DESY Zeuthen
Summer Student Programme 2008

Final report

IceTop: Variations in decay times of waveforms and influence of snow

Sarah Becker

Georg-August Universität Göttingen
sarahbecker@uniqueobject.com

Abstract

The IceTop air shower detector is an array of Cherenkov tanks at the south pole. Here, two topics concerning the evaluation as well as the simulation of the detector are analysed: Firstly, the shape of the tank signals, the so-called waveforms, are investigated. Their decay time is found to depend significantly on the reflectivity of the tank liner. Secondly, the influence of snow on top of the tanks on the DOM signal is examined. Its influence is seen to be distinct, but it does not account for all variations of the DOM signals.

Contents

1 Introduction

1.1 Cosmic rays

Cosmic rays are high energetic particles from outer space hitting the earth's atmosphere. Their origin is still not fully understood. Here is a brief summary of the known facts, following [1]:

Shape of the energy spectrum. The energy spectrum, that is, the count rate per time, area, incident solid angle and energy interval vs. the energy of the incoming particle, shows the following features: In the first range, up to approximately $10^{14}\,\mathrm{eV}$, the spectrum follows a power law $\sim E^{-\gamma}$ with $\gamma \approx 2.7$. At the so-called knee the slope increases to $\gamma \approx 3.0$. At around $10^{19}\,\mathrm{eV}$, the so-called ankle, the slope decreases again. For even higher energies, the count statistics are so low that results about the specific shape are not yet conclusive. However, there is the prediction of the GZK(Greisen-Zatsepin-Kusmin)-cutoff of the spectrum at $10^{20}\,\mathrm{eV}$, because at these energies space is expected to become opaque for protons due to interactions with the microwave background radiation.

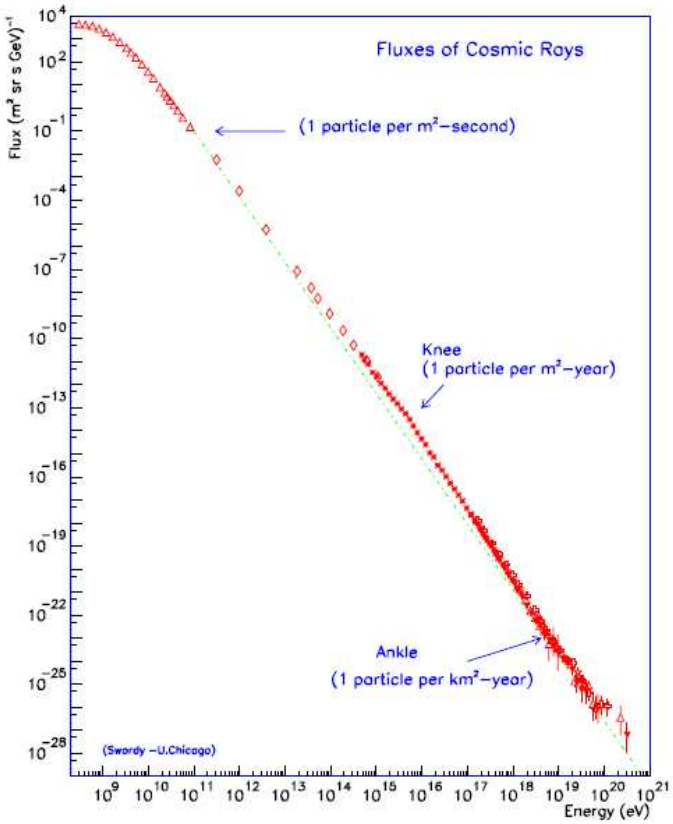

Figure 1: The energy spectrum of cosmic rays. Taken from [1].

Air showers. When measurements are taken on the surface of the earth, one has to keep in mind that due to the shielding effect of the earth's magnetic field, particles with lower energies (up to some 10 GeV) practically cannot reach the ground. Additionally, the primary particles will in general undergo interactions with the atmosphere, cause shower cascades and can thus only be measured indirectly. Such a shower originates when a nucleus collides with an air molecule. In the collision, typically pions are produced which decay into muons, photons and neutrinos. The photons in turn start an electromagnetic cascade, the soft component of the shower. The muons are more penetrating; some hit the earth without undergoing further reactions before. The pions themselves form the main part of the hadronic component of the shower.

Origin of the particles. Particles with energies of some GeV are believed to stem mainly from the sun. For higher energies, the particles can be subdivided into galactic particles originating from inside the galaxy and extragalatic ones. Particles of energies above 10^{18} eV are probably of extragalactic origin because the magnetic field of the galaxy is not able to confine them inside anymore.

Chemical composition. Cosmic rays consist to 98% of nuclei, the rest are basically electrons. Protons make up the vast majority of the nuclei; their abundance is about 87%. Twelve more percent are α-particles, and the remaining 1% are heavier nuclei occuring with about the same frequencies as in the solar system.

1.2 Set up of the experiment

Since there are still a lot of open questions concerning cosmic rays, such as the origin of ultra high energy particles, the reason for the knee or the precise acceleration processes, they are investigated in various experiments. One of them is IceCube with its surface detector IceTop at the south pole. When completed (expected for 2011), it will consist of 80 strings reaching up to 2450 m deep into the antarctic ice, each equipped with 60 digital optical modules (DOMs) distributed at equal distance along the lowest 1000 m. To each such string belongs an IceTop station on top with two tanks filled with ice and provided with another two DOMs each. The DOMs are essentially photo multipiers together with a processing unit in a transparent sphere. In each IceTop tank, one of them is run in sensitive mode

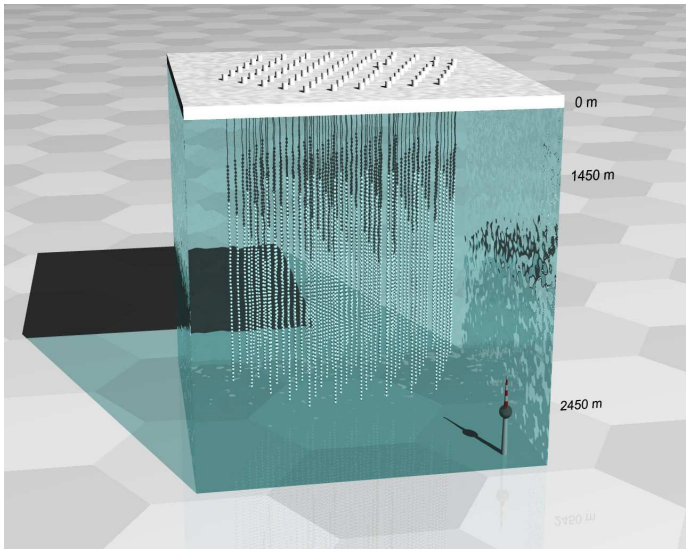

Figure 2: Setup of the IceCube and IceTop experiments. Taken from Adam Lucke's webpage, http://www.ifh.de/~alucke.

(high gain), the other in the less sensitive, but not as quickly saturated low gain mode.
The in-ice part serves primarily as a neutrino detector. The neutrinos can be detected via the Cherenkov light generated by particles hit by them in the ice. The low neutrino scattering cross-section is the main reason for putting the DOMs into the ice: A large detector volume – in this case 1 km^3 of ice – is needed to reach a sufficiently large interaction probability. Neutrinos are a particularly interesting part of the cosmic rays since they – in contrast to the charged component – do not get deflected by magnetic fields and thus preserve their direction information. Therefore, it is expected to be able to find sources of extragalactic cosmic rays in this way.
IceTop is looking at air showers caused by cosmic ray nuclei in the atmosphere. From the composition of the shower particles one tries to infer precise information about the chemical composition of cosmic rays. This is possible because heavier nuclei will produce a higher amount of muons.
The muons (and potentially also the neutrinos) produced in the pion decays of an air shower can also

trigger the in-ice part of the detector. But they are not the extragalactic neutrinos it is trying to detect. Here, the IceTop information can be used to distinguish primary neutrinos from secondary ones.

1.3 Evaluation and simulation

Most physical processes involved in the experiment – from the evolution of an air shower to the response of the detectors – are much too complex to be described with a few simple formulae. Nevertheless, for the evaluation of the data it is necessary to know what is expected; what the things searched for are supposed to look like. Therefore simulations are used. There are two basic types: Firstly, the ones in which all physical processes are described as accurately as possible, and secondly those in which measured signals or similar quantities are randomly generated according to known distributions. The latter is much faster, but it has the disadvantage that the distributions are needed as an input – which are only attainable with the first kind of simulations or from experiments.
In this sense, it is not only the evaluation of experimental data that need the simulations to have something to compare with, but also the simulations that need the experimental data in order to be as realistic as possible. During the weeks of my stay, I was evaluating some aspects of the data in order to feed it into simulations.

2 Decay time of waveform

2.1 Waveform

The waveform is the basic observable in a digital optical module (DOM). It is the measured DOM signal, i.e. the course of the output voltage in time. When a tank is struck by a particle of sufficient

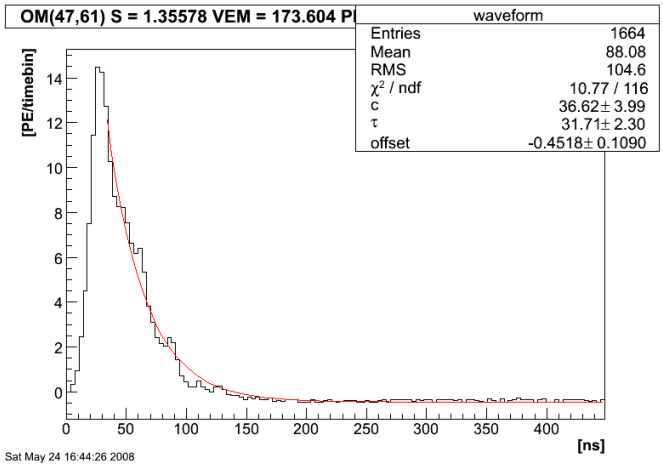

Figure 3: A typical waveform with fitted exponential. Taken from the IceCube wiki.

energy, Cherenkov radiation is produced. This light is then reflected at the tank walls and gradually absorbed in the ice. As both effects cause a decay of the number of photons proportional to the number of photons present, this leads to an overall exponential decay. A typical waveform (see fig. 3) starts thus with a steep rise which is then followed by an exponential decay.
For the further data processing, only the integrated charge, converted to so-called vertical muon equivalents (VEMs, see below), and the incidence time of the hit are used. Even though its decay constant is hence disregarded in the evaluation process, it is needed for a realistic simulation, since the reaction of the discriminator of the DOM (that is, its trigger) is sensitive to the height of the voltage and not to its integral, and thereby to the specific shape of the signal.

2.2 Decay times for different DOMs

One of the main characteristics of the waveform is its decay time, i.e. the decay constant of the exponential tail. It is obtained by fitting an exponential to the data.
Due to differences in the contact between the ice and the DOM, the quality of the ice in the tanks (cracks or bubbles etc.) and possibly other effects, the DOMs have different decay times for the waveform. In particular, the tanks that belong to the first four stations (21, 29, 30 and 39) deployed in 2005 are endowed with a liner of higher reflectivity (Tyvek) than the newer ones (zirconium). To investigate this effect quantitatively, the distributions of decay times for both kinds of tanks were plotted separately, see fig. 4, and the peak positions were determined. In these plots, a cut on the measured total amount of light was applied in order to exclude very small events with correspondingly large fluctuations, namely, the total number of photo electrons was required to be larger than 1 VEM. As can be seen in fig. 4, the most probable value (i.e. the middle of the bin at the peak) of the

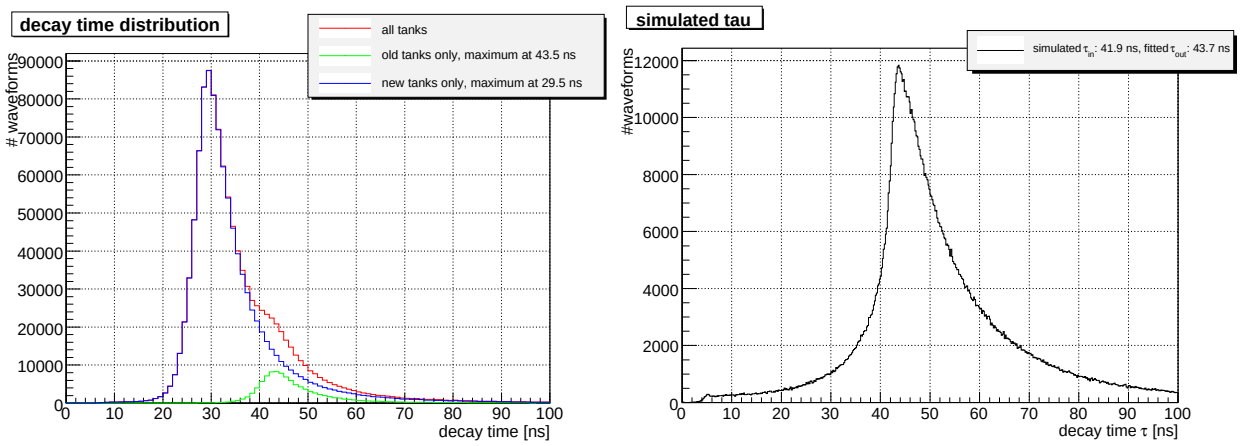

Figure 4: Left: Distribution of decay times for all tanks, 2005 tanks ("old tanks") with the Tyvek liner and the newer ones with the zirconium liner separately. It is obvious that the 2005 tanks have a significantly larger decay time due to the higher reflectivity of the Tyvek. Data from three days in July and September 2007. Right: Decay time distribution in a simulation. The overall shape of the distribution is in reasonable agreement with the experimental data.

experimental decay time for the 2005 tanks is 43.5 ns, while for the newer ones it is 29.5 ns.
As a result of this finding, the simulation was adjusted. The behaviour of the Cherenkov photons in the tank is modelled by a software called I3ArrayShowerTracer. One of its input parameters is the most probable decay time τ_{in}. This is the decay constant of the distribution of the photo electrons arriving at the PMT. Its signal is further processed by programs simulating the reaction of the photo multiplier, the processing units in the DOM and the trigger. Then, they are passed on to the programs which also process the experimental data. One of their fit parameters is the most probable decay time of the DOM's voltage signal τ_{out}. Due to the influence of the electronics in the DOM, it is in not equal to τ_{in}. For a realistic simulation, it is therefore necessary to find τ_{in} such that τ_{out} is equal to the measured value in the experiment. As the new tanks clearly dominate the statistics (simply because there are more of them installed by now), it had been assumed that all tanks have their most probable decay time and accordingly in the simulation, τ_{in} was set to a value such that τ_{out} for the new tanks was obtained.
Now the input time τ_{in} had to be determined such that the output τ_{out} time would be equal to the decay time as measured for the old tanks $\tau_{\text{old tanks}}$. Therefore, simulations were run with three different τ_{in} in the vincinity of the desired $\tau_{\text{old tanks}}$ and a linear regression was performed which is shown in fig. 5. Our result is that τ_{in} has to be set to

$$\tau_{\text{in, old tanks}} = 41.9\,\text{ns}$$

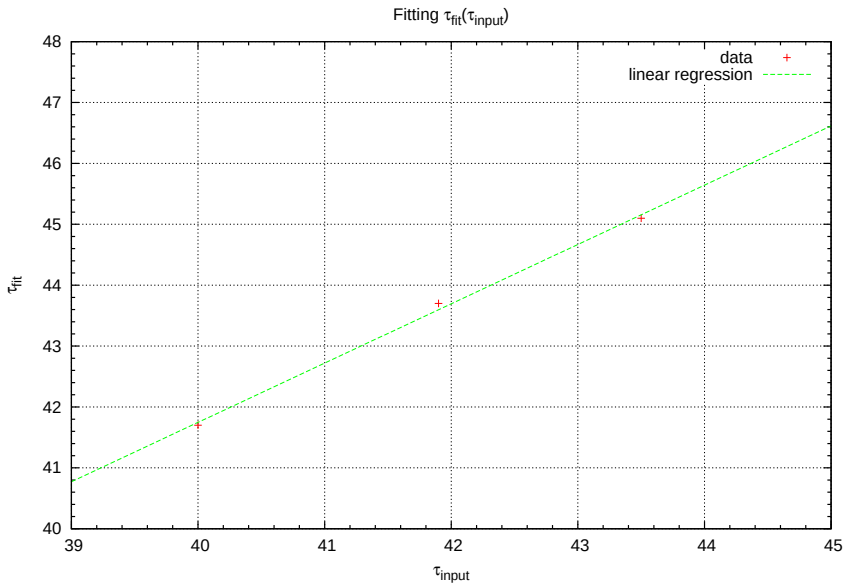

Figure 5: Determining the necessary τ_{in} in order to obtain $\tau_{\text{out}} = \tau_{\text{old tanks}}$ via linear regression.

in order to get $\tau_{\text{out}} = \tau_{\text{old tanks}} = 43.5\,\text{ns}$, which is by now already implemented in the simulation. To improve the simulation further, one could determine τ_{in} as a function of the desired τ_{out} over a range from ca. 20 ns to 50 ns. With this information, it would be possible to implement different decay times for each DOM individually. In this way, one could account for the smaller, but still considerable differences in the decay times of the DOMs among the "old" resp. "new" ones (see fig. 6).

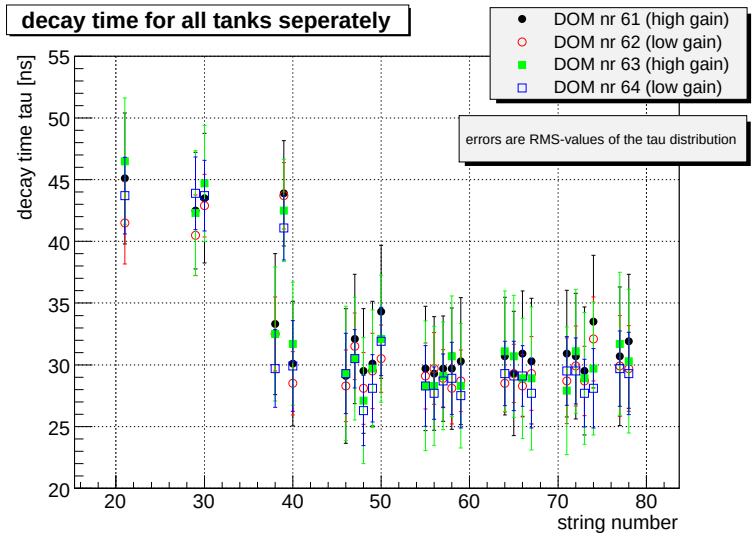

Figure 6: Decay times for all DOMs separately. As the differences between individual DOMs among the Tyvek liner (station number 21, 29, 30, 39) resp. zirconium liner tanks (all others) are still distinct (ranging over 6 ns for the Tyvek liner tanks and over 8 ns for those with zirconium liner), one could further improve the simulations by implementing different decay times for each DOM. Data from three days in July and September 2007.

3 Influence of snow

As a second task, I investigated the influence of the snow on top of the tanks. It is an intuitively very obvious reason for signal deviations from expectation. Since its height ranges from some cm (for instance on the tanks at station 38) to 1.7 m (on tank A at station 57), one expects the effects to be quite large, especially on the low energetic electromagnetic component. When examining it closer, it turns out, however, that the relations are not as plain and that there must be other effects accounting for deviations in the measured signals.

3.1 Deviations in the VEM calibration

3.1.1 Calibration procedure

To allow for comparison between different DOMs, the number of photo electrons generated in a tank is converted to vertical muon equivalents (VEMs). One VEM is the average amount of photo electrons created in the DOM when its tank is vertically traversed by a muon. Using this unit has the advantage that it is (approximately) independent of the special features of a certain tank or DOM, such as small deviations in the settings of the photomultiplier or the quality of the ice in the tank. To obtain the conversion factor, a calibration has to be performed regularly, which is done about once every ten days. It consists basically of recording a photo electron spectrum of every DOM. A typical such spectrum can be seen in fig. 7. For the calibration run, all coincidence triggers are deactivated because one is

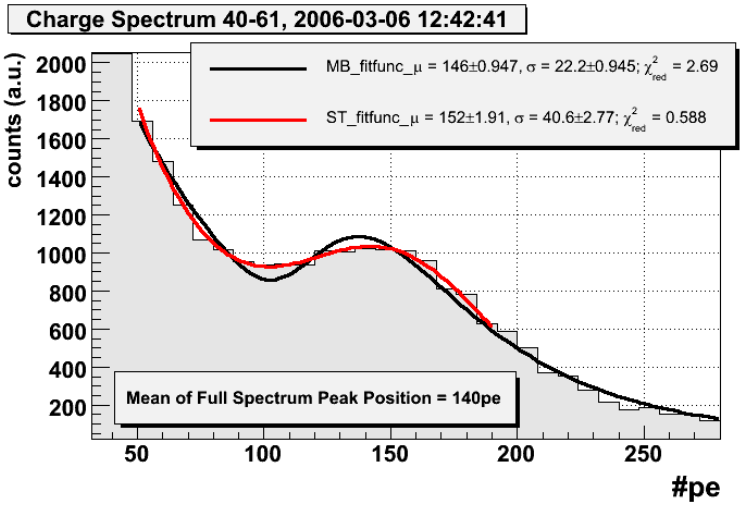

Figure 7: An energy spectrum of a DOM from a calibration run. To obtain the VEM peak value, two functions are fitted to the histogram and the maxima of the fit functions are determined. See text for more detailed explanations of the fit procedure. Taken from the IceCube wiki.

only interested in the reaction of single DOMs to single particles and not looking for shower fronts. Additionally, the high voltage at the low gain DOMs is increased in order to make single muons visible for them (otherwise, one muon would not produce enough photo electrons to give a distinct signal in a low gain DOM). To limit the amount of data (which would otherwise increase because the coincidence is switched off), the discriminator level of each DOM is raised [2].
The shape of the distribution of the number of photo electrons (which is in good approximation proportional to the energy deposit in the tank [3]) is dominated by an approximately exponentially decaying electromagnetic part and a peak caused by muons. They create a peak because their energy deposition $\mathrm{d}E/\mathrm{d}x$ in the ice is almost constant over a large energy range. This means that the amount of energy deposited (and thus photo electrons created) by a muon is essentially proportional to its path length in the tank [3].

In order to determine the VEM value, the following two functions are fit to this spectrum:
MB fit function:

$$f_{\mathrm{MB}}(x) = \frac{p_0}{p_2\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-p_1}{p_2} + \exp\left\{\frac{x-p_1}{p_2}\right\}\right)\right] + \exp\left(p_3 + p_4 x\right) + \exp\left(p_5 + p_6 x\right) \quad (1)$$

ST fit function:

$$f_{\mathrm{ST}}(x) = p_0 \exp\left[-\frac{1}{2}\left(\frac{x-p_1}{p_2}\right)^2\right] + \exp\left(p_3 + p_4 x\right) \quad (2)$$

The peak term in the MB fit function is the Moyal parametrisation of the Landau distribution [4] which describes the fluctuations of the energy deposit of charged particles in matter. The ST fit function assumes the peak to be Gaussian.
The VEM value is now obtained like this: The peak positions of both fit functions are numerically calculated, the average of both values (with equal weight) is taken and finally multiplied by 0.95 [2]. The last step is done because one wants to obtain the *vertical* muon equivalent, and simulations showed that cuts on small zenith angles lead to a shift to the left of the peak.
This shift is easily understandable because the path in the tank is longer for slanting muons. If you measure the spectrum of each DOM seperately, however, there is no way to know the zenith angle of an incoming particle and thus all muons contribute, causing the muon peak to lie slightly above the VEM value. Another reason for this shift to higher light production is the fact that the muon can be accompanied by electromagnetic particles which also contribute to the number of generated photo electrons.

3.1.2 Influence of snow on electromagnetic component

Since it was assumed that the snow would damp the electromagnetic part of the shower while the muons pass through it unaffected, the ratio of the integral over the electromagnetic component of the fit function to the integral over the peak component was calculated for both fit functions. This is an approximation, since, as described in the above paragraph, the muon peak always includes an electromagnetic component as well, which was in this choice of normalisation assumed to be negligible. For both functions, the integrals ranged over the width of the peak, from the peak position minus p_2 to the peak position plus p_2 (cmp. eq. (1) and (2)). The error of these quantities was calculated via propagation of uncertainty from the errors of the fit parameters. The correlation between the normalised electromagnetic background and the depth of the snow on the tanks was found to be less pronounced than expected, ranging between -0.49 and -0.14. To test this result, we tried to find other measures for the electromagnetic background such as altering the integration ranges or just dividing the number of events at a certain pe value by the height of the muon peak, but could not find a more distinct correlation to the snow height for these quantities either. Additionally, it is questionable whether it is justified to use them since their choice is quite arbitrary.

3.2 Deviations of total signal intensity from lateral fit expectation

In another approach, the deviation of the measured signals per DOM in an air shower was compared to the expectation from a fit of the lateral distribution of the shower energy.

3.2.1 Lateral Fit

An air shower measured on the ground will be laterally extended around the straight continuation of the trajectory of the originally incoming particle, the shower core. From simulations the actual distribution of the VEM values was inferred and parametrized by a lateral distribution function. This function is now fitted to the data in order to determine the width of the shower and its core position. A detailed description of the fit procedure and related topics can be found in [5]. From the result of the

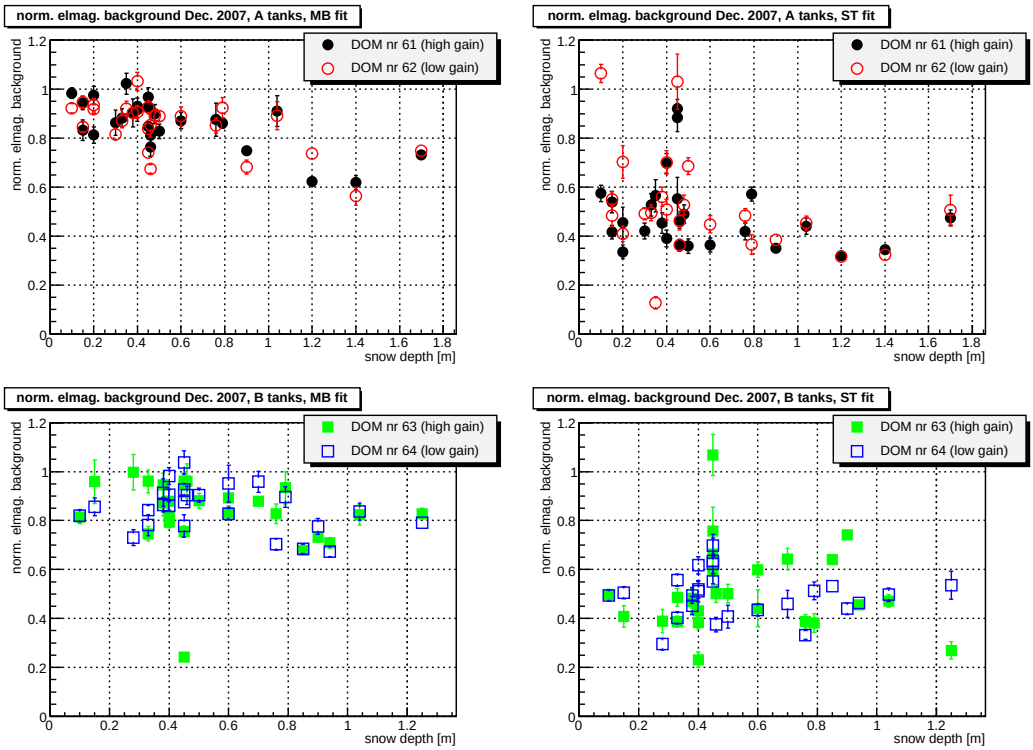

Figure 8: Normalised electromagnetic background as a function of the snow depth on the tanks. The correlations are -0.40 ± 0.05 for the high gain DOMs and -0.49 ± 0.06 for the low gain ones in the MB fit and -0.14 ± 0.01 for high gain and -0.16 ± 0.02 for low gain in the ST fit. The errors were calculated via propagation of error from the errors of the data points. There seems to exist an influence of the snow, but it is clearly not accounting for all variations in the electromagnetic background.

fit the expected number of photo electrons in a certain DOM can in turn be calculated and compared to the measured value.
To determine the direction of the shower, the arrival times of the pulses are used. There are two slightly different fit procedures: The first assumes the shower front to be even, which is rougher but also simpler and thus more stable. The other is more sophisticated as it takes the curvature of the shower front into account. Some of the data used here was processed with the curvature fit and some without. Since both fit methods yield quite similar results, my findings should remain unaffected by these changes.

3.2.2 Deviations caused by snow

For the deviations of the measured DOM signal from the expectation value obtained from the lateral fit a dependency on the thickness of the snow layer on top of the tank is also expected. In this case, a correlation of around -0.6 (depending on the particular fit function used and the measure for the deviation chosen) is found.
In detail: At first the linear deviation between expected and measured energy (in VEM) was calculated, that is, the difference ($\mathrm{VEM}_{\mathrm{meas}} - \mathrm{VEM}_{\mathrm{exp}}$). As this is a statistically fluctuating quantity, this was done for a lot of events (data from 2007) and the deviations were filled into a histogram. The mean and the most probable value of these were read out. If all tanks were the same and the fit was perfect, both values should be zero. Since both conditions are not completely fulfilled, one can see a slight asymmetry and an offset.

To have a normalised measure, secondly the normalised difference of the logarithms

$$\frac{\log_{10}(\mathrm{VEM_{meas}}) - \log_{10}(\mathrm{VEM_{exp}})}{\sigma_{\log_{10}}(\mathrm{VEM_{exp}})}$$

was regarded. Here, $\sigma_{\log_{10}}(\mathrm{VEM_{exp}})$ is the width of the distribution of the fluctuations which is

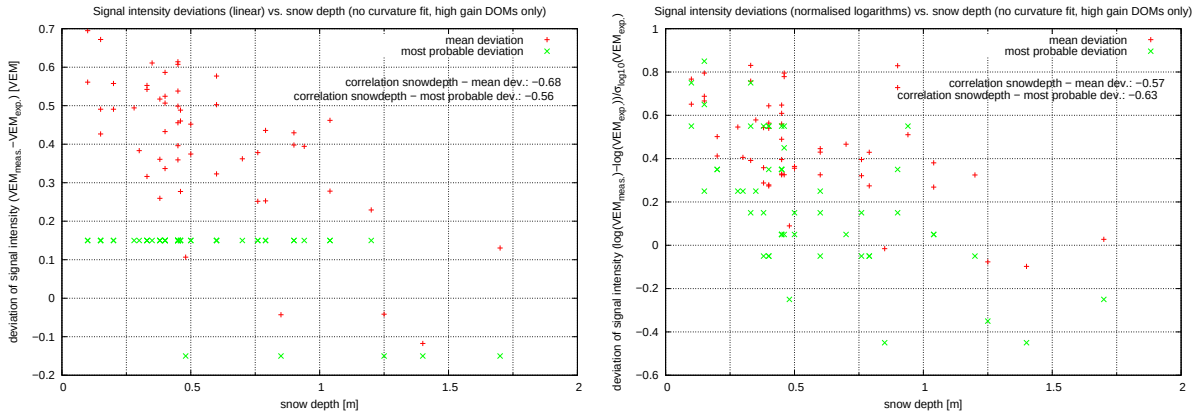

Figure 9: Deviations between measurement and expectation from fit vs. snow depth. Although the snow is obviously not the only parameter determining the deviations, there is a clear correlation between the two quantities.

approximately Gaussian. It depends on the mean value $\mathrm{VEM_{exp}}$ [6]. Mean and most probable value for these deviations were determined as well.
Linear regression yields the dependencies listed in table 1: Since the fluctuations of the distributions

method	deviation per snow depth	error (of linear regression)
linear & mean	-0.35 VEM/m	0.05 VEM/m
linear & most probable	-0.15 VEM/m	0.03 VEM/m
norm. logarithms & mean	-0.38 1/m	0.08 1/m
norm. logarithms & most probable	-0.58 1/m	0.10 1/m

Table 1: Dependence of sinal intensity deviations of snow

of the deviations is high for the low gain DOMs, they were neglected for the comparison to the snow height.
As a further improvement, one could try to think of a reasonable measure for the errors of the mean and most probable deviations and take these into account for the calculation of the correlations.

References

[1] H. Kolanoski. Einführung in die Astroteilchenphysik, 2007. http://www-zeuthen.desy.de/~kolanosk/astro0607/skript.html; online; accessed 10 September 2008.

[2] L. Demirörs. IceTop VEM calibration. IceCube wiki. http://wiki.icecube.wisc.edu/index.php/IceTop_VEM_calibration; online; accessed 10 September 2008.

[3] M. Beimforde. *Calibration of air shower signals in the IceTop detector using cosmic ray muons.* Diploma thesis, Humboldt-Universität zu Berlin, 2007.

[4] J. E. Moyal. Theory of ionization fluctuations. *Philosophical magazine letters*, 46:p, 1955.

[5] S. Klepser. *Reconstruction of extensive air showers and measurement of the cosmic ray energy spectrum in the range of 1-80 PeV at the south pole.* Dissertation, Humboldt-Universität zu Berlin, 2008.

[6] F. Kislat. *Study of charge and time fluctuations of signals in the IceTop detector.* Diploma thesis, Humboldt-Universität zu Berlin, 2007.

Acknowledgements

First of all, I would like to thank Fabian Kislat for his explanations and his steady advice. He has been a never-run-dry source of suggestions for improvements. Without him my work would not have been half as good.
I am also indebted to Hermann Kolanoski for giving me the opportunity to work in his group, and for his guidance and encouragement.
I want to thank various other people for various reasons: Adam Lucke for his help, the very nice pictures and his relaxed spirit. Bernhard Voigt for supplying the espresso machine (without which this work would also have been impossible). Konstancja Satalecka for the Sunday at the lake. Everybody who brought cake and other food to our office. And everybody else who made it a nice place to be.
Thanks a lot to the organisers of the summer student programme, especially Sabine Baer, Silvia Rückert, Thomas Naumann and Karl-Heinz Hiller who did all the administration, and Karl Jansen who organised and accompanied our trip to Hamburg.
And, last but not least, to all the other summer students!