is assumed as a simplified view of the interaction. A nucleus with atomic number A and total energy E_p is assumed to behave similar to A protons, each with energy E_p/A. The shower maximum is then located at a depth of $X_{max} \propto \lambda_r \ln(\frac{E_p}{A\,E_c^e})$.

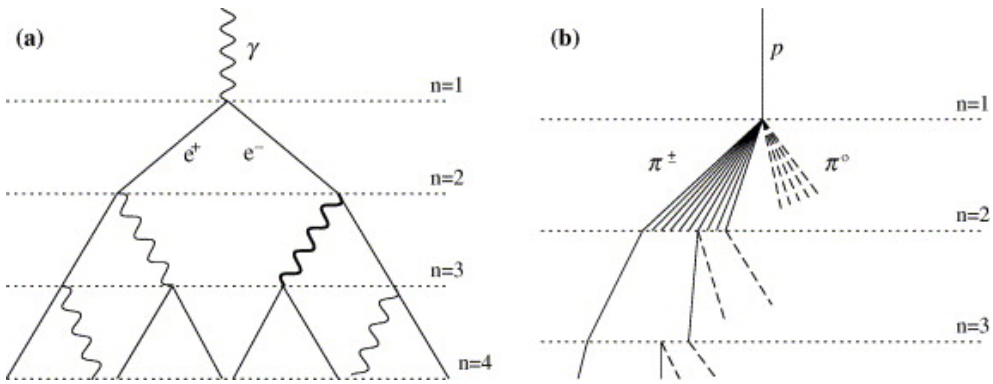


Figure 4: Schematic views of (a) an electromagnetic cascade and (b) a hadronic shower. In the hadron shower, dashed lines indicate neutral pions which do not re-interact, but quickly decay, yielding electromagnetic subshowers (not shown). Not all pion lines are shown after the n = 2 level. Neither diagram is to scale. Adapted from [12].

1.3.1 Light emission

The particles in the shower move with a velocity larger than the speed of light in air. Therefore Cherenkov light is produced. The direction of the light depends on the speed of the particle. The angle θ between the direction of the particle and the light cone has the following velocity dependence: $\cos\theta = \frac{c}{nv}$ with c the speed of light in vacuum, n the refractive index of the medium and v the velocity of the particle. The anisotropy of the light emission has the disadvantage, that it can only be detected in a cone around the shower axis.
Next to Cherenkov light, fluorescence light ($\lambda = 300 - 400\,\mathrm{nm}$) is produced in the atmosphere. The electrons of the electromagnetic shower excite the nitrogen of the atmosphere. When the nitrogen falls back to the ground state, light is produced. This fluorescence light is emitted isotropically and can be seen from all directions. The amount of light depends highly on the number of electrons. Therefore, the shower development can be reconstructed using knowledge about the time dependent amount of light produced.

1.3.2 Radio emission

The charged particles produced in the shower are influenced by the geomagnetic field of the Earth. Due to the Lorentz force, the positive and negative particles move in opposite directions. This force is perpendicular to the magnetic field and the direction of the particle ($\mathbf{F} = \mathrm{q}\,(\mathbf{v} \times \mathbf{B})$), and the strength of the force is $\mathrm{F} = \mathrm{qvB}\sin\alpha$ where α the angle between the direction of the particle and the magnetic field. It is assumed, that the particles on average move in the direction of the shower axis, which in turns is the direction of the primary cosmic ray. If the shower axis is parallel to the magnetic field, no charge separation takes place. The Lorentz force is maximal, when the particles move perpendicular to the magnetic field. Due to the motion of charges, an electric dipole moment is created, which causes electromagnetic radiation. This effect is called geomagnetic emission. An other mechanism that causes radiation in the MHz regime is due to charge excess. The electrons travel with the shower front towards Earth and knock out further electrons in the atmosphere, while the positrons are annihilated, so that there is also a charge separation along the shower axis. The contribution of charge excess to the radiated electric field is small (of order 10 %) compared to the geomagnetic effect.
The strength of the electric field depends on shower parameters. First, there is the $\sin\alpha$ dependence as described above. A large factor in the power of the emission is due to the energy of the shower. As described by the Heitler model, the number of particles produced depends linearly on the energy. A larger number of particles leads to a larger electric current and therefore to a larger electric field. The position of the shower maximum for vertical showers is lower in the atmosphere as for showers originating from slanted cosmic rays. The atmosphere can absorb more electrons and positrons for not vertical showers so that a smaller electric field can be measured on Earth. A $\cos\theta$ dependence is therefore introduced, with θ being the angle between the shower axis and the vertical. An extensive air shower is a local phenomenon with most particles near the shower axis, therefore the strength of the field is (observer) distance dependent.

1.4 Shower parameters

The air showers are characterized by different parameters. First there is the direction. The direction is given by two coordinates, the zenith angle θ, which is the the angle between the vertical and the shower axis, and the azimuth angle ϕ, which is the angle in the horizontal with respect to the east. The direction is calculated

from the time differences of measurements of the arrival of the signals in different stations. The position where the shower axis hits the ground is called shower core. The shower core, given in x and y coordinates, is reconstructed using the measurements of the signal size. The position where the signal size should be maximal is taken to be the shower core. The uncertainties on the core position are typically in the order of several tens of meters. The perpendicular distance between shower axis and a station is called R. The parameters are all shown in figure 5.

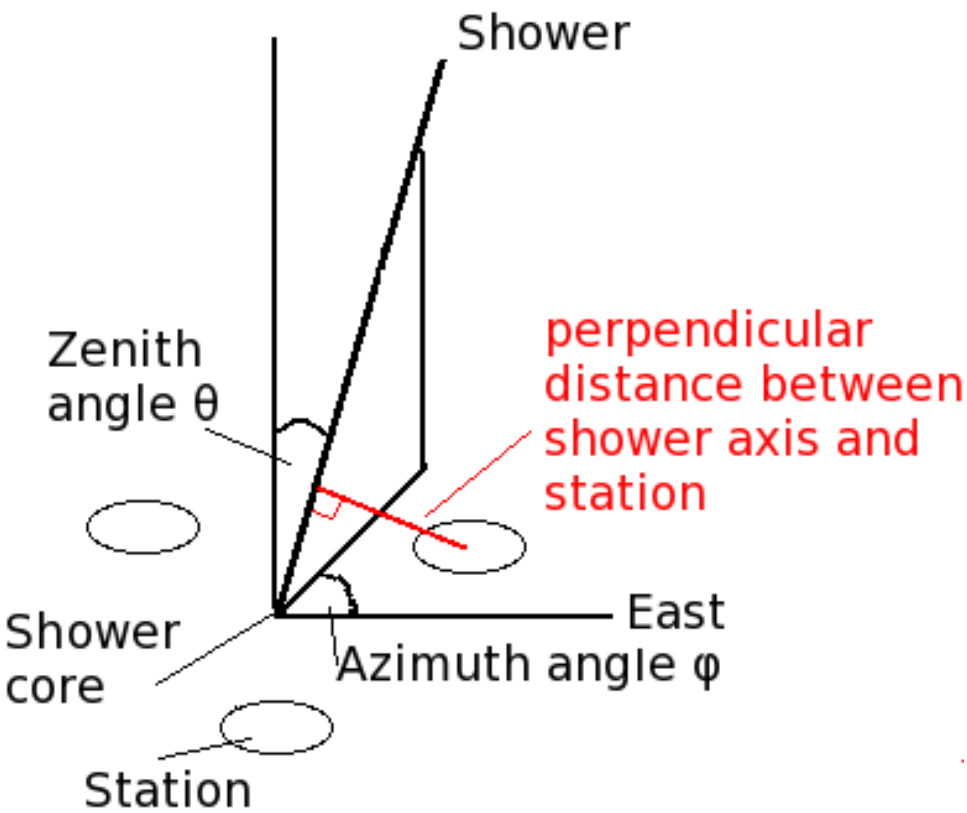


Figure 5: The picture shows an incoming shower with parameters that are reconstructed.

2 Pierre Auger Observatory

The Pierre Auger Observatory has been built to detected cosmic rays of the highest energies. The flux of these cosmic rays is so low (1 per km^2 per century), that an observatory of a large size is necessary to collect a reasonable number of particles. This Observatory, built in Malargüe, Argentina, covers an area of 3000 km^2.
Extensive air showers are measured with a hybrid detector. That means that a shower is measured with different techniques at the same time. An overview of the observatory can be seen in figure 6. The dots symbolize water tanks that form the Surface detector (SD). The lines from the border of the SD to the middle of the array symbolize the field of view of the Fluorescence detector (FD). Near the FD station Coihueco the Auger Engineering Radio Array (AERA) is placed.

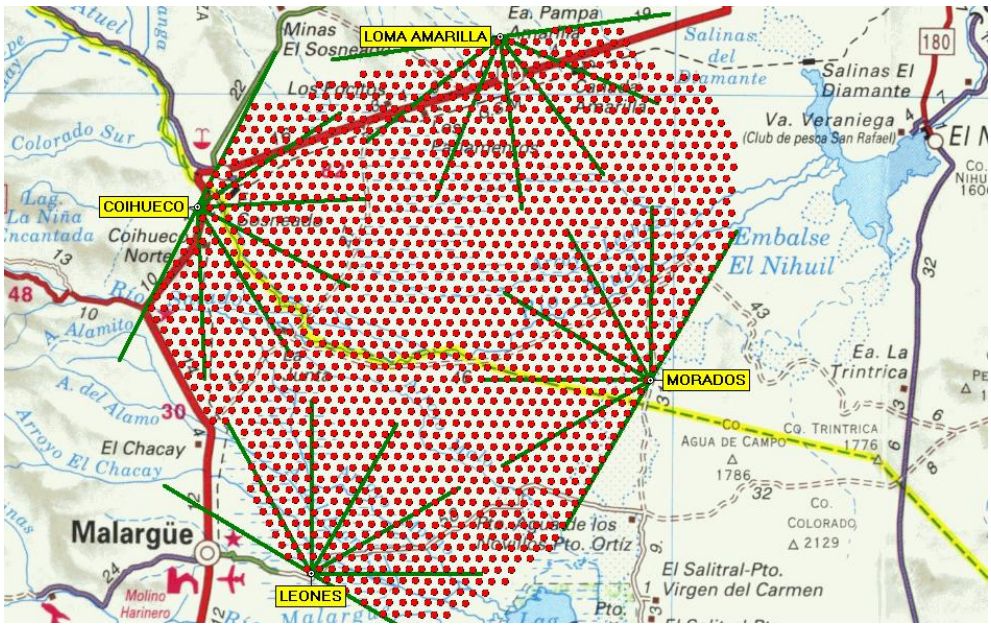


Figure 6: This figure gives an overview of the observatory. Each dot sympolises a water tank, the lines show the field of view of the Fluorescence detector. Each lines corresponds to 20 km.

2.1 Fluorescence detector - FD

The Fluorescence detector is built to measure the fluorescence light, that is emitted by the nitrogen of the atmosphere. These measurements are performed using 24 detectors in four buildings, that overlook the area of the Surface detector. There are three further telescopes at Heat (High Elevation Auger Telescopes) near Coihueco. Each telescope has a field of view of 30° × 30°, the six telescopes in one station are

arranged such that the field of view of a station is $180° \times 30°$. A picture of a SD building and one fluorescence detector is shown in figure 7. Behind each of the doors a telescope is located.
The light is focused on a camera that contains 440 pixels, which are digitized every $100\,ns$. From the amount of light in each pixel, the number of electrons and the shower development is reconstructed. A disadvantage of this detection method is, that the the amount of light is so small, that measurements are only possible in clear moonless nights (10% of the time). A detailed description can be found in [13].

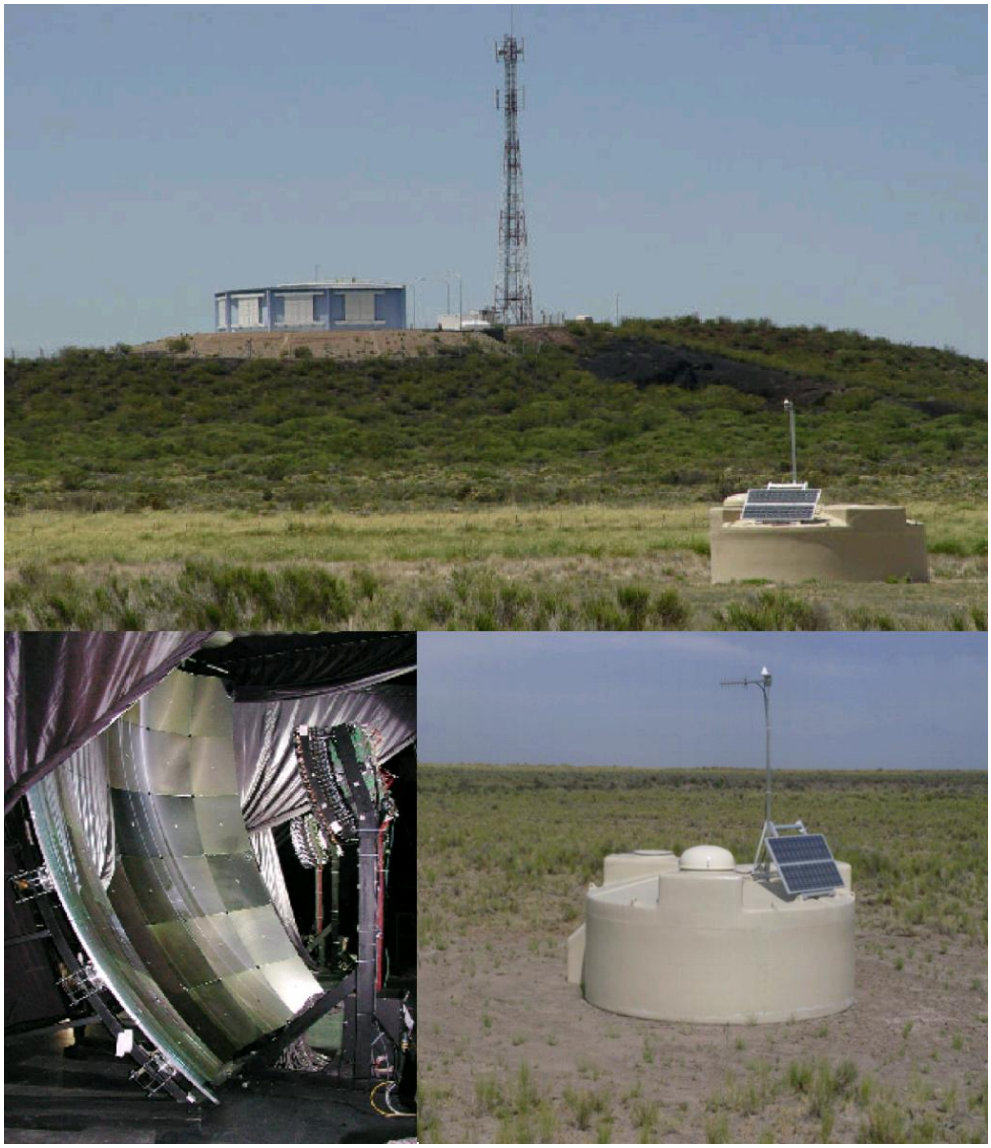


Figure 7: The top picture shows a FD station and a surface detector. The picture on the bottom left shows one of the fluorescence telescopes with mirror and camera, and the one on the bottom right a closer view of a SD station with solar panel and antenna. [14]

2.2 Surface detector - SD

Over the area, of 3000 km^2, 1600 water tanks are distributed in a triangular structure. The distance between the tanks is 1500 m. Each tank consists of 12 m^3 pure water and three photo multiplier tubes (PMT), pointing downwards.
When the particles of the air shower enter the tank, Cherenkov light is produced and measured by the PMT's. From the light, the number of Cherenkov particles is reconstructed. The number of particles measured in each tank provides a footprint of the shower.
Each station has a solar panel for electricity and an antenna to communicate with the central data acquisition system.
In the north-west corner, near Coihueco the density of the tanks is enlarged. This area is called the infill area. In this area a tank is placed between two tanks to get a distance of 750 m between the stations. This area has a hexagonal shape and an area of 23.5 km^2 [15]. Due to the higher density of the detectors, also smaller showers can be measured. This lowers the energy threshold of Auger to an energy below 10^{18} eV. As this is the range where the radio antennas are sensitive, AERA is placed there. Further information over the SD can be found in [16].

2.3 Auger Engineering Radio Array - AERA

The radio stations (figure 8) measure the electric field induced by an air shower in two polarization directions north-south and east-west. The frequency range of the measurements is between 30 MHz and 80 MHz.
AERA is in its first stage. Twenty-four stations are placed in the infill area, in a hexagonal structure with a spacing of 150 m. The layout of AERA can be seen in figure 9. The first stage has a dense core (red points), while spacing between the stations (blue and green) will be larger for the other stages. The big hexagon shows the infill area, with the yellow dots representing the water tanks. The tank in the middle of the

Figure 8: The photo shows a radio station, with the antenna in the north-south and east-west direction, the solar panel and the electronic box.

first stage is called Kathy Turner.
A advantage of the detection with radio is, that the duty circle is nearly 100%. A detailed description about the future of AERA, can be found in the proposal [17]. Events that are recorded by AERA and the surface detector are called coincident events. In this thesis, they are are also called radio events. Only coincidences are used for the thesis.

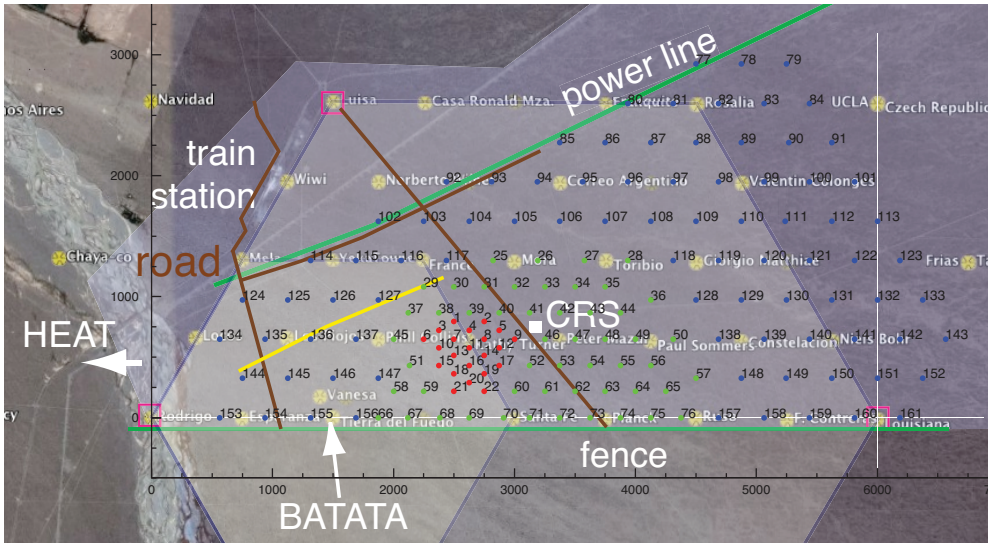


Figure 9: The figure shows the position of the radio antennas in the infill area. The infill is marked with the big hexagon and each water tank is indicated by a yellow dot. Until know, only the first stage, the red dots, are built. The blue and green dots symbolize AERA phase two and three.

2.4 Reconstruction of the air shower

There are two software packages to reconstruct the events measured with the detectors. They are called CDAS [18] and Offline [19]. The typical parameters that are reconstructed are already explained in the previous chapter and shown in figure 5. Additionally the energy of the primary particle is estimated. A comparison of the parameters, reconstructed with both packages is shown in figure 32 and also done in [20].

2.4.1 CDAS

CDAS, the Central Data Acquisition System, can reconstruct measurements of both SD and FD online. It decides, whether an event is stored on disk or not. When comparing RD data with SD data, not all events measured with the surface detector

are necessary. Therefore a specialized CDAS reconstruction is used, that is running offline over the stored data. This reconstruction makes only use of SD data. The following parameters are written to file: AugerId, number of SD stations, the energy reconstructed with three different algorithms, the arrival time of the shower in UTC, the date and the arrival time of the shower, the shower core with uncertainties and the zenith and azimuth angles with uncertainties. Only showers, where the distance (R) between shower axis and station Kathy Turner is less than 2500 m, are written to file. The events that fulfill this condition are then divided into two groups. One file consists of the coincidence events, while the other file contains the events that could have been measured with AERA. For the analysis, the energy estimation of the NEWFD reconstruction is used. This reconstruction is based on the latest calibration with the fluorescence detector.

2.4.2 Offline

A reconstruction with Offline of the parameters of the air shower is possible with data of all measurement techniques. Therefore Offline is necessary to reconstruct the data taken with AERA.
Based on SD data Offline provides, beside the parameters described above, also the radius of the shower front and the signal at 450 m from the shower core. An event display of a shower based on SD data is shown in figure 10.
From the AERA data, the electric field in both measured polarization directions and in the vertical are reconstructed. Using the time stamps when the maxima of the electric field is recorded, a reconstruction of the arrival direction is made. Using the maximal amplitudes, a shower core is reconstructed. The event display of a radio event is shown in figure 11.

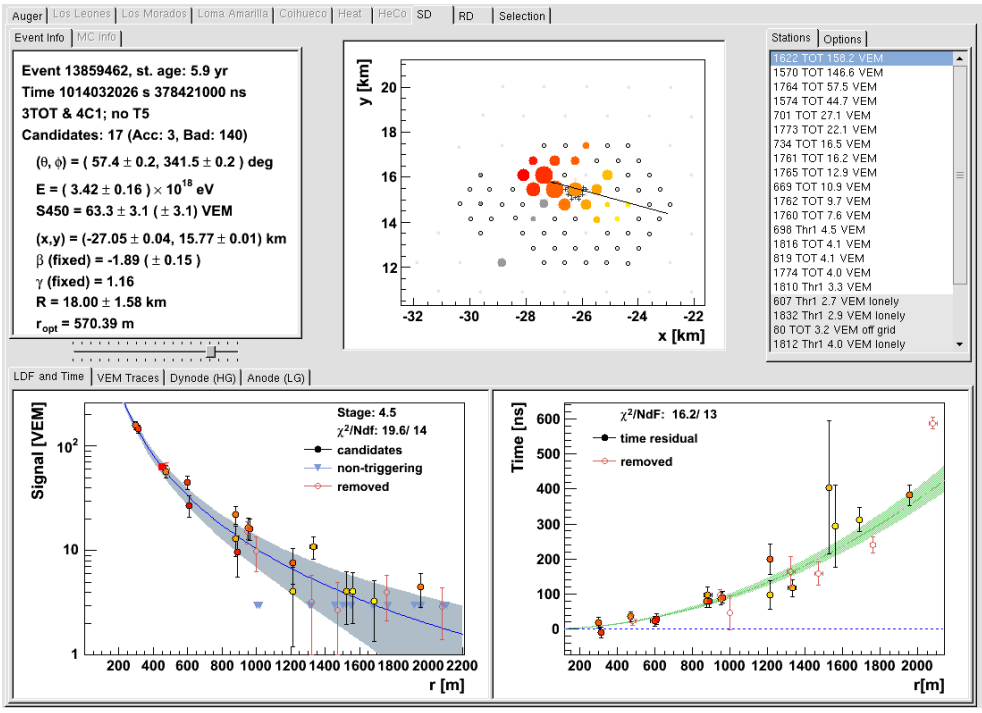


Figure 10: This picture shows the event display of an air shower measured with SD. In top left corner are the reconstructed parameters printed. The picture in the top middle shows the infill area. The colored dots are the tanks that have measured the signal, the color code gives the arrival time, the size of the dots the amount of energy deposit. The line corresponds to the calculated shower core and azimuth angle. In the top right corner is a list of the stations, that have measured the signal. The bottom left plot shows the lateral distribution function, the signal as function of the distance. The bottom right plot shows the arrival time as function of distance, with respect to the tank that has measured the first signal.

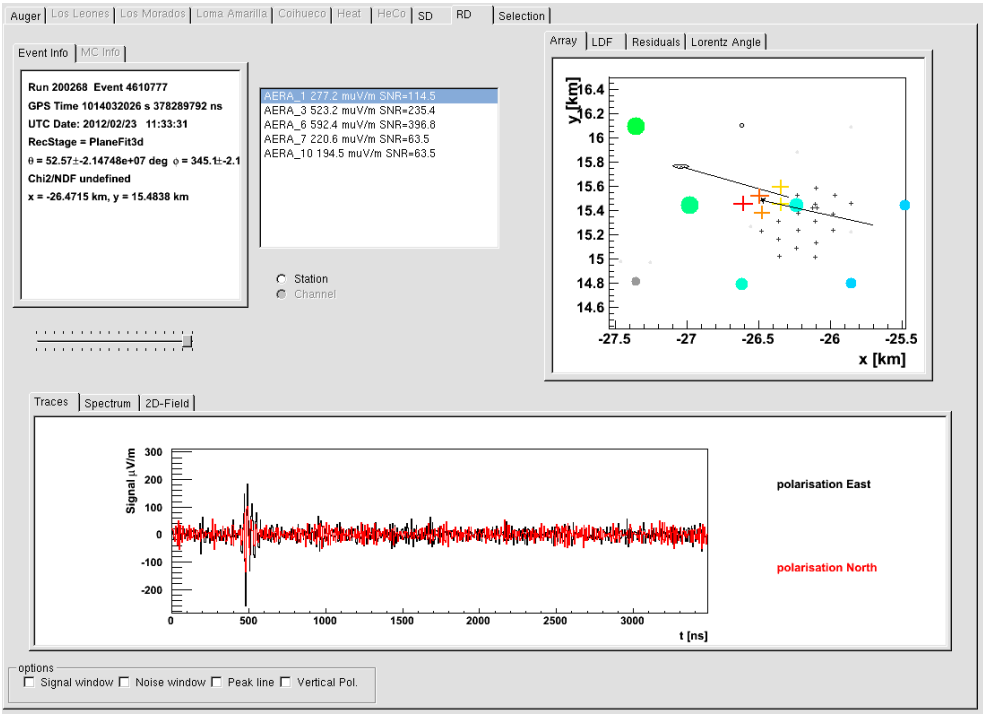


Figure 11: In the top left corner, the properties of the shower are given, in the middle are the stations given, that have measured the shower. The right plot shows the array. The stations that have measured the shower are marked with color. In the bottom graphic, the trace of the north-south and east-west signal is shown for the first station, in this picture station 1.

3 Data set

For the analysis, two data sets of coincident events are used. Coincident events are showers, that are measured with SD and with radio. Sometimes they are called radio events, but they are always measured with SD as well. A cut on the core position is applied for the reconstructed events. In the following chapter, the reason for this selection criteria is explained. Figure 12 shows the area, of allowed shower core positions.

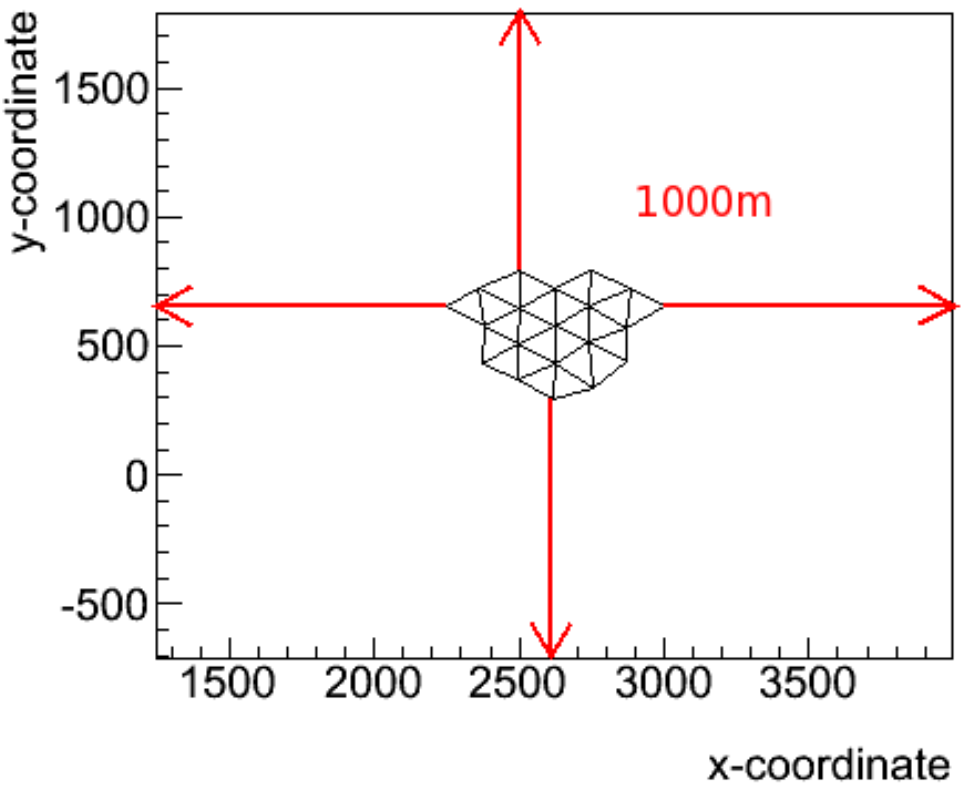


Figure 12: This figure shows the position of the first 20 radio station and the possible core positions after the cut. The lines indicate the hexagonal structure of the array. The red arrows have a length of 1000 m.

The first data set contains coincidences recorded between the 15th of April 2011 and the 2nd of June 2011. The radio setup is such that only two radio stations need to be triggered simultaneously. In this setup 20 stations are used. A list of coincident events is given in table 1. The second data set contains measurements recorded between the 11th of November 2011 and the 29th of February 2012. These coincidences are measured using only half of the array (station 1-12) requiring at least three radio stations to be triggered simultaneously. A list of these coincident events is in table 2.

The CDAS reconstruction is used for the calculation of the efficiency and the energy spectrum. The Offline reconstruction is used to find a new parameterization. The choice for this package is made, because Offline provides direct access to the measured radio signal. The study of the radio parameterization uses all coincidences recorded between April 2011 and February 2012. For this analysis additional selection criteria are that the SD event is recorded in at least 5 SD stations, the zenith angle of the event is less than 60°, the shower core is located in the predefined area and the angular difference between the radio and SD reconstruction is less than 15°. Furthermore, some events are excluded due a thunderstorms condition (marked with a $*$). During thunderstorms, the natural electric field of the Earth is much larger, which leads to a larger induced electric field and therefore to events that would not be recorded during normal conditions.
Additionally, there are two data sets of pure SD events, reconstructed with CDAS. One for the period from April - June 2011 and one for November 2011 - February 2012. Also on these data sets, a cut on the core position is made. Furthermore, a cut on the energy is applied to the second data set, as events with an energy less than 0.1 EeV would not be measured.

SD event number	AERA run	stations	Date
11527860	2150	8,14	April 15 2011 16:03
11528374	2150	12,14,16,20	April 15 2011 18:57
11531810	2150	11,13,16,18,20	April 16 2011 14:09
11535629	2150	12,17	April 17 2011 11:10
11542306	2150	2,5	April 19 2011 01:11
11551431	2150	1,4	April 20 2011 21:58
11554089	2154	1,3,4	April 21 2011 12:53
11563807	2155	4,6,11,16,18	April 23 2011 12:26
11570943	2155	12,13,14,16,18,19	April 25 2011 05:09
11598758	2161	15,16	April 28 2011 17:10
11599064	2161	2,5	April 28 2011 18:15
11601652	2161	8,14,17	April 29 2011 03:55
11611756	2161	12,14	April 30 2011 08:44
11614136*	2161	14,16,19,20	April 30 2011 18:54
11629745	2162	4,7,8,11,12,14,16,18	May 2 2011 17:34
11638288	2162	19,20	May 3 2011 13:06
11638812	2162	18,20	May 3 2011 15:37
11638865	2162	2,4,8	May 3 2011 15:52
11638937	2162	2,4,5,8,9,12,17	May 3 2011 16:11
11645902	2162	3,5,9,12,13,16,19	May 4 2011 7:49
11657029	2162	10,16	May 5 2011 11:38
11665769	2162	12,17,19	May 6 2011 19:45
11711427	2169	2,4,5,7	May 13 2011 3:17
11722810	2169	13,14,16,20	May 15 2011 9:47
11733401	2169	14,16,19	May 17 2011 20:26
11761728	2174	2,4,11,14	May 23 2011 14:40
11775343	2178	3,4,10,11	May 25 2011 17:37
11856712	2184	4,5,8,9,12	June 2 2011 15:13

Table 1: The event information of the coincident events, denoted as data set 1, measured between April and June 2011.

SD event number	AERA run	stations	Date
13147215	200025	1,4,6,7	Nov 11 2011 16:58
13153808	200034	1,2,4,5,8,10,11,12	Nov 12 2011 16:25
13202346	200165	1,3,7,10	Nov 19 2011 10:33
13219438	200200	1,2,4	Nov 21 2011 14:39
13227380	200200	4,7,8,10,11,12	Nov 22 2011 10:11
13242712	200206	10,11,12	Nov 24 2011 04:26
13285718	200206	1,3,6,10	Nov 28 2011 22:18
13294618	200207	3,6,10	Nov 29 2011 23:40
13309533	200207	1,2,4	Dec 1 2011 21:09
13332163	200207	1,2,4,7,8,12	Dec 5 2011 05:23
13385283	200216	3,7,8,10	Dec 14 2011 15:57
13388539	200216	6,7,10	Dec 15 2011 06:09
13441393	200227	1,2,3,4,5,6,8,9,10,11,12	Dec 22 2011 08:36
13457775	200236	2,4,8	Dec 24 2011 14:40
13471815	200236	1,4,6,10,11	Dec 26 2011 11:20
13514610	200241	1,2,4,7	Jan 2 2012 22:02
13525668	200241	2,7,11,12	Jan 4 2012 22:57
13528765	200241	1,7,10	Jan 5 2012 13:08
13531665	200241	1,3,6,7,10	Jan 6 2012 01:04
13563187	200243	2,4,5,11,12	Jan 12 2012 06:22
13574562*	200250	1,4,7	Jan 15 2012 00:52
13591140*	200252	1,2,3,4,6,7,11	Jan 18 2012 22:14
13624724*	200257	1,2,3,4,5,7,8,10,11,12	Jan 23 2012 21:37
13665696	200259	1,4,8,9,12	Jan 28 2012 18:05
13733292	200266	2,4,8	Feb 6 2012 06:33
13754952	200266	1,3,7	Feb 9 2012 16:11
13783364	200268	1,2,5,4,7,11	Feb 13 2012 11:23
13792138	200268	4,7,8	Feb 15 2012 07:58

SD event number	AERA run	stations	Date
13793053	200268	2,3,5,8,9,10	Feb 15 2012 12:56
13804976	200268	7,8,11	Feb 17 2012 00:58
13850080	200268	1,2,4,5,6,7,9,10,11,12	Feb 22 2012 11:49
13859462	200268	1,3,6,7,10	Feb 23 2012 11:33
13873939	200270	1,2,3,4,5,7	Feb 25 2012 01:25
13881125	200270	5,9,12	Feb 25 2012 21:33
13881591	200270	5,9,12	Feb 25 2012 23:11
13896507	200270	3,4,6,7,10,11	Feb 27 2012 07:45
13899064	200270	3,6,7,10	Feb 27 2012 14:52
13908159	200270	1,2,3,4,5,6,8,10	Feb 28 2012 14:27
13917612	200270	1,2,3,4,5,6,7,8,9,10,11,12	Feb 29 2012 12:05

Table 2: The event information of the coincident events, denoted as data set 2, measured between November 2011 and February 2012.

4 Parameterization of cosmic ray induced radio signals

The first observations of radio signals, produced by air showers were made by Jelley [21] in 1964. One of the first parameterizations for the radio pulses was made by Allan [22] in 1971. He provided the following parameterization for the signal in $\mu\mathrm{Vm}^{-1}\mathrm{MHz}^{-1}$.

$$\varepsilon_\upsilon = 20\left(\frac{\mathrm{E_p}}{10^{17}\mathrm{eV}}\right)\sin(\alpha)\cos(\theta)\exp\left[-\frac{\mathrm{R}}{\mathrm{R_0}}\right] \tag{1}$$

with $\mathrm{E_p}$ the primary energy of the cosmic ray, α the angle between the direction of the incoming cosmic ray and the geomagnetic field, R the distance from the detector to the shower axis and the falloff distance $\mathrm{R_0}$ was fixed to 150 m.

A different parameterization was found by Huege and Falcke [23], also in $\mu\mathrm{Vm}^{-1}\mathrm{MHz}^{-1}$.

$$\begin{aligned} |\mathrm{E}| &= \mathrm{f}\,\mathrm{E}_\theta\left(\frac{\mathrm{E_p}}{10^{17}\mathrm{eV}}\right)^{0.96}\exp\left[-\frac{200\mathrm{m}\,(\alpha(\mathrm{Xmax})-1)+\mathrm{l}}{\mathrm{l}_\theta\,\alpha(\mathrm{Xmax})}\right] \\ &\quad \exp\left[\frac{10-\upsilon\,(\mathrm{MHz})}{47.96\,\exp[-\mathrm{l}/\mathrm{b}_\theta]}\right] \end{aligned} \tag{2}$$

with the fudge factor $\mathrm{f}=1$, $\mathrm{E_p}$ the primary energy of the cosmic ray, $\alpha(\mathrm{X_{max}}) = 1.00636\left(\frac{\mathrm{X_{max}}}{631\,\mathrm{gcm}^{-2}}\right)^{-1.50519}$, l the perpendicular distance between shower axis at the station and υ the frequency in MHz. 55 MHz is used, as this is the middle of the frequency range of AERA. $\mathrm{X_{max}}$ is the atmospheric depth with the maximum number of particles, which is calculated as follows: $\mathrm{X_{max}} = 150 + 10\left(\log_{10}\left(\frac{\mathrm{E_p}}{10^{10}}\right)\right)^2$. A fit for the parameters E_θ, l_θ and b_θ was performed by Timmermans [24] with the following result:

$$\mathrm{E}_\theta = 6.1194 - 20.43\cos(\theta) + 26.546\cos^2(\theta)$$

$$\mathrm{l}_\theta = 238.34 - 515.9/\cos(\theta) + 416.67/\cos^2(\theta)$$

$$\mathrm{b}_\theta = 370.74 - 410.11/\cos(\theta) + 259.72/\cos^2(\theta)$$

The two formulae are very different in their amplitude. For an event with $\mathrm{E_p} = 10^{18}$ eV, $\theta = 45°$ and $\phi = 0$, $\varepsilon_\upsilon(\mathrm{R})$ (black triangles) and $|\mathrm{E(l)}|$ (red dots) are displayed in figure 13. It can be seen that both functions fall as a function of distance. The maximum amplitude of the Huege Falcke curve is six times smaller than the maximum of the Allan curve. The difference between the curves decreases with

increasing distance.

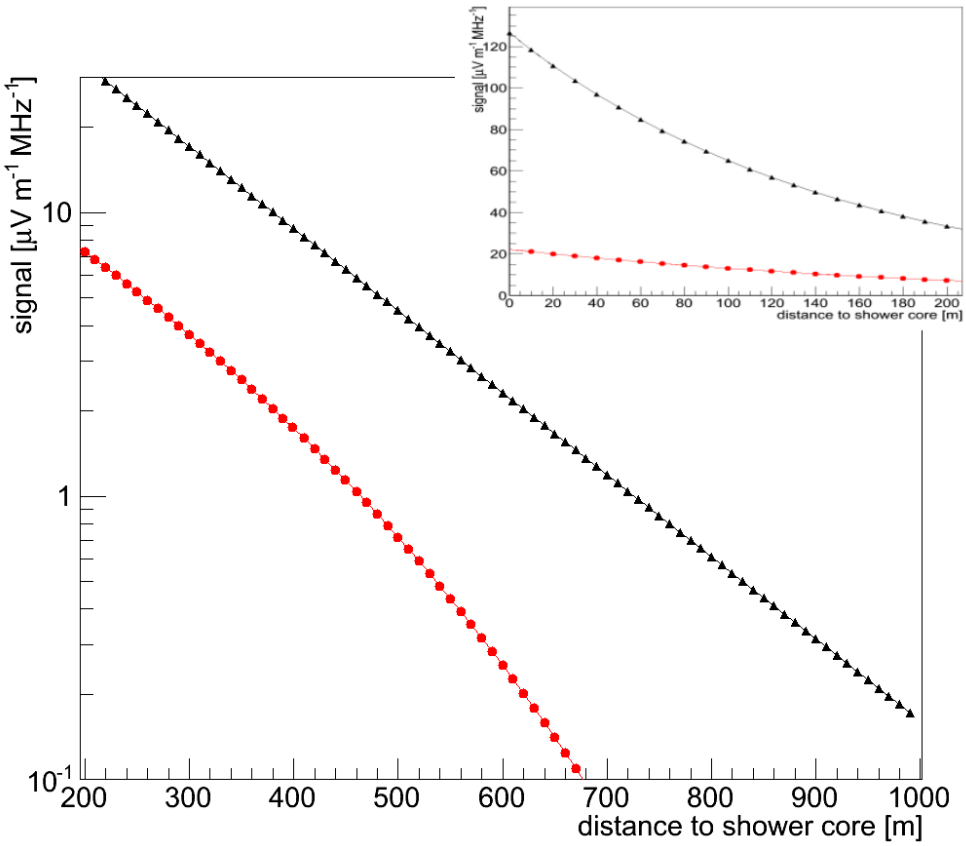


Figure 13: The signal, depending on the distance between station and shower axis for the formula by Allan (black triangles) and Huege and Falcke (red dots). For an event with the following parameters: $E_p = 10^{18}$ eV, $\theta = 45°$ and $\phi = 0$.

To get a feeling for the formulae, the signal is calculated for different angles for an energy of 10^{18} eV at a distance of 100 m between shower axis and observer in figure 14. The signal in color code is shown for the formula by Allan on the left hand side. There is a minimum when the shower axis is parallel to the magnetic field (in the north) and a maximum when it is perpendicular. The signal depending on θ is shown for the formula of Huege and Falcke as graph on the right hand side. The azimuth angle is only necessary for the calculation of the distance dependence, but the distance is fixed for this plots, so that there is no azimuth dependence. It is clear, that the signal calculated with Allan is much larger than the one calculated with the formula by Huege and Falcke.

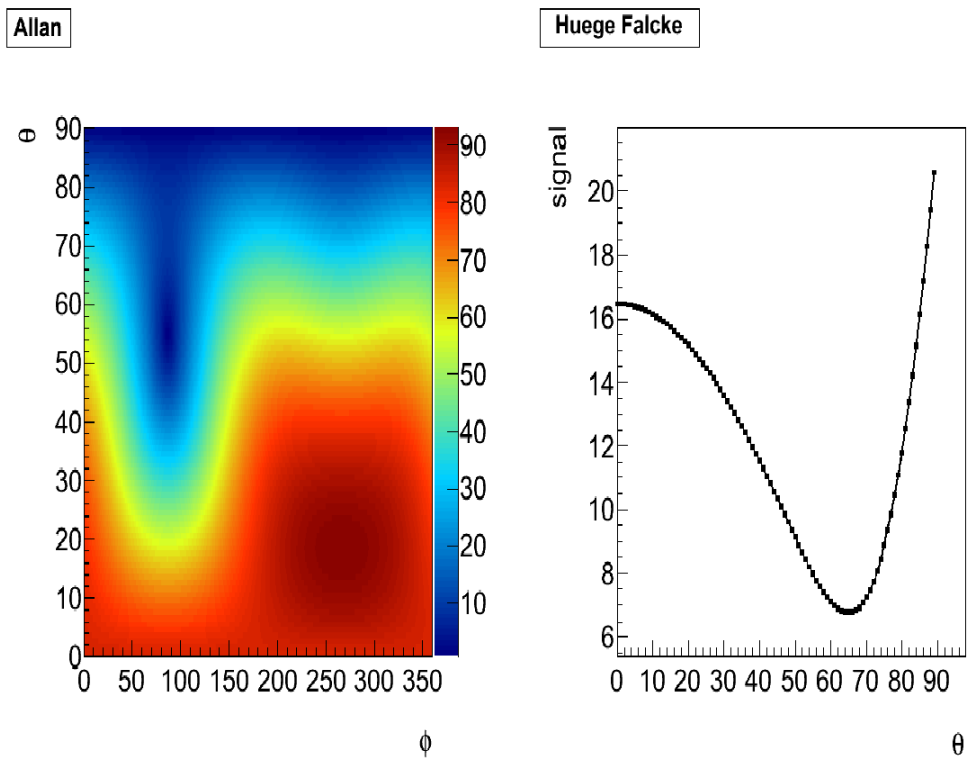


Figure 14: On the left side is the signal in color code according to the formula of Allan, depending on the azimuth and zenith angles. The right side shows the signal, according to formula 2, depending on the zenith angle, there is no direct azimuth angle dependence. A fixed distance of 100 m and an energy of 10^{18} eV is used.

4.1 Simulation of the effective area

To simulate the effective area, 1 million events are generated for different energies. The events are generated uniformly in cos θ between 0.5 and 1 and in ϕ between 0 and 360 degrees. The distribution of θ leads to more events with a larger angle and only a few events coming from the vertical. Also x and y are chosen uniformly, within the predefined surface as explained in the previous chapter. The energy is generated according to the energy spectrum given by Nagano and Watson [25].
The radio signal strengths are calculated in $\mu\mathrm{Vm}^{-1}\mathrm{MHz}^{-1}$, however the signals are measured in the stations in μV. Therefore, the signals according to the formulae are multiplied with 1 m and 50 MHz. The property 1 m corresponds roughly to the dimensions of the antenna and the 50 MHz corresponds to the frequency range, in which the signals are measured.
The signal must exceed a threshold to be measured. The threshold is fixed by components of the antenna and the trigger settings. A typical threshold is around 200 ADC which corresponds to a voltage output of the antenna of 66.66 μV. The formulae for the signal strength and the threshold are now translated into AERA specific formulae. The signals are as follows:

$$\varepsilon_v = 50 \times 20 \left(\frac{\mathrm{E_p}}{10^{17}\mathrm{eV}}\right) \sin(\alpha)\cos(\theta)\exp\left[-\frac{\mathrm{R}}{\mathrm{R_0}}\right]\mu\mathrm{V} \tag{3}$$

$$\begin{aligned} |\mathrm{E}| &= 50\,\mathrm{E}_\theta \left(\frac{\mathrm{E_p}}{10^{17}\mathrm{eV}}\right)^{0.96} \exp\left[-\frac{200\mathrm{m}\,(\alpha(\mathrm{Xmax})-1)+\mathrm{l}}{\mathrm{l}_\theta\,\alpha(\mathrm{Xmax})}\right] \\ &\quad \exp\left[\frac{10-\upsilon\,(\mathrm{MHz})}{47.96\,\exp[-\mathrm{l}/\mathrm{b}_\theta]}\right]\mu\mathrm{V} \end{aligned} \tag{4}$$

An event is considered to be measurable, if the calculated signal, using either formula 3 or 4 is above the threshold of 66.66 μV in at least three stations. This corresponds to a trigger requirement of three stations recording an event simultaneously to reconstruct the arrival direction.
The surface is divided into bins of 10 m $\times$ 10 m. The efficiency in each bin is defined as is the number of measurable events that have their core in this bin divided by the total number of events in the bin. Next the effective area $\mathrm{A_{eff}}$ is calculated as $\sum_{allbins} \mathrm{A} \times$ efficiency with A the area of one bin (100 m^2). The effective area is energy dependent. In figure 15 and 16 the array and the efficiency (in color code) in the different bins are shown for different energy ranges, for the Allan and Huege-Falcke formula respectively. Events with a low energy are only measured when the core falls in the array, and even then the efficiency is not large. With increasing energy,

the efficiency increases. Events with a core outside the array get a large probability to be measured. The effective area, depending on the energy and different areas, is shown for both formulae in figure 17.

When comparing the efficiency per bin and the effective area, it is clear that the larger signal of the formula by Allan (compared to the signal of Huege Falcke) leads to a larger area. Furthermore the events at the border of the surface are not measured by the array for energies below 10^{18} eV for Allan. Events with an energy above this value occur less frequently, so that the probability to measure such an event is small. Therefore, the cut on the core position is made on the reconstructed CDAS data. In the simulation with the Huege Falcke formula, the highest efficiency is in the center of a triangle. The distance from this point to one of the three corners is roughly the same for the three stations for vertical showers and the probability of a large signal in three stations is given. There are no events with an energy less then $10^{16.5}\,eV$ that have a signal in three stations above the threshold.

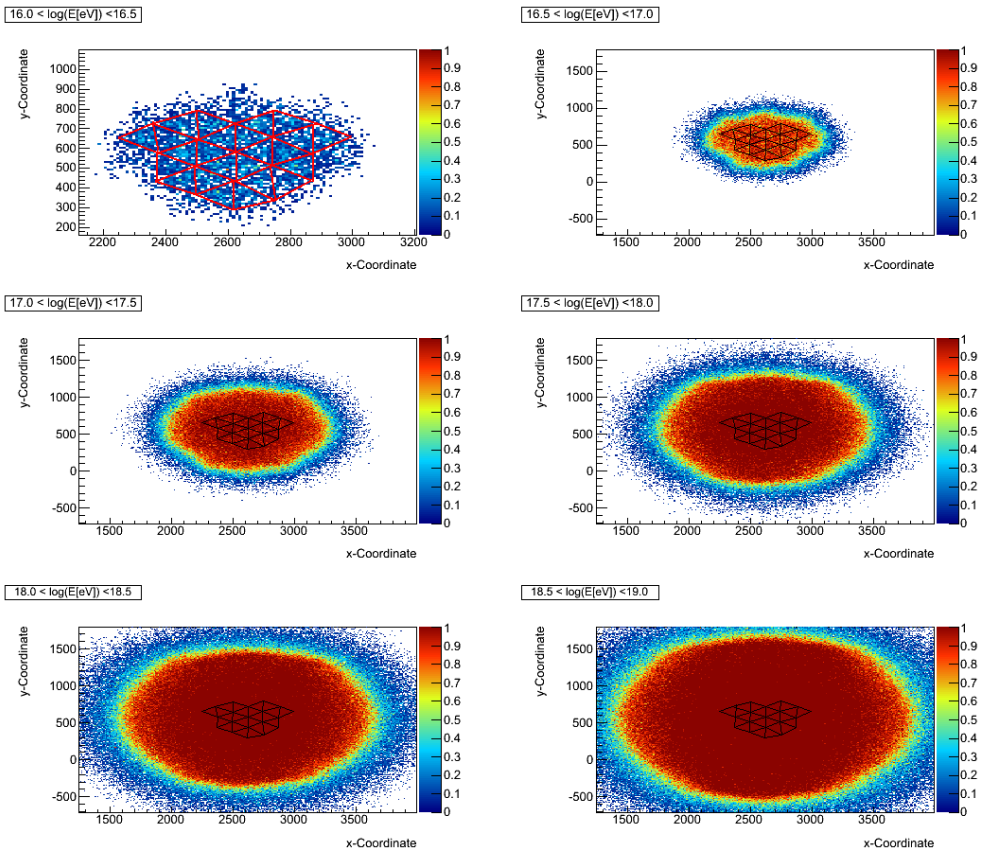


Figure 15: For each histogram, 1 million events are generated. When the signals, calculated with formula 3, of at least three stations, are above the threshold, the event is marked as measurable. The color code gives the number of measurable events, divided by the total number of events for each bin of 10 m × 10 m. The different histograms correspond to different energies. The efficiency grows with an increase of energy.

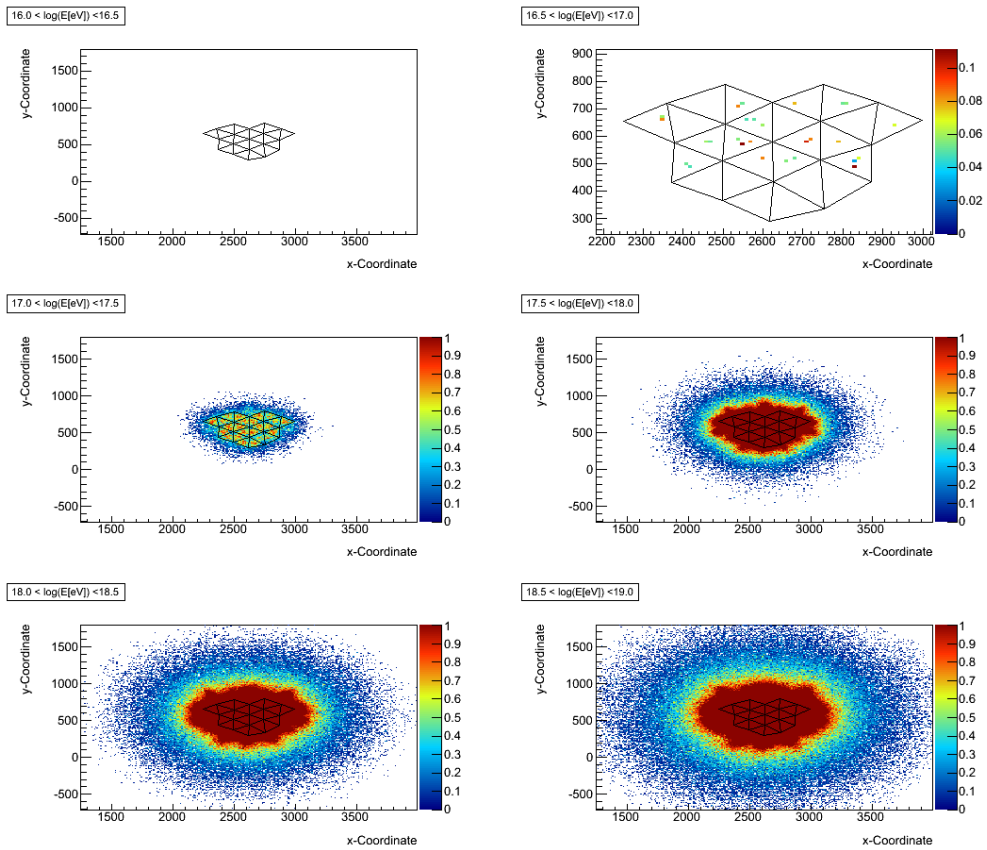


Figure 16: For each histogram, 1 million events are generated. When the signals, calculated with formula 4, of at least three stations, are above the threshold, the event is marked as measurable. The color code gives the number of measurable events, divided by the total number of events for each bin of 10 m × 10 m. The different histograms correspond to different energies. The efficiency grows with an increase of energy.

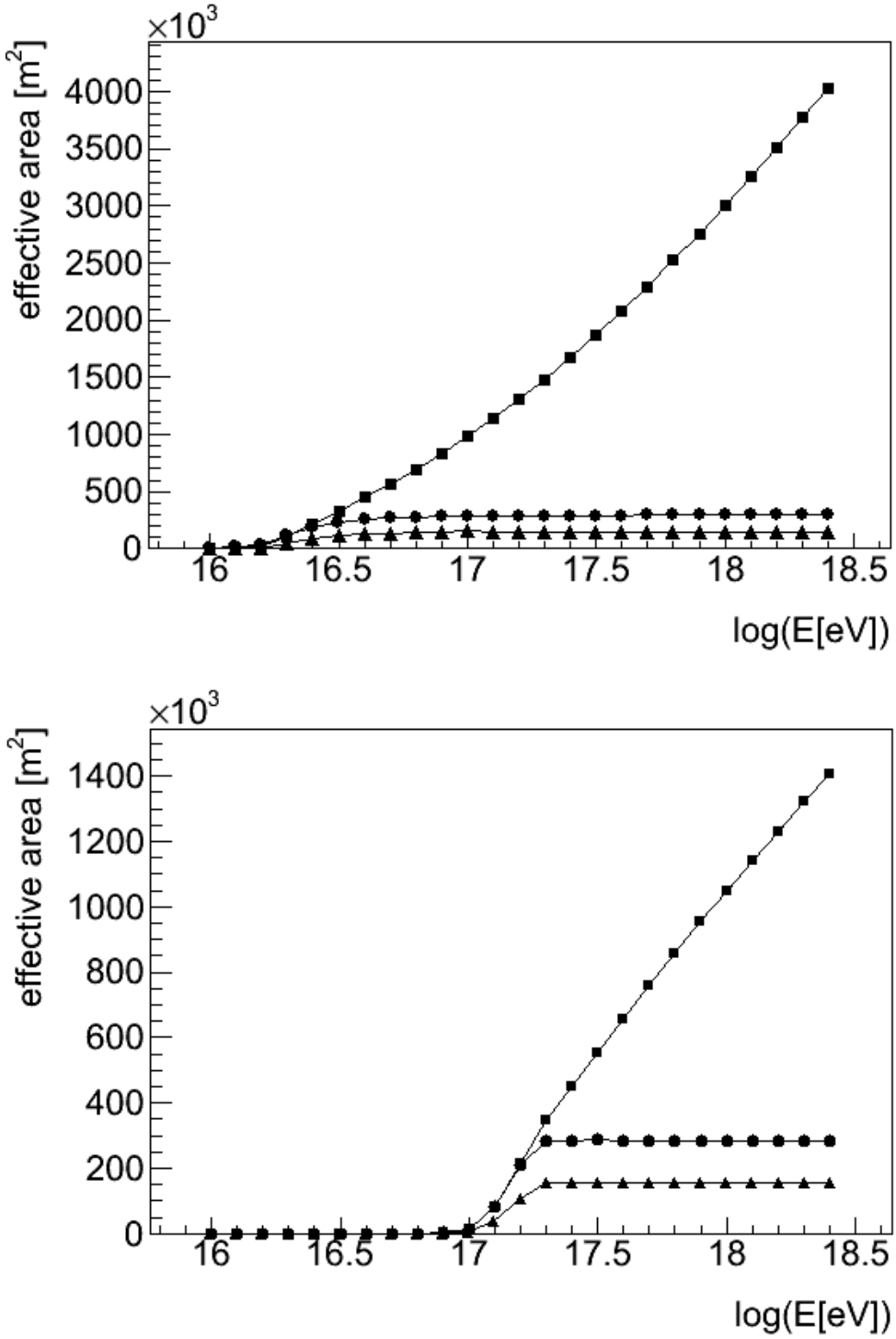


Figure 17: The effective area, calculated with formula 3 for the top plot and with formula 4 for the bottom plot. The effective area decreases, when decreasing the area. The effective area is calculated, taking into account the area as defined in section 3 (squares), or shower core positions that are within 150 m of three stations considering an array of the first 12 (triangles) or 20 (circles) stations. For each data point 1 million events are generated, with an uniform distribution of the shower core in the corresponding area, $0.5 < \theta < 1.0$ and $0 < \phi < 2\pi$. The threshold was 66.66 μV.

5 How to count SD events that should be measured by AERA

In this chapter, the data taken with the surface detector are compared to the data taken with the radio stations. The data period between the 15th of April and the 2nd of June 2011 is chosen. During the given period 11001 events are recorded with the surface detector only, while 28 events are recorded with the radio antennas and SD .
From the detected events the core position, θ, ϕ and the energy are known. For the 11001 events the radio signal strength is calculated using the formula 3 and 4. According to the calculation with the Allan formula 2151 events should be measurable with three or more radio stations. The events are also calculated with the Huege Falcke formula. According to this calculation 306 events should be measurable. Figure 18 shows the energy distributions of the events. The top left histogram shows the energy distribution of the 11001 events, while the middle and bottom histograms show the energy distribution of the measurable events, using the Allan and Huege/Falcke formula respectively. The efficiency, the number of measurable events divided by total number, for each bin, is shown on the right hand side.

5.1 AERA random trigger data

The arrival time of the measurable events is known. The next step is to investigate, why these events are not measured with AERA.
The random trigger data provides information on the status of each station. This data is recorded every 10 seconds. This means that 12 triggers are expected in 2 minutes (chosen length of a time bin). This is not the case for all time bins. With this information, the probability that an event is measured with a given station at a given time is calculated as the number of random triggers in the corresponding time bin divided by 12. If there are no random triggers, it can be assumed that the station was off.
For 1250 events, measurable according to the Allan formula, and 166 events, for the formula of Huege Falcke, random trigger data is taken (or available) and the probability that these events are measured can be calculated.
From the calculation it is known, which stations should record an event. Using the status of these stations the total probability to measure an event with three or more stations can be calculated. All possible combination off three or more stations above the threshold need to be taken into account. If p_i denotes the probability that

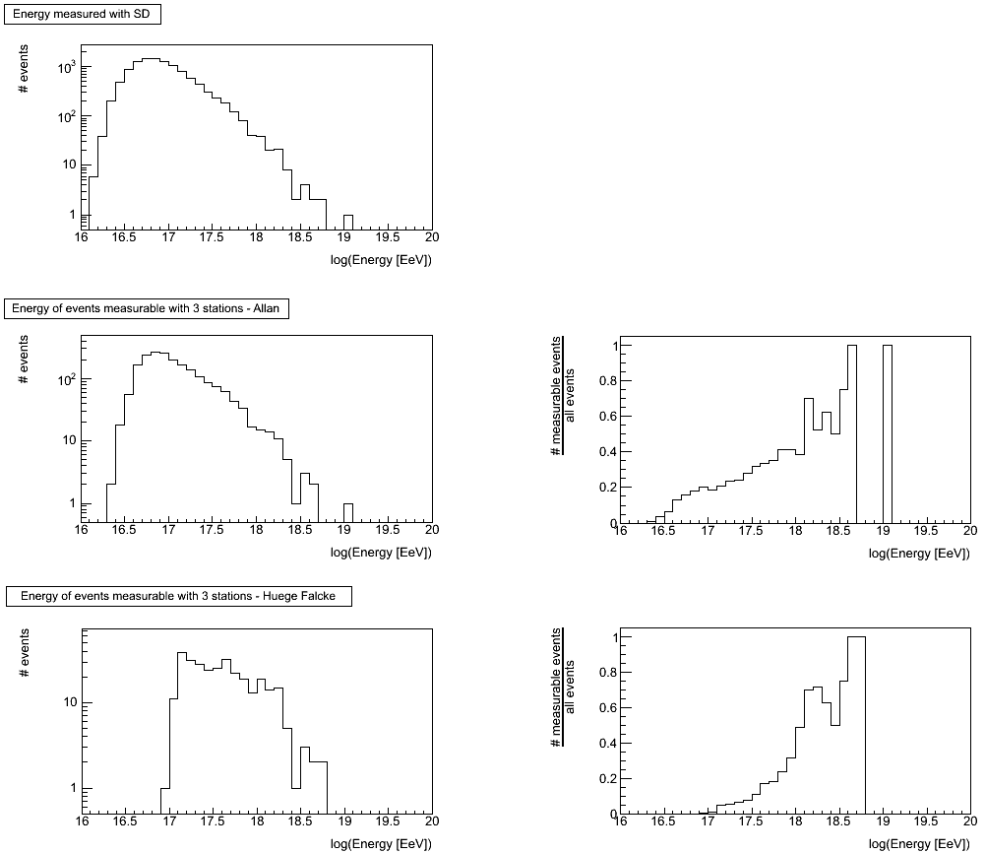


Figure 18: Between 15 April and 2 June 2011, 11001 events are measured with SD only, where the distance between shower axis and Kathy Turner is less than 2500 m. The energy distribution of these is shown in the top histogram. The middle left histogram shows the distribution of the events, that pass the threshold value using the Allan formula. The middle right plot shows the corresponding efficiency (the histogram to the left divided by the top histogram). The energy distribution from the events that would be measured according to the formula by Huege and Falcke is shown in the bottom left histogram, the corresponding efficiency in the right diagram.

station i has measured an event, $(1 - \mathrm{p_i})$ gives the probability that this station has not measured at this moment. $\mathrm{p_i}$ is only calculated for the stations that should have recorded a signal above the threshold for a certain event. The probability that a certain event is measured with three or more stations is then given by the formula:

$$
\begin{aligned}
\mathrm{P(m)} = \; & \prod_{\mathrm{n=1}}^{\mathrm{m}} \mathrm{p_n} \\
& + \sum_{\mathrm{a=1}}^{\mathrm{m}} (1 - \mathrm{p_a}) \frac{\prod_{\mathrm{n=1}}^{\mathrm{m}} \mathrm{p_n}}{\mathrm{p_a}} && \text{if } \mathrm{m} \geq 4 \\
& + \sum_{\mathrm{a=1}}^{\mathrm{m-1}} \sum_{\mathrm{b=a+1}}^{\mathrm{m}} (1 - \mathrm{p_a})(1 - \mathrm{p_b}) \frac{\prod_{\mathrm{n=1}}^{\mathrm{m}} \mathrm{p_n}}{\mathrm{p_a p_b}} && \text{if } \mathrm{m} \geq 5 \\
& + \sum_{\mathrm{a=1}}^{\mathrm{m-2}} \sum_{\mathrm{b=a+1}}^{\mathrm{m-1}} \sum_{\mathrm{c=b+1}}^{\mathrm{m}} (1 - \mathrm{p_a})(1 - \mathrm{p_b})(1 - \mathrm{p_c}) \frac{\prod_{\mathrm{n=1}}^{\mathrm{m}} \mathrm{p_n}}{\mathrm{p_a p_b p_c}} && \text{if } \mathrm{m} \geq 6 \\
& \ldots
\end{aligned}
$$

where m is the number of stations that should have recorded a signal above the threshold. The number of 'added sums' depends on m. If m = 3 non of the sums has to be taken into account. For m = 4 the first sum needs to be added. The last term that is added has to contain m − 3 summations. The first term gives the probability, that the event is measured with all stations that should have recorded that event. The sums give the probability, that the signal is unexpectedly not seen by a certain number of stations (the number of the sum symbols), but that all other stations (at least three) do measure the event.

The distribution of the probability, that at least three stations, that should have seen the event, were operational, of the 1250 events, according to the Allan formula, is shown in figure 19. The calculation with the formula of Huege and Falcke is shown in figure 20. There are 740 events for the calculation with the formula by Allan and 80 events for the Huege Falcke formula, that have a probability of greater than 0.99, that at least three stations, with a signal above the threshold, were operational.

The same calculation is done for the 28 events that are really measured with the radio stations. Due to the incomplete data set, the probability of 25 (Allan) and 21 (Huege Falcke) events is calculated. For 19 events calculated with the Allan formula and 14 events calculated with the Huege Falcke formula the calculated probability is greater than 0.99. This serves as a cross check of the efficiency calculation, because for all of these events the probability of recording the events is 1, as they are actually recorded.

5.2 Missing high probability events

The events that are measured with a probability greater than 0.99 according to the two formulae are investigated in more detail, as it is still not clear, why so many events are missing. Therefore the perpendicular distance between the shower and

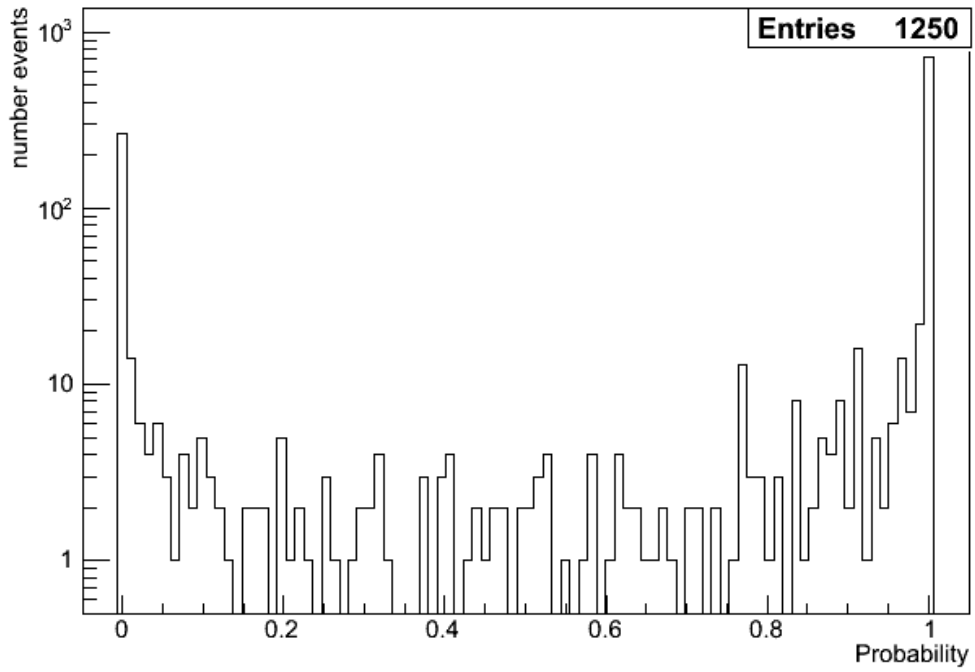


Figure 19: Probability that at least three RD stations, that should have measured a signal above the threshold according to the Allan formula, were operational. The random trigger data is not for all events available, so that the number of entries is different from the one in figure 18.

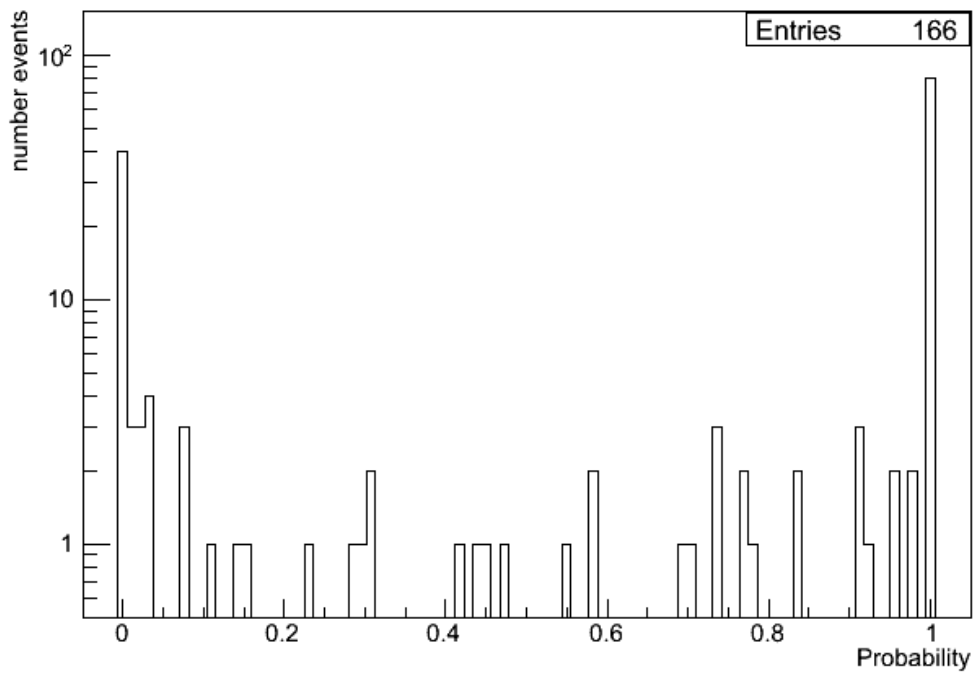


Figure 20: Probability that at least three RD stations, that should have measured a signal above the threshold according to the Huege Falcke formula, were operational. The random trigger data is not for all events available, so that the number of entries is different from the one in figure 18.

the radio station that should have recorded the event is plotted versus the logarithm of the energy of the event. This is done for events that are measured with SD only (left side), and for the events that are measured also with the radio stations (right side). Due to the limited statistics for the measurements, it is chosen to show the distance to all stations that should have measured an event, and not only to a single specific station. These histograms using the Allan formula are shown in figure 21. It is clearly visible, that according to the calculation most SD measurements are at an energy less than 10^{17} eV while the energy of the measured radio events is not in this range. Furthermore the calculation shows events that should have been measured by stations that are more than 600 m away from the shower axis. This is not seen in the radio data. The formula as used until now does not provide similar results for the SD events and the measured radio events. Therefore it is necessary to modify the formula.

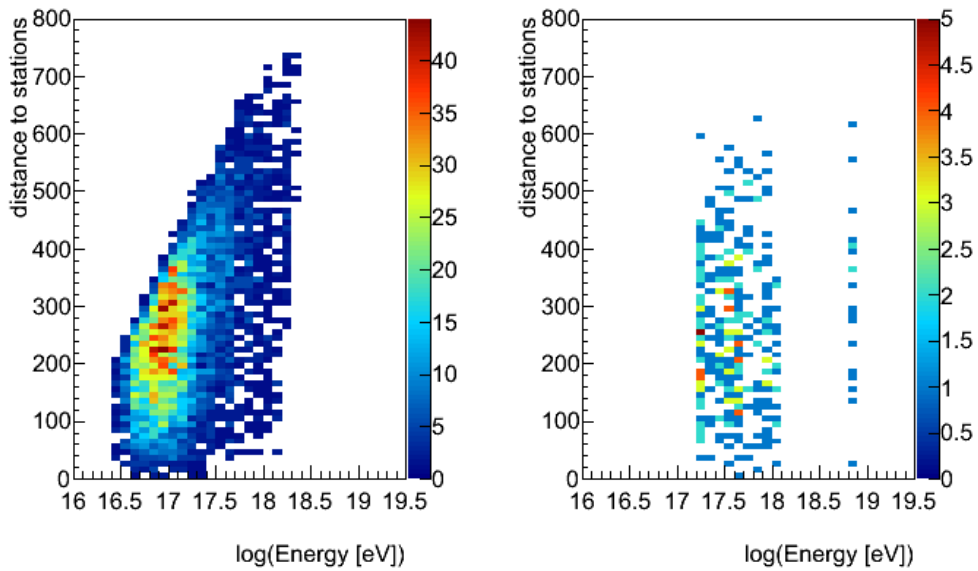


Figure 21: For the events that have a probability > 0.99, the perpendicular distance between shower axis and station is plotted versus the energy of the event. This is done for all stations, that should have a signal (Allan) above the threshold. The color code indicates the number of entries per bin. The left plot shows the expected SD events, while the right one shows the measured RD events.

The same comparison is made using the formula according to Huege and Falcke. The result is plotted in figure 22. It can be seen that also radio stations that are more than 1 km away would have a signal above the threshold. It is clear that also this formula should be modified to match the data, as the shape of the two histograms is not the same.

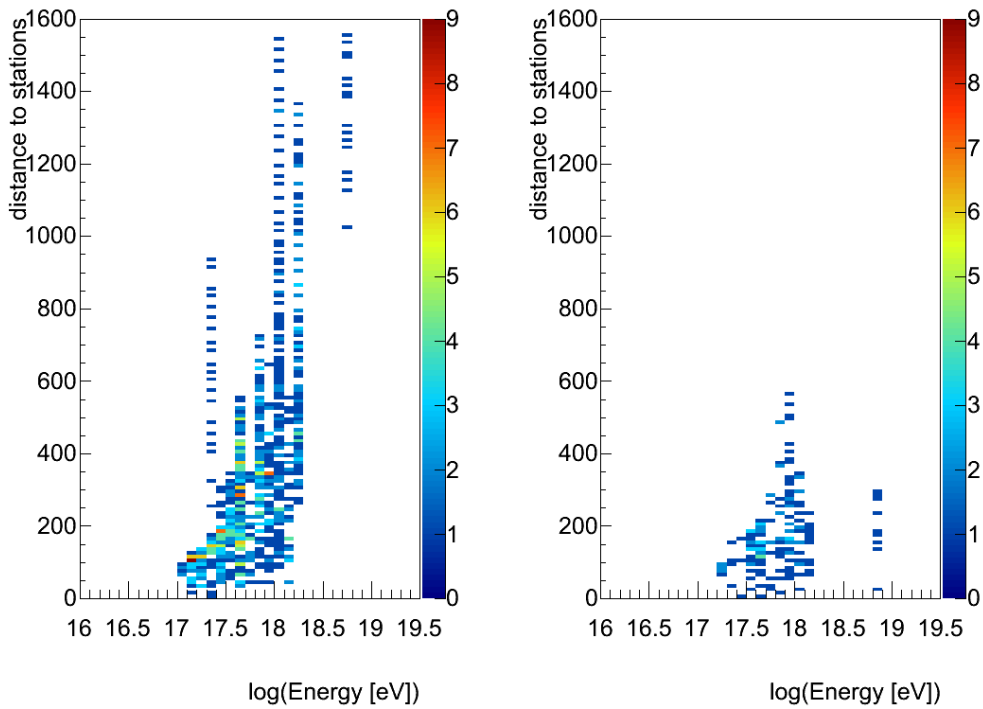


Figure 22: For the events that have a probability > 0.99, the perpendicular distance between shower axis and station is plotted versus the energy of the event. This is done for all stations, that should have a signal (Huege and Falcke) above the threshold. The color code indicates the number of entries per bin. The left plot shows the expected SD events, while the right one shows the measured RD events.

5.3 Modification of the parameterizations

The formula of Allan (formula 3) can be changed at different positions. First, the normalization can be changed into an other numerical constant, the energy dependence can be changed, by using a power of $E_p/10^{17}$ eV and R_0 can be changed to get an other distance dependence.
With a change of these three parameters it is possible to get a formula that seems to be more realistic. In order to remove low energetic events, that are hardly seen by the array, a power of $E_p/10^{17}$ eV is used. The power is increased from 1.0 to 1.5. An lower increase did not have a proper effect. An decrease of the fall of distance has the effect, that stations further away will not be simulated to have a signal above the threshold, as the signal of events is now larger than before, when $E > 10^{17}$ eV. R_0 is changed to 80 m (also 100 m was tried). Calculations with a modification of these parameters show also that the normalization needs to be changed due to the fact that the calculated signal was much stronger than the threshold of 66.66 μV. The numerical constant is changed from 20 to 5.5. Also slightly different modifications are tried, but it seems, that this fits best. The formula that is used to simulate the signal is therefore:

$$\varepsilon_\upsilon = 50 \times 5.5 \left(\frac{E_p}{10^{17}\text{eV}}\right)^{1.5} \sin(\alpha)\cos(\theta)\exp\left[-\frac{R}{80\text{m}}\right]\mu\text{V} \tag{5}$$

Using these modifications, figure 21 changes to figure 23. There are now 53 expected SD events and 19 measured events with a probability > 0.99. The energy distribution of these events is shown in figure 24.
Using the 53 events that are expected and not measured and the 19 events that are actually measured with radio, the efficiency of the array is calculated. There are in total $53 + 19 = 72$ events that should be measured according to the calculation while 19 are really measured, the average efficiency is therefore 26%. Also for this parameterization, the effective area is calculated, depending on the energy with a fixed threshold. And displayed in figure 25.

The formula by Huege and Falcke was of the form:

$$\begin{aligned} |E| &= 50 \times E_\theta \left(\frac{E_p}{10^{17}\text{eV}}\right)^{0.96} \exp\left[-\frac{200\text{m}\,(\alpha(\text{Xmax}) - 1) + C_1 * l}{C_2 * l_\theta\,\alpha(\text{Xmax})}\right] \\ &\quad \exp\left[\frac{10 - \upsilon(\text{MHz})}{47.96\,\exp[-C_3 * l/b_\theta]}\right]\mu\text{V} \end{aligned} \tag{6}$$

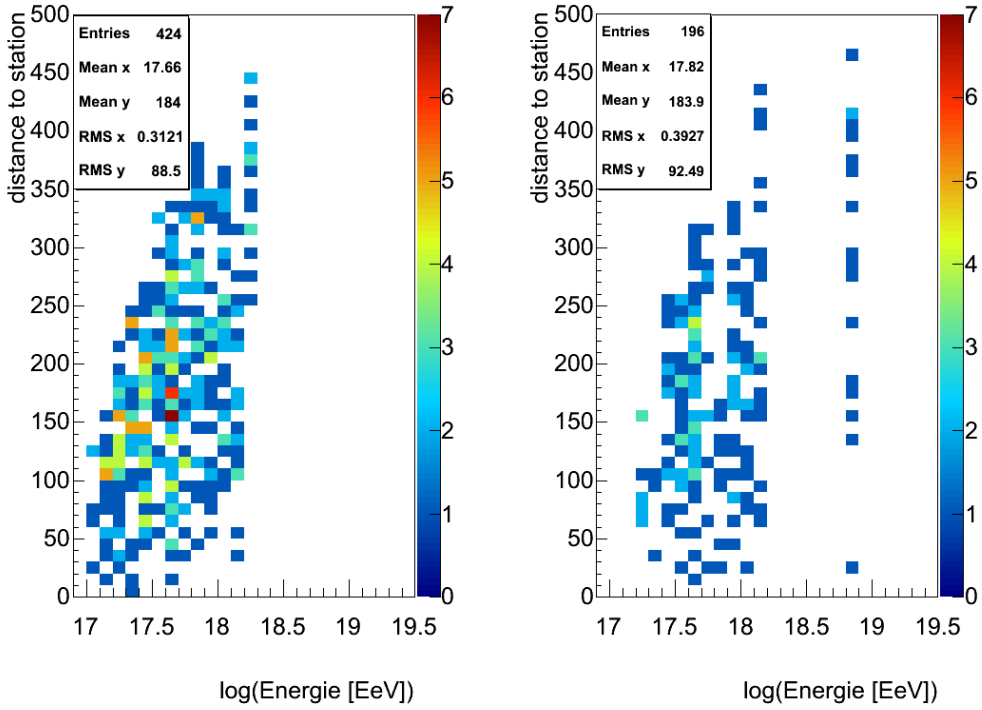


Figure 23: For the events that have a probability > 0.99, the perpendicular distance between shower axis and station is plotted versus the energy of the event. This is done for all stations, that should have a signal, according to formula 5, above the threshold. The color code indicates the number of entries per bin. The left plot shows the expected SD events, while the right one shows the measured radio events.

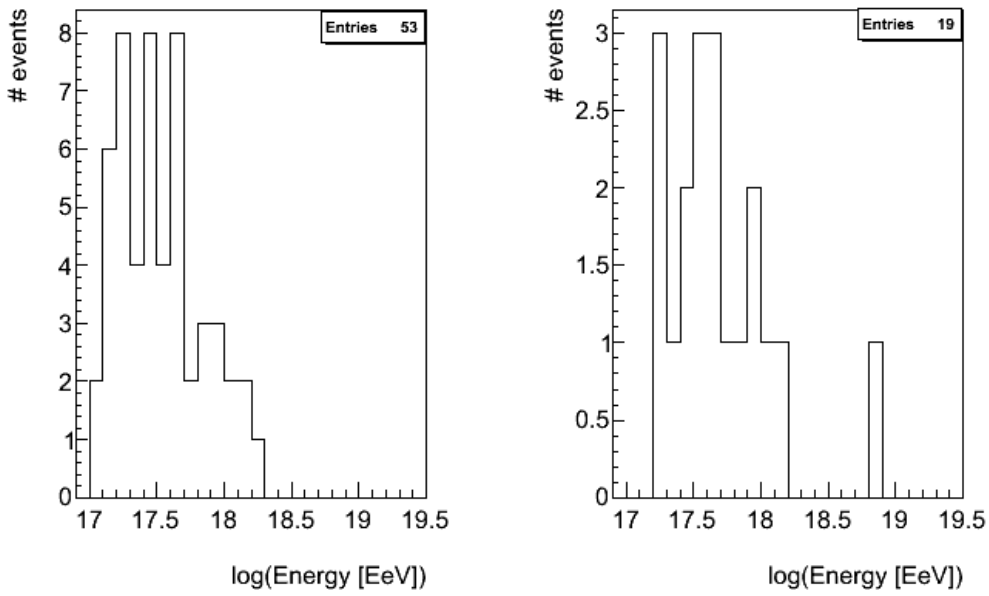


Figure 24: The energy distributon corresponding to the calculation with the new Allan formula for the SD / coincident events on the left/right hand side.

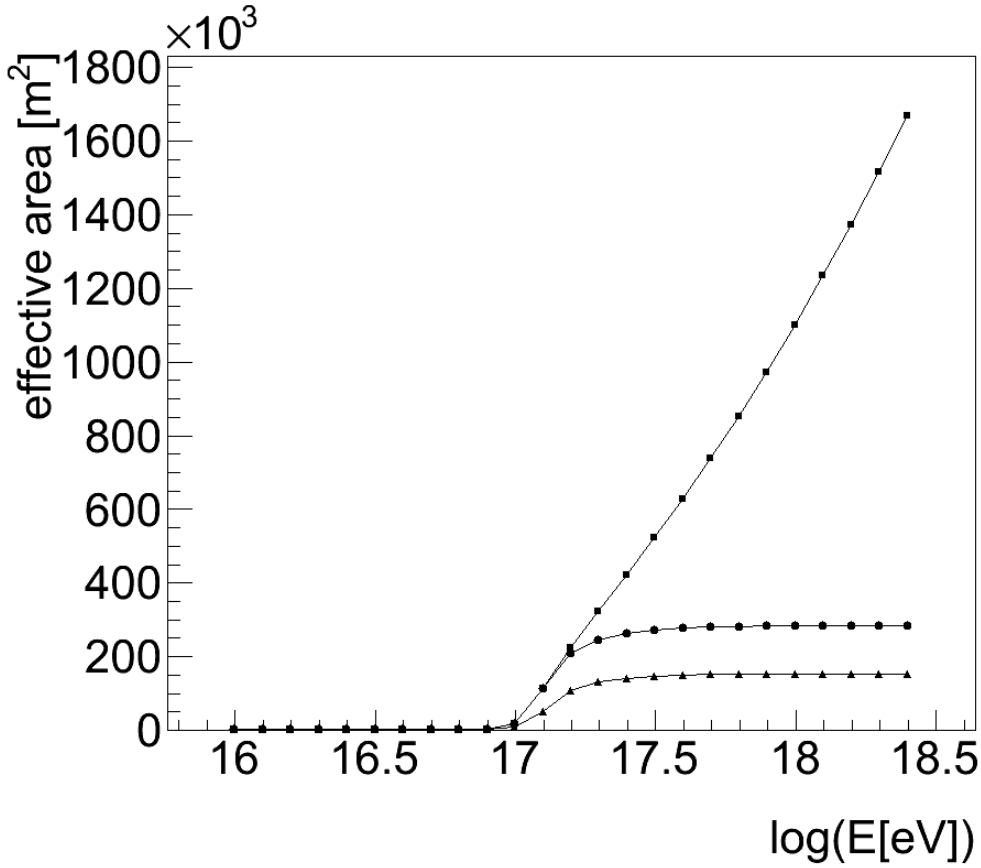


Figure 25: The effective area, calculated with formula 5. The effective area decreases, when decreasing the area. The effective area is calculated, taking into account the area as defined in section 3 (squares), or shower core positions that are within 150 m of three stations considering an array of the first 12 (triangles) or 20 (circles) stations. For each data point 1 million events are generated, with an uniform distribution of the shower core in the corresponding area, $0.5 < \theta < 1.0$ and $0 < \phi < 2\pi$. The threshold was $66.66\,\mu$V.

with $C_1 = 1$, $C_2 = 1$ and $C_3 = 1$. In order to remove events with an energy below 10^{17} eV the power of the energy dependence is changed to 1.35. To get the same shape for the SD events and the radio events in the energy versus distance histogram, different modifications of the numerical constants 200 m, C_1, C_2 and C_3 and the exponent of the energy are tried. Unfortunately it did not succeeded to get rid of SD events, where the distance between shower axis and station exceeds the 500 m. The only effect of the modification was a reduction of the number of expected radio events. This only leads to lower statistics.
It is not possible to easily change the formula of Huege and Falcke in such a way, that the distributions of the SD events, for which a radio counterpart is expected, matches with the one for the radio events.

6 Second data set and a time dependent threshold

Using the previous calculation, it was assumed, that the radio threshold was fixed. However, due to high fluctuations of the noise level, it was more efficient to vary the threshold as a function of time, at each station. The threshold of four stations vs the time is shown in figure 26. The threshold values, that are used in this and the following chapter are from November 2011 - February 2012 and are averaged for a period of 30 minutes. A station is triggered on the pulse in the east-west channel. That means, that the recorded signal must be above the threshold for this channel. For one event, there are different signals measured and reconstructed. First there is the raw trace in east-west direction, which is used for triggering. Then there is a reconstructed trace (Offline, also EW) and a signal that is calculated according to formula 5. Comparisons of the maximum of the reconstructed trace and the calculated signal as well to the maximum of the raw trace is shown in figure 27, for the November - February data.

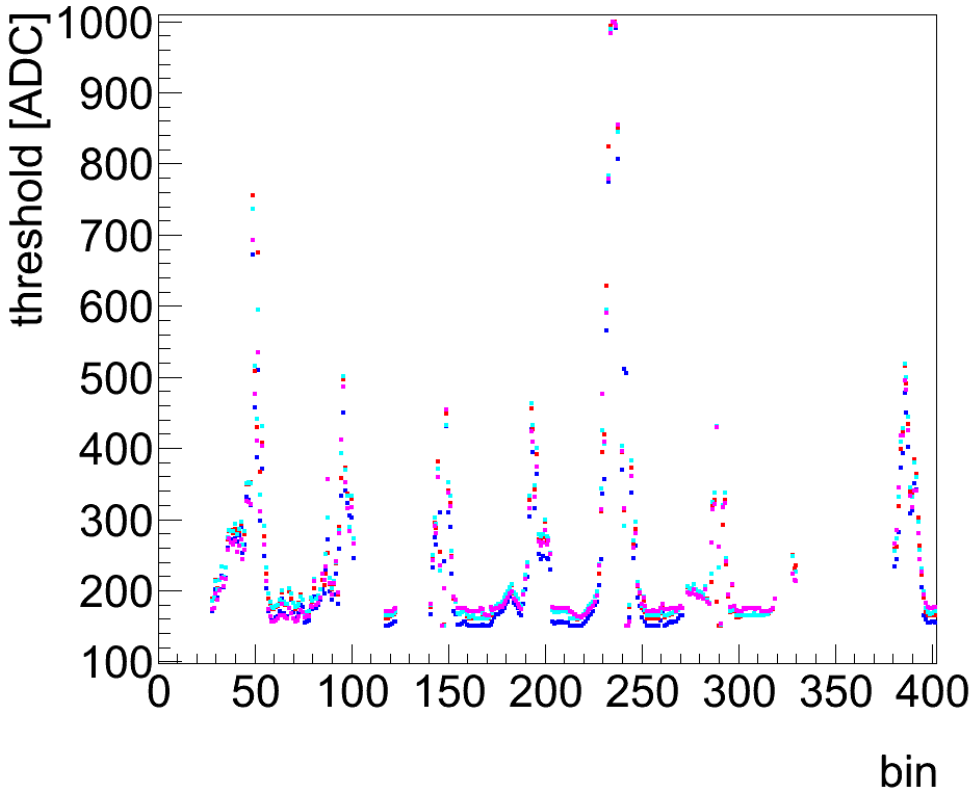
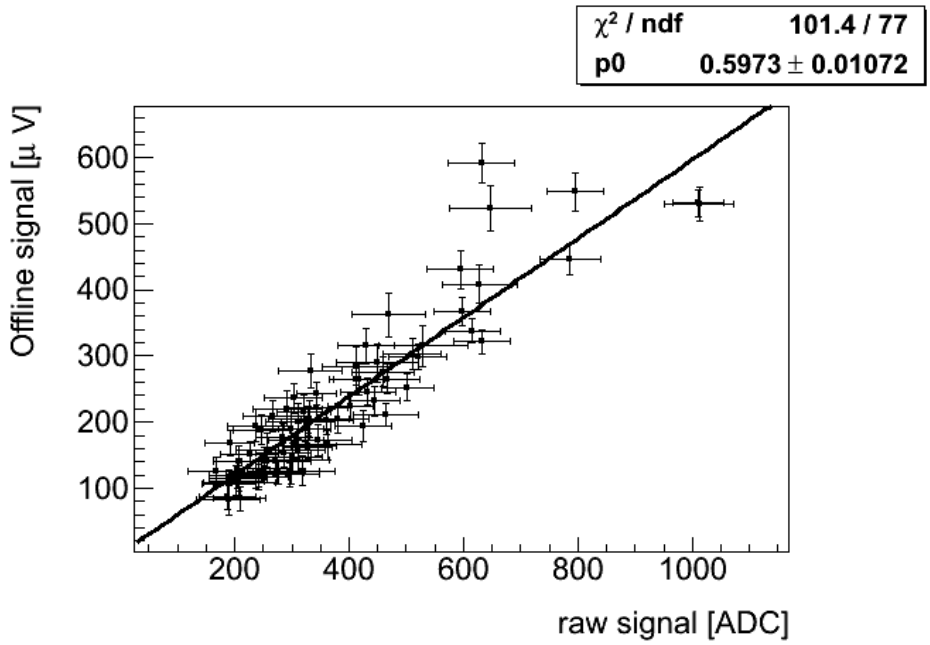


Figure 26: The threshold in east-west direction, in bins of 30 minutes, depending on the time for four stations (showing all stations would make the plot unreadable). Bin zero corresponds to the first 30 minutes of the 11th of November 2011.



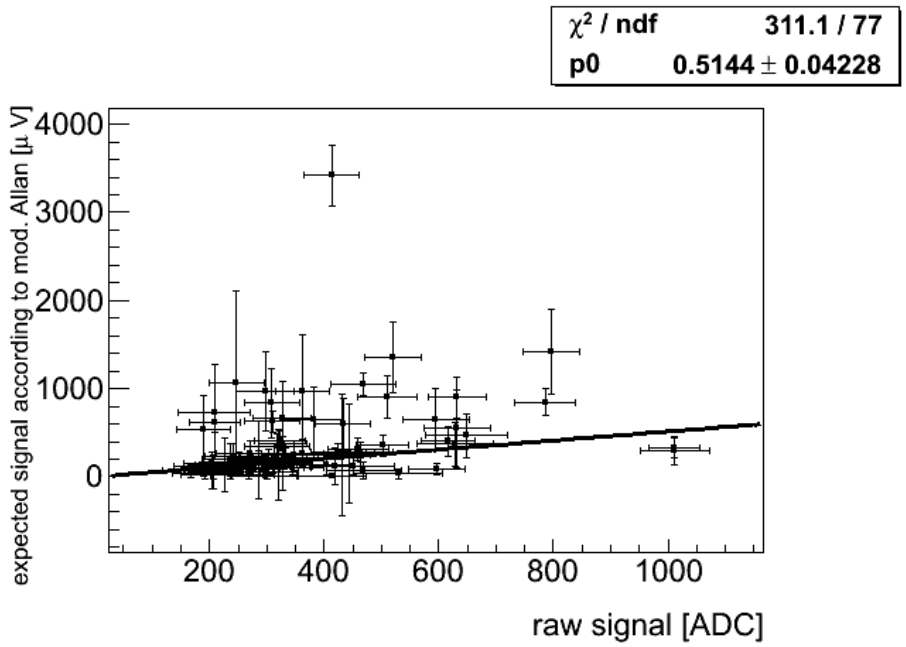


Figure 27: For each coincident event, the maximum amplitude of the raw signal and the maximum amplitude of the from Offline reconstructed signal in east west direction is known. Furthermore the expected signal is calculated according to formula 5. In the plots above a linear fit is performed for calibration. Coincidence events between November 2011 and February 2012 are used.

To convert the expected signal to the raw signal size, a linear fit is performed resulting in: $\text{Signal}_{\text{raw}}\,[\text{ADC}] = 1.944 * \text{Signal}_{\text{formula 5}}\,[\mu\text{V}]$. The measured RMS values of the noise for each event are used as the uncertainties of the raw and the Offline signal. The uncertainties of the expected signals result from the uncertainties in the shower geometry and the energy. The band around 1000 μV corresponds to stations, where the distance between station and shower axis is small. At these distances, the parameterization is bad.
A comparison of the raw signal to the threshold value is shown in figure 28. The threshold value shown corresponds to the average 30 minutes interval and is therefore not necessarily the true value at the moment of triggering. The raw signal and the threshold are both in ADC counts and therefore directly comparable. The station is triggered, when the raw signal exceeds the threshold. The one-to-one connection is indicated with the line. There is one point below the line. This can be explained to originate from the rough time binning of the threshold value.

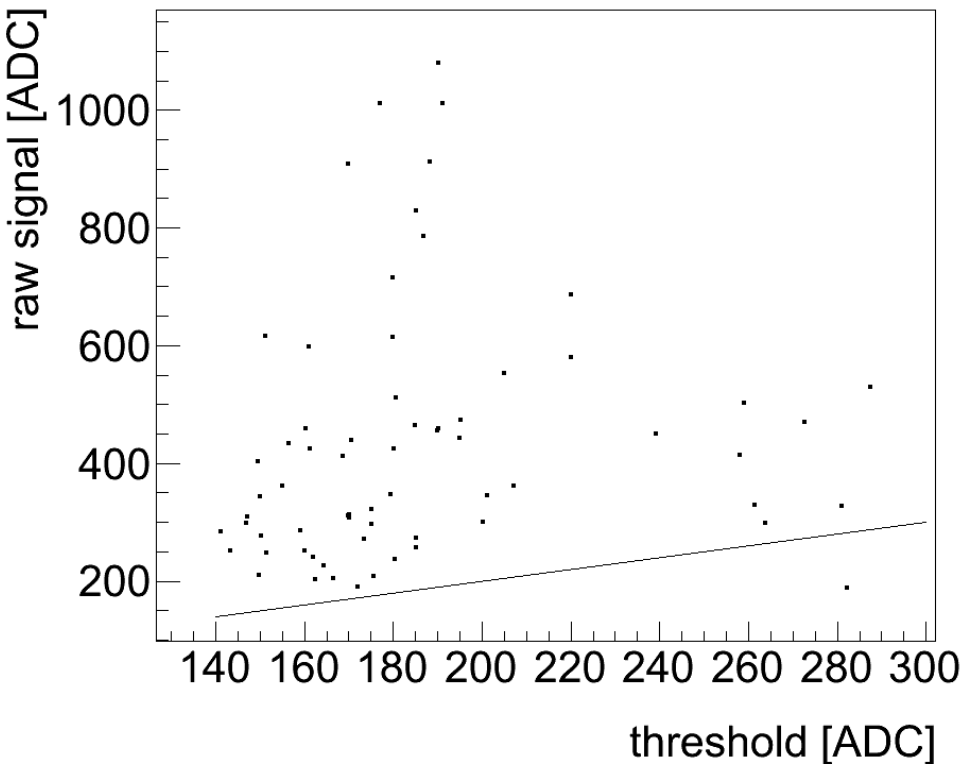


Figure 28: The maximum amplitude of the raw signal is shown versus the average threshold value. The line corresponds to the condition where the raw signal equals the threshold value and shows, that all data point, besides one are above the threshold.

In the previous chapter, the calculated signal was compared to a threshold value of 66.66 μV, which was stated to correspond to a threshold of 200 ADC. Between

these values is a factor three. The correlation found in this chapter between μV and ADC for the signals is almost a factor two. It seems, that the normalization factor chosen for the modified Allan formula is not correct.
Therefore, figure 23, is redrawn with the April - June data in figure figure 29. For the histograms signals in ADC and a fixed threshold of 140 ADC, as this is the minimum threshold, are used. The histograms have still the same shape. However, the wrong normalization factor is corrected with the fit between the raw signal and the calculated signal.

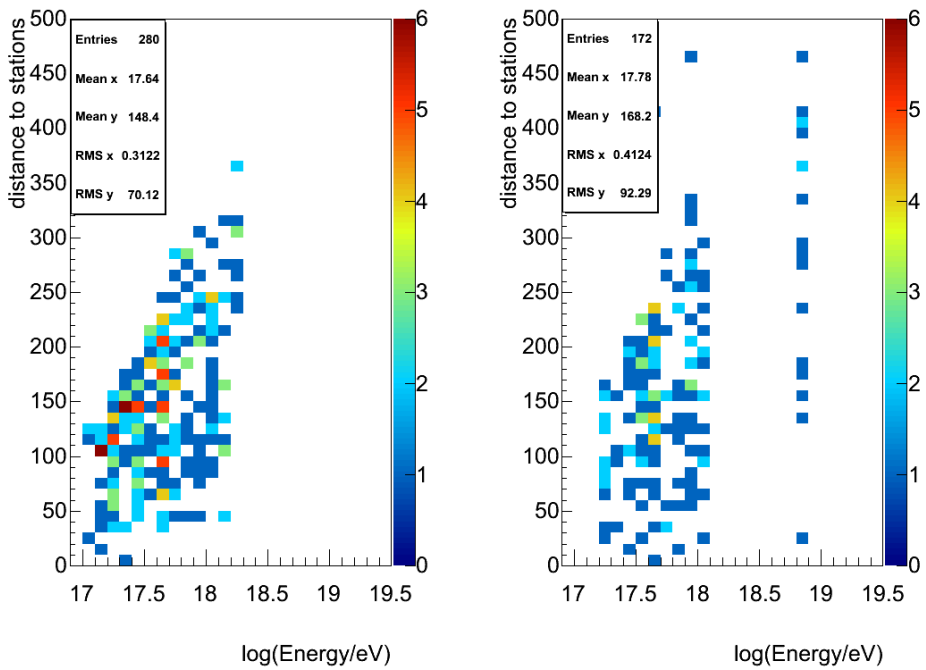


Figure 29: For the events that have a probability larger than 0.99, the perpendicular distance between shower axis and station is plotted versus the energy of the event. This is done for all stations, that should have a signal above the threshold of 140 ADC. The color code indicates the number of entries per bin. The left plot shows the expected SD events, while the right one shows the radio measured events between April and June 2011.

6.1 Comparison April-June data to November-February data

For the pure SD events, measured between November 2011 and February 2012, the expected signals are calculated in ADC and compared to a fixed threshold of 140 ADC. From the 9655 events (with an energy above 10^{17} EeV), 319 events are

	April-June	November-February
total number SD events	1101	9655 (with Energy cut)
number of stations	20	12
number of days	49	111
number of events expected to be measured with three or more stations with fixed threshold of 140 ADC	204	319
number of events with probability $>$ 0.99	45	161
number of events with probability equal to zero	117	65
number of events expected to be measured with three or more stations with variable threshold		181
number of events with probability $>$ 0.99		104
number of events with probability equal to zero		9

Table 3: A comparison of the number of events of the SD data of April-June and November- February.

expected to be measurable. Of these events, 161 have a probability that at least three stations, that should have seen the event were operational, which is greater then 0.99. When taking into account the time dependent threshold the number of events, with a probability greater than 0.99, reduces to 104. A direct comparison of the numbers, corresponding to the two data sets is shown in table 3. Also the signals of the April-June data set are calculated in ADC and the probability is recalculated.

For the second data set only 12 stations are available ($\sim$ half of the array), while the measured period is about twice as long. From this, the number of expected events should roughly be the same for both data sets. When taking only the fixed threshold and no random trigger data, the number of events is 204 for the April data set and 319 for the November data set. When taking into account the random trigger data, there are 117 and 65 events, for the April and November data respectively, that have a probability of zero to be measured. There are two reasons, that can lead to this probability. First, the stations have measured no random triggers and the stations are out. A second reason is, that random trigger data is not available for the whole period, which is the case for the April- June data. When the station is

	April-June	November-February
total number coincidence events	28	39
number of stations	20	12
number of days	49	111
number of events expected to be measured with three or more stations with fixed threshold of 140 ADC	24	32
number of events with probability $>$ 0.99	17	25
number of events with probability equal to zero	3	1
number of events expected to be measured with three or more stations with variable threshold		29
number of events with probability $>$ 0.99		24
number of events with probability equal to zero		0

Table 4: A comparison of the number of coincident events of the data set of April-June and November- February.

out, also the threshold is not available. Therefore, less events have a probability of zero in the time dependent calculation. The same comparison is made for the coincidence events, the result is summarized in table 4. Comparing the number of pure SD events and radio events, 24 events are measured, while 104 are not. The average efficiency is therefore 18% for the second data set.

6.2 Energy spectrum

With the information about the time dependent threshold, the number of SD and RD events measured in the given time, and a parameterization for the signal strength, it is possible to calculate the effective area and then the flux. The coincident events between November 2011 and February 2012 are used for this.
First, a cut on the core position and a cut on the zenith angle ($< 60°$) is performed. The core of the event must be within 150 m of three stations in the horizontal plane. The reason for the core position is, that the effective area is parameterization dependent and does not work properly for all events and distances. However it is shown, that the parameterization is working in the array. Outside the array, the effective area is strongly energy dependent.
The expected signals of the SD and coincident events are calculated for all events in the first 12 stations (in ADC counts) and compared to the threshold in the corresponding time bin. Events that have three signals above the threshold are taken into account for further calculations, where only the information of the stations is used, that have a signal above the threshold. The difference between signal and threshold can be seen in the top left histogram of figure 30 for the coincident events and in the top right histogram for all events (SD and RD). Both histograms are integrated from large differences to lower differences. The integration of the top left plot is divided by the integration of the top right plot. The result is shown in the bottom histogram. This gives the fraction of signals that is measured with RD, depending on the difference between signal and threshold. It can be seen, that only a small fraction of the events is measured with radio. Even if the difference between threshold and signal is large, the probability is not one, and air showers are not measured. A fit is done on the bottom histogram, to give each signal in the following calculation a probability to be measured. For the first part, a linear fit is performed. For the second part, a horizontal line if fitted. Resulting in the probabilities given in formula 7.
To calculate the flux, in $\mathrm{m^{-2}\,s^{-1}\,sr^{-1}\,eV^{-1}}$, the effective area as a function of time must be calculated. First, the energy range is divided into different energy bins. The first bin contains events with $0.1\,\mathrm{EeV} < \mathrm{E} < 0.25\,\mathrm{EeV}$, the second one events

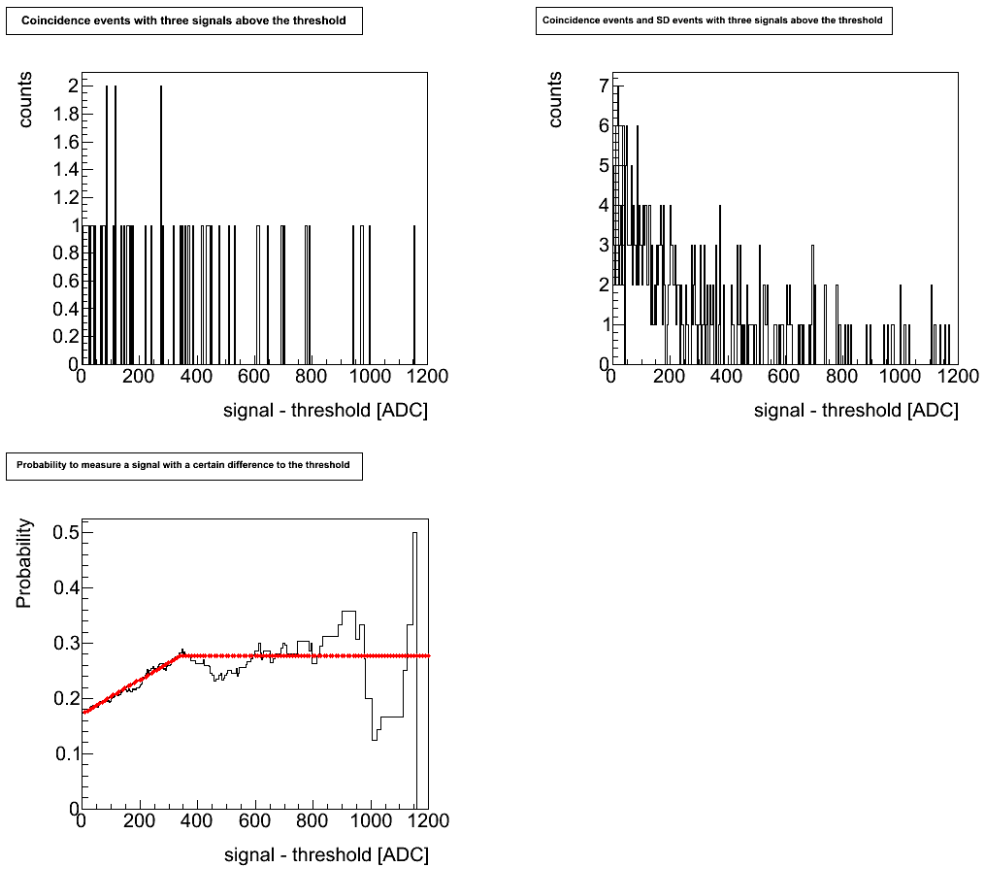


Figure 30: The top histograms show the difference between signal and threshold on the x-axis and the number of counts on the y-axis. The top left plot corresponds to coincidence events, the top right plot to SD events. The bottom plot shows an integration (from high to low) of the top left plot divided by the top right one. With a linear fit for the first part.

with $0.25\,\mathrm{EeV} < \mathrm{E} < 0.5\,\mathrm{EeV}$, the third one events with $0.5\,\mathrm{EeV} < \mathrm{E} < 1.0\,\mathrm{EeV}$ and the last one with $1.0\,\mathrm{EeV} < \mathrm{E} < 1.5\,\mathrm{EeV}$. The bins do not have the same width. Other bins are not investigated, due to the fact, that no RD events remain with higher or lower energies than the ones studied.
1 million events are generated for each energy bin, with a core within 150 m of three stations, and an uniform distribution of the angles with $0.5 < \cos\theta < 1.0$ and $0 < \phi < 2\,\pi$. Then, the signals are calculated for the 12 stations in ADC counts and compared to each threshold setting. For the given period, 5328 settings, each valid for 30 minutes exists.
To calculate the effective area for each threshold setting, the area is divided into bins of $5\,\mathrm{m} \times 5\,\mathrm{m}$. For each surface bin, a local efficiency(per area bin) is calculated. The total effective area (for one setting) is then the sum of all local efficiencies multiplied by $25\,\mathrm{m}^2$.
The local efficiency is the number of events that would be measured divided by the total number of events, that have their shower core in this bin. If the array would record all signals that are above the threshold, an event is counted if it has three signals above the threshold. But it is known, that the probability to measure an event is not one. The bottom histogram of figure 30 gives each signal a probability to be measured. The probability is :

$$\mathrm{P} = \begin{cases} 0.0 & \text{if signal-threshold} < 0.0 \text{ ADC} \\ 0.1699 + 0.0003141 \times \text{difference} & \text{if } 0.0 < \text{signal-threshold} < 340 \text{ ADC} \\ 0.277 & \text{if signal-threshold} > 340 \text{ ADC} \end{cases} \tag{7}$$

With this, the probability that an event is measured with three or more stations is calculated as in formula 5.1. The local efficiency is therefore calculated as $\sum_{\text{all events}} \mathrm{P(m)}$ divided by the total number of events (again for the events that have their core in that bin).
The total effective area per time, for the whole period is the sum of the effective areas of all threshold setting multiplied by 1800 s, due to the fact, that each threshold setting was valid for 30 minutes.
The flux is calculated as the number of RD events, that would be measured with at least three stations according to the formula, divided by the effective area per time, the steradian corresponding to a zenith angle of 60°, and the width of the energy bin. In the first two energy bins 4 coincident events are recorded, while there are 2 in the third one and one event in the last bin. The resulting spectrum is shown with the red squares in figure 31. The uncertainties on the x -axis correspond to the uncertainty on the energy range, while the uncertainty on the y-axis comes from the uncertainty of the number of measurements ($\sqrt{\mathrm{N}}$). The black squares with much

smaller uncertainties correspond to the energy spectrum calculated with the AMIGA infill area data [26]. The spectrum shown corresponds to the events with an zenith angle of less then 30°. A spectrum for larger angles is similar, but has no data points for energies smaller than $10^{17.55}$ eV. Both spectra, the one based on radio data and the one based on the AMIGA data, are the same within uncertainties.

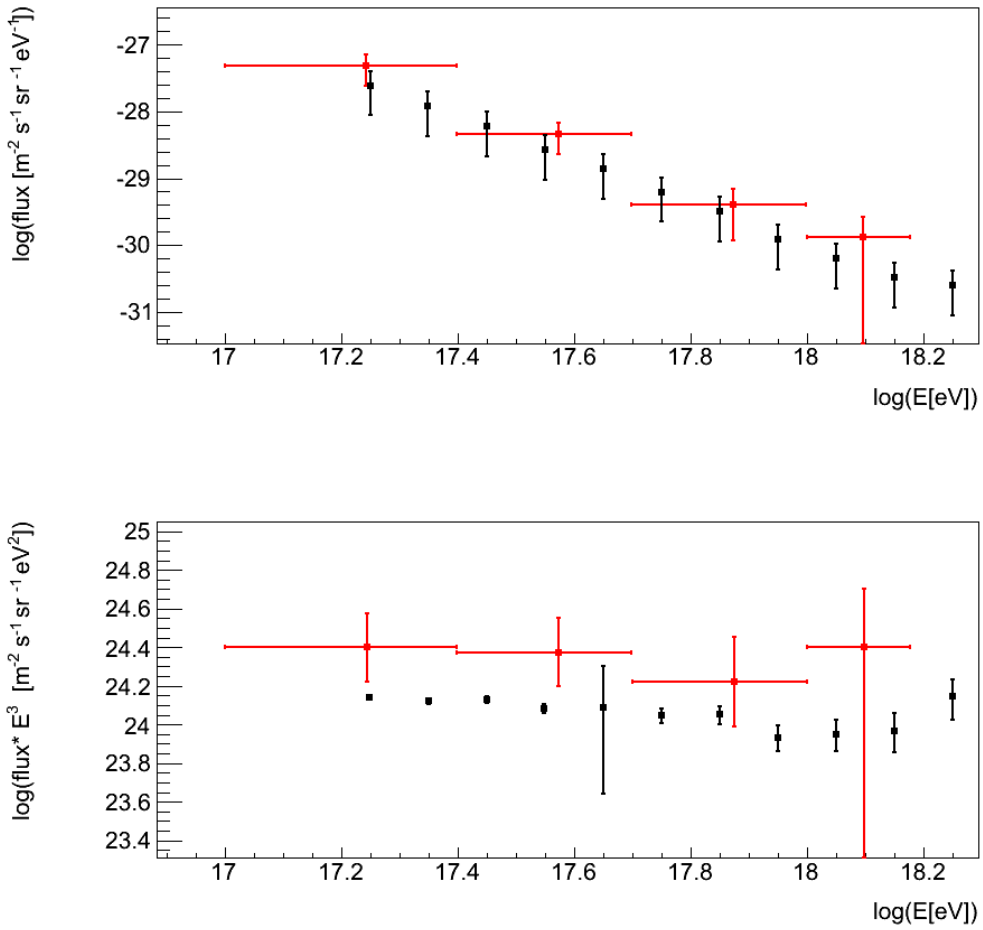


Figure 31: The top plot shows the logarithm of the flux, depending on the energy. For the bottom plot the flux is multiplied with E^3. The red data points corresponds to the flux calculated on the radio data, the black data points are obtained from AMIGA.

However it must be mentioned, that the obtained spectrum is not SD independent. The surface detector data is necessary, to calculate a probability, that an air shower is seen with a certain station. Therefore, the effective area depends on SD information. Not taking a probability into account, the effective area wound be much larger, resulting in a lower flux. Furthermore the energy estimate of the SD reconstruction is used.

An other fact that must be mentioned is, that the raw signal - expected signal dependence is calculated with the results of the Offline reconstruction, while the expected

signals for figure 30 are calculated with the parameters given from the CDAS reconstruction. Due the different energy, core and angle reconstruction, the expected signal is different. The expected signal, according to formula 5, is calculated for both reconstructions and plotted in figure 32.

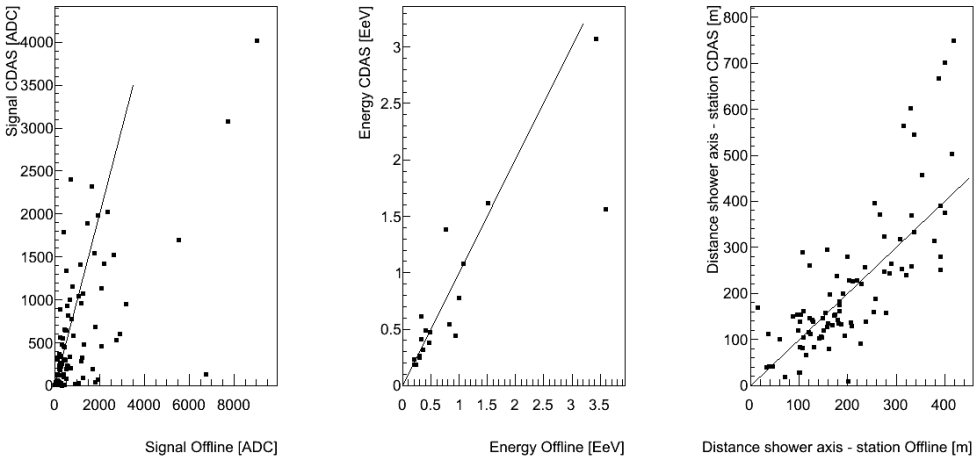


Figure 32: The left plot shows the expected signal, calculated with the parameters for Offline and CDAS versus each other. The middle plot shows the the energy estimated according to the two reconstructions for the same events. The right plot shows the distance between shower axis and station (only stations that have measured the shower in reality). The lines show the one - to - one correspondents.

7 A data based parameterization

Using the coincident events between April 2011 and February 2012, I have tried to find a new parameterization for the radio signal strength. The data set contains events that are reconstructed with Offline and that have no indication for thunderstorms. Other conditions are, that the events are measured with at least with five or more SD stations and that the angular difference between the radio and the SD reconstructed direction is less than 15.0°. Furthermore a cut on the zenith angle ($< 60°$) is applied. After this cut, 28 events with in total 126 measurements/stations remain for finding a new parameterization. Due to the fact, that the data is triggered on the signal in the east-west polarization, it was tried to find a parameterization for this signal. Therefore the maximum amplitude of the reconstructed (Offline) electric field in east west direction is used.
Due to the geometry of the problem a sin α, cos θ, energy and distance dependence (station - shower axis) is expected. The big questions are the shape of the energy and distance dependence. It is possible to divide the data set into bins, either energy or distance bins, depending on the parameter, that should not be taken into account in the first step. The results of the division into distance bins is shown in this chapter. A division into energy bins is problematic. All stations of an event will be in the same bin, and there are bins with only one event. There is not enough data to find a good parameterization.

7.1 Division into distance bins

The first step is to divide the data set into different distance bins. The first bin contains measurements, where the distance between shower axis and station is less than 80 m, and the final bin contains events with a distance greater than 320 m. In between each bin has a range of 40 m. In figure 33 the dependence of the signal on the energy, E, is shown. A linear fit through the origin is made. It can be seen, that the signal increases with the energy. A linear fit does not seem not to be the best choice for all plots. The figures in the appendix B show the signal versus the energy multiplied with sin α cos θ for the different distances. Also for this plots, a linear fit is performed. From the linear fit, we find an energy sin α cos θ dependence in each bin. The slope is plotted in figure 34, depending on the distance. The distance is the average distance of the data points in each bin. From this figure, we decided to perform an exponential fit on the slope as a function of the distance. With the fit for the slope the distance dependence can be found. The formula for the signal

becomes than:

$$S = e^{(f+gR)} \times \frac{E}{10^{18}} \times \sin\alpha \times \cos\theta\,\mu V \tag{8}$$

with $f = 7.17 \pm 0.16$ and $g = -0.0046 \pm 0.0011$ the parameters of the fit of the slope, R the distance between shower and station in meter and E the energy in eV.
The differences between the signal calculated with this formula (Signal_c) and the Offline signal (Signal_m) can be seen in figure 35. The line in the top plot represents the optimal position of the data points, if both quantities are equal. It can be seen, that the points are clustered around the line. The large uncertainties are due to the uncertainties in the core position and energy, which are taken into account. The bottom figure shows the difference between the calculated and the measured signal, normalized to the Offline signal. This is distributed round zero. A Gaussian fit is performed. Dependencies between the normalized difference and the different parameters are not found. A disadvantage of this formula is, that the calculated signal strength is in the order of the threshold level, even for low energetic events. This results in a maximum effective area, as can be seen in figure 36, when the probability to measure an event is one, when having three signals above the minimum threshold. Therefore I conclude that the formula is a good parameterization for the measured data, but does not describe the signal strength of low energetic events, that are not measured with AERA.
The formula can be written in the same form as the Allan formula. The normalization factor would be 1299.8 and the fall of distance would be 217.4 m. The decrease of the signal strength due to the distance dependence is compensated by the large normalization factor.

7.2 Energy spectrum

With the same method, as described in the previous chapter, the energy spectrum is calculated. The recalculation from the signal to the raw signal is : $\text{Signal}_{raw}[\text{ADC}] = 2.73 * \text{Signal}[\mu V]$ (corresponding plot is in appendix B). The probability that an event is measured by a station is given as follows:

$$P = \begin{cases} 0.0 & \text{if signal-threshold} < 0.0 \text{ ADC} \\ 0.1416 + 0.0004929 \times \text{difference} & \text{if } 0.0 < \text{signal-threshold} < 262 \text{ ADC} \\ 0.2631 & \text{if signal-threshold} > 262 \text{ ADC} \end{cases} \tag{9}$$

Also the figure corresponding to this fit is in the appendix.
The effective area is not so different from the one calculated with the other formula, when taking into account a signal dependent probability. A plot with all spectra,

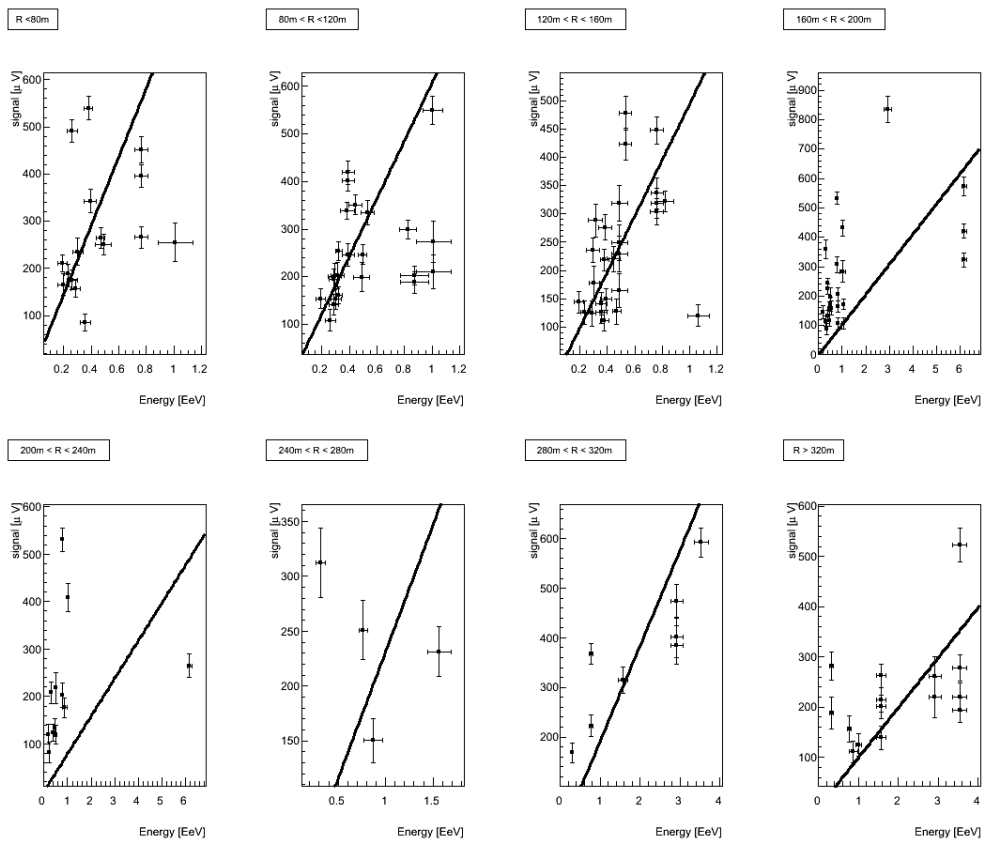


Figure 33: The maximum amplitude of the signal in east west direction (Offline) versus the logarithm of the energy in EeV, for different distances.

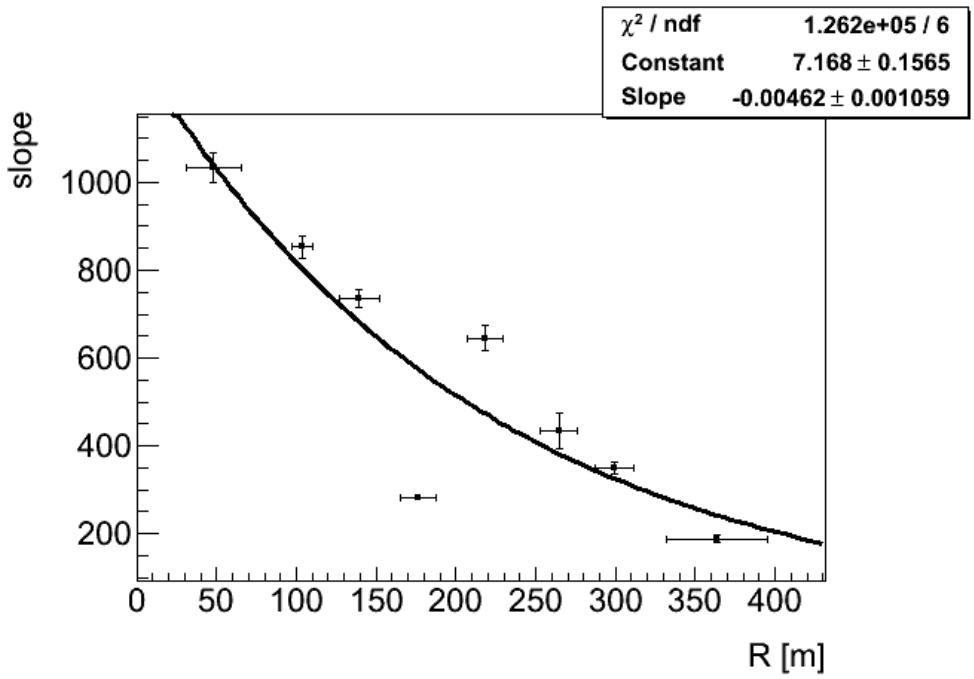


Figure 34: The slope, of the linear fit of figure 44 - 47 depending on the average distance.

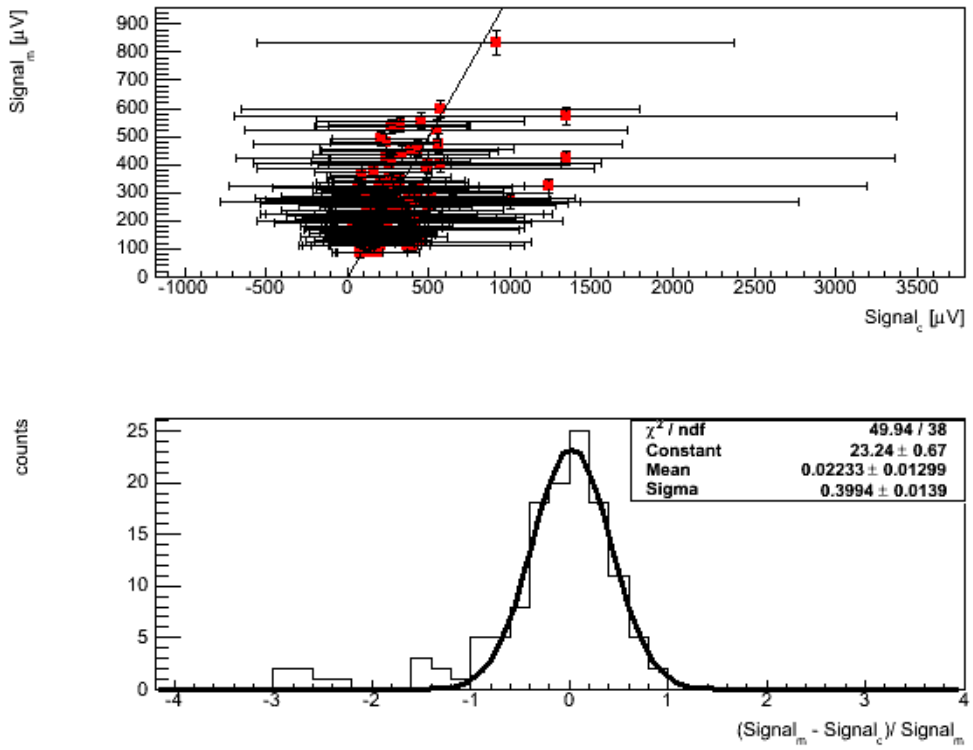


Figure 35: The Offline reconstructed signal (Signal_m) is plotted versus the calculation according to formula 8. The line in the top plot is the one - to - one connection. For the bottom plot a Gaussian fit is performed.

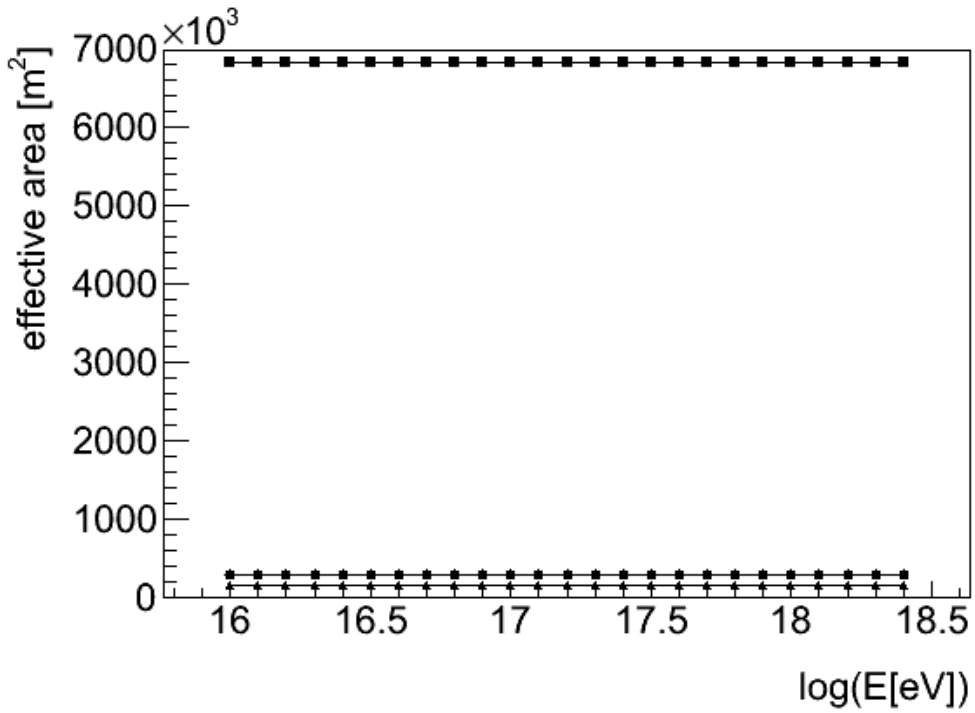


Figure 36: The effective area calculated with the new formula, taking into account the area as defined in section 3 (squares), or core positions that are within 150 m of the first 20 (circles) or 12 (triangles) stations. For each data point 1 million events are generated, with an uniform distribution of the shower core in the corresponding area, $0.5 < \theta < 1.0$ and $0 < \phi < 2\pi$. The threshold was 140 ADC

black for the spectrum of the infill area, red for the spectrum using formula 5 to simulate the radio signal and blue for the calculation with this formula are shown in figure 37. The number of radio events is slightly different, as there are now 5 events in the first bin. The flux corresponding to the second and third energy fits better to the infill spectrum than the one from the previous calculation.

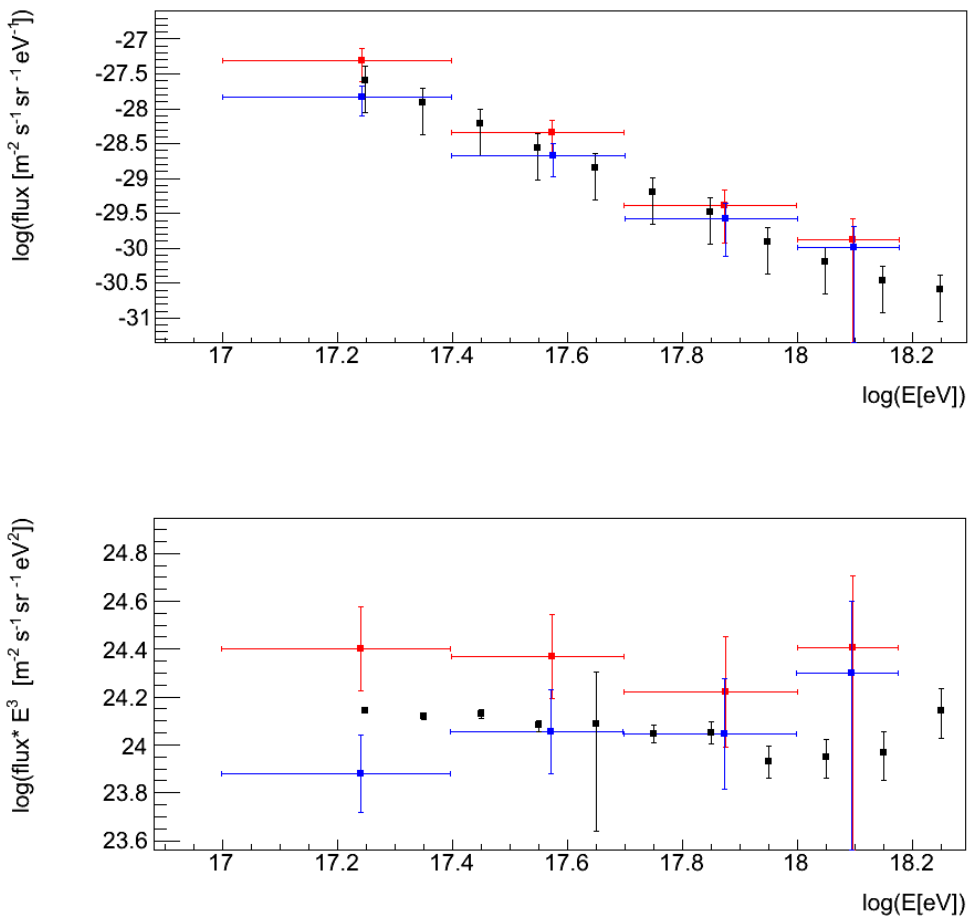


Figure 37: The energy spectrum obtained with the new formula, blue data points. The spectrum calculated in the previous chapter is shown with the red markers, and the spectrum of the infill area with the black markers.

7.3 Comparison of the two formulae

For a comparison of the formulae, the plots, as in figure 35, are made for the parameterization according to formula 5 and the new formula, with respect to the raw signal. The signal in μV is transformed to ADC with the corresponding factors. The plots are shown in figure 38 and 39 for the modified Allan and the new formula respectively. The Gaussian fit for the calculation with formula 5, has a mean of 0.03,

the mean of the fit of the new parameterization is is at 8.3, on average is formula 5 the better parameterization. The Gaussian fit on the comparison of the new formula with the Offline reconstructed signal was with a mean of 0.02 much better.

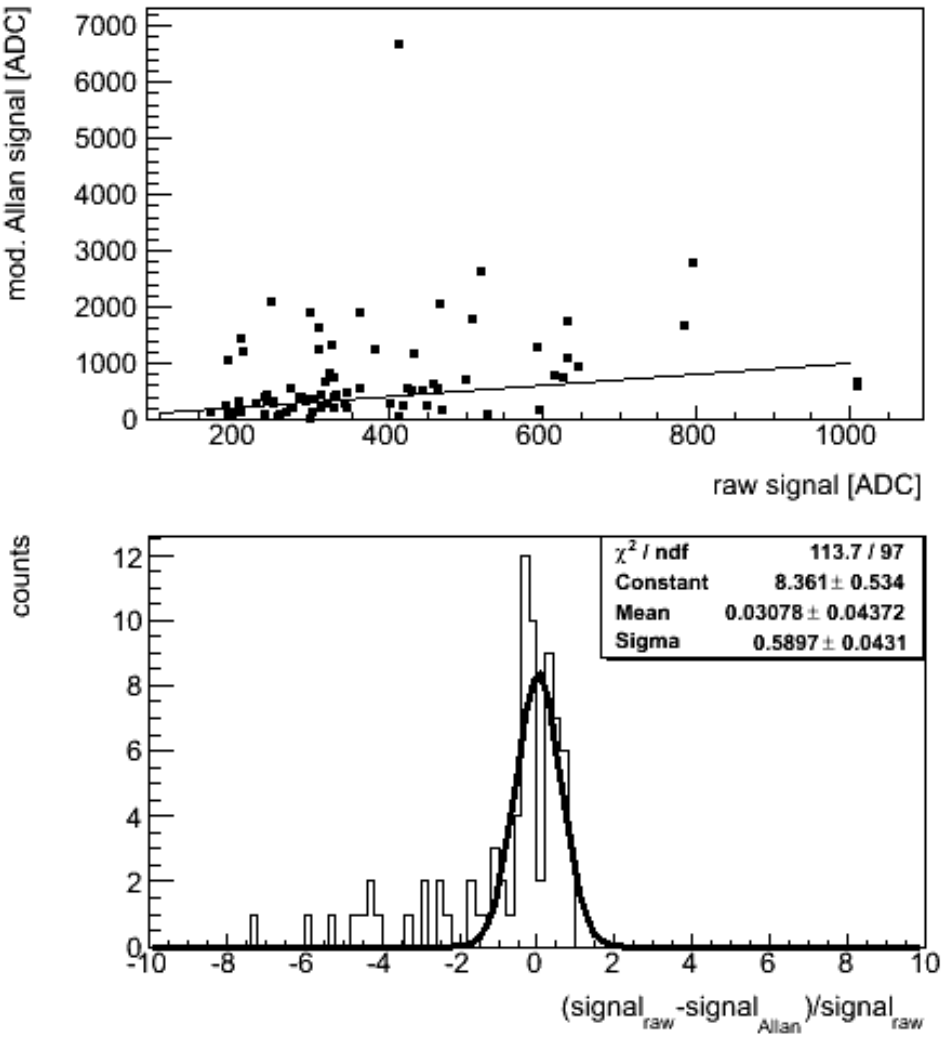


Figure 38: The Offline raw reconstructed signal ($\mathrm{Signal_m}$) is plotted versus the calculation according to formula 5. The line in the top plot is the one - to - one connection. For the bottom plot a Gaussian fit is performed.

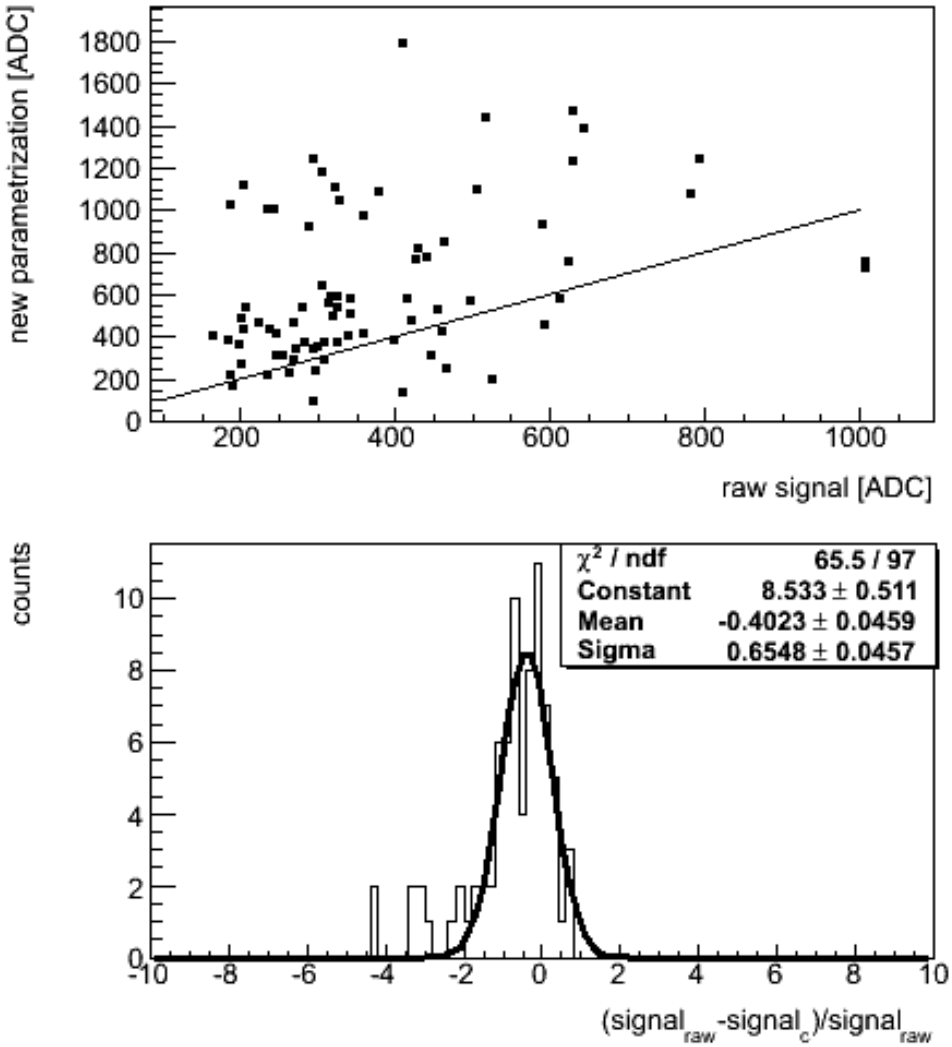


Figure 39: The Offline raw reconstructed signal ($\mathrm{Signal_m}$) is plotted versus the calculation according to formula 8. The line in the top plot is the one - to - one connection. For the bottom plot a Gaussian fit is performed.

8 Conclusion

The results of chapter 5 and 6 show, that it is possible to calculate an energy spectrum based on AERA data (figure 31). The spectrum obtained and the one from the infill area agree within the uncertainties. The large uncertainties arise from low statistics.
Important is to notice, that the formulae as introduced in chapter 4, for the signal strength of cosmic ray induced air showers, show different behavior for events that are measured with AERA and events that should have been measured according to the calculated signal strengths. A change in the formula of Allan, leads to a similar behavior for both groups of events.
Using this parameterization and the uptime for each station, the probability to measure an air shower is calculated. The average efficiency to measure air showers is in the order of 20%. The fact that the obtained energy spectrum is of the right order of magnitude shows, that the parameterization and the calculated efficiency are reasonable.
Unfortunately the energy spectrum is not SD independent, due to the calculation of the efficiency, using the missing SD events. Furthermore, the energy calculated with the SD data is used, due to the fact that there is no reliable radio based energy reconstruction. The spectrum covers only a small energy range, due to low statistics. While 39 events are measured with AERA, only 11 events have their core position located in the array and the calculated signal strength above the thresholds. Due to changes in the array, an extension of this data set is not possible. A new run, with more stations over a longer period could give a better spectrum. Furthermore, an increase of the efficiency would lead to more statistics.
The events of the second data set are of a higher quality than the ones of the first data set, where also events measured with only two radio stations are included. This shows the improvement of the array. Using less stations in a period as twice as long, more events are detected with a higher quality. Also events with energies, larger than used in the obtained spectrum are measured. Unfortunately, their core positions were located outside the array.
The comparison of the raw signal with the calculated signal shows, that the used parameterization is not an ideal description. Especially for stations that are close to the shower core. Therefore, it was tried to find a new parameterization in section 7. The result is a formula (formula 8) that is only slightly different from the one given by Allan. Only the distance dependence and the normalization are modified. Also in the new parameterization, the signals falls off exponentially as function of the distance. The problem of finding a new parameterization is, that the signal should

be zero, when the energy is zero. This leads to the linear fits through the origin as shown in figure 33 and 44 - 47. The second restriction is, that the signal should be zero far away from the shower core. This is the motivation for the exponential behavior.

Formula 8 is a good parameterization for the signals measured with AERA. For signals, corresponding to low energetic events, the expected signal is to large, leading to a high expected detection efficiency for events with energies, that are never measured with AERA.

Even though the differences in structure is small between the formulae 5 and 8, a better energy spectrum, in comparison with the spectrum from the infill area, is obtained (figure 37).

Section A shows the effective area for different energies and the parameterizations used. The most realistic parameterization is formula 5, represented by the red markers. For this parameterization, a threshold dependent effective area is calculated. Above an energy of 1 EeV all events should be measurable, independent of the arrival direction. For energies above $10^{17.5}$ eV, 85% of the air showers should be measurable.

A Energy dependent effective area - a comparison of the different formulae

The effective area is energy dependent. The effective area is calculated for different formulae and core positions. The first plot shows the effective area, when the core is in an area as defined in figure 12. Figure 41 shows the effective area, when the shower core is within 150 m of three stations, taking an array of 20 stations into account. For figure 42, only the first 12 stations are taking into account. For each energy, one million events are generated, the angles and the core position are chosen uniformly. An event is counted as measured, if there are three stations with a signal above the threshold.
For the formula by Allan, the magenta data points, the signal is calculated in μV and compared to a threshold of 66.66 μV. The same holds for the calculation with the formula of Huege and Falcke, represented by the green markers. With the modified formula of Allan (red) and the new parametrization (blue), the signal is calculated in *ADC* and compared to the minimum threshold of 140 ADC.

The modified formula of Allan, formula 5, describes the signal best. It gives a reasonable energy spectrum and a reasonable effective area for low energetic events. Therefore, the effective area is calculated for different energies and different thresholds and normalized to the maximum possible area. Core position within 150 m of the first 12 stations are used. The plot is shown in figure 43. As expected, the effective area is largest for small thresholds. For energies above 1 EeV, all events should have three signals above the threshold and should be measured. Even at lower energies, round $10^{17.5}$ eV, the effective area is more than 86% of the maximum effective area.

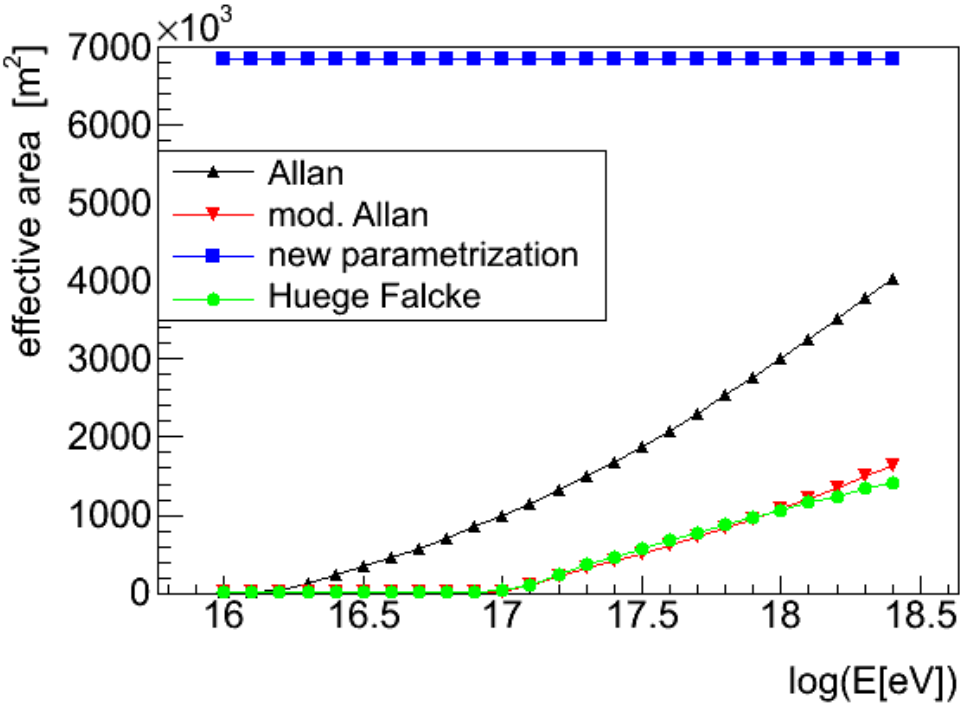


Figure 40: Effective area, depending on the energy for different parameterizations, when an area, as indicated in figure 12 is taken into account.

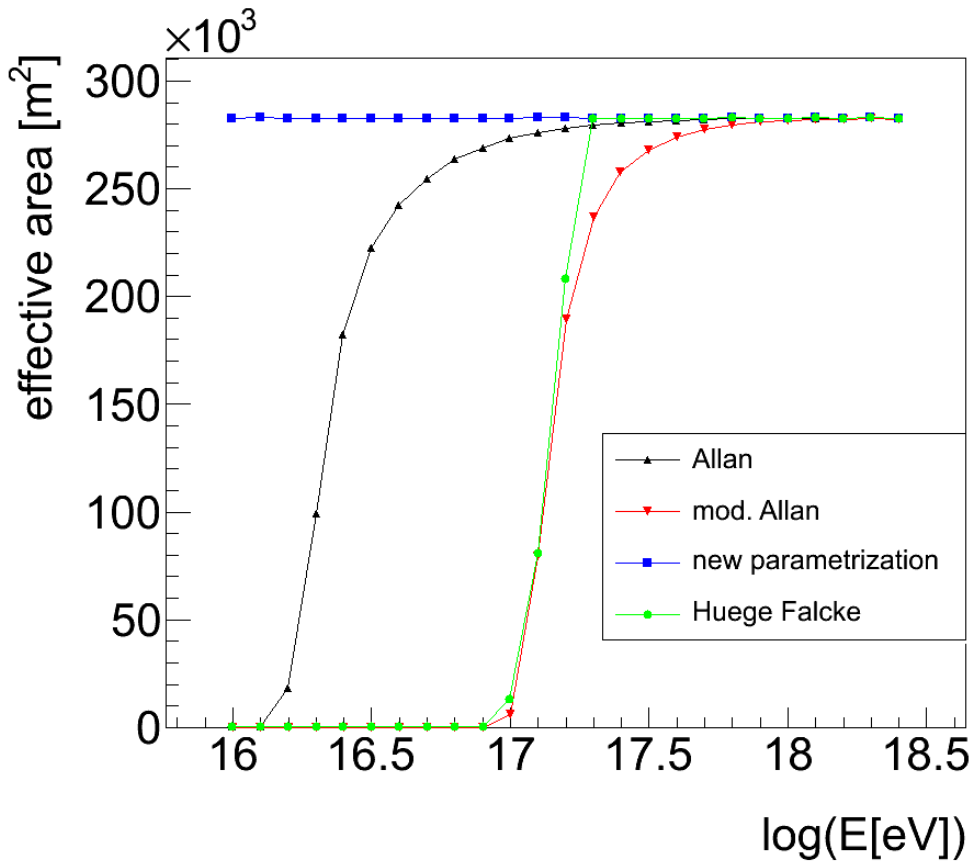


Figure 41: Effective area, depending on the energy for different parameterizations, when taking into account core positions within 150 m of three stations, when the first 20 stations are available.

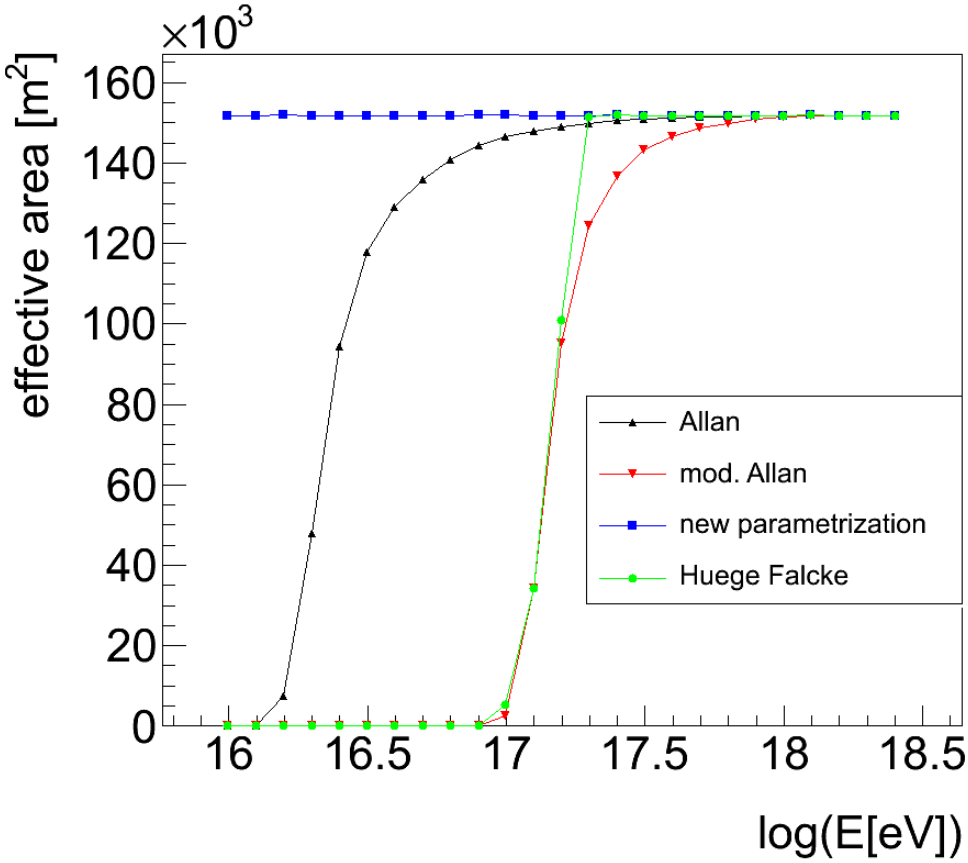


Figure 42: Effective area, depending on the energy for different parameterizations, when taking into account core positions within 150 m of three stations, when the first 12 stations are available.

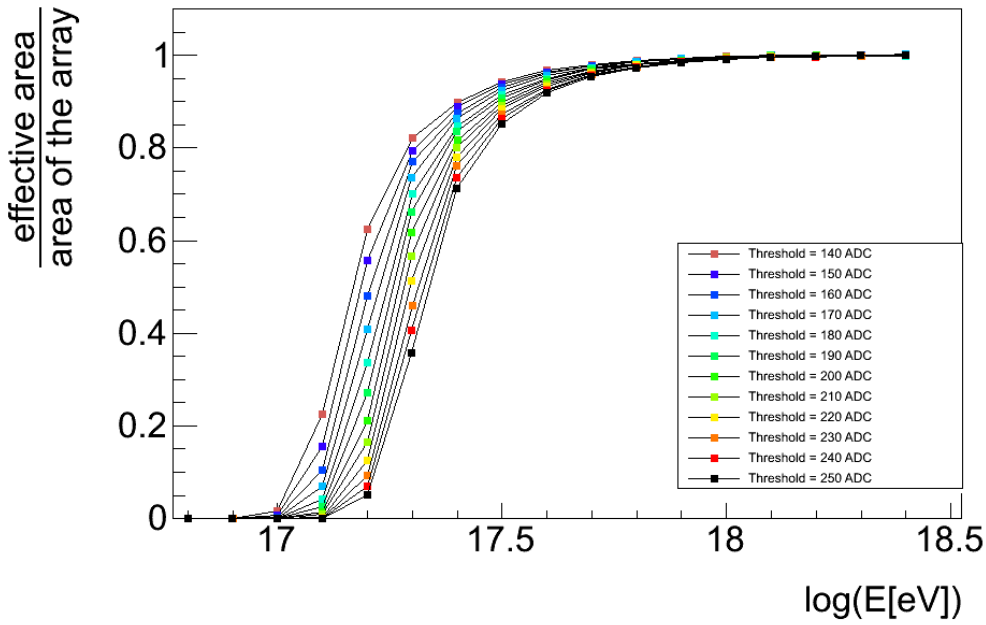


Figure 43: Effective area, depending on the energy and threshold (different graphs and colors) for a parameterizations with formula 5, when taking into account core positions within 150 m of three stations, when the first 12 stations are available.

B Additional plots for the data based parameterization

Here additional plots for chapter 7 are placed.

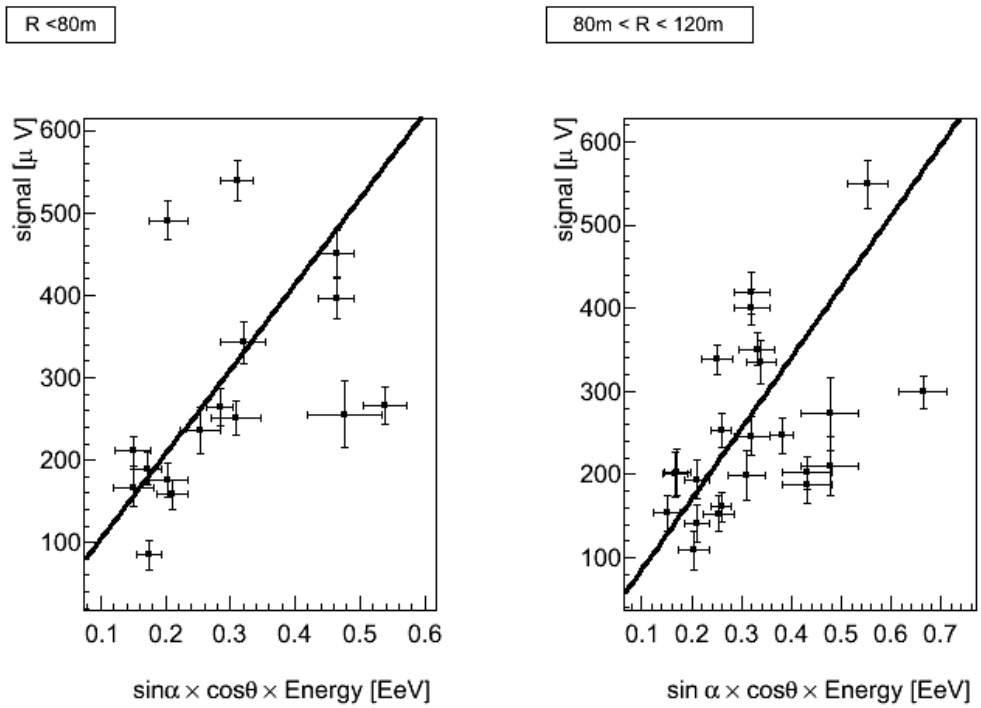


Figure 44: The maximum amplitude of the signal in east west direction (Offline) versus the cos θ $\times$ sin α $\times$ energy in EeV, for different distances.

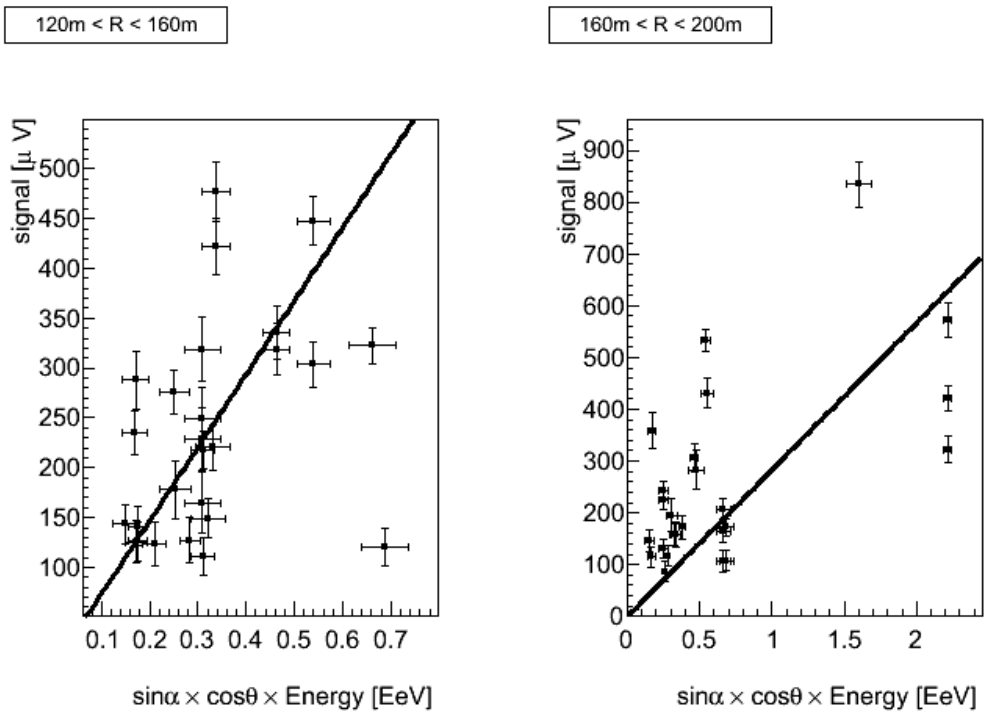


Figure 45: The maximum amplitude of the signal in east west direction (Offline) versus the cos θ $\times$ sin α $\times$ energy in EeV, for different distances.

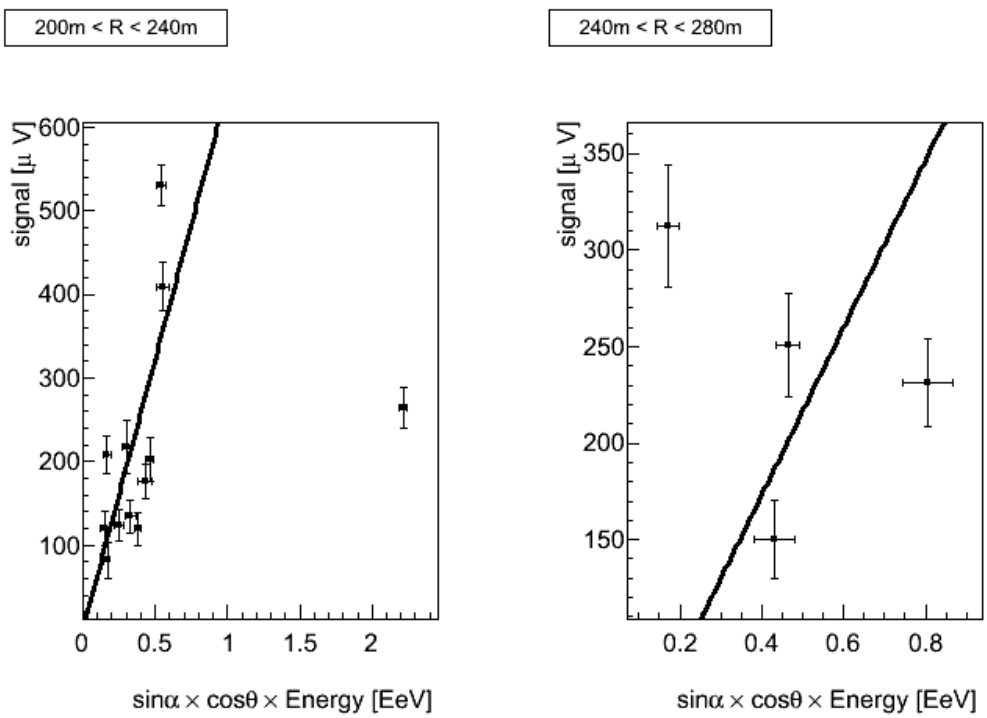


Figure 46: The maximum amplitude of the signal in east west direction (Offline) versus the cos θ $\times$ sin α $\times$ energy in EeV, for different distances.

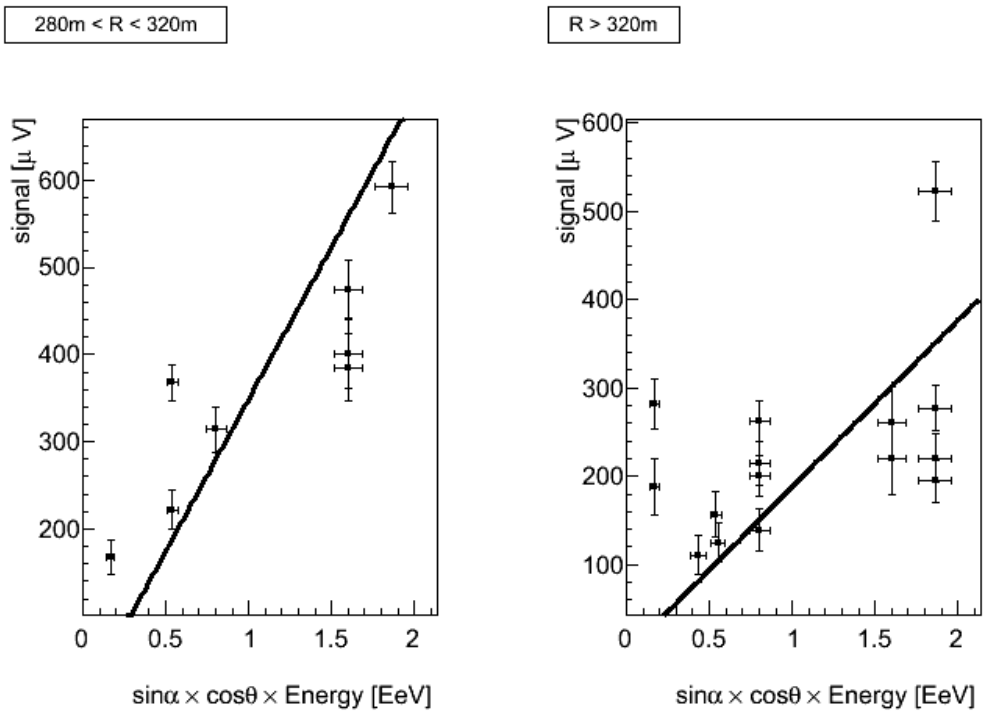


Figure 47: The maximum amplitude of the signal in east west direction (Offline) versus the cos θ $\times$ sin α $\times$ energy in EeV, for different distances.

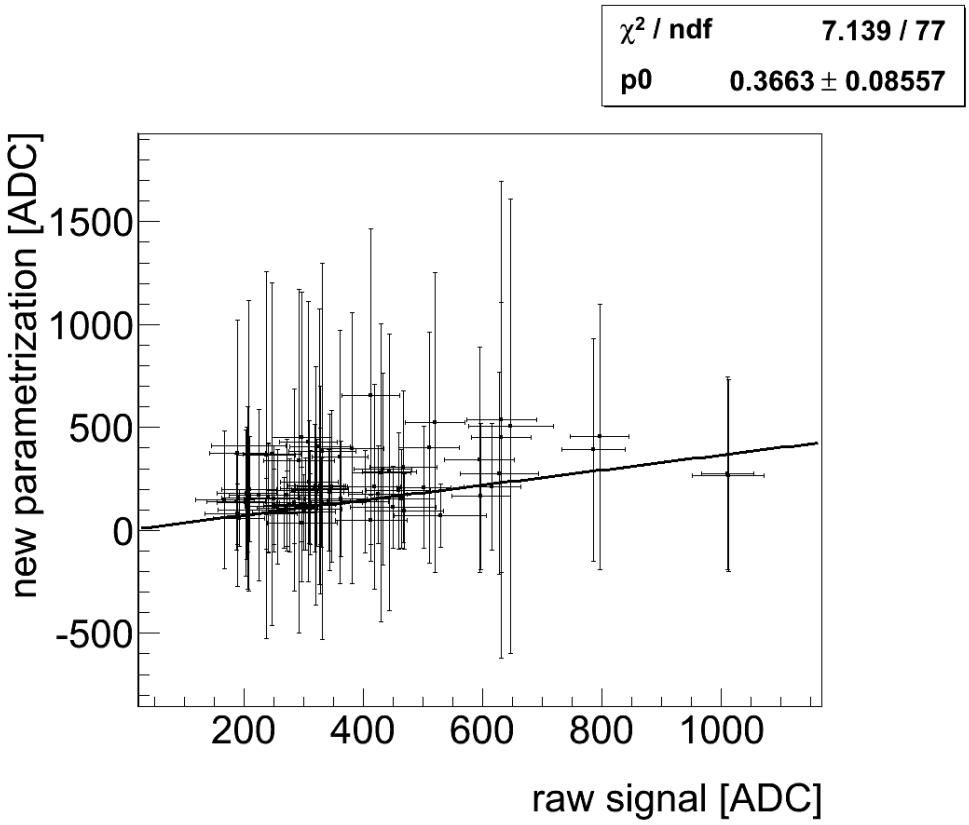


Figure 48: The dependence between the signal of the new parameterization and the maximum of the raw trace.

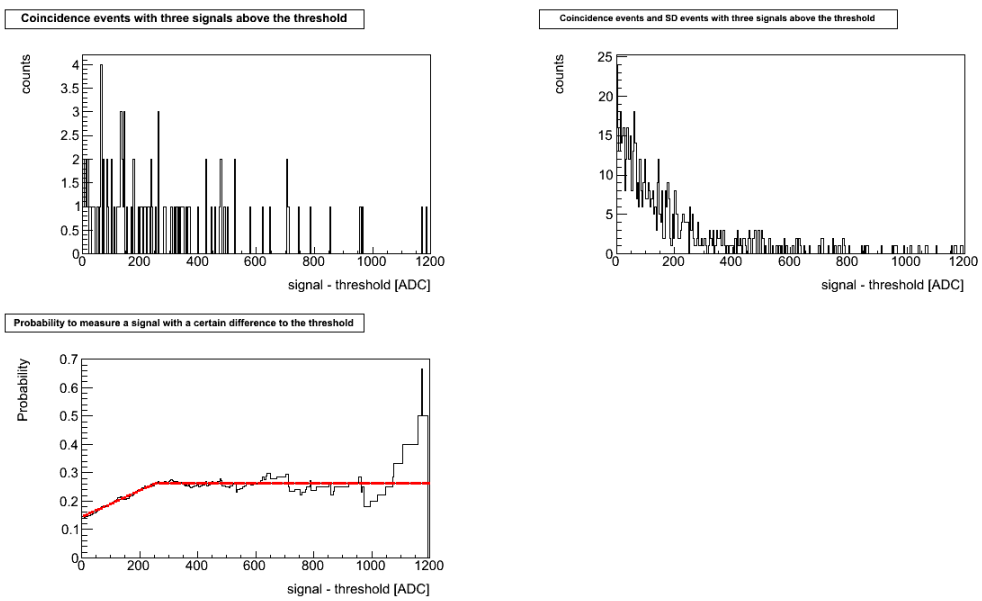


Figure 49: The calculation of the probability to measure an signal, depending on the difference between signal and threshold.

References

[1] T.Stanev, "High energy cosmic rays", Springer 2004, 2nd ed., page 10.

[2] $http://public.web.cern.ch/public/en/LHC/Facts-en.html$ accessed on May 2, 2012

[3] T.Wulf, "About the radiation of high penetration capacity contained in the atmosphere", Physikalische Zeitschrift, Vol. 5, 1910, pages 152-157.

[4] V.F.Hess, "Durchdringende Stahlung bei sieben Ballonfahrten", Physikalische Zeitschrift, Vol. 8, 1912

[5] R.A.Millikan and G.H.Cameron, "High frequency rays of cosmic origin", Physical Review, Vol. 28, No. 5, November 1926.

[6] P. Auger, "Extensive cosmic-ray showers", Reviews of Modern Physics, Vol. 11, July-October 1939.

[7] K. Greisen, "End to the Cosmic-Ray Spectrum?", Physical Review Letters, Vol. 16, No 17, April 1966.

[8] G.T. Zatsepin and V.A.Kuzmin, "Upper Limit of the Spectrum of Cosmic Rays", Journal of Experimental and Theoretical Physics Letters, Vol. 4, p.78, August 1966.

[9] J. Blümer et al., "Cosmic rays from the knee to the highest energies", Progress in Particle and Nuclear Physics 63, 2009.

[10] The Pierre Auger Collaboration, "Measurement of the energy spectrum of cosmic rays above $10^{18}eV$ using the Pierre Auger Observatory", Phys. Letters B 685, 2010.

[11] W. Heitler, "The Quantum Theory of Radiation", Courier Dover Publications, 1954.

[12] J. Matthews, "A Heitler model of extensive air showers", Astroparticle Physics, Vol. 22, Issues 5 - 6, January 2005.

[13] The Pierre Auger Collaboration, "The Fluorescence Detector of the Pierre Auger Observatory", Nucl.Instrum.Meth., 2009.

[14] $http://www.auger.org/observatory/pics/$ accessed on May 2, 2012.

[15] A. Etchegoyen, for the Pierre Auger Collaboration, "AMIGA, Auger Muons and Infill for the Ground Array", Proceedings of the 30st ICRC, Mérida, 2007.

[16] The Pierre Auger Collaboration, "The Surface Detector System of the Pierre Auger Observatory", Nucl. Instrum. Meth., 2008.

[17] The AERA group, "AERA proposal for the construction of the 20 km^2 Auger Engineering Radio Array at the Southern Auger Observatory", Internal PAO publication, Gap Note 2009-172.

[18] http://www.auger.org.ar/CDAS-Public/

[19] D.Veberič and M.Roth, "Offline Reference Manual- SD Reconstruction", Internal PAO publication, Gap Note 2005-035.

[20] K.van Norel, "Is the shape of the air shower front useful for researching the nature of cosmic rays?", Bachelor thesis, 2012, HEN-473, *http* : *//www.ru.nl/ehef/research/publications/*.

[21] J.V.Jelley, J.H.Fruin, N.A.Porter, T.C.Weekes, F.G.Smith and R.A.Porter, "Radio Pulses from Extensive Cosmic-Ray Air Showers", Nature, Vol. 205, page 327 - 328, 1965.

[22] H.R.Allan, "Radio emission from extensive air showers", Prog. Elem. Part. Cosmic Ray Physics, 10, page 171 - 296, 1971.

[23] T.Huege and H.Falcke, "Radio emission from cosmic ray air showers: Simulation results and parametrization", Astroparticle Physics, Vol. 24, page 116-136, 2005.

[24] C.Timmermans, "A Radio Detection test setup", Internal PAO publication, Gap Note 2006-009.

[25] M.Nagano and A.A.Watson, "Ultrahigh-energy cosmic rays", Reviews of Modern Physics, Vol. 72, No. 3, July 2000.

[26] D. Ravignani et al., "Observation of the spectrum with the AMIGA infill", Internal PAO publication, Gap Note 2011-010.