Diplomarbeit

Muon Energy Reconstruction for ANTARES in the Low Energy Range using Artificial Neural Networks

Gerold von Lachemair

November 16, 2011

Erlangen Centre for Astroparticle Physics
Friedrich-Alexander-Universität Erlangen-Nürnberg
Department für Physik

Abstract

Neutrino astrophysics is a relatively new and interesting field of science. The observation of neutrinos in large Cherenkov neutrino telescopes like IceCube at the South Pole, which is already completed, the planned KM3NeT detector in the Mediterranean Sea and its predecessor ANTARES opens a vast field for potential discoveries in both astronomy and physics. The most prominent among them is the goal to find distant point sources of cosmic neutrinos from galactic as well as extragalactic sources, but also other associated fields of science are of interest, such as the observations of Gamma Ray Bursts, of Dark Matter decay reactions or exotic particles like Magnetic Monopoles. For many of these targets of research, a precise reconstruction of the energy of the muons emerging from a neutrino's interaction with ambient matter is necessary. There are various approaches to such a reconstruction. The one presented here is the use of Artificial Neural Networks (ANNs) for pattern recognition in the the neutrino event signature of the ANTARES neutrino telescope. Although ANNs are already used for energy reconstruction in ANTARES, this technique is extended in this work to the low energy range for secondary muon energies below 100 GeV. To this end various sets of input parameters for the ANN, which are especially adapted to the features of low energy muons, are tested.

Contents

Chapter 1

Neutrino Astronomy

1.1 Neutrinos as astrophysical messengers

Neutrinos are fermions with no electric charge and small, but non-zero mass. They are only subject to weak interactions and gravitation and were initially proposed by Pauli in 1930 in his theoretical explanation of the beta decay spectrum [2]. The first experimental discovery was performed by Clyde L. Cowan and Frederick Reines in 1956 with neutrinos produced by a nuclear reactor [2]. There are three types of neutrinos, the electron, muon and tau neutrino.

Due to their small cross section for interaction with atoms, compared to charged particles and photons, neutrinos can travel through the universe nearly without scattering by matter or being influenced by galactic magnetic fields. In contrast to that, charged cosmic ray particles except of them with the largest energies are deflected by magnetic fields so that their incident direction gives no information about their point of origin. γ-rays are absorbed or scattered by matter. This makes neutrinos an interesting messenger to detect and analyse distant high energetic astrophysical objects and processes.

1.2 Generation of High-energy Neutrino

Highly energetic cosmic neutrinos ($E_\nu > 1$ GeV) are produced by pion and kaon decay reactions:

$$\pi^\pm, K^\pm \rightarrow \mu^\pm + \bar{\nu}_\mu(\nu_\mu) \tag{1.1}$$

In astrophysical sources in non thermal equilibrium, charged particles like protons and electrons could be accelerated by reflection on relativistic shock fronts, which is called the Fermi mechanism [17]. If the particles are confined long enough by magnetic fields and pass the shock front multiple times, they can reach extremely high energies up to the ZeV range (Ultra High Energy, UHE) [23]. If mainly electrons are accelerated, they undergo electromagnetic processes such as inverse Compton scattering, bremsstrahlung or synchroton radiation [23]. This leads to the emission of highly energetic γ-radiation, but no neutrinos are produced. On the other hand, highly energetic protons undergo processes such as e.g. $p + \gamma \rightarrow \pi^+ + n$ which produce short-lived mesons like pions or kaons which then decay further in reactions like the decay described by equation1.1. Sources in which these hadronic processes take place are emitting high energetic neutrinos in addition to γ rays. Although of the spectra of γ rays give indications which processes take place in a source, neutrino telescopes could help to further confirm which process really takes place. High energetic charged particles can also leave the sources, but due to reasons mentioned above their origin cannot be determined without great uncertainties except for the highest energies $E_{CR} > 10^{19}$ eV [23].

1.3 Possible Candidate Neutrino Sources

In the following section a brief overview of the possible sources of high energy astrophysical neutrinos is presented.

Supernova Remnants

In the final stage of a massive ($> 8M_{\odot}$) star's evolution, its core collapses to a neutron star or a black hole. This causes the outer shells of the star to explode, a phenomenon that is called supernova. The core collapse generates a very large amount of neutrinos in the MeV energy range. These neutrinos can be measured and are so far the only type of astrophysical neutrinos originating from outside the solar system that are experimentally confirmed [3]. After a supernova, a shell-type supernova remnant is left behind. In these objects, particles can be accelerated by the Fermi mechanism, and the strong variable magnetic fields of a possible neutron star can accelerate already highly energetic particles even further [14]. The accelerated charged particles could then cause reactions such as mentioned above where neutrinos are generated. Recent measurements by FERMI LAT, however, show that leptonic processes are favoured [8].

Pulsar Wind Nebulae

Pulsar Wind Nebulae are like shell-type SNRs, with a pulsar inside that drives a relativistic wind that interacts with the surrounding SNR. This interaction is a possible source for high energy neutrinos [14].

Microquasars

Microquasars are binary systems which consist of a giant star and a compact object, e.g. a black hole. When the companion star expands over its Roche lobe, matter can flow over the Lagrange point and accrete on the compact object. A large amount of energy is released and jets are formed. In the jets diffuse shockwave acceleration occurs and high energy charged particles can react with photons or ambient matter and subsequently produce high energy neutrinos [14].

Active Galactic Nuclei

An active galactic nucleus contains a supermassive (10^6 - 10^9 solar masses) black hole. Matter accretes on that black hole forming highly energetic jets. They are one of the most likely extragalactic sources of high energetic cosmic rays and also a highly favoured candidate neutrino source [14].

Gamma Ray Bursts

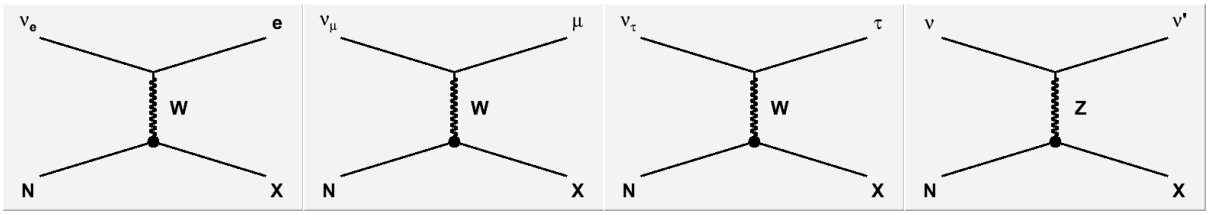
Gamma Ray Bursts are powerful events, although they typically last only for a couple or few tens of seconds. They are characterised by a high intensity eruption of gamma rays and an X-ray and visual afterglow. Their origin is still not clear, but it is supposed to be a merger of two neutron stars or a core collapse supernova of a super massive star [14].

Diffuse Neutrino Flux

Even if it is not possible to identify point sources, there could be a chance to see the diffuse neutrino background for large areas such as the galactic plane, galaxy clusters or large star-forming regions [23].

Dark Matter

Some theories suggest that Dark Matter, an unknown component of galactic halos, which is seen only through its gravitational interaction with stellar matter, consists of Weakly Interacting Massive Particles (WIMPs) which could annihilate into high energy neutrinos [21]. Such WIMPs could loose energy in dense matter such as the Sun's or the Earth's core due to scattering, and become gravitationally bound inside of these objects. Due to the annihilation of the WIMPs, the Sun and the Earth's core eventually could be detected as high energy neutrino sources. Different theories suggest that the mass of the lightest Dark Matter particle lies in the 10 GeV to TeV scale [21], and due to that and energy conservation their decay particles and subsequent neutrinos would have energies in the same scale. This is a motivation for developing an accurate energy reconstruction in the low energy scale.



(a) ν_e CC Interaction (b) ν_μ CC Interaction (c) ν_τ CC Interaction (d) ν NC Interaction

Figure 1.1: Neutrino interactions in matter where N denotes a nucleon and X an hadronic particle shower

1.4 Detection Methods for Neutrinos

1.4.1 Neutrino Interactions in Matter

If a high energy neutrino interacts with matter, two basic types of interaction can occur, charged current interaction (CC) and neutral current interaction (NC). In case of NC interaction a Z^0 boson is interchanged between the neutrino and a nucleon of the matter and the neutrino is not changed, but it looses energy, and a hadronic shower is produced. In case of CC interaction, a $W^\pm$ boson is interchanged and besides a hadronic shower the lepton which belongs to the same flavour as the neutrino is produced. The high-energy electrons cause an electromagnetic shower when they travel through matter and are stopped within a few meters, while muons and taus are able to travel greater distances before they finally decay. As a tau lepton decays after a very short time ($2.9 \cdot 10^{13} s$) [11], only muons can travel a sufficient distance to reconstruct their direction. Therefore usually muons are the particles used for astronomy. The average angle between the neutrino track and the muon track is:

$$\theta_{\nu\mu} \leq \frac{0.6^\circ}{\sqrt{E_\nu[TeV]}} \tag{1.2}$$

That gives reasonably small angular deviations for neutrino energies above 100 GeV, so that the muon track can be used to point back to the neutrino source [24][14].

1.4.2 Cherenkov Light Production

When a muon or other charged particle travels trough an insulator medium it polarizes the molecules of the medium. If the particle has a velocity greater than the speed of light in that medium, an overall dipole moment is created and light is emitted as a consequence. The light is coherent and is emitted in a cone with an angle θ_C given by

$$\cos\theta_C = \frac{1}{n\beta}, \qquad \beta = \frac{v_\mu}{c} \tag{1.3}$$

which yields 42° for sea water with $n = 1.33$[18].

1.4.3 Neutrino Telescopes

The basic idea for a neutrino telescope is equipping a large volume of a transparent medium with photo detectors, and then reconstruct the muon tracks by precise measurement of the arrival times of the Cherenkov photons of muons at the detectors. Due to the small flux of neutrino induced muons at the energies of interest and small interaction cross sections, large detection volumes are necessary [14]. Such large volumes cannot be artificially generated, so natural media such as ice (AMANDA, IceCube) or water (BAIKAL, ANTARES) will be instrumented. The geometry of the detector, the number of photo detectors, and the overall size of the experiment is chosen with respect to the expected neutrino yields (and budget).

Chapter 2

The ANTARES Detector

2.1 Layout of ANTARES Detector

The ANTARES Neutrino Telescope is located in the Mediterranean Sea about 25 kilometres south of the Southern coast of France. It consists of 12 strings with optical modules (OMs) , the so called "lines" and one instrumentation string. The lines are anchored on the sea floor at a depth of 2500 m. A buoy on top of each flexible string holds them upright. The strings are connected with an underwater junction box which provides power and data connection to the shore station via a commercial grade telecommunication cable. All lines have 25 storeys each, which have a vertical distance of 14.5 m. The first storey is placed 100 m above the sea ground. The distance between the strings is about 70 m. Each storey consists of three OMs and an electronics container, the Local Control Module (LCM). The OMs are arranged in a regular triangle around the LCM. Each OM contains a 10 inch Hamamatsu photomultiplier tube (PMT) which looks downward at an angle of 45°. The analogue PMT signals are digitialized by Analog-Ring-Samplers (ARSs). Each PMT has two of them to reduce dead times. The readout electronics as well as sensors for positioning are also included in the LCM [1].

2.2 Data Acquisition

Each time a photon hits the PMT, a specific charge is produced. If this charge exceeds a value equivalent to 0.3 photo electrons, the ARS is read out and the charge value and its exact arrival time is sent to shore. These values describe the so called L0 hits. There is no trigger algorithm running offshore except this 0.3 pe threshold. This is called the "All Data to Shore" concept. On shore, a computer farm runs the various trigger algorithms to filter the data. All L0 hits in a time window of 2.2 μs before and after the triggered hits are stored as an event. That is called event building. [16].

2.3 Trigger Algorithms

Trigger algorithms are used to extract the interesting physics events out of the raw data which is sent to the onshore computer farm to reduce the amount of data to be stored to a reasonable value, which is about 2 GB for a data taking period of 2 hours.

L1 Hit Selection

The L1 Hit Selection is applied before all other triggers. Its criterion is met when there are two or more L0 hits on one floor within a 20 ns window. Also, all L0 hits larger than 2.5 pe are L1 hits. All L1 hits are passed to the other triggers [16].

3N Trigger

The 3N Trigger is the standard physics trigger. It first looks for time correlated hits corresponding to a muon from any direction and then performs a scan of directions for a muon arriving from these directions.

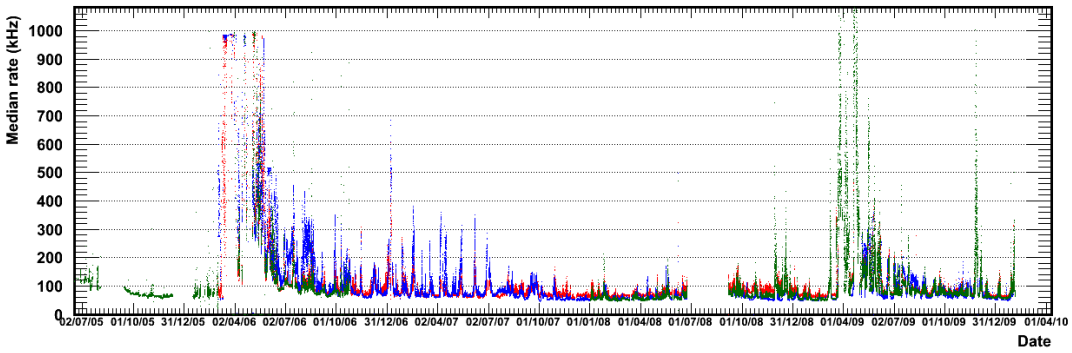


Figure 2.1: Plot of median rate per OM for the last years. The spikes are caused by variation of biolumimiescence. [15]

It has the ability to reduce random background by a large factor [16].

2T3 Trigger

The 2T3 Trigger criteria is met when there are 2 T3 clusters. A T3 cluster are 2 L1 hits on adjacent stories or storeys next to adjacent storeys are in coincidence. The counting is exclusive, so a L1 hit belonging to one T3 cluster can not belong to the other. Thus a minimum of 4 L1 hits is necessary [13].

2.4 Background

Like in any physics experiment, there is always background present in ANTARES. We have to distinguish between the muon and neutrino background from the atmosphere and the optical background generated in the detector volume.

2.4.1 Optical Background

There are two main processes contributing to the optical background, ^{40}K decay and bioluminescence. ^{40}K is an isotope of potassium which is present in sea water as potassium chloride. ^{40}K is radioactive and decays to ^{40}Ca in an β^- decay or ^{40}Ar by electron capture. It causes a decay rate of 13000 Bq per m^3 of sea water. The electrons produced by the β decay and electrons caused by secondary reactions of electron capture are producing Cherenkov light which is detected by the PMTs. This, together with the PMT's dark current, causes an average rate per OM of about 40 kHz. Bioluminescence is produced by both bacteria and larger lifeforms. The light from bacteria usually increases the rate per OM to 60 to 80 KHz, sometimes the rate lies even at 100 kHz or more. The larger animals create "bursts" of light, which last for several seconds and cause rates usually up to the MHz range. With all these effects considered, the average rate per OM is usually around 60 kHz, but sometimes rises much higher. The bioluminescence activity seems to be correlated with the sea current and water properties like temperature and salinity. The hits caused by HE particles have to be sorted out of the large sample of background hits by event building and hit selection algorithms [7][20].

2.4.2 Muon Background

When a high energetic charged particle hits the Earth's atmosphere, it reacts with nuclei of the air atoms, generating high energetic showers of secondary particles, including muons. The livetime of these muons is energy dependent due to time dilation and can be large enough to reach the ANTARES detector 2200 m under sea level. They are called downgoing muons. The atmospheric muons from the other side of the Earth are shielded by rock. That means that muons coming from the ground are produced within the material under the detector. By now, the only process known for that is neutrino interaction. Thus a

relatively simple upgoing/downgoing decision algorithm can sort out the atmospheric muons. However, the ratio of atmospheric muons to neutrino induced muons is 10^6 : 1, such that the distinguishing algorithm has to be efficient [24].

2.4.3 Atmospheric Neutrinos

The reactions of charged HE particles in the upper atmosphere are also producing neutrinos. These neutrinos also can react with a water or rock atom and generate a muon in the ANTARES detector, and produce an irreducible background in the end of the reconstruction chain. They can only be removed by identifying neutrino sources through large statistics, or by the fact that in the highest energy range the possibility of having an atmospheric neutrino decreases due to the different energy spectra of atmospheric neutrinos ($E^{-3.7}$) and cosmic neutrinos from Fermi acceleration (E^{-2}) [14]. Therefore a precise energy reconstruction in the high energy range is necessary. But for other fields of astroparticle physics like the search for Dark Matter also an energy reconstruction in the low energy range is necessary.

Chapter 3

Software and Reconstruction Tools

3.1 Monte-Carlo-Simulations

The physical processes in the ANTARES detector are simulated using Monte-Carlo-Simulation. The high energy neutrino interaction events are usually simulated with the GENHEN[9] software package. Atmospheric muons and shower events which generate multiple muons are simulated with Corsika and the subsequent muons are propagated to the detector with the MUSIC code. Atmospheric muons can also be simulated using the MUPAGE software, a parametric simulation that consumes less CPU power than Corsika's full Monte-Carlo-Simulation. If a particle hits a specific volume around the detector, a cylinder with a diameter of 476 m around the detector center and an height of 592 m, called 'the can', it is passed to the KM3[10] code, which simulates the muon propagation in the can, Cherenkov light propagation and photo-electron generation at the photo-cathodes. KM3 uses the MUSIC code to propagate the particle, and pre-existing photon tables to calculate the number of photons arriving at each PMT. The Cherenkov photons caused by other secondary particles than muons are calculated by the GEASIM code.

3.2 *SeaTray*

SeaTray is the software framework in which nearly all of the data handling and reconstruction tools in the ANTARES collaboration are executed. The tools it provides are called modules, most of them are written in C++. Measurement and simulation data files are divided into so called frames, where one frame contains one event in the most cases. The frames are passed from module to module through the chain of modules that is defined in the python steering file. For a new reconstruction or test one has to decide if the existing modules are sufficient, then only a steering file has to be written, or a new module has to be created.

3.3 Cluster Hit Selection

The Cluster Hit selection is a hit selection algorithm written for low energy track reconstruction in[21]. It loops over all hits and then adds for each hit all hits within a certain time and distance window to a set of hits. If the set exceeds a given hit number, the direction of the lines in the time-versus-height plane connecting the hits with the origin hit of the set is calculated. If the Root-Mean-Square (RMS) of the distribution of the directions exceeds a certain value, the hit which has largest deviation from the mean is removed from the set. This is repeated until the RMS is satisfactory or the number of hits in the set falls below the given minimal hit number. For the first case the original hit is selected, otherwise not. Out of this hit selection, hits over a given amplitude threshold are stored in another selection. Three of these selections are done, each of them with different parameters for time window, distance, hit number, maximum RMS and big hit threshold. If a hit is at least in two of the basic selections or in one and in a sample of the large hits, it is added to the final hit selection. According to [21], this hit selection

method has a high purity of 85 % and averagely selects 78 % of signal hits. It is especially designed for low energy reconstruction.

3.4 *Posidonia*

Posidonia is a track reconstruction algorithm specially designed for the reconstruction of low energy muons. It is based on a maximum-likelihood scan fit. There are also other track reconstruction algorithms which also perform maximum-likelihood fits in most cases, but use different approaches. The reconstruction used in this work slightly differs from the original *Posidonia* used in [25] as there was an other type of prefit used.

Pre-hit-selection The pre-selection of hits is not part of the *Posidonia* reconstruction, but it is necessary to have a relatively pure hit selection as input to *Posidonia*. In this work a hit selection performed with the above mentioned cluster hit selection merged with the triggered hits was used, as this is the usual hit selection for *Posidonia* which was used in [25].

Event classsification *Posidonia* distinguishes the events in single-string events, where only one line has triggered hits, and multi-string events, where two or more lines are hit. The reconstruction chain then differs between these event classes. In this work, the two cases were also separated in the further reconstruction.

Prefit A prefit is first applied to the multi-string events. A maximum-likelihood-fit is performed. A maximum-likelihood-fit varies the parameters of the track until the maximal value of the likelihood, which is a function of track parameters and arrival times of the photons at the PMTs, is reached. For the maximum-likelihood-fit a probability-density-function (PDF) specialised for multi-string events is used. The PDF is the function for the probability that a photon comes from a certain track and is used to calculate the likelihood. For the prefit, a scan over a set of different starting tracks is performed, due to the need for a starting track close enough to the real track to avoid the fit running into a local maximum. The fit with the largest maximum is then considered as the prefit track. For the prefit, the triggered hits are used as input.

Hit selection A hit selection is then used to find the hits that are compatible with the prefit from the preselected hits. All hits that have a length of the calculated Cherenkov photon path lower than 100 m and a time residual $t_{cherenkov} - t_{measured}$ in an interval of -10 and 40 ns are selected.

Final fit A fit with the same technique as the prefit is made, but instead of scanning over different tracks the prefit track is used and the input hits are the hits from the optimized hit selection.

Single-string string selector For the events classified as single-string events, a string selector is applied which selects the most hit string. A collection of causally connected hits is then made for this string.

Single-string Final Fit A fit is made in the same way like the multi-string prefit with scanning over a set of starting tracks. Here a specialised PDF for single-string events is used [25].

Chapter 4

Muon Energy Loss Phenomenology in the ANTARES Detector

For various types of physics analyses it is crucial to know the energy of the neutrino. However, there is a systematic difficulty, as for most events the distance from the muon interaction vertex to the detector is unknown, so one can only know a minimal energy of the neutrino even if the muon's energy in the detector is precisely measured. Nevertheless, the muon energy reconstruction is important. Therefore, from now on the energy reconstruction always refers to the energy of the muon.

4.1 Muon Energy losses in matter

When a muon travels through matter, it can loose energy through various processes. The two main components are ionization losses and radiative processes.

4.1.1 Ionisation

The ionisation losses are described by the Bethe-Bloch equation which denotes the average energy loss of a heavy ($m_{part} >> m_e$) particle .

$$-\frac{dE}{dX}[MeVg^{-1}cm^2] = \frac{\kappa z^2}{\beta^2}\frac{Z}{A}[\frac{1}{2}\ln\frac{2\gamma^2\beta^2c^2m_eT_{max}}{I^2} - \beta^2 - \frac{\delta(\beta)}{2} - \frac{C}{Z}] \tag{4.1}$$

where $\kappa = 4\pi e^4 m_e/c^2 N_A$, T_{max} the maximum kinetic energy transferable to an electron (= E for the relativistic case [11]), I the ionisation energy, Z and A the order and mass number of the absorber material, z the charge of the projectile, and $\delta(\beta)$ an energy dependent density correction. This formula is valid in the energy range $0.02 < \beta\gamma < 1000$. At larger energies, radiative corrections are important. From figure 4.2 it can be seen that for $\beta\gamma < 3$ the energy loss decreases with energy, and there is a broad minimum for $\beta\gamma = 3.5$. Particles in that minimum are called "minimally ionizing particles". After that minimum the energy loss increases slowly, until radiative effects become important [11].

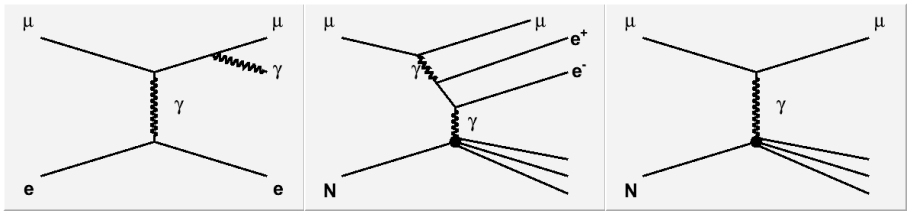
4.1.2 Radiative Processes

In the high energy range, larger than about 500 GeV for muons in water, radiative losses gain importance. At the so called "muon critical energy", $E_{\mu c}$, ionization losses and radiative losses are equal. The value of $E_{\mu c}$ is about 780 GeV in water. [24] The total energy loss is described by

$$-\frac{dE}{dX} = a(E) + b(E)E \tag{4.2}$$

where $a(E)$ denotes the ionization losses and $b(E)E$ the radiative losses. There are three radiative processes that contribute to that. Their Feynman diagrams are shown in figure 4.1.

Bremsstrahlung The muon interacts with the electromagnetic field of a nucleus or an electron and emits a photon.



(a) Bremsstrahlung (b) Pair production (c) Photonuclear interaction

Figure 4.1: Possible ways for high energy muons to loose energy

Pair Production The muon interacts with the electromagnetic field of a nucleus and produces an electron-positron pair.

Photonuclear Interaction The muon interacts with a nucleus through deep inelastic scattering and a hadronic shower is produced. The high energetic secondary electrons and photons also undergo radiative processes, which leads to the emission of more particles, so that electromagnetic cascades are formed. The hadronic particles from photonuclear interactions cause hadronic showers and subsequent electromagnetic subshowers on the same length scale. All particles, as long as their velocity is greater than that of light in water, will cause Cherenkov radiation. All of these processes have a stochastic nature, so that a reconstruction of the particle energy from the Cherenkov light cannot be done in a pure analytical way with maximum accuracy. [11][24]

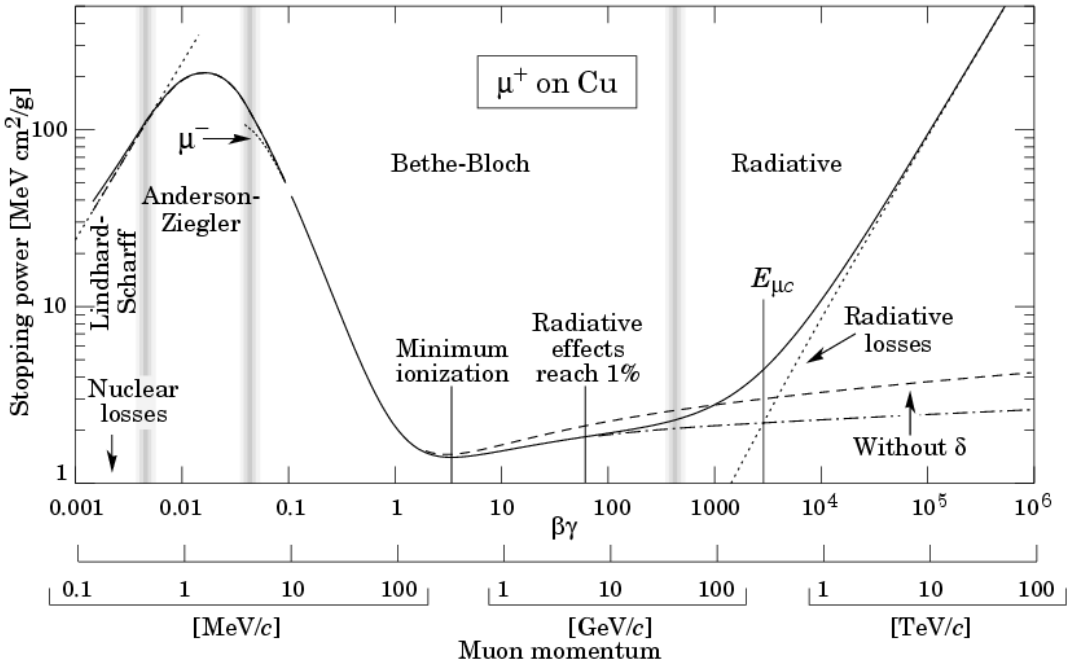


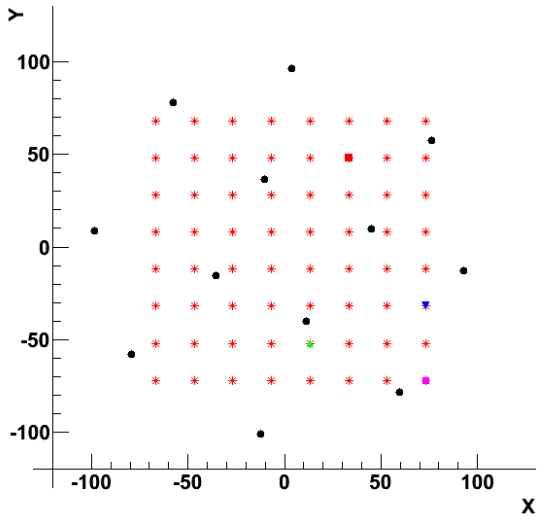
Figure 4.2: Energy losses of muon in a thin layer of Cu dependent on muon energy [5]

4.2 Detector Response in the Low Energy Range

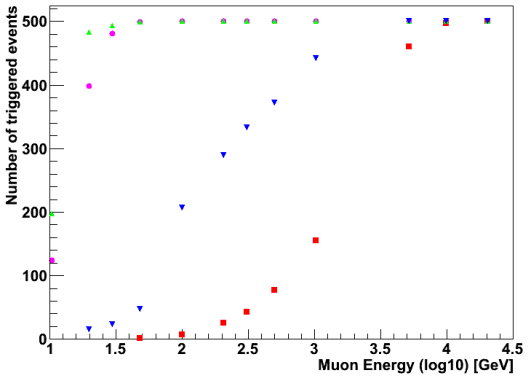
As can be seen from the sections above, we can classify the particle energy loss into two different ranges, one where ionization is dominant and one where radiative processes are dominating. In the latter case the amount of light the particle emits per unit distance is dependent on particle energy. In the ionisation regime, however, this is not the case, which means that two tracks with different energies are not distinguishable when only a section of track is detectable. Only the length of the whole track shows

energy dependence. That means that it is only possible to reconstruct a muon's energy with reasonably small uncertainties if the detector is large enough to cover the whole length of the track. As a muon in ionisation regime looses about 210 MeV per meter of track length[19], and the instrumented volume of the ANTARES detector covers a surface of roughly 290000 m^2 and a height of 348 m, muons with energies up to about 100 GeV can be contained in the ANTARES instrumented volume. If the neutrino's interaction vertex is outside of the detector, only the energy of the muon as it enters the instrumented volume can be reconstructed as the distance between vertex and detector is not known. This is also true for particles in the radiative energy regime.

To examine the detector response to muons of different energies, a special Monte Carlo simulation was performed with discrete muon energies and starting positions. The muon start positions were located on a horizontal plane at approximately the z-coordinate of the first detector floor and in a grid of 8x8 starting points with distances of 20 m between two neighbouring points. The center of the grid was at the detector center. In figure 4.3a the start positions are marked with red squares, while black squares denote the OM positions. The muons were propagated upwards with zenith angle of 180°. The simulation was repeated 500 times for each muon energy $E_i \in \{10, 20, 30, 50, 100, 200, 300, 500, 1000, 5000, 10000, 20000\} GeV$ and starting position. There was no optical background added and standard ANTARES PMT simulation [1] was performed and standard 3N and 2T3 Triggers were applied. Figure 4.3b shows the response of the trigger in relation to different track positions. The four graphs show how many of the 500 simulated events meet one of the trigger criteria. It immediately becomes clear that the trigger efficiency is dependent on the distance of the track position to the neighbouring lines.



(a) OM positions (black dots) and starting positions of simulated muons (red stars). Start positions of tracks shown in 4.3b are denoted with same colour here.



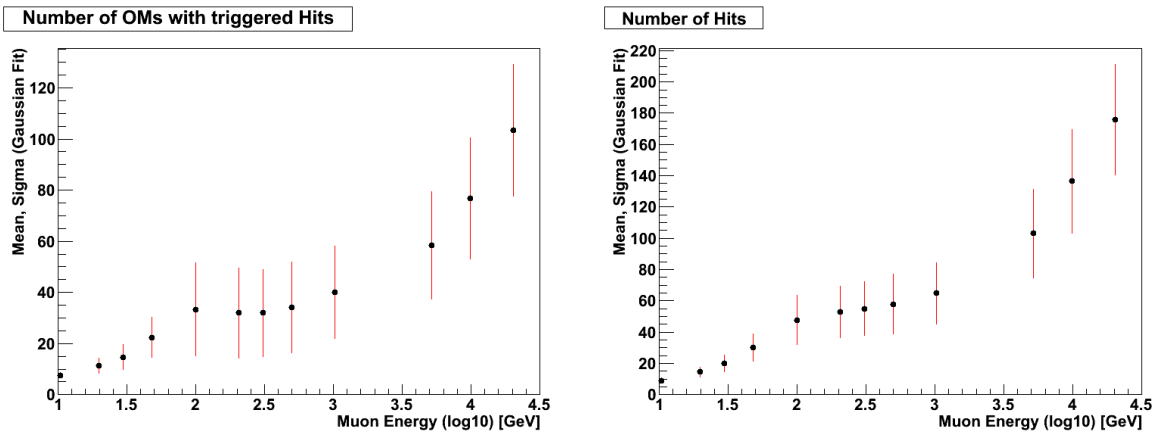
(b) Number of triggered events out of 500 simulated events for different track positions

Figure 4.3: Monte Carlo track positions and trigger

The first approach to an energy reconstruction is to use the total amount of emitted light which characterises the energy loss. Plots 4.4a and 4.4b show the energy dependence of two parameters (Number of OMs with triggered hits and total number of hits), which give an indication for the total amount of light emitted in the detector. The plots show the expected behaviour, at low energies up to 100 GeV the amount of light rises with energy, in the intermediate energy range it is nearly constant, and above 1 TeV it increases rapidly. The second rise corresponds to radiative effects taking place. The low energy behaviour is corresponding to a varying track length in the detector. Plot 4.5 shows the Monte Carlo track lengths inside the "can" (cylinder with height 592 m, diameter 476 m around detector center) from the simulation. Above an energy of about 100 GeV the track is long enough to pass through the whole

[1] PMT simulation parameters: integration time 40 ns, deadtime 250 ns, ARS threshold 0.3 photoelectrons

detector, and due to that the detectors response is similar for all tracks with this or higher energies until radiative processes take place.



(a) Number of OMs with triggered hits against Monte Carlo energy

(b) Number of hits against Monte Carlo energy

Figure 4.4: Plots of energy dependence of two parameters

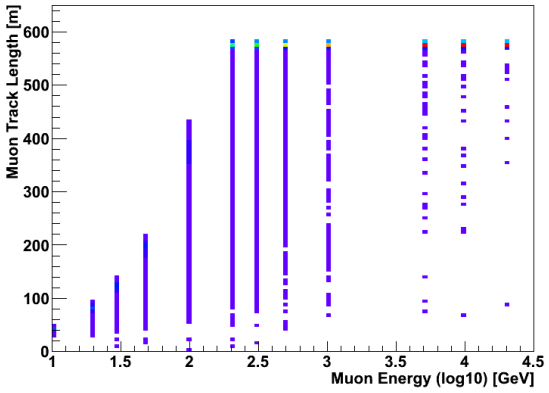
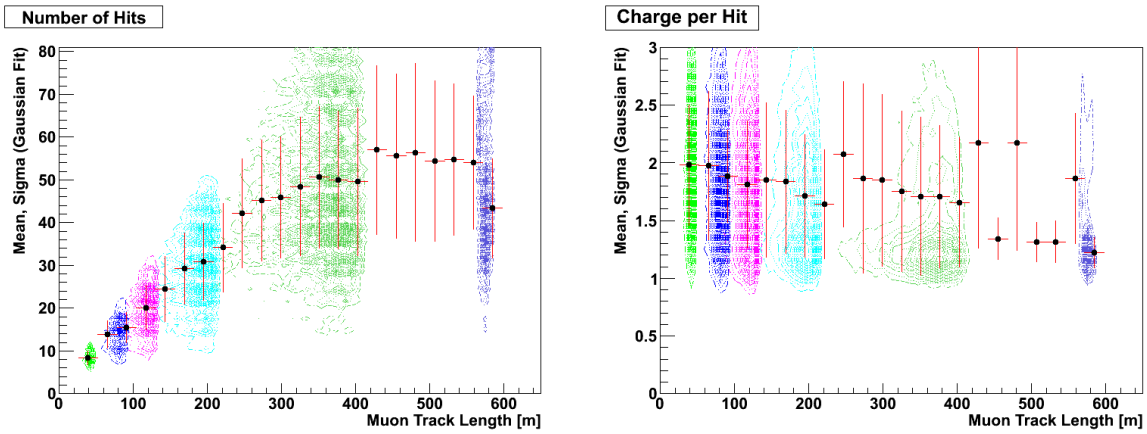


Figure 4.5: Monte Carlo track lenghts for different energies

To find a possible solution for reconstruction of energies below 100 GeV, one can try to use parameters that are dependent of the track length inside the detector. Figures 4.6a to 4.7b show four possible parameters. In Figure 4.6a it can be seen that the number of hits is increasing up to 100 GeV where the muon travels through the whole instrumented volume. In figure 4.6b, which shows the average charge per hit, a tendency to 1 photoelectron per hit is seen for higher energies. This is probably due to the fact that for low energies the events with tracks with greater distance to the next line meet the trigger criterion less likely and so tracks with small distance to the next line are overrepresented. Also hits $>$ 2.5 photoelectrons are contributing more to meet trigger criterion. In figure 4.7a the number of floors with triggered hits (hits in coincidence to a trigger criterion) is shown. A clear dependency on the track length can be seen, which is no surprise as the tracks are entering the instrumented volume perpendicular to the floors so the number of floors is strongly dependent on track length. It can also be seen that for 100 GeV and 200 GeV the distribution is nearly the same, as it is expected due to the muon crossing the whole instrumented volume at about 100 GeV. Figure 4.7b shows the number of strings with triggered hits. There is much less dependency on the track length here, due to the fact that most of the events are single string events as the tracks are vertical. Again here the trigger filters out events that have shorter tracks if they are far away from the next string. This explains the lack of two and three string events for

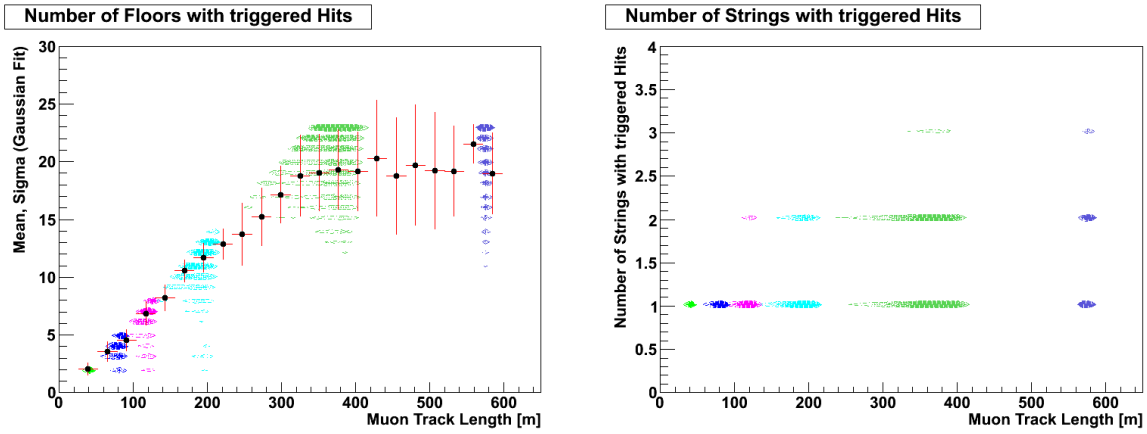
lowest energies.



(a) Number of hits in detector versus Monte Carlo track length

(b) Average charge per hit

Figure 4.6: Black dots denote the mean value for a Gaussian fit over the energy range up to 200 GeV. The red error bars denote the sigma of the fit. The coloured contour plots show the value distribution for a specific energy. Light green stands for 10 GeV, blue for 20 GeV, magenta for 30 GeV, cyan for 50 GeV, dark green for 100 GeV and violet for 200 GeV. Some tracks also have lengths between the displayed bunches



(a) Number of floors with triggered hits versus Monte Carlo track length

(b) Number of strings with triggered hits versus Monte Carlo track length

Figure 4.7: Black dots denote the mean value for a Gaussian fit over the energy range up to 200 GeV. The red error bars denote the sigma of the fit. The coloured contour plots show the value distribution for a specific energy. Light green stands for 10 GeV, blue for 20 GeV, magenta for 30 GeV, cyan for 50 GeV, dark green for 100 GeV and violet for 200 GeV. The second Plot has no Gaussian fit due to the small number of possible values makes the fit strange and possibly misleading

Of course in reality the track vertices are distributed over the whole medium and track zenith and azimuth angle is not restricted. In addition there is the probability that the detector gets light from the vertex shower of the neutrino interaction if the vertex is inside or at the edge of the instrumented volume. Also there is an optical background. Due to that the extraction of the energy or the track length out of the accessible parameters is a multidimensional and relatively complex problem. In such a case the application of a neural network may help to extract the energy dependent features out of the raw data.

Chapter 5

Artificial Neural Networks

5.1 Biological Neural Networks

The understanding of the brain as a whole seems to be relatively weak, but scientists have determined the basic concepts of it. It is clear that a brain consists of a network of interconnected cells, called neurons. Although there are some complex circumstances considering the neuron, its basic functionality is to sum up the information from other neurons and act as a decision element. As it can be seen in figure 5.1, a neuron consists of its cell body, some dendrites and an axon. The neuron receives the information from the axons of other neurons which end on its cell body or dendrites, via electrochemical interactions. The gaps between the axons and the dendrites or cell surface in which the interaction takes place are called synapses. Synapses can have varying strengths. The excitation of all synapses is summed up and if it exceeds a threshold at a certain time and and place an action potential is created and travels through the axon to the synapses of the connected neurons.

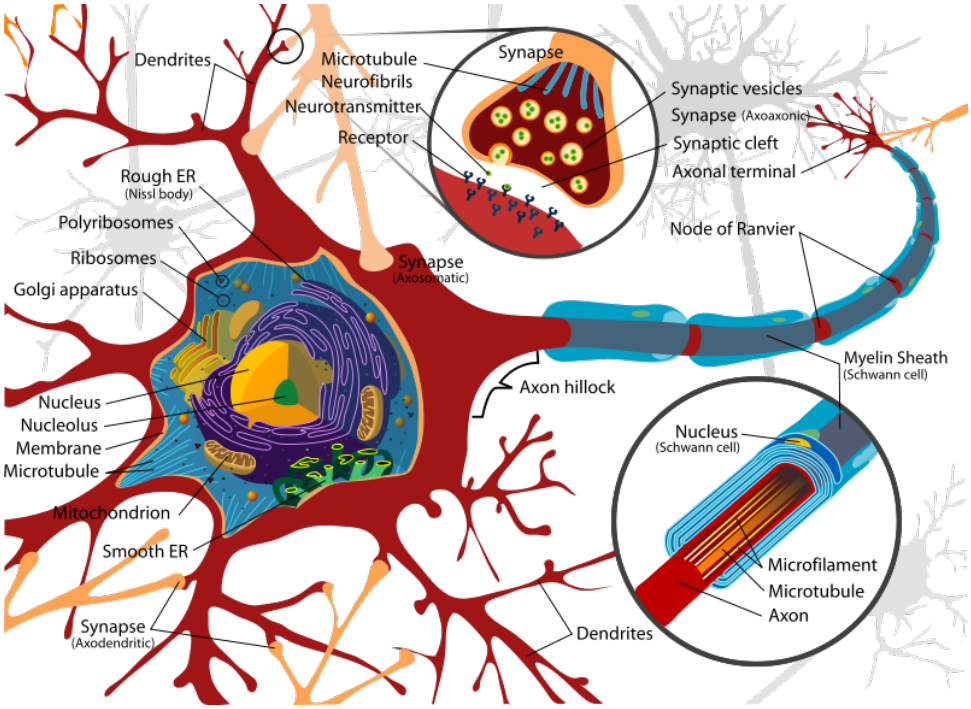


Figure 5.1: Diagram of a biological neuron [4]

5.2 Neural Networks in Computing

The basic principle of an artificial neuron is the same as in biology. An artificial neuron is an object, in most cases implemented by software, that has multiple inputs and one output. In the most simple case, the output is the result of a step function which is evaluated for the sum of the neuron inputs. To model the strength of the synapses of the biological neurons, a "connection weight" is multiplied with each input of the neuron. So for the most simple case the output y_j of neuron j is

$$y_j = g(\sum_i w_{ij} x_i) \tag{5.1}$$

where g is the step function, x_i the input values of the node and w_{ij} are the connection weights for each input.

If a continuous output instead of a yes/no decision is necessary, like for energy reconstruction, a continuous function is used instead of a step function. An artificial neuron that uses such a function is called perceptron. The activation function used here is the sigmoid symmetric function $g_{sigsym}(a)$.

$$g_{sigsym}(a) = tanh(a \cdot s) = \frac{2}{1 + exp(-2s \cdot a)} - 1 \tag{5.2}$$

Here a is the input of the function and s is a parameter called "activation steepness" that controls the steepness of the function. With growing s, the function nears a step function, while smaller s are used for continuous problems. In this work, s was fixed at 0.3.

For this work, feed-forward networks were used. A feed-forward network is a network where the information travels only from input to output. The basic topology is shown in figure 5.2. The decision making neurons are in the hidden and output layers. The connection weights are the parameters which were adapted to the specific problem for which the neural net should be used. This is done by the training process.

5.3 Network Training

In the training process the w_{ij} are adjusted. To minimize the error of the neural network, a standard algorithm called error backpropagation is used. This process is described here briefly in adoption from [24]. If $\mathcal{E}^N(\mathbf{y})$ is an error function for an output vector $\mathbf{y}$ of a neural net where N is the number of data points in data set, its dependence on an weight w_{ij} can be written as

$$\frac{\partial \mathcal{E}^N}{\partial w_{ij}} = \frac{\partial \mathcal{E}^N}{\partial a_j} \frac{\partial a_j}{\partial w_{ij}} \tag{5.3}$$

with $a_j = \sum_i w_{ij} z_i$. Due to definition, $\frac{\partial a_j}{\partial w_{ij}}$ can be replaced by z_i and $\frac{\partial \mathcal{E}^N}{\partial a_j}$ is defined to δ_j so that

$$\frac{\partial \mathcal{E}^N}{\partial w_{ij}} = \delta_j z_i \tag{5.4}$$

Then for the output layer

$$\delta_k = \frac{\partial \mathcal{E}^N}{\partial a_k} = g'(a_k) \frac{\partial \mathcal{E}^N}{\partial y_k}. \tag{5.5}$$

For the hidden layers the chain rule has to be applied with $a_k = \sum_j w_{kj} g(a_j)$ giving

$$\delta_j = \frac{\partial \mathcal{E}^N}{\partial a_j} = \sum_k \frac{\partial \mathcal{E}^N}{\partial a_k} \frac{\partial a_k}{\partial a_j} = g'(a_j) \sum_k w_{kj} \delta_k. \tag{5.6}$$

With that formula, called back-propagation formula, all δ can be calculated starting from the output layer. From the δ the derivatives $\frac{\partial \mathcal{E}^N}{\partial w_{ij}}$ can be obtained. Based on the summed up error for the whole data set, the w_{ij} are adjusted by minimizing it with a standard minimizer [12]. For the implementation of ANNs and training algorithms the FANN library was used [22].

In this work the training process is done in cycles, called epochs. After each epoch, the mean-square-error $\mathcal{E}^N_{MSE}$ is calculated.

$$\mathcal{E}^N_{MSE} = \sum_{n=1}^{N} (y(\mathbf{x}^n, \mathbf{W}) - E^n_\mu)^2 \tag{5.7}$$

N denotes the number of data points in the data sample.

Two independent data samples are used for the training process, one of them is called "training" sample what is used to train the neural net in each epoch. The second one, called "validation" sample, is used to evaluate the training process. Both training and validation mean-square-errors are calculated after each training epoch and the validation MSE is used to end the training process if it does not decrease over 5 subsequent epochs. Due to training and validation set being independent specialisation of the neural net to the training set is avoided [24].

The number of nodes used in each net is set to two times the number of input parameters with a minimum of 112 nodes for the first hidden layer and one quarter of the first hidden layer for the second hidden layer. These numbers were considered as best in [24] and should be sufficient. An increase of node numbers does not show any effect on the results.

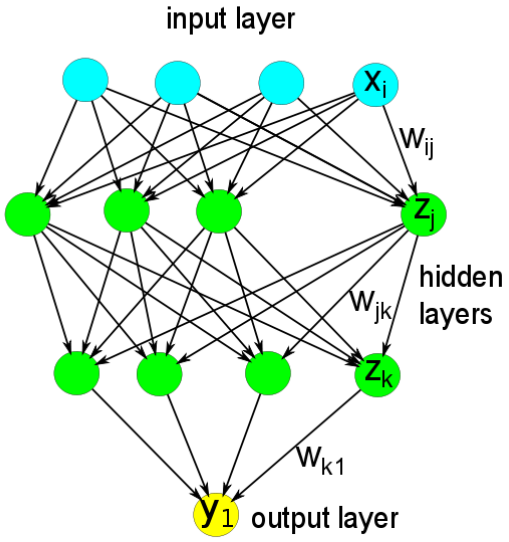


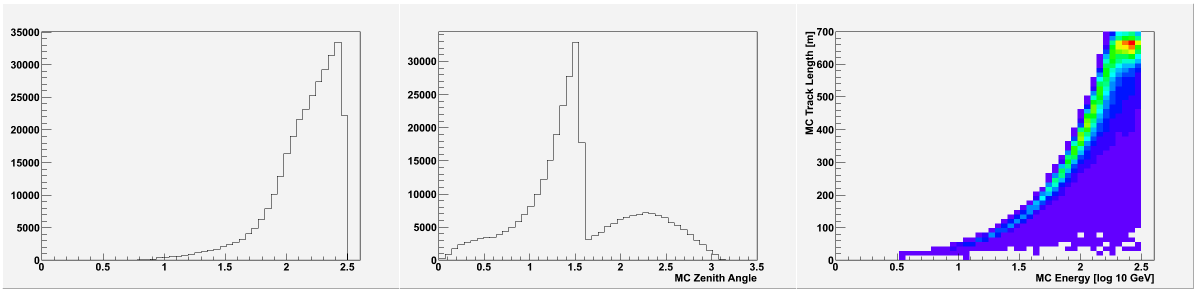
Figure 5.2: Scheme of an artificial neural network

Chapter 6

Testing of Muon Energy Reconstruction with ANN

6.1 Used Data sample

For this work pre-existing Monte-Carlo simulations were used. They were selected from a set of muon neutrino MC files. The files contain low energy neutrino and anti neutrino as well as up- and downgoing events. The KM3 simulation was already performed. The dataset was then converted to a suitable file format for the use in *SeaTray*, and all events with muon energy greater than 300 GeV were cut away due to the results from chapter 4. Also events where the particle with highest energy in the can was not a muon were discarded. In this work, all steps that handle raw data are done in *SeaTray*. 60 kHz white noise was then added and a standard PMT simulation was performed. After L1 hit selection and 3N and 2T3 trigger were used on the sample, the triggered events were written to files for the following reconstruction steps. The files contain the hit-maps as well as information from the Monte-Carlo simulation process for all events. A hit map is a map that contains a for each OM a list of hit times and charges of all hits on that OM. For this work only information that would be also available also from real data is used in further steps. This is at first the hit map that contains all hits that are causally connected with the hits that meet the trigger condition. They are called triggered hits. A second hit map contains all hits of the event (2.2 μs before and after trigger). Another map contains the status of each OM e.g. working or not, and also the detector geometry is used. For this work a standard detector geometry was used as well as 885 functional OMs. In order to test the reconstruction ability of the trained nets, a second data sample was created the same way using different Monte-Carlo files, called test sample.



(a) Energy distribution (b) Zenith angle distribution (c) MC track length over energy

Figure 6.1: Energy and zenith angle distribution and MC track length of MC simulated muons in the test sample

6.2 The Energy Reconstruction Chain

To reconstruct the muon energy, a basic reconstruction chain was used in this work. It follows the reconstruction chain as used in [24]. An overview over the complete work-flow starting with the Monte-Carlo sample is shown in figure 6.2. In the first step, pre-existing reconstruction algorithms were performed to generate more advanced selections of hits or other datasets that describe the event to use in the energy reconstruction from raw data. After that the parameters that will be used are extracted from that data. A subset of the events is selected as the training set and another subset as the validation set. The parameters of the training set are then scaled to $[-1, ..., 1]$ to give all parameters the same range what is necessary for the next steps. The function

$$x_i^{scaled} = \frac{4}{\pi} \arctan(\frac{x_i^{unscaled}}{\alpha}) - 1, \tag{6.1}$$

where x_i denotes the input parameters, is used to scale the parameters. The scaling parameter is evaluated by repeatedly scaling the set and checking the distribution of the scaled values. In the next step a principal component analysis (PCA) is performed. The PCA calculates the covariance matrix with entries

$$cov(x_i, x_j) = \frac{1}{N} \sum_{n=1}^{N} (x_{i,n} - < x_i >)(x_{j,n} - < x_j >) \tag{6.2}$$

and its eigenvalues and eigenvectors. The training set is transformed to feature space using these eigenvectors in order to extract the independent features of the data which are then rescaled for input of the NN. Scaling and transformation parameters are stored. The validation set parameters are preprocessed in the same way using the same scaling and transformation. The neural net is then trained using training and validation sets. After that training, the NN and again the scaling and PCA parameters are used to reconstruct the test data sample, and the result is written to a file for analysis.

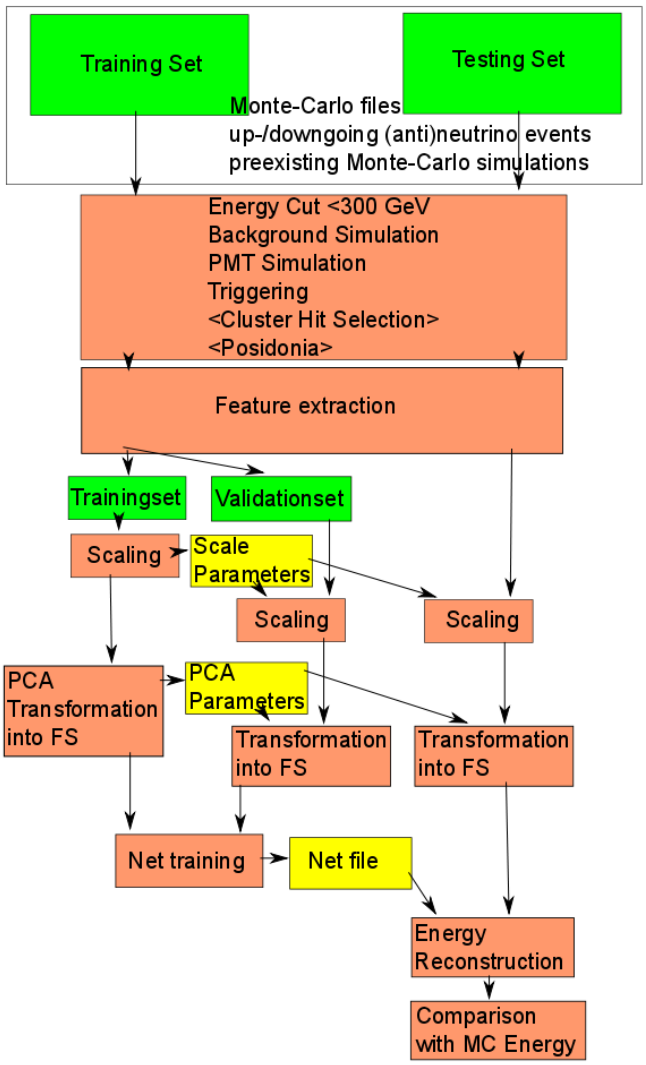


Figure 6.2: Scheme of Energy reconstruction testing chain. Orange: algorithms, Green: training and test data, Yellow: scaling and PCA parameters and NN

6.3 Test of the Simple Estimator

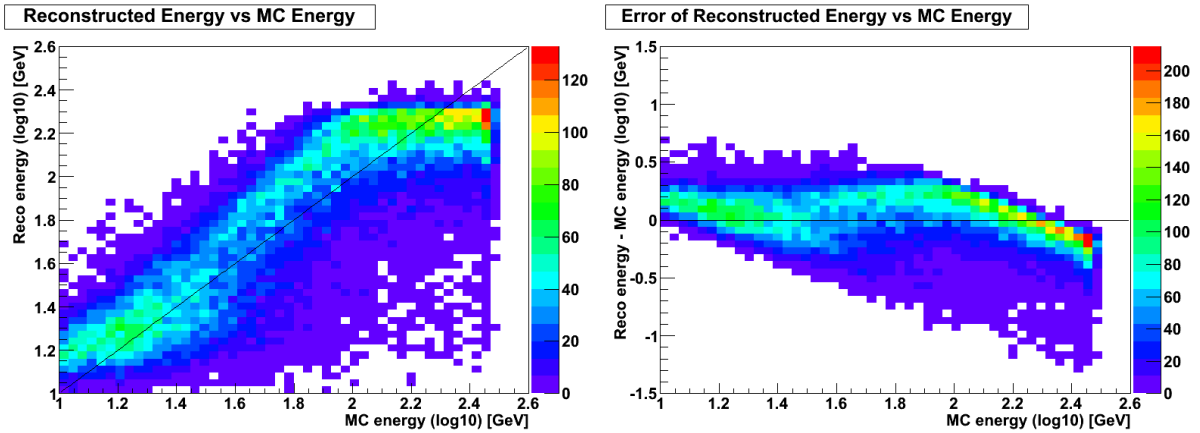
The Simple Estimator uses 20 input parameters for the Neural Net. A slightly different version was introduced in [24]. They work with the basic event hit distributions mentioned above. No hit selections or track reconstructions except the trigger are used before extracting them. The parameters that are extracted from the raw data are:

- Number of OMs with triggered hits. Triggered hits means all hits coincident with the hits that meet trigger conditions.
- Number of triggered hits
- Number of all hits: the number of all hits in a time window from 20 ns before first triggered hit to 300 ns after last triggered hit
- Number of hits in background window: This is the number of hits in time windows with a duration of 1000 ns, 700 ns before and 700 ns after first and last triggered hit. The parameter is necessary to pass information about the background conditions to the net.
- Number of working OMs: Some OMs could be out of order.
- Number of hit repetitions: the average number of hits per OM in an time window of 500 ns after the first hit on the OM
- Average charge per triggered hit
- RMS of the charge distribution of triggered hits
- Average charge per hit
- RMS of the charge distribution of all hits
- Number of triggered hits on first hit string
- Number of storeys with triggered hits
- Number of strings with triggered hits
- Distance between first and last triggered hit
- Distance between center of gravity (CoG) of the event and CoG of the detector
- Approximate zenith angle: The zenith angle is approximated by the zenith angle of the line between the first and last triggered hit
- Time difference between first and last triggered hit
- Average time of triggered hits after first triggered hit
- RMS of the time distribution of hits after first triggered hit
- Sum of inverse hit velocity: The inverse hit velocity is defined as the time of a hit over the distance of a hit to the first triggered hit. It is summed over all triggered hits

The distribution of all input parameters with respect to the muon energy distribution is given in the appendix.

After performing of the preprocessing, the neural network was trained, and then the test data sample was reconstructed by the trained net. The training and validation samples are not included in the test sample, so that an eventual specialisation of the NN to the training sample, although it should be excluded by the training process, causes no effect in the evaluation of reconstruction error.

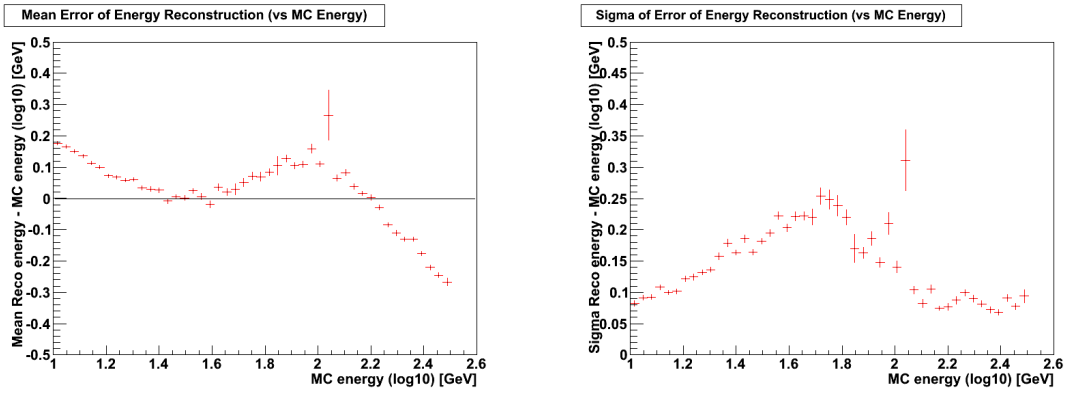
The results of the reconstruction are then compared with the Monte-Carlo-energies. Figure 6.3a shows the 2D-plot of the reconstructed energy over the MC energy. Figure 6.3b shows the difference between the reconstructed energy and the MC energy over MC energy. For each bin in x-axis of this



(a) Reconstructed energy over MC energy

(b) Difference of reconstructed energy and MC energy over MC energy

Figure 6.3: Energy reconstruction of the test sample over MC energy



(a) Mean of Gaussian fit of energy difference through MC energy bins

(b) Sigma of Gaussian fit of energy difference through MC energy bins

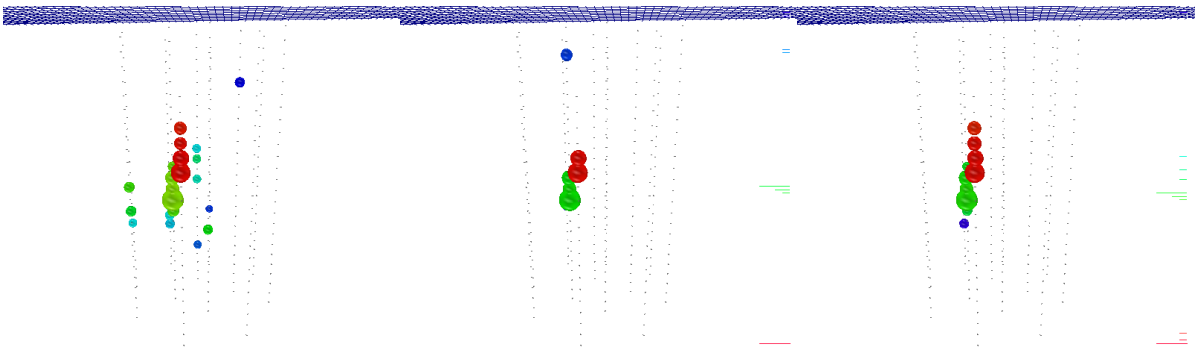
Figure 6.4: Evaluation of reconstruction quality by Gaussian fit of reconstructed energy distribution

plot a Gaussian fit was performed and the mean and sigma values were plotted in figures 6.4a and 6.4b. It has to be mentioned that the sigma value of the Gaussian fit can show a low value in the case there is a huge and thin spike anywhere in the distribution. In this case, the sigma value does not necessary tell much about reconstruction quality. The number of events in each bin in MC energy was limited to 1000 to ensure comparability between the bins. The plots for the difference of MC and reconstructed energy over the reconstructed energy can be found in the appendix.

As it can be seen from figures 6.3a to 6.4b, the reconstruction is relatively accurate for energies lower than 100 GeV. In that range the maximum mean of the difference to MC energy is +0.2, and the maximum sigma is 0.25, in logarithmic units, which would be 50% and 80% in non logarithmic units. Above 100 GeV it is obvious from figure 6.3a that energies cannot be distinguished any further. This is exactly as it is expected from chapter 4. From plot 6.1c it can also be seen that at an energy of 100 GeV the track length of muons are about 350 m, which is in the same range as the size of the detector. So it should be in fact an upper limit for energy reconstruction. Any effort to improve reconstruction in that range will probably be unsuccessful. Nevertheless, one can try to improve energy reconstruction in the range below 100 GeV.

6.4 Test of Cluster Hit Selection

One approach to improve the energy reconstruction is to find a input hit selection which is more specific for the muon related hits than the original triggered hit sample. As the Cluster Hit selection Algorithm was designed for low energy track reconstruction, it was considered. An inspection of some example events with *GLshovel*, a *SeaTray* module for displaying hit distributions in the detector, shows that the hit distribution from Cluster Hit selection is indeed closer to the original Monte-Carlo hit sample than the triggered hits. One example event is shown in figure 6.5.



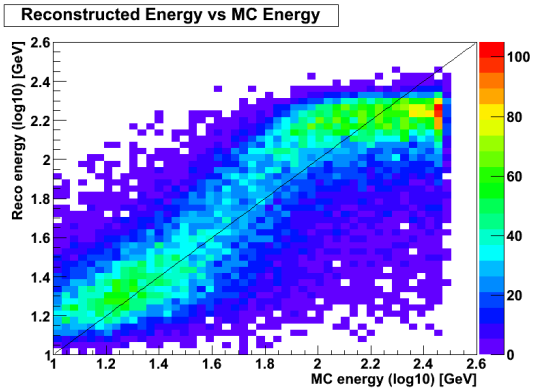
(a) Monte-Carlo hits (b) Triggered hits (c) Hits from Cluster Hit selection

Figure 6.5: Exemplary comparison of triggered hits and hits from Cluster Hit selection

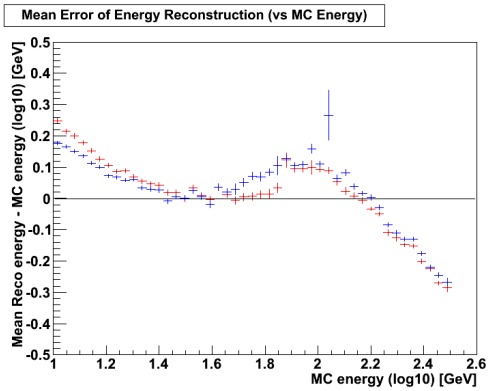
The hit distribution, which is the output of the Cluster Hit selection, was used as an input for the Simple Estimator. The cluster hits replace the triggered hits which were used before. The resulting plots are shown in Figures 6.6a to 6.7. More plots are shown in the appendix.

Although it seems from the example events that the Cluster Hit selection selects more hits from the original muon than the trigger, using it before the energy reconstruction does not improve it. It was expected that it would not help at energies over 100 GeV, but also for lower energies there is no improvement. In the range lower than 25 GeV the mean and sigma values of the error even increase (max 0.06 in log units resp. max 0.05 in log units), while in the range 50 to 80 GeV the mean decreases (max 0.08 in log units).

As a result it can be concluded that a better hit selection has little effect on the energy reconstruction with the Simple Estimator. This can possibly be explained with the fact that the Simple Estimator uses parameters retrieved from all hits in the time window in addition to the triggered hits, so that a hit not



(a) Reconstructed energy over MC energy



(b) Mean of Gaussian fit of energy difference through MC energy bins (red). Blue markers show the result from Simple estimator without hit selection as reference.

Figure 6.6: Evaluation of the quality of energy reconstruction for Simple estimator with Cluster Hit selection

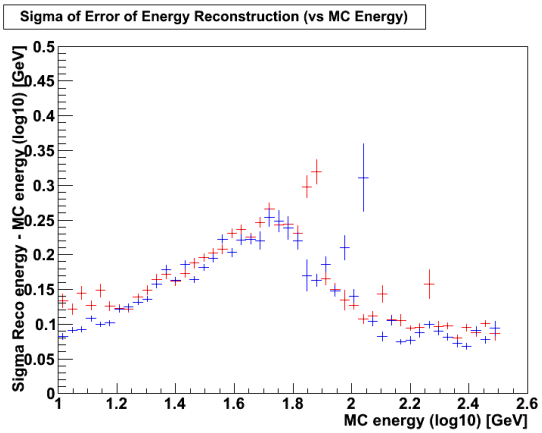


Figure 6.7: Sigma of Gaussian fit of energy difference through MC energy bins (red). Blue markers show the result from Simple estimator without hit selection as reference.

included in the triggered hits is nevertheless recognised by the net. Also some parameters like strings with triggered hits do not differ much between Cluster Hit selection's selected hits and triggered hits.

6.5 Test of Cherenkov Estimator with *Posidonia* Track Fitter

Another approach to improve energy reconstruction is to pass information about the muon track to the neural net. For this task a track reconstruction algorithm is used before the energy reconstruction, in which a specialised version of the energy estimator is used, which can process the additional information. This is the so called Cherenkov estimator, which was introduced in [24] and extended in this work by implementing additional parameters, which were partly used in [26]. The Cherenkov estimator needs a reconstructed muon track in addition to the hit maps the simple estimator uses. Using the Cherenkov estimator has already achieved good results for reconstructing energy in the high energy range (Fig. 6.8).

- Number of OMs with triggered hits (Triggered hits means all hits coincident with hits which meet the trigger condition.)
- Number of triggered hits
- Number of hits in the time window between 20 ns before the first and 300 ns after the last triggered hit (From now on called "timewindow"
- Average Charge of triggered hits
- Average Charge of hits in time window
- RMS of charge distribution of triggered hits
- Number of stories with triggered hits
- Number of strings with triggered hits
- Effective track length

The effective track length is calculated as the distance between the first and last point on the muon track where the photons that are detected would have left the track if they are emitted under the Cherenkov angle of 42° and not scattered. All triggered hits with time residuals from -10 to +250 are considered. The time residual is the difference of the measured photon arrival time at the OM and the time it should arrive if it would start at the muon track under the Cherenkov angle and no scattering takes place.

- Zenith angle of muon
- Distance of closest approach of the muon track to detector COG
- Average Cherenkov distance of hits in time window
- RMS of Cherenkov distance distribution for hits in time window

The Cherenkov distance describes the distance a photon would travel if it is emitted at the muon track under Cherenkov angle.

- Average time residual of hits in time window
- RMS of distribution of time residuals of hits in time window
- Event duration (time difference between first and last triggered hit)
- Average time residual per Cherenkov distance for hits in time window
- RMS of distribution of time residuals per Cherenkov distances for hits in time window
- Fraction of hits in time window with time residuals < 40 ns
- Fraction of hits in time window with time residuals < 250 ns
- Fraction of hits in time window with time residuals > 600 ns
- Average Amplitude of hits in time window with time residuals < 40 ns

- Average Amplitude of hits in time window with time residuals < 250 ns
- Track length in an cylinder around the detectors COG with height and radius of the detector
- Number of hits in background window (duration 2x 1000 ns, ending 700 ns before first and starting 700 ns after last triggered hit)

The six following parameters were used separately for each $x \in \{20, 40, 60, 80, 100\}$. Angular acceptance is a value that parametrizes the amount of light the OMs could receive from a certain muon track.

- Number of working OMs in a radial distance between $x - 20$ m and x m around the muon track
- Sum of angular acceptance in a radial distance between $x - 20$ m and x m around the muon track
- Number of OMs with hits in a radial distance between $x - 20$ m and x m around the muon track and hit time residual < 40 ns
- Number of OMs with hits in a radial distance between $x - 20$ m and x m around the muon track and hit time residual > 40 ns, < 250 ns
- Charge of hits in a radial distance between $x - 20$ m and x m around the muon track and hit time residual < 40 ns
- Charge of hits in a radial distance between $x - 20$ m and x m around the muon track and hit time residual > 40 ns, < 250 ns
- Number of working OMs

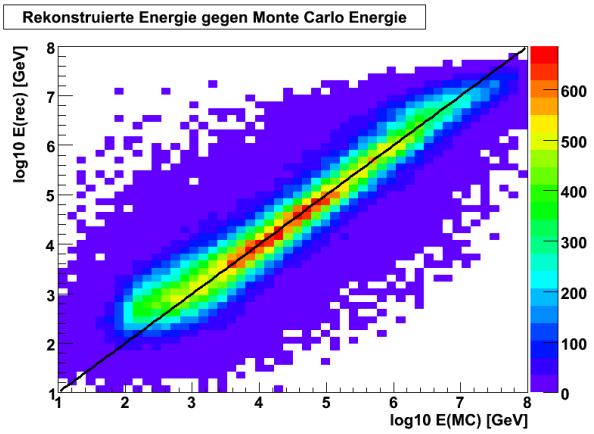
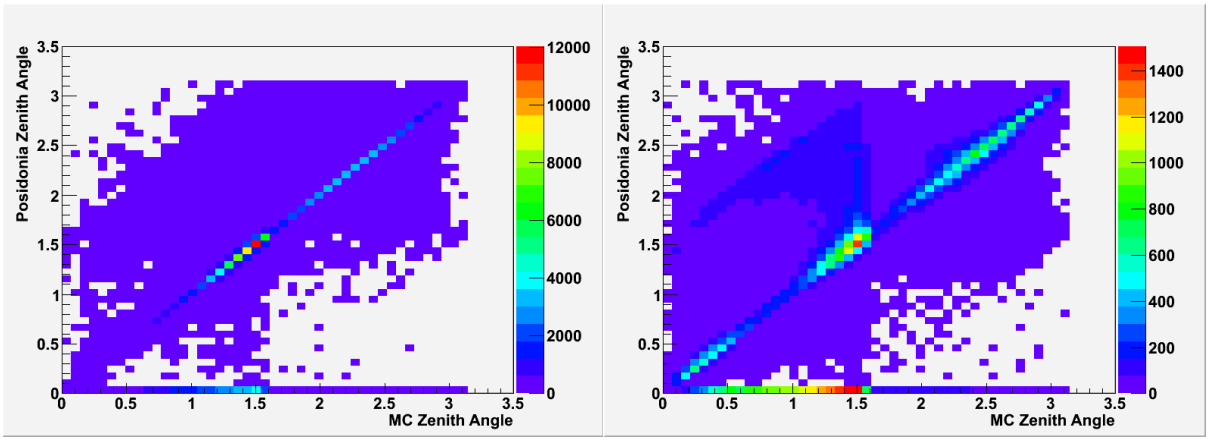


Figure 6.8: Behaviour of the Cherenkov estimator in high energy reconstruction

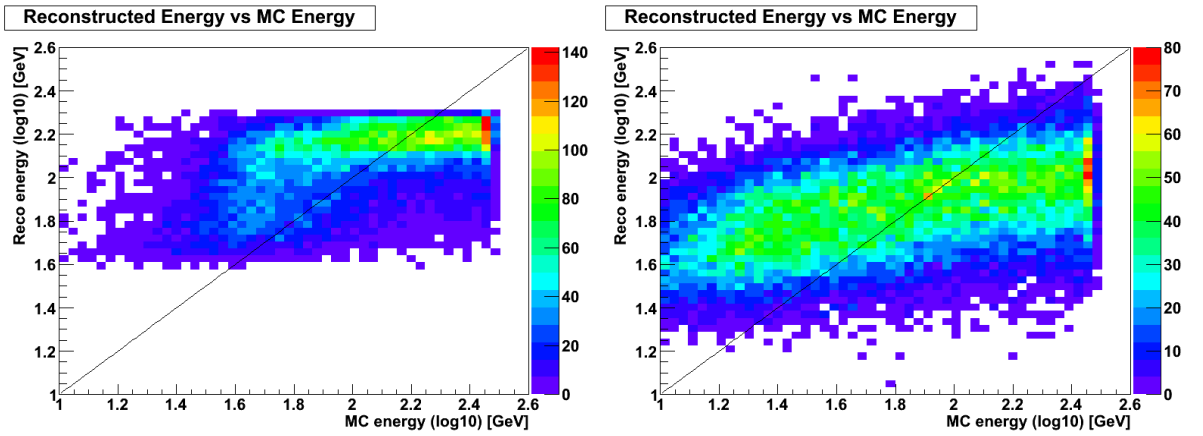
As the usual ANTARES track reconstruction is designed for high energies, for reconstructing the muon track for low energies a special track reconstruction, *Posidonia*, was used for this work. It is described in chapter 3. Figures 6.9a and 6.9b show the accuracy of the *Posidonia* reconstruction for the example of the zenith angle. It can be seen that most of the muons that are reconstructed have a relatively accurate zenith angle, especially for the multi-line events. For single-line events there is a mirror solution for some tracks that causes some down-going muons to be reconstructed as up-going. For some events *posidonia* is not able to reconstruct them, they are displayed here with zenith angle 0. For single-string events, 20690 out of 112268 events are not reconstructed, for multi-line events 35429 out of 203912.

The Cherenkov Estimator used the reconstructed muon track in addition to the raw hit samples. As *Posidonia* distinguishes events that have hits on multiple lines and single-string events, different NNs were trained for both cases. The results are shown here, in different plots for multi-string and single-string events. Figures 6.10a to 6.12a show multi-string, figures 6.10b to 6.12b single-string events. Events which could not be reconstructed with *Posidonia* were cut away before plotting.



(a) for multi-string events (b) for single-string events

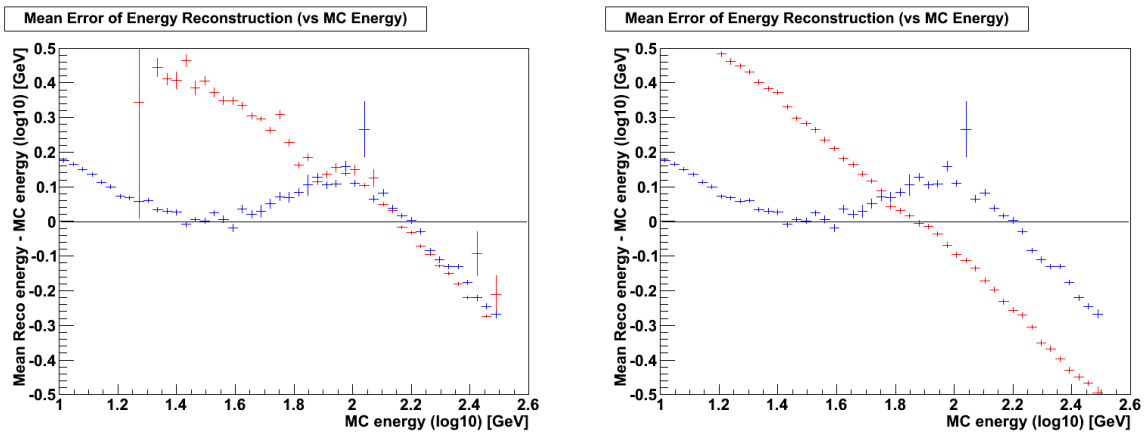
Figure 6.9: Zenith angle of track reconstructed with *Posidonia* over MC zenith angle. Muons not reconstructed by *Posidonia* are displayed with zenith angle = 0



(a) for multi-string events (b) for single-string events

Figure 6.10: Energy reconstructed with Cherenkov Estimator used with Posidonia track reconstruction shown over Monte-Carlo energy

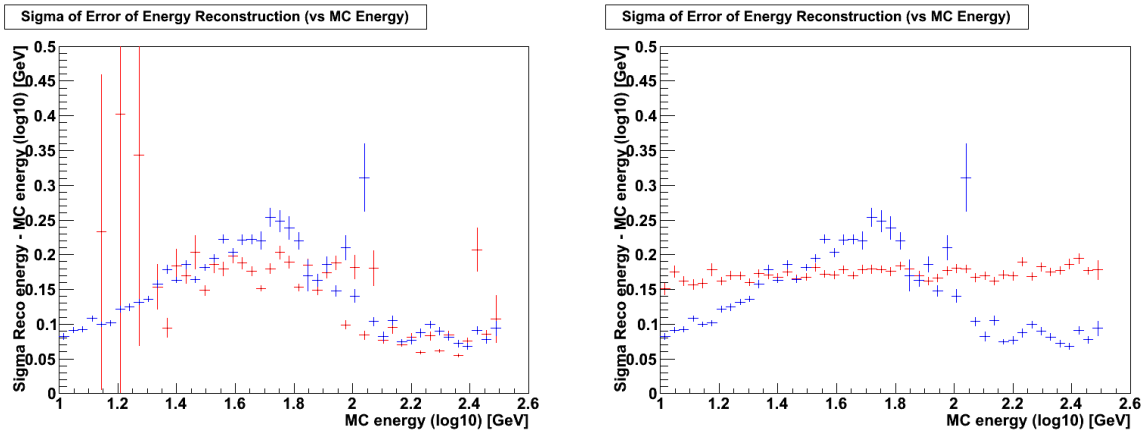
Despite of the high accuracy of the Cherenkov Estimator for the reconstruction of high energies, namely energies in an range where radiative energy losses are important, and the high accuracy of the track fitted with *Posidonia* for most of the events, the performance of the Cherenkov Estimator used with *Posidonia* tracks in the low energy range is significantly lower than the Simple Estimator's performance. This is easily seen from figures 6.10a and 6.10b in comparison to 6.3a. The sigma values of the Gaussian fit are sometimes lower than for the Simple Estimator version, but this is obviously a statistical effect. For the case of multi-string events figure 6.10a clearly shows that the vast majority of that events have energies above 100 GeV where, as seen from chapter 4, the detector's response is not dependent on energy any more, which explains the failure of the NN to accurately reconstruct the energy. For the case of single-string events there are also many events below 100 GeV, but the quality of the reconstruction is nevertheless poor. An explanation is that many parameters in the parameter set are designed to be dependent on the Cherenkov light intensity per unit of track, and not to the track length.



(a) for multi-string events

(b) for single-string events

Figure 6.11: Mean values from Gaussian fit of reconstruction error along MC energy for energy reconstructed with Cherenkov Estimator used with Posidonia track reconstruction (red). Blue markers show the result from Simple estimator without hit selection as reference.



(a) for multi-string events

(b) for single-string events

Figure 6.12: Sigma values from Gaussian fit of reconstruction error along MC energy for energy reconstructed with Cherenkov Estimator used with Posidonia track reconstruction (red). Blue markers show the result from Simple estimator without hit selection as reference.

Chapter 7

Test of Advanced Parameters

It can be tried to improve the quality of energy reconstruction by enlarging the set of parameters that are passed to the neural net. In this work, some pre-existing algorithms, that extract features out of the hit maps, are tested whether they are providing useful information for energy reconstruction with neural networks.

7.1 CalcCluster

CalcCluster is an algorithm which first sorts all the input hits by their charge and then adds them to clusters, by checking if they are causally connected to the other hits of the cluster. Therefore, the equation

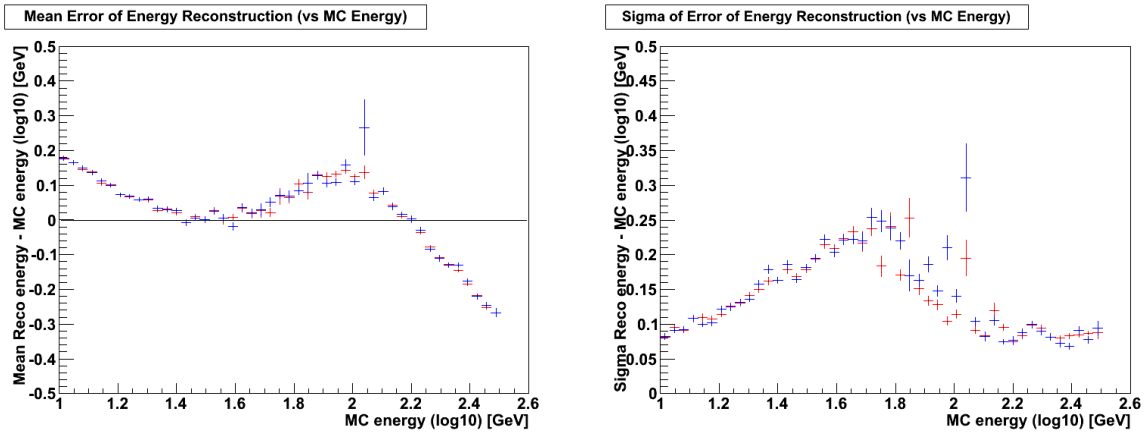
$$\Delta t <= n/c \cdot \Delta x + t_s \tag{7.1}$$

with Δt the time difference between two hits and Δx the distance between them must be fulfilled for all the hits in cluster. If it is not, a new cluster is generated. As input hits the L1 hit selection is used. Then, four parameters are calculated.

- Number of Clusters
- Maximum Size of Cluster
- Average Size of Clusters
- Overlap between Clusters: This is the number of all hits in all clusters over the total number of hits

This is performed for each t_s = 0 ns, 10 ns, 20 ns and 50 ns, so that 16 parameters are generated. t_s stands for the travel time of photons which is assumed for scattering [6]. The 16 parameters are then used in addition to the 20 parameters as input for the NN. The preprocessing and net training is done in the same way as above.

The results are shown in figures 7.1a and 7.1b. The blue indicators show the values of the simple estimator reconstruction as a reference. As it can easily be seen, the use of the CalcCluster does not improve the energy reconstruction. That means that either the new parameters do not provide extra information to the neural net, or the net cannot adapt to the information due to the increase of the number of parameters.



(a) Mean value

(b) Sigma value

Figure 7.1: Use of CalcCluster: Mean and sigma values of Gaussian fit of difference between MC and reconstructed energy over MC energy. Blue indicators are for Simple Estimator as a reference.

7.2 CalcDensity

The CalcDensity classifier divides the detector in 20 subvolumes, 5 divisions in the vertical (similar to the sectors) and 4 in the horizontal plane. The event is then divided into timeslices of 50 ns each. A dataset which contains the data of one subvolume and one timeslice is called a subunit. For each subunit the charge density, meaning the total charge in the subunit over the number of active OMs in the subunit is calculated. Also the pulse size, meaningy2 the total charge per subvolume over the number of hits in the subvolume, is calculated. Then some statistical values about the set of subunits are calculated and returned [6].

- Difference of the mean charge density of the subvolume with the largest mean charge density and the subvolume with the second largest
- Difference of the mean charge density of the subvolume with the largest mean charge density and the subvolume with the third largest
- Relation between the mean charge density of the subvolume with the second largest mean charge density and the subvolume with the largest
- Relation between the mean charge density of the subvolume with the third largest mean charge density and the subvolume with the largest
- Difference of the largest pulse size per subvolume and the second largest
- Difference of the largest pulse size per subvolume and the third largest
- Relation between the second largest pulse size per subvolume and the largest
- Relation between the third largest pulse size per subvolume and the largest
- Number of subunits that contain 75% of the total charge
- Number of subunits that contain 90% of the total charge
- Number of subunits that contain 100% of the total charge
- Fraction of subunits that contain 75% of the total charge
- Fraction of subunits that contain 90% of the total charge
- Fraction of subunits that contain 100% of the total charge

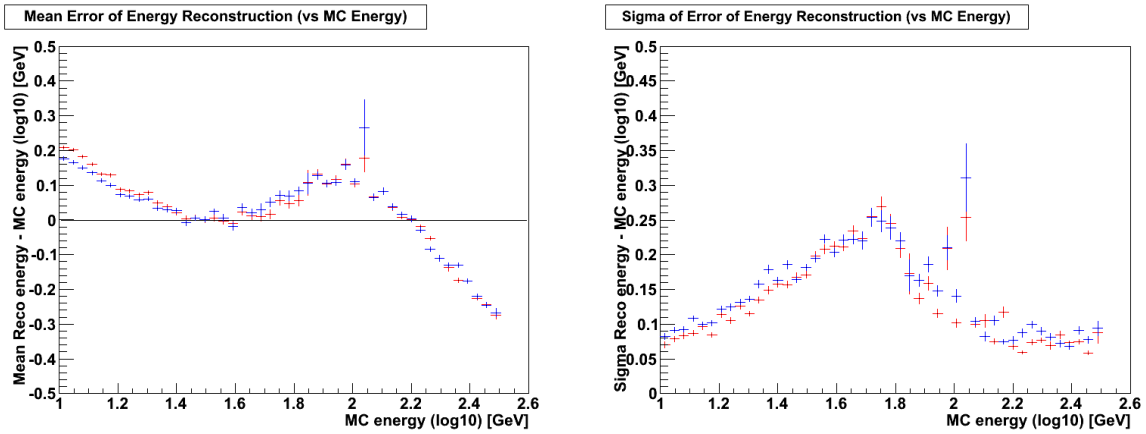
- Number of subunits that contain a charge density larger than the charge density of noise
- Number of subunits that contain a charge density larger than the double of charge density of noise
- Fraction of subunits that contain a charge density larger than the charge density of noise
- Fraction of subunits that contain a charge density larger than the double of charge density of noise

Fraction of subunits means here the number of subunits in the condition over the total number of subunits. The charge density of noise is a constant value and is predefined as 0.02. In addition to that, some other statistical parameters about the event are returned.

- Number of storeys with hits
- Number of storeys with hits in the timeslice with the largest number of hit storeys(the timeslice is now 100 ns)
- Fraction of storeys with hits in the timeslice with the largest number of hit storeys
- Difference between the number of storeys with hits in the timeslice with the largest number of hit storeys and the number of storeys with hits in the timeslice with the second largest number of hit storeys
- Relation between the number of storeys with hits in the timeslice with the second largest number of hit storeys and the number of storeys with hits in the timeslice with the largest number of hit storeys
- Total charge of all hits
- Average charge per hit
- Time from first to last hit
- Time from first to last triggered hit
- The constant charge density of noise

The noise density parameter is not passed to the neural net because it is constant for all events. The time from last to first triggered hit and the average charge per hit are also not passed as they already are input parameters from the Simple Estimator parameter set that is passed to the net in addition to the new parameters.

The results are shown in figures 7.2a and 7.2b. Although the set of parameters from CalcDensity is significantly larger than the set of the Simple Estimator there is no improvement of the energy reconstruction's accuracy seen from the comparison. The mean error of the reconstruction even increases slightly (about 0.02 in logarithmic units) in the range of very low energies up to 15 GeV. There are three possible explanations for that fact. First it could be possible that the the new parameters do not provide energy dependent information or information that cannot be used to improve energy reconstruction, due to their not being dependent on any intrinsic parameter of the energy loss mechanism. The second possibility is that the parameters provide such information, but due to their high number the NN also has more difficulties to extract the relevant information, and the third possibility is that the energy loss mechanism and the detector's response to that do not have more intrinsic energy dependent information at all.



(a) Mean value

(b) Sigma value

Figure 7.2: Use of CalcDensity: Mean and sigma values of Gaussian fit of difference between MC and reconstructed energy over MC energy. Blue indicators are for Simple Estimator as a reference.

7.3 CalcTensor

The CalcTensor algorithm first calculates the centre of gravity of the hits, both weighted with charge and time, and the unweighted one. Then the event's tensor of inertia and its quadrupole tensor are calculated for the charge weighted and the unweighted version. For the tensors of inertia the eigenvalues are stored. The central principal axis is calculated using the weighted inertia matrix, and the orientation of the particle and the zenith is estimated from that. At last, the quadrupole moment along the central principal axis is calculated and the following parameters are returned [6]:

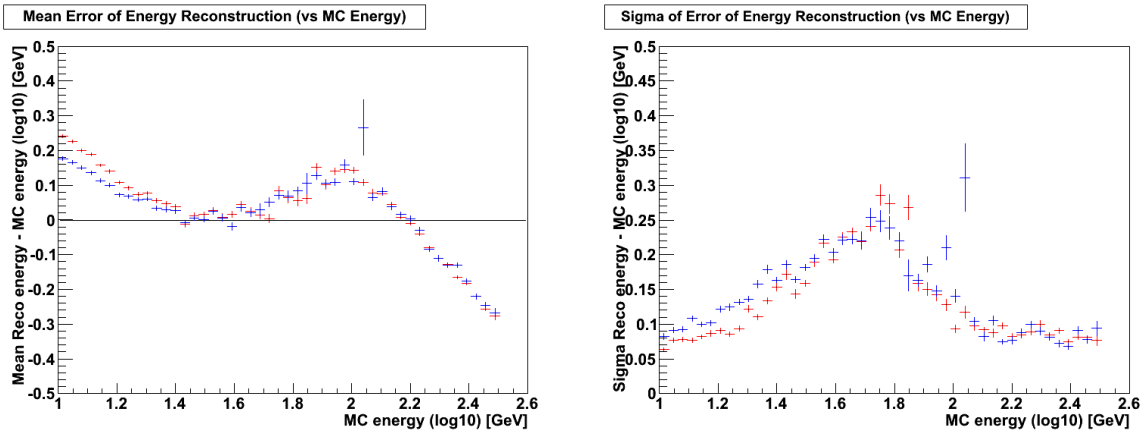
- Number of hits the classifier used
- Sum of charge if used hits
- Number of OMs
- Coordinates of center of gravity (COG)

Three sets of coordinates are written out, for the unweighted, weighted by charge and weighted by time COG. Each X,Y,Z coordinate is a single parameter.

- Sum of non-diagonal elements of the quadrupole tensor weighted by charge
- Normalized sum of non-diagonal elements of the quadrupole tensor weighted by charge
- Sum of non-diagonal elements of the unweighted quadrupole tensor
- Normalized sum of non-diagonal elements of the unweighted quadrupole tensor
- Normalized sum of X*Y*Z weighted by charge
- Eigenvalues of tensor of inertia

The three eigenvalues of the tensor are ordered by their values. This is done for both the unweighted tensor and the tensor weighted with charge. In addition, the normalized values are given. The eigenvalues of the unweighted tensor are normalized with the hit number, those of the weighted tensor with the total charge.

- Decision parameter for track orientation calculated by the unweighted values
- Decision parameter for track orientation calculated by the weighted values
- Calculated zenith angle



(a) Mean value

(b) Sigma value

Figure 7.3: Use of CalcTensor: Mean and sigma values of Gaussian fit of difference between MC and reconstructed energy over MC energy. Blue indicators are for Simple Estimator as a reference.

- Quadrupole moment along track, unweighted
- Quadrupole moment along track, unweighted, normalized
- Quadrupole moment along track, weighted by charge
- Quadrupole moment along track, weighted by charge, normalized

The preprocessing, net training and reconstruction was executed as in the cases above and the figures 7.3a and 7.3b show their results.

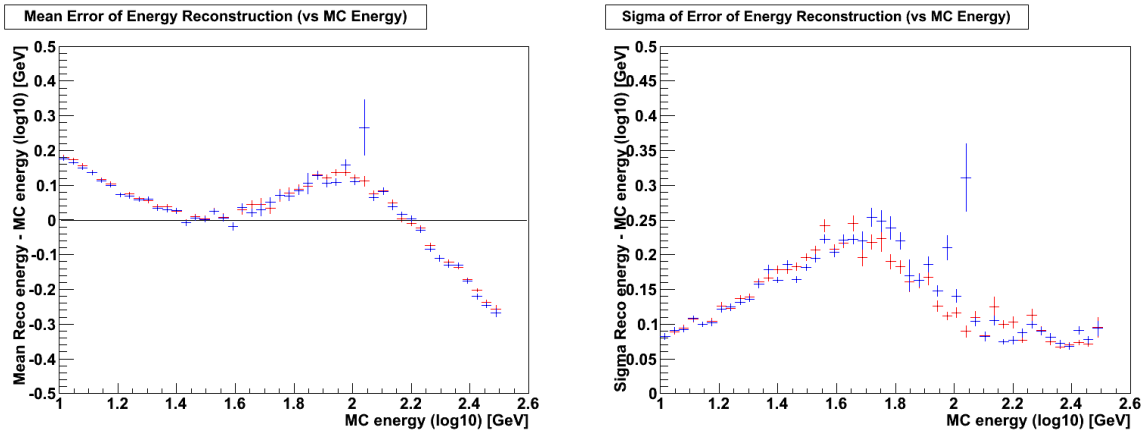
The result is nearly similar to the result of the use of CalcDensity. The increase of the mean error is even a bit larger for the lowest energies (about 0.06, log units), while the sigma shows a small decrease for the same range.

7.4 CalcShape

The CalcShape classifier tests if modules with hits are adjacent, i.e. no line or storey lies between them. All groups with adjacent hits are considered to be one cluster. For the largest cluster its total charge and the size of its envelope, i.e. the number of storeys lying on its border is now calculated. A set of six parameters is then returned [6]:

- Number of clusters
- Charge of largest cluster [1]
- Number of modules in largest cluster [2]
- Number of modules on boarder of largest cluster [3]
- Compactness parameter 1: $\sqrt{[3]^3/[2]^2}$
- Compactness parameter 2: $\sqrt{[3]^3/[1]^2}$

In figures 7.4a and 7.4b the results of the reconstruction with the CalcShape parameters are displayed in comparison to the Simple Estimator reconstruction. They do not differ from each other, which indicates that the output parameters of CalcShape also do not provide additional information about the energy loss mechanism to the NN.



(a) Mean value

(b) Sigma value

Figure 7.4: Use of CalcShape: Mean and sigma values of Gaussian fit of difference between MC and reconstructed energy over MC energy. Blue indicators are for Simple Estimator as a reference.

All four tested feature extraction algorithms could not provide parameter sets that improve the energy reconstructions with NNs. As they all use different approaches to extract features from raw data, it has to be considered that the seen uncertainty is caused by the transformation of the energy dependent track length, which is, as seen in chapter 4, the only energy dependent value available to the experimentalist, to the response of the detector. Nevertheless, some circumstances have to be checked which could decrease the accuracy of the NN energy reconstruction, such as the existence of parameters which are strongly correlated to each other or unimportant at all.

7.5 Methods for parameter discarding

To take a decision which parameter to discard, several possibilities exist to test the parameter. None of them can give a definite answer, and usually several of them are tried before the decision is made. The simplest way to test a parameter is to look directly on its dependence on the muon energy. This can be done by looking at the parameter-versus-energy plot by eye, or, in a more formal way, by calculating the covariance between the parameter and energy. The limit of this method is that even a parameter which is energy independent, e. g. the zenith angle, can nevertheless be important for the energy reconstruction.

Another value which can help to decide if a parameter should be discarded is its importance for the trained NN. The importance is defined as the product of the connection weights along one path from the output of the NN to the considered input parameter, summed up for all possible ways through the net that lead to this parameter. Again this method is limited, as it does not consider the influence of the activation function. The importance values for the NNs in this work had to be transformed back from feature space to the parameter space which was done using the inverse transformation of that used in the preprocessing.

The dependence of the NN to a specific parameter can also be a hint on how important this parameter is. After a NN was trained, a test data sample is reconstructed several times, for which the parameter to test is fixed for each reconstruction run and varied in steps between its maximum and minimum between the runs. The dependence is then calculated from the variation of the mean error of the reconstruction for the sample in respect to the parameters variation. This method is also called Holdback InPut Randomization method (HIPR). The limit of this method is that by varying one fixed parameter it can give parameter sets to the NN that are not possible in reality. The dependences also had to be transformed from feature space to parameter space.

It is also possible that two or more parameters which are given to a NN are highly correlated or anti-correlated to each other, so that passing only one of them to the net would nevertheless pass the same information to it. As the ability of the NN training process to adept to the data worsens with

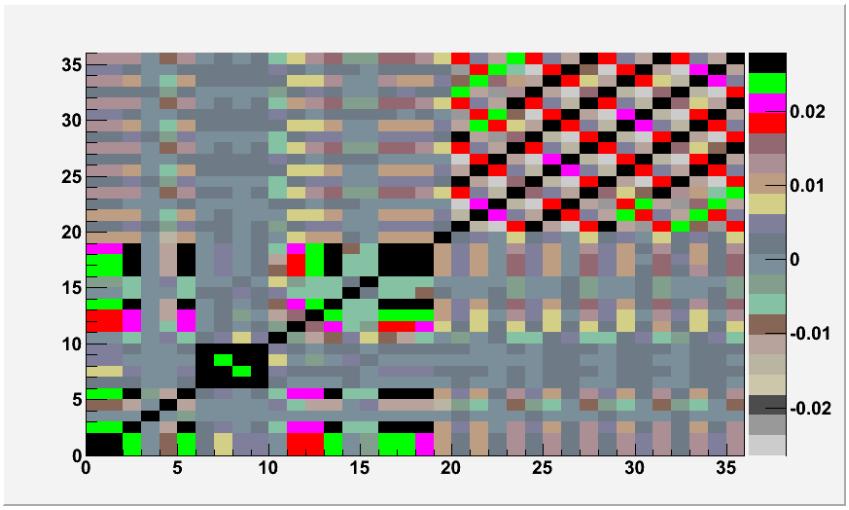


Figure 7.5: Covariance matrix for parameters with use of CalcCluster. Black denotes highly correlated parameters, white large anti-correlations

more parameters, such redundancies have to be avoided. To check the parameter set for correlations, the covariance matrix, which is already calculated by the PCA, can be used. An example of this matrix is given in figure 7.5.

All four parameter sets from the different data extraction algorithms have been tested with these methods. As both Dependence and Importance methods have their limitations, it was decided only to discard a parameter when it would have a low value for both methods and is not considered as energy dependent from the value-energy plot. The examination of the Importance and Dependence values showed some correlations of low values, but many of them were considered to be energy dependent from the plots. This could be a hint for the existence of correlations between parameters. On that fact the decision was made to base the discarding of parameters on eliminating redundancies. Only one parameter was discarded at this step, the number of functional OMs from the simple estimator dataset. This parameter has a constant value of 885 in this work, since no defective OMs were considered the in detector simulation. In the case of a test of the reconstruction's response to different detector conditions, this parameter should of course be included again.

7.6 Discarded parameters

Analysis of the covariance matrix for the CalcCluster parameters shows that there is no strong correlation between CalcCluster's output and the original Simple Estimator parameters. Inside these two parameter sets, however, there were strong correlations. To reduce redundancy between the different t_s steps, it was decided only to use the steps $t_s = 0$ and $t_s = 50$ ns. Also the parameter "Average Size of Clusters" was discarded as it was strongly anti-correlated to the number of clusters. For the parameter set from Simple Estimator, the following parameters were excluded:

- Number of triggered hits: correlated to Number of OMs with triggered hits
- Number of working OMs: always 885 for this tests
- Number of hit repetitions: correlated to Number of all hits
- RMS of the charge distribution of triggered hits: correlated to Average charge per triggered hit
- Average charge per hit: correlated to Average charge per triggered hit
- RMS of the charge distribution of all hits: correlated to Average charge per triggered hit
- Distance between first and last triggered hit: correlated to Number of all hits

- Time difference between first and last triggered hit: correlated to Number of all hits
- Average time of triggered hits after first triggered hit: correlated to Number of all hits
- RMS of the time distribution of hits after first triggered hit: correlated to Number of all hits

The same analysis was performed for the parameters extracted by CalcDensity. In this set, there were correlations between them and Simple Estimator parameters, in addition to correlations inside the CalcDensity set. The Simple Estimator parameters that were discarded above are also discarded here. Further discarded parameters were:

- Difference of the mean charge density of the subvolume with the largest mean charge density and the subvolume with the second largest: correlated to Fraction of subunits that contain a charge density larger than the double of charge density of noise: correlated to Number of subunits that contain a charge density larger than the double of charge density of noise
- Difference of the mean charge density of the subvolume with the largest mean charge density and the subvolume with the third largest: correlated to Fraction of subunits that contain a charge density larger than the double of charge density of noise: correlated to Number of subunits that contain a charge density larger than the double of charge density of noise
- Relation between the mean charge density of the subvolume with the second largest mean charge density and the subvolume with the largest: anti-correlated to Number of subunits that contain a charge density larger than the double of charge density of noise
- Relation between the mean charge density of the subvolume with the third largest mean charge density and the subvolume with the largest: anti-correlated to Number of subunits that contain a charge density larger than the double of charge density of noise
- Difference of the largest pulse size per subvolume and the third largest: correlated to Difference of the largest pulse size per subvolume and the second largest
- Relation between the second largest pulse size per subvolume and the largest: correlated to Difference of the largest pulse size per subvolume and the second largest
- Relation between the third largest pulse size per subvolume and the largest: correlated to Difference of the largest pulse size per subvolume and the second largest
- Fraction of subunits that contain 90% of the total charge: correlated to Fraction of subunits that contain 75% of the total charge
- Difference between the number of storeys with hits in the timeslice with the largest number of hit storeys and the number of storeys with hits in the timeslice with the second largest number of hit storeys: correlated to Fraction of storeys with hits in the timeslice with the largest number of hit storeys
- Time from first to last hit: correlated to Number of all hits

For the output of the CalcShape algorithm the two compactness parameters were discarded as they were anti-correlated to the number of modules in the largest cluster and the charge in the largest cluster. The Simple Estimator parameter set was also reduced in the testing.

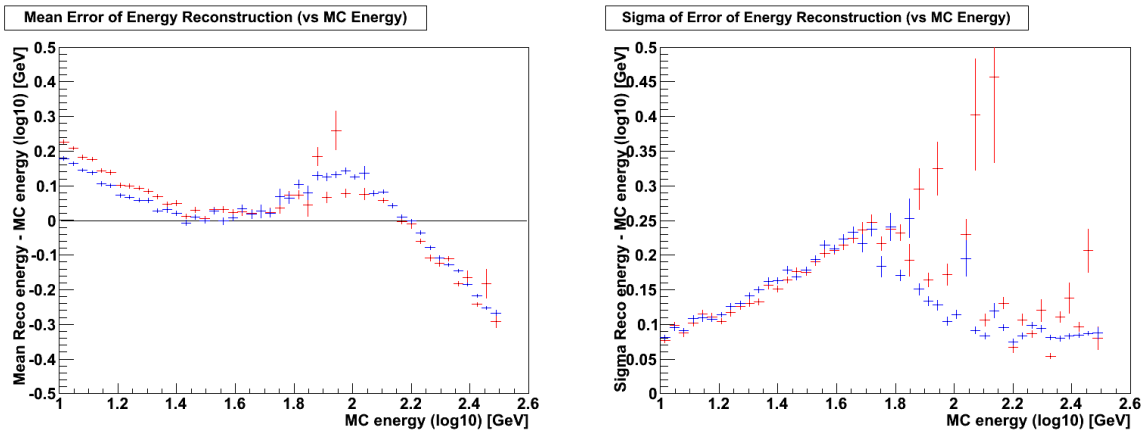
The output parameters of the CalcTensor algorithm were not correlated to Simple Estimator parameters but had some correlations between themselves. To reduce them, the following parameters were discarded:

- Number of OMs: correlated to Number of hits the classifier used
- Normalized sum of non-diagonal elements of quadrupole tensor weighted by charge: correlated to Sum of non-diagonal elements of quadrupole tensor weighted by charge

- Normalized sum of non-diagonal elements of the unweighted quadrupole tensor: correlated to Sum of non-diagonal elements of the unweighted quadrupole tensor
- Eigenvalues (EV) of tensor of inertia:
- $EV_2^{unweighted}$: correlated to $EV_1^{unweighted}$
- $EV_3^{unweighted}$: correlated to Number of hits the classifier used
- $EV_2^{unweighted,normalized}$: correlated to $EV_1^{unweighted,normalized}$
- $EV_2^{amplitude}$: correlated to $EV_1^{amplitude}$
- $EV_2^{amplitude,normalized}$: correlated to $EV_1^{amplitude,normalized}$
- Calculated zenith angle: anti-correlated to the y position of the unweighted COG
- Quadrupole moment along track, unweighted: correlated to $EV_1^{amplitude}$
- Quadrupole moment along track, unweighted, normalized: correlated to $EV_1^{amplitude,normalized}$
- Quadrupole moment along track, weighted by charge: correlated to $EV_1^{unweighted}$

7.7 Results of test with reduced number of parameters

The results of the tests with discarded parameters are displayed in figures 7.6a to 7.9a.

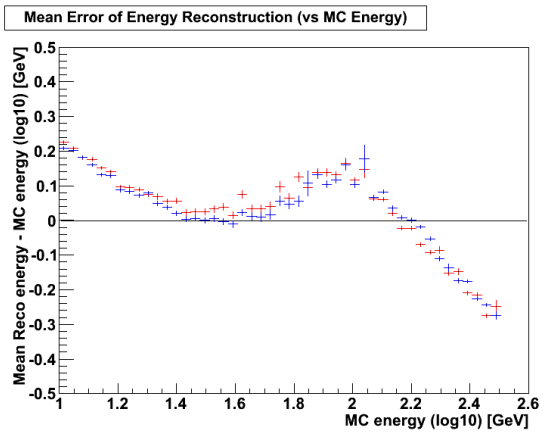


(a) Mean value

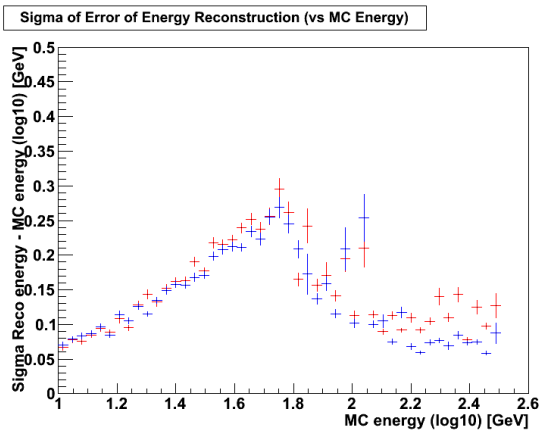
(b) Sigma value

Figure 7.6: Comparison of mean and sigma values from Gaussian fit through difference of reconstructed energy and Monte-Carlo energy over Monte-Carlo energy for the parameter sets extracted with Calc-Cluster. The values for the set with all parameters included are displayed in blue, the values for the test after parameter discarding are displayed in red.

Despite of some small changes (< 0.02 for sigma and mean values in regions where the Gaussian fit seems to be accurate) , there is no real difference in the reconstruction accuracy between tests with reduced number of parameters and tests with the full parameter sets. This strongly suggests that the inaccuracy seen in the tests before does not arise from redundancies in the parameter sets.

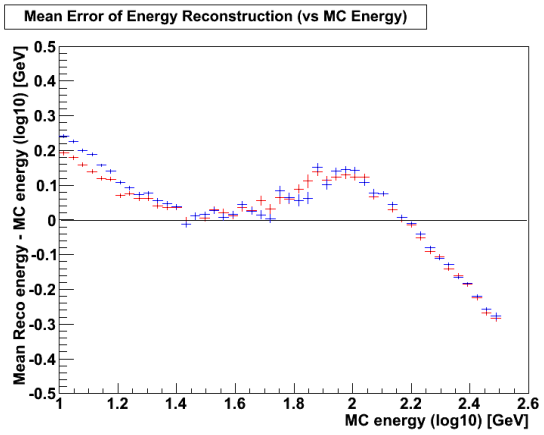


(a) Mean value

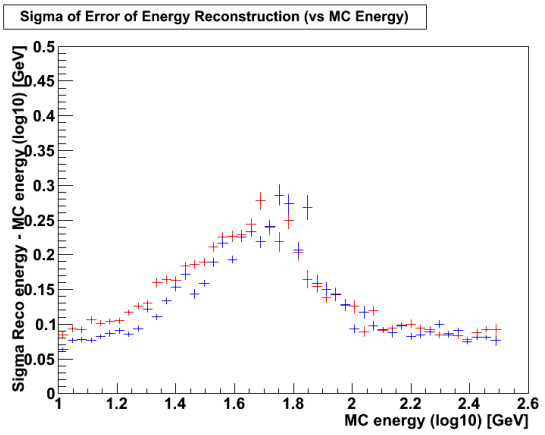


(b) Sigma value

Figure 7.7: Comparison of mean and sigma values from Gaussian fit through difference of reconstructed energy and Monte-Carlo energy over Monte-Carlo energy for the parameter sets extracted with CalcDensity. The values for the set with all parameters included are displayed in blue, the values for the test after parameter discarding are displayed in red.

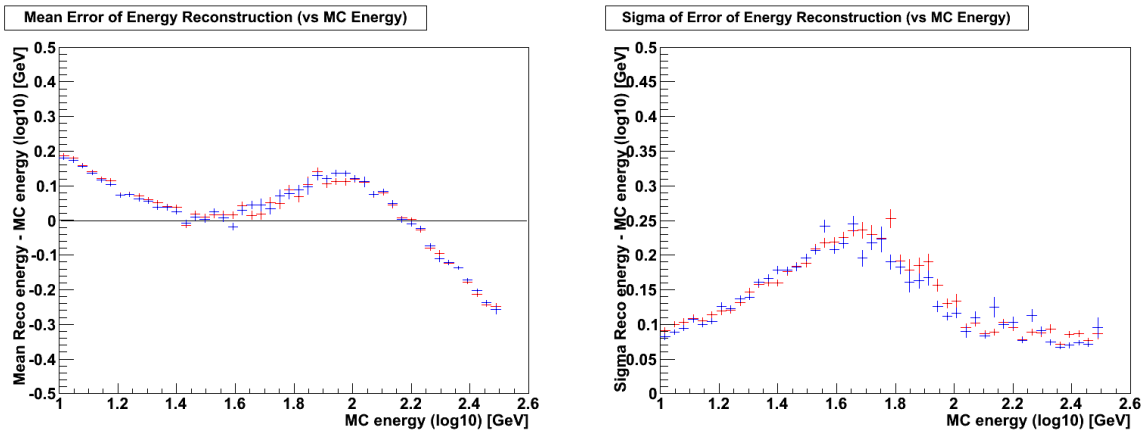


(a) Mean value



(b) Sigma value

Figure 7.8: Comparison of mean and sigma values from Gaussian fit through difference of reconstructed energy and Monte-Carlo energy over Monte-Carlo energy for the parameter sets extracted with CalcTensor. The values for the set with all parameters included are displayed in blue, the values for the test after parameter discarding are displayed in red.



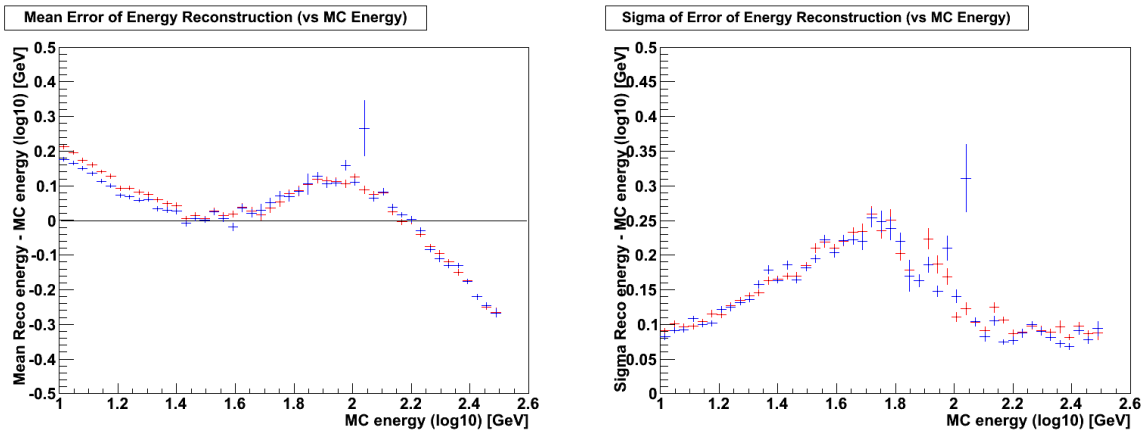
(a) Mean value

(b) Sigma value

Figure 7.9: Comparison of mean and sigma values from Gaussian fit through difference of reconstructed energy and Monte-Carlo energy over Monte-Carlo energy for the parameter sets extracted with CalcShape. The values for the set with all parameters included are displayed in blue, the values for the test after parameter discarding are displayed in red.

7.8 Combination of parameter extraction algorithms

To get the advantage of all parameter extraction algorithms, it was tried to create a "best of" net. To do so, the parameters of CalcShape, CalcTensor, CalcCluster and CalcDensity were put together. The parameters which were already discarded above were not included to have a reasonable size of the parameter set. As the covariance matrix showed almost no strong correlations, it was decided not to exclude any more parameters.



(a) Mean value

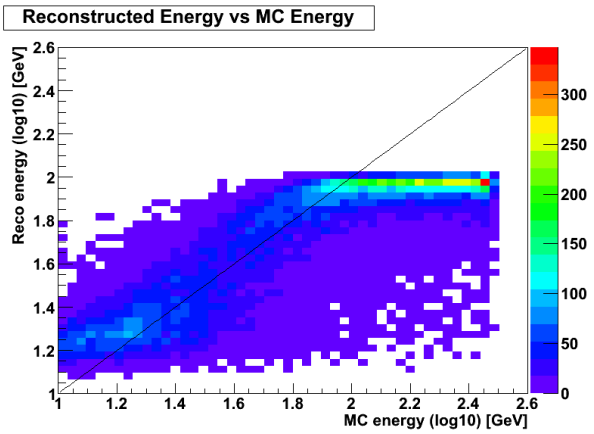
(b) Sigma value

Figure 7.10: Comparison of mean and sigma values from Gaussian fit through difference of reconstructed energy and Monte-Carlo energy over Monte-Carlo energy for the combined parameter sets of all four extraction algorithms (red). For comparison, the same values are plotted with blue markers for the Simple Estimator.

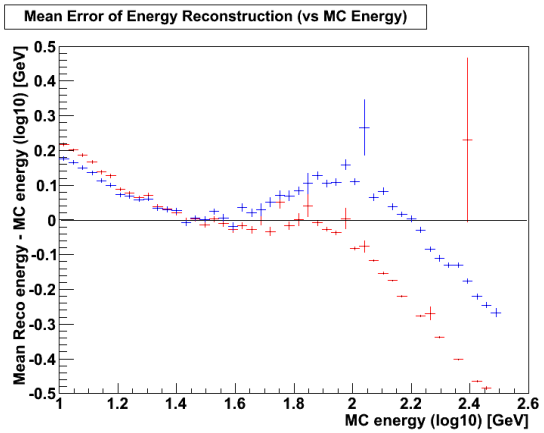
Unfortunately even the combined parameters show no improvement against the initial simple estimator parameters. It should be again mentioned here that for energies above 100 GeV this was expected due to the reasons highlighted in chapter 4. For energies below 100 GeV, the result again seems to indicate that the parameters of the Simple Estimator are representing the low-energy energy loss functionality in the detector with sufficient accuracy.

7.9 Tests with reduced energy range of training set

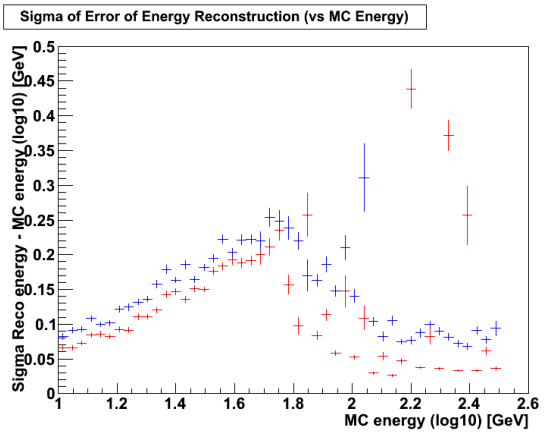
All previous tests show an overestimation of the muon energy in the range of a Monte-Carlo energy below 100 GeV while energies above this energy are underestimated. As the tests confirm the assumption which was made in chapter 4, that it is not possible to reconstruct energies if their corresponding track length is larger than the dimensions of the instrumented volume, and due to the fact that the training process for an ANN tries to minimize the overall error $\mathcal{E}^N_{MSE}$ for the whole training set, it seems to be a good choice to reduce the maximal energy in the training sample to 100 GeV. This was done for the Simple Estimator and the results are shown in figures 7.11a to 7.11c. The test of the reconstruction was done with the usual testing sample.



(a) Plot of reconstructed energy over MC energy



(b) Mean value



(c) Sigma value

Figure 7.11: Evaluation of reconstruction quality for the Simple Estimator with a reduced energy range for the training set. Gaussian fit through the error of the reconstruction along the Monte-Carlo energy. Values for the Simple Estimator with original training set are displayed in blue for reference

From figures 7.11a to 7.11c it can be seen that the overestimation of the energy in the range from 25 to 100 GeV indeed vanishes, but this is not the case for energies below 25 GeV. All energies above 100 GeV are reconstructed at about 100 GeV, which was expected as the net was not trained for energies above 100 GeV. For the range below 25 GeV there are probably too few hits in the detector for the track length to be estimated.

Chapter 8

Summary and Outlook

In this work an energy reconstruction for muons with energies in the low energy range (below $E_{\mu C} \approx 780 GeV$) was developed using Artificial Neural Networks. As it had been shown in chapter 4 a, successful energy reconstruction in the energy regime of nearly constant ionisation losses can only be based on parameters which are dependent on the total length of the muon track. The most important restriction for the reconstruction of muon energy in the low energy regime is given by the fact that the instrumented volume of the detector has to be large enough for the track to end within it. This restricts the maximum energy of this method to about 100 GeV for the ANTARES detector.

The development of the reconstruction was based on the work already done in [24]. The pre-existing Simple Estimator proved already to deliver a sufficient energy reconstruction with errors comparable to those of the high energy regime in [24]. Various tests were performed to further improve the quality of the energy reconstruction in the low energy regime.

The application of a hit selection with higher purity than the afore used to the input hits of the energy reconstruction however failed to improve the results of the reconstruction. The conclusion was made that the hit selection performed with the application of the trigger criterion is already suitable for the use within the Simple estimator. Therefore it was aimed for an enhancement of the quality of the simple estimator by enlarging its parameter set. Parameters that were extracted from raw data by four separate feature extraction algorithms, and which had been developed for oscillation analysis, were tested for their capability to add new features to the ANN data set. The result of these tests was such that they in fact do not improve the quality of the energy reconstruction significantly. As the four extraction algorithms are based on different approaches where one uses causally connected hits, another calculates parameters from the density of hits and charge in detector sub-volumes and time-slices, yet another calculates the event's tensor of inertia and the last one writes out parameters describing the shape of the illuminated volume in the detector, the conclusion was made that the parameters which the Simple Estimator provides are already describing the energy loss mechanism and the detector's response for low energies very well. To address possible issues which could cause poorer reconstruction quality if the parameter set gets too large, the input parameter set was analysed for redundancies and parameters which were weakly used in ANN training. These parameters were then discarded. Unfortunately even the discarding of redundant parameters did not provide a reconstruction of better quality than that for the Simple Estimator alone. The test of the usual track-related energy reconstruction for high energies together with a low-energy track fitter was also performed, but with little success as reconstruction quality significantly worsens in comparison to the previous tests.

A way to address the mean offset of the reconstruction in future work would be to reduce the maximal energy in the training sample to 100 GeV, as this is roughly the maximum energy for which the track is contained in the instrumented volume of ANTARES. As the training process evaluates the error for the whole training sample, exclusion of all energies above that value, can serve to enhance reconstruction quality for the whole range of the net.

The maximum accuracy of the energy reconstruction achieved in the low-energy range was calculated by a Gaussian fitt in the distribution of the logarithmic reconstruction error, which yielded a mean offset of maximally 0.2 and a sigma of maximally 0.25 for the range 10 GeV to 100 GeV. This corresponds to

factors of 1.5 resp. 1.8 in non logarithmic values. This is quite good compared to factors of 3 usually achieved with ANNs for energy reconstruction in the high energy range.

The next step that should be followed after this work is to test the low-energy estimator against mis-reconstruction of events with higher energies. Also the response to different detector conditions such as higher background rates or defective OMs has to be evaluated. It is still possible that these tests would yields a different relevance for reconstruction quality of one or more of the tested parameters, which were behaving similarly in ANNs trained with low background rates.

Bibliography

[1] http://antares.in2p3.fr/overview/detector.html.

[2] http://de.wikipedia.org/wiki/neutrino.

[3] http://de.wikipedia.org/wiki/supernova 1987a.

[4] http://de.wikipedia.org/w/index.php?title=datei:complete_neuron_cell_diagram_de.svg &filetimestamp=20080501125630.

[5] http://pdg.lbl.gov/2005/reviews/contents_sports.html#expmethetc.

[6] personal correspondence Frederike Schoek.

[7] The antares telescope reveals unexpected intense and variable deep-sea bioluminescence. ANTARES Internal Note, 2007. ANTARES-Draft/2007-001.

[8] A. A. Abdo and Fermi LAT Collaboration. Observations of the young supernova remnant rx j1713.7-3946 with the fermi large area telescope. *ApJ, 734, 28, 2011*, 2011.

[9] David J.l. Baily. Genhen v5r1: Software documentation, 2002.

[10] David J.l. Baily. Km3 v2r1: User guide, 2002.

[11] H. Bichsel, D. E. Groom, and S. R. Klein. Passage of particles through matter. *Journal of Physics G Nuclear and Particle Physics*, 2010.

[12] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1994.

[13] J Carr, S Escoffier, and D Zaborov. Proposition for an alternative trigger based on the t3 cluster trigger. ANTARES Internal Note, 2007. ANTARES-SOFT-2007-016.

[14] T. Chiarusi and M. Spurio. High-energy astrophysics with neutrino telescopes. *The European Physics Journal*, 2010.

[15] C. Curtil. Median rate of measured single photon counts for the milom, line 1 and il07 since july 2005 until june 2009.

[16] M De Jong. The antares trigger software. ANTARES Internal Note, 2005. ANTARES-SOFT-2005-005.

[17] Florian Gorges. Astronomie mit Teilchen.

[18] John David Jackson. *Klassische Elektrodynamik*. Number ISBN 3-11-018970-4. Walter de Gruyter, Berlin, 2006.

[19] S. Klimushin, E. Bugaev, and I. Sokalski. Precise parametrizations of muon energy losses in water. In *Proceedings of ICRC 2001: Copernicus Gesellschaft*, 2001.

[20] Herv é lafoux. K40 simulation, 1999. ANTARES-Opmo/1999-002.

[21] Holger Martin Motz. *Dark Matter Search with ANTARES*. PhD thesis, Friedrich Alexander Universität Erlangen Nürnberg, 2011.

[22] Steffen Nissen. Implementation of a fast artificial neural network library (fann), 2003.

[23] Elisa Resconi. Neutrino astronomy (summer semester 2011, Universität Erlangen-Nürnberg).

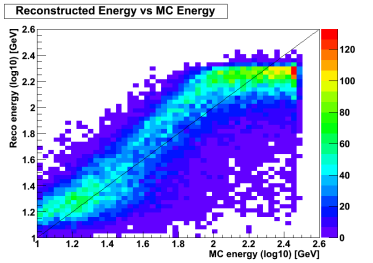
[24] Jutta Schnabel. Muon energy reconstruction for the neutrino telescope antares using neural networks. Diplomarbeit, Friedrich Alexander Universität Erlangen Nürnberg, October 2010.

[25] Frederike Schoek. *Studies on atmospheric neutrino oscillations with the ANTARES neutrino Telescope*. PhD thesis, Friedrich Alexander Universität Erlangen Nürnberg, 2011.

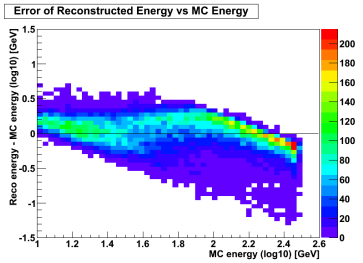
[26] Guus Wijnker. NNerG energy reconstruction using neural networks.

Additional Comparison Plots

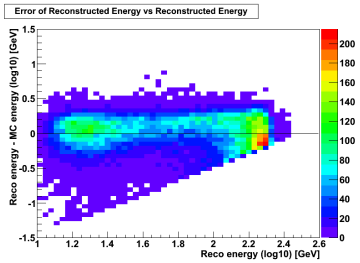
Simple Estimator



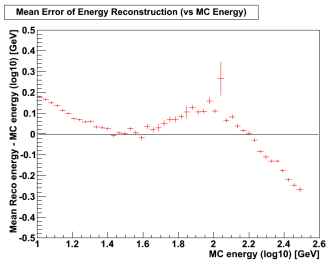
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



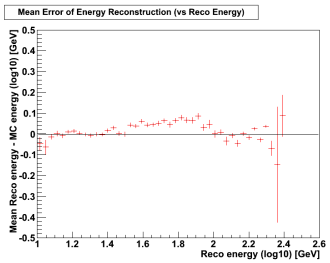
(b) $E_{rec} - E_{MC}$ over E_{MC}



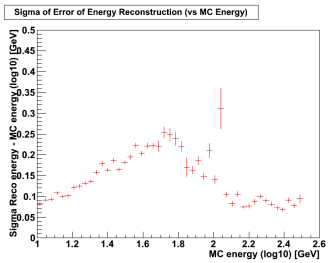
(c) $E_{rec} - E_{MC}$ over E_{rec}



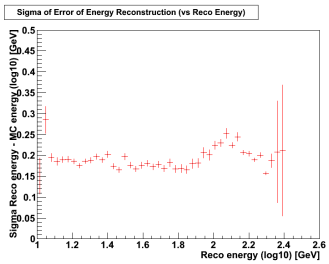
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



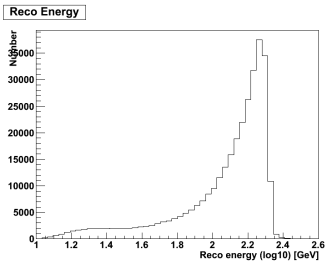
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



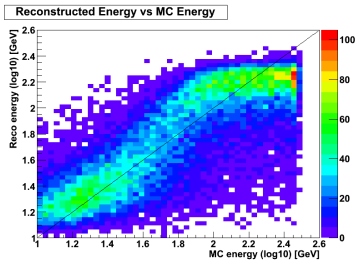
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



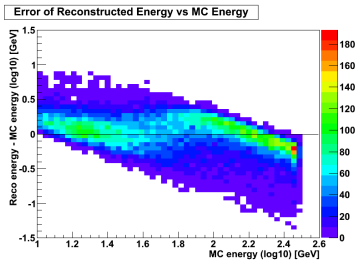
(b) Distribution of E_{rec}

Simple Estimator with Cluster Hit selection

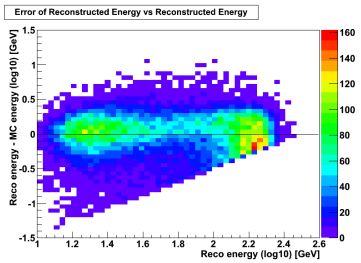
Blue markers denote Simple estimator values as reference.



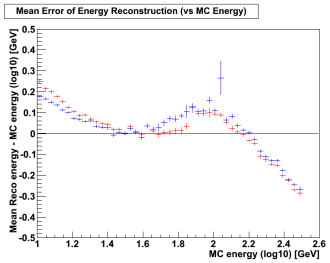
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



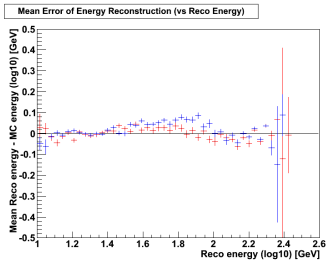
(b) $E_{rec} - E_{MC}$ over E_{MC}



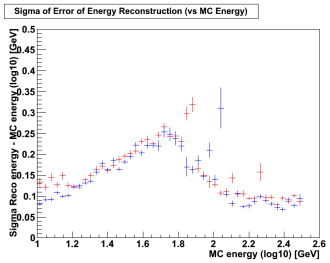
(c) $E_{rec} - E_{MC}$ over E_{rec}



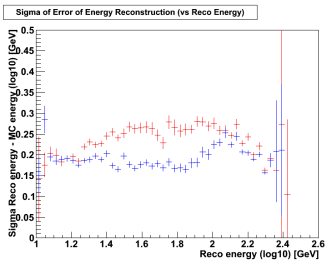
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



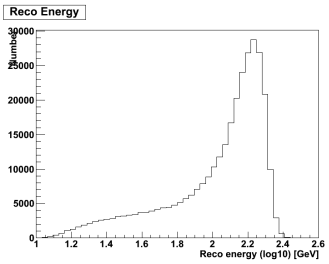
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



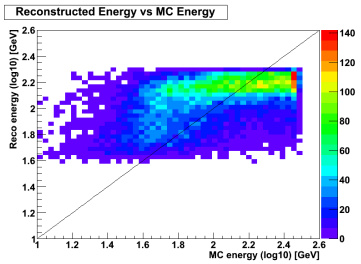
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



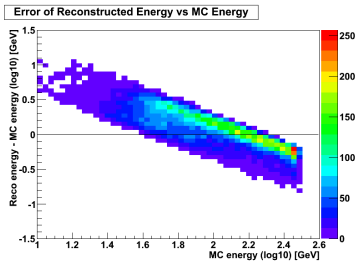
(b) Distribution of E_{rec}

Cherenkov Estimator with Posidonia (3D)

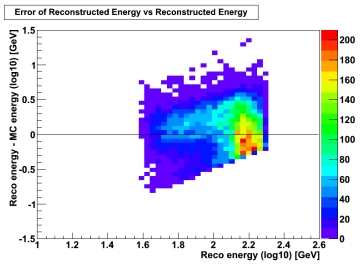
Blue markers denote Simple estimator values as reference.



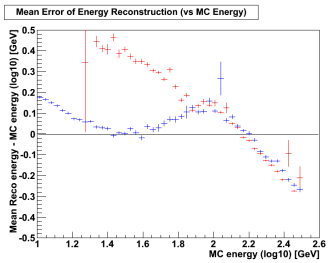
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



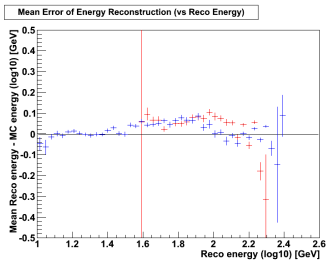
(b) $E_{rec} - E_{MC}$ over E_{MC}



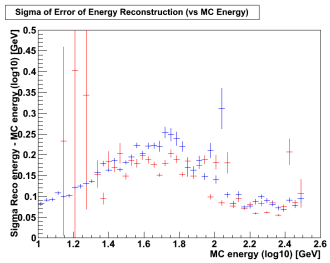
(c) $E_{rec} - E_{MC}$ over E_{rec}



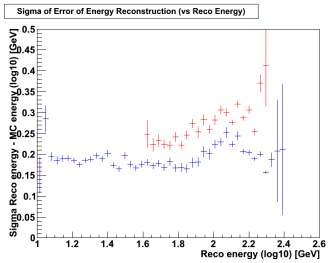
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



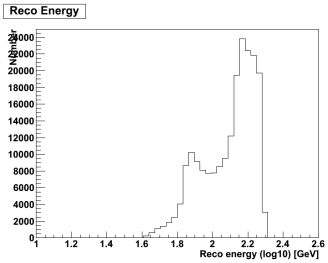
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



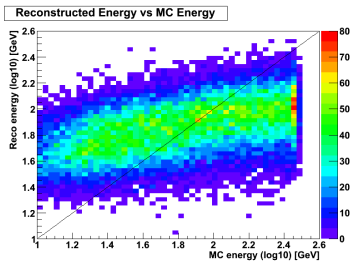
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



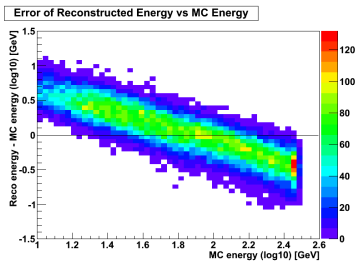
(b) Distribution of E_{rec}

Cherenkov Estimator with Posidonia (1D)

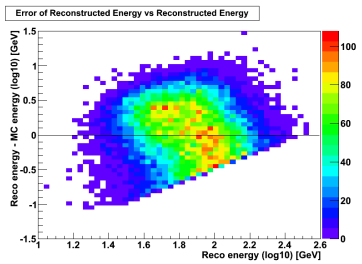
Blue markers denote Simple estimator values as reference.



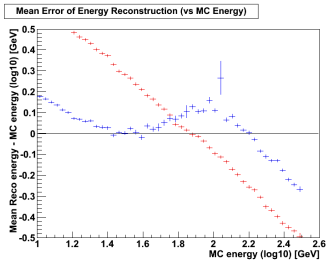
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



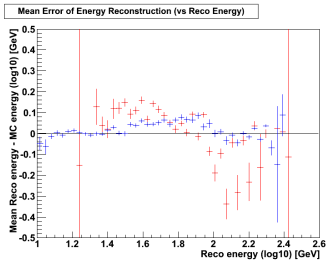
(b) $E_{rec} - E_{MC}$ over E_{MC}



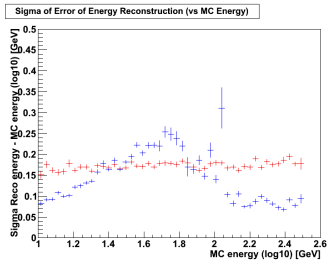
(c) $E_{rec} - E_{MC}$ over E_{rec}



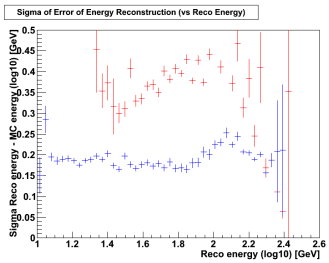
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



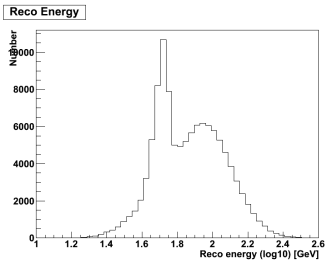
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



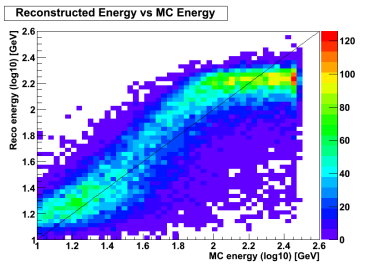
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



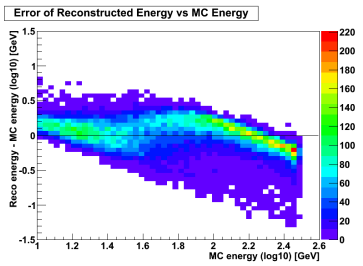
(b) Distribution of E_{rec}

CalcCluster

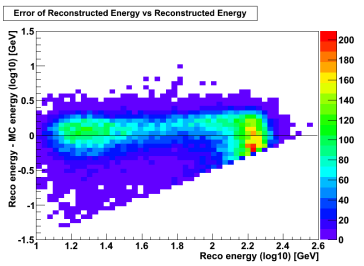
Blue markers denote Simple estimator values as reference.



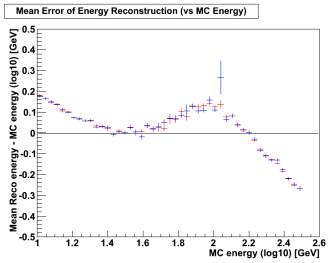
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



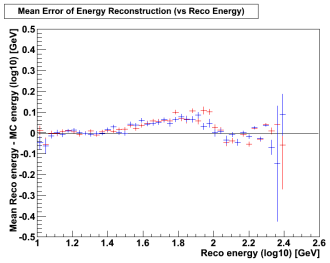
(b) $E_{rec} - E_{MC}$ over E_{MC}



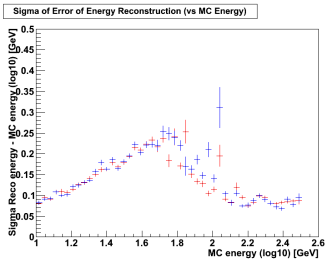
(c) $E_{rec} - E_{MC}$ over E_{rec}



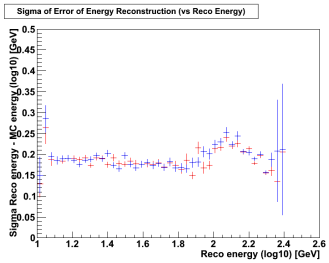
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



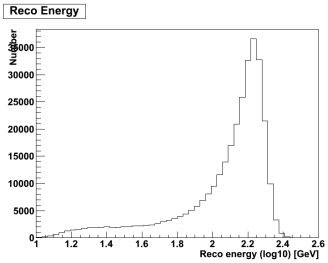
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



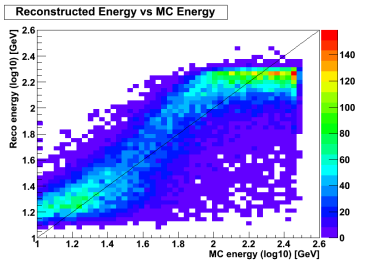
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



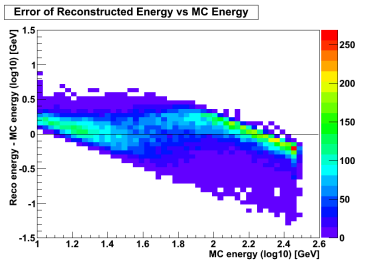
(b) Distribution of E_{rec}

CalcDensity

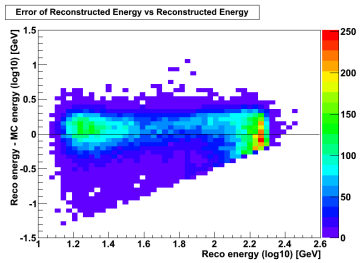
Blue markers denote Simple estimator values as reference.



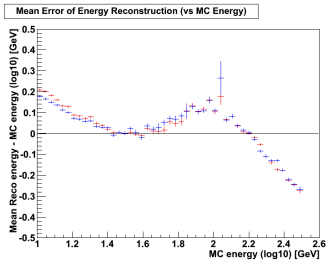
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



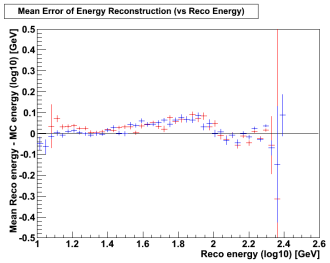
(b) $E_{rec} - E_{MC}$ over E_{MC}



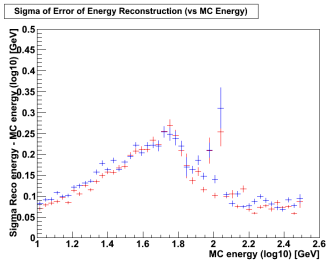
(c) $E_{rec} - E_{MC}$ over E_{rec}



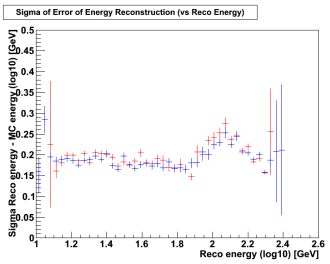
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



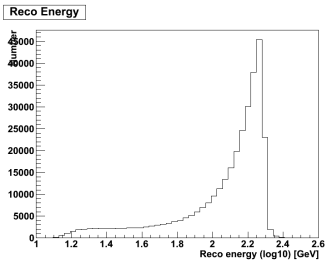
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



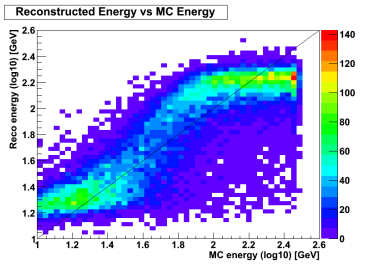
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



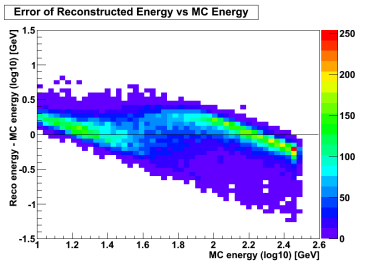
(b) Distribution of E_{rec}

CalcTensor

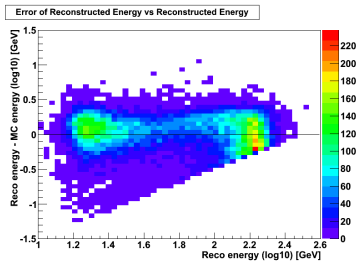
Blue markers denote Simple estimator values as reference.



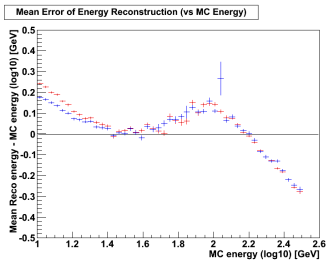
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



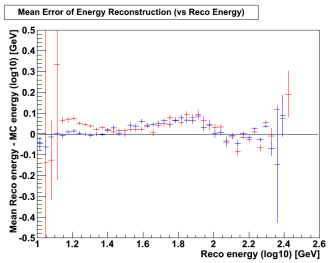
(b) $E_{rec} - E_{MC}$ over E_{MC}



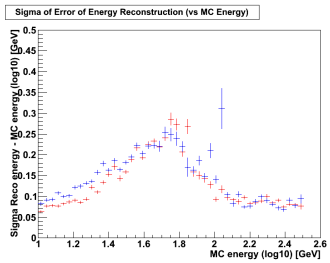
(c) $E_{rec} - E_{MC}$ over E_{rec}



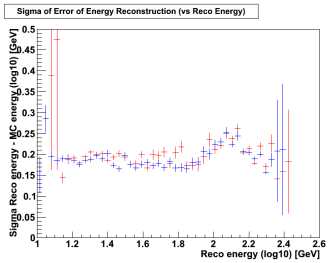
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



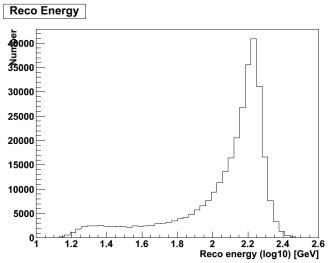
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



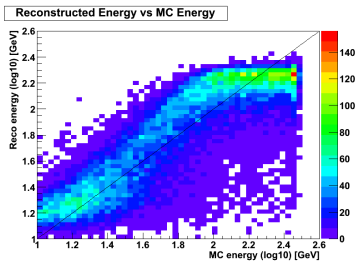
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



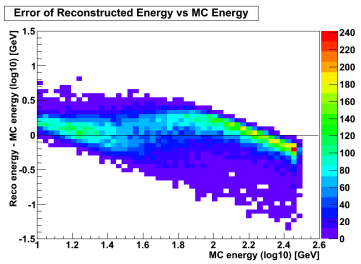
(b) Distribution of E_{rec}

CalcShape

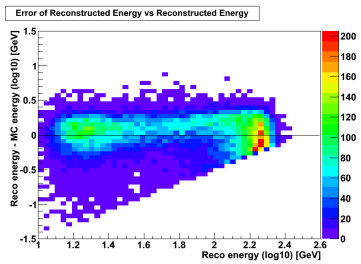
Blue markers denote Simple estimator values as reference.



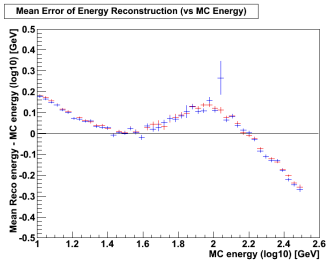
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



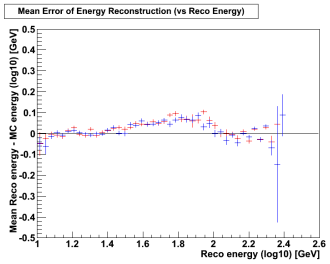
(b) $E_{rec} - E_{MC}$ over E_{MC}



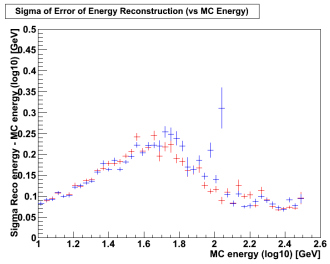
(c) $E_{rec} - E_{MC}$ over E_{rec}



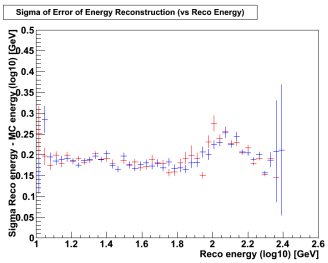
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



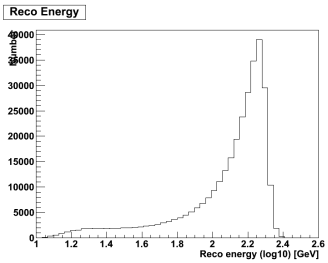
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



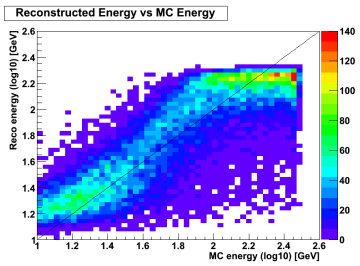
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



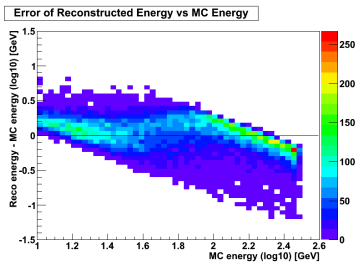
(b) Distribution of E_{rec}

CalcCluster (Reduced number of parameters)

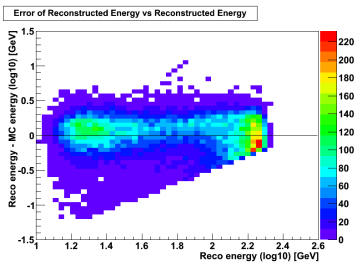
Blue markers denote values with full parameter set as reference.



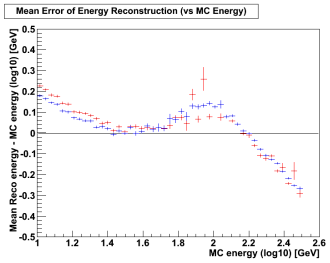
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



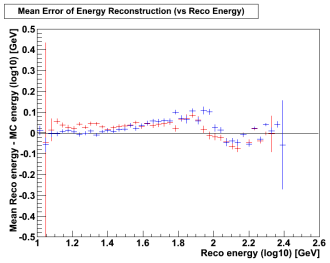
(b) $E_{rec} - E_{MC}$ over E_{MC}



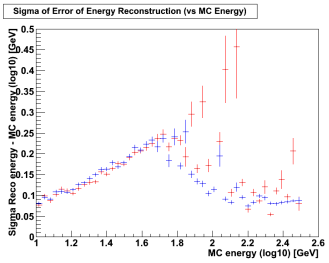
(c) $E_{rec} - E_{MC}$ over E_{rec}



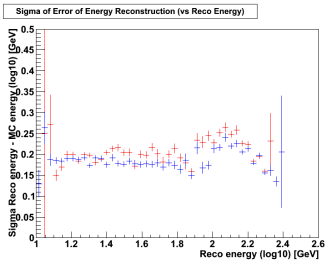
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



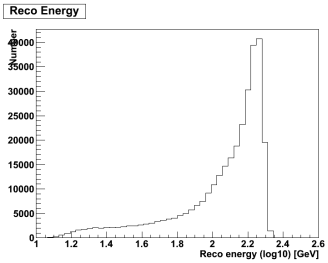
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



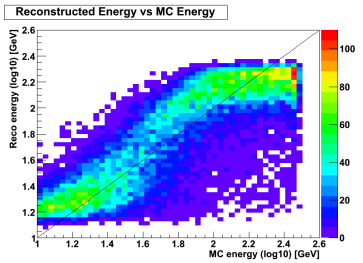
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



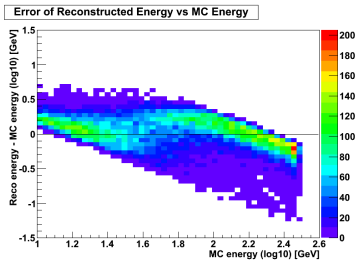
(b) Distribution of E_{rec}

CalcDensity (Reduced number of parameters)

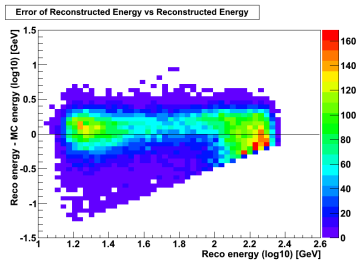
Blue markers denote values with full parameter set as reference.



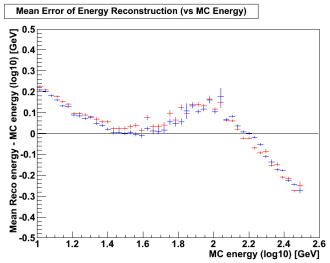
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



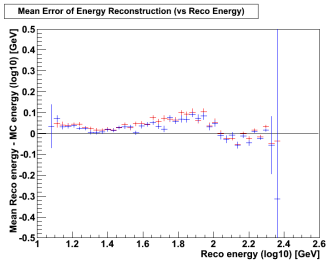
(b) $E_{rec} - E_{MC}$ over E_{MC}



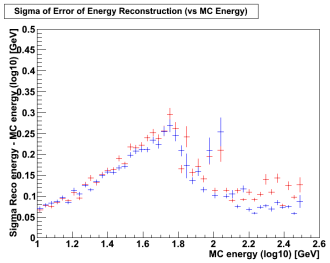
(c) $E_{rec} - E_{MC}$ over E_{rec}



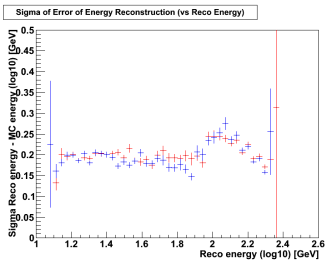
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



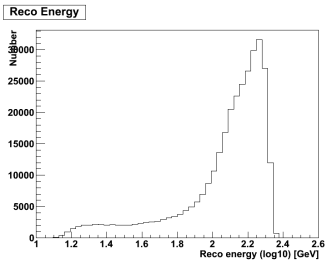
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



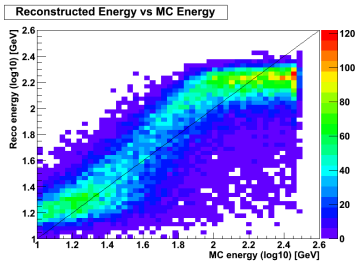
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



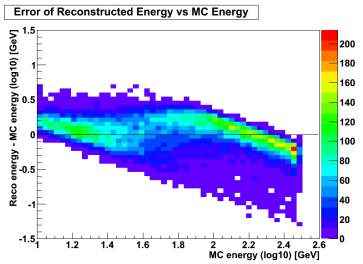
(b) Distribution of E_{rec}

CalcTensor (Reduced number of parameters)

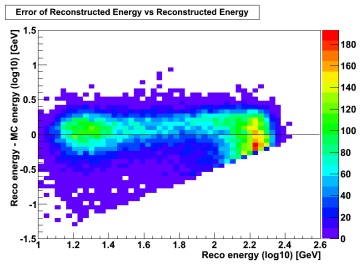
Blue markers denote values with full parameter set as reference.



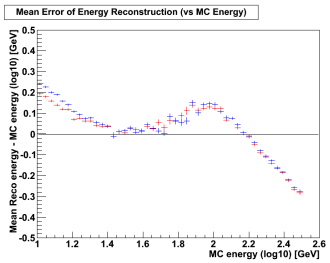
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



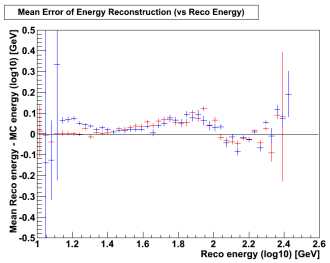
(b) $E_{rec} - E_{MC}$ over E_{MC}



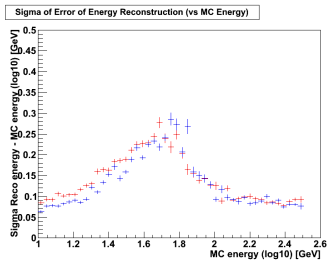
(c) $E_{rec} - E_{MC}$ over E_{rec}



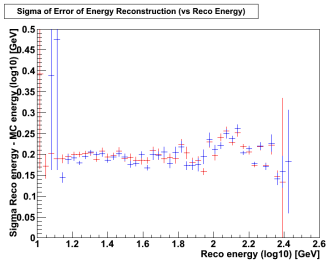
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



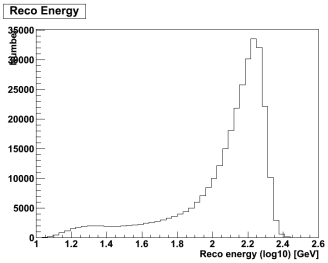
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



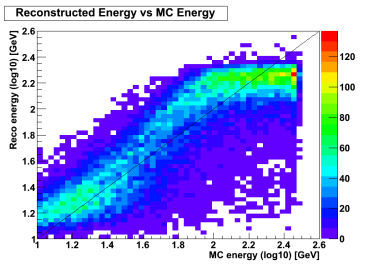
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



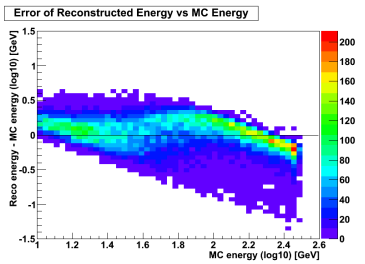
(b) Distribution of E_{rec}

CalcShape (Reduced number of parameters)

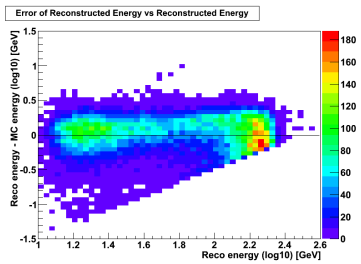
Blue markers denote values with full parameter set as reference.



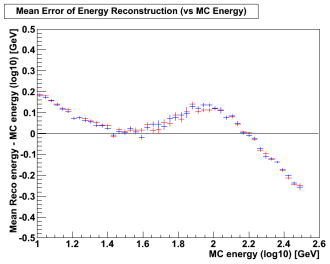
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



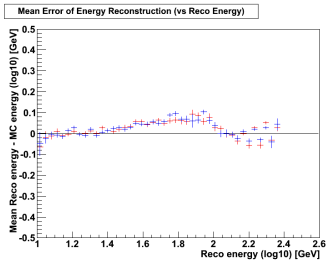
(b) $E_{rec} - E_{MC}$ over E_{MC}



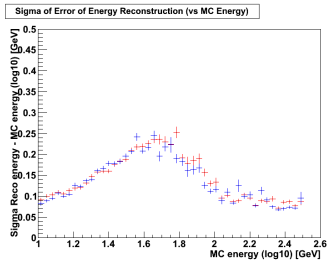
(c) $E_{rec} - E_{MC}$ over E_{rec}



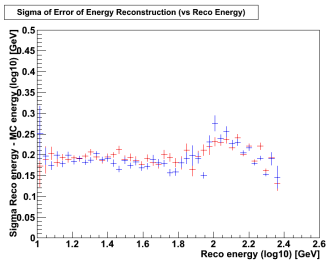
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



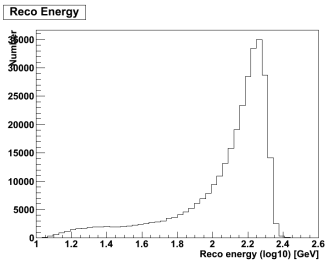
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



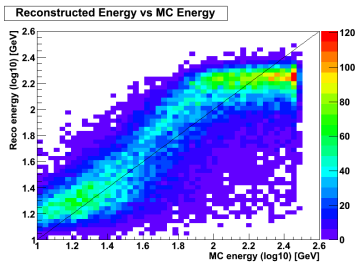
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



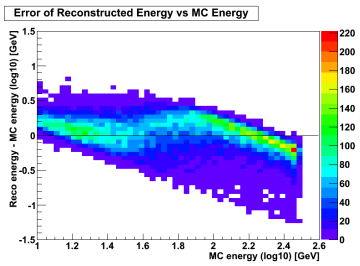
(b) Distribution of E_{rec}

Combination of parameters

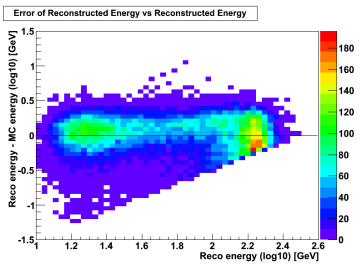
Blue markers Simple Estimator values as reference.



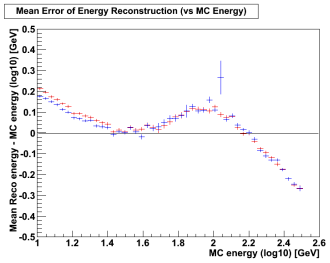
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



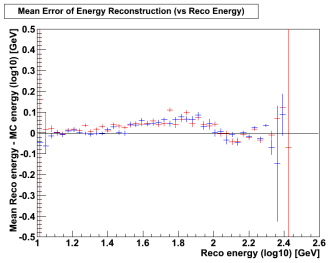
(b) $E_{rec} - E_{MC}$ over E_{MC}



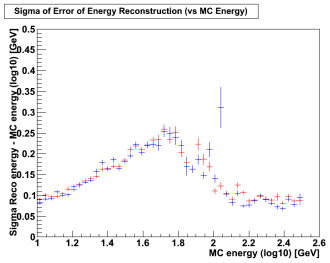
(c) $E_{rec} - E_{MC}$ over E_{rec}



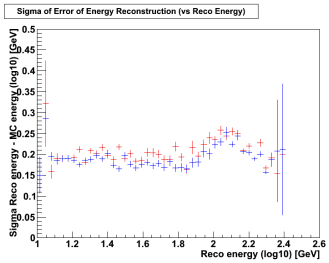
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



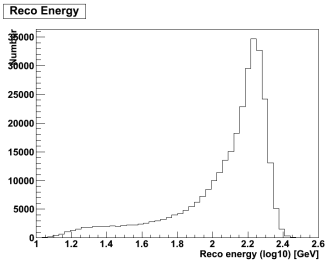
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



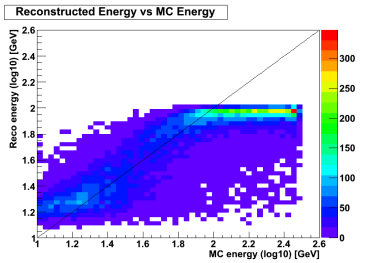
(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



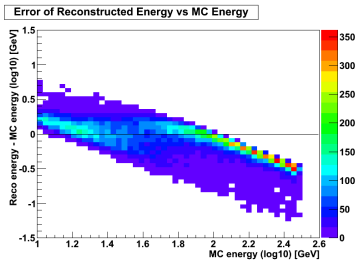
(b) Distribution of E_{rec}

Simple Estimator with reduced range training

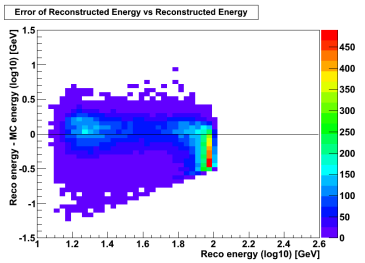
Blue markers Simple Estimator values for full training set as reference.



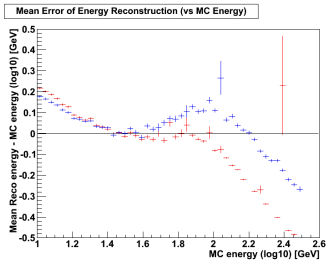
(a) Reconstructed E_{rec} energy over MC energy E_{MC}



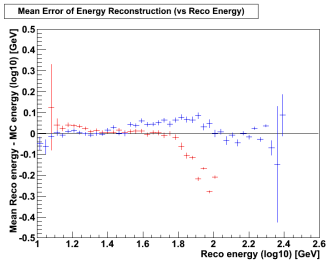
(b) $E_{rec} - E_{MC}$ over E_{MC}



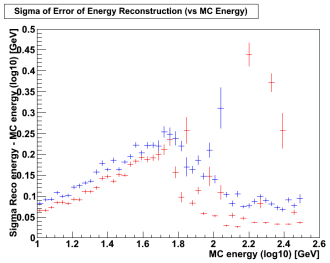
(c) $E_{rec} - E_{MC}$ over E_{rec}



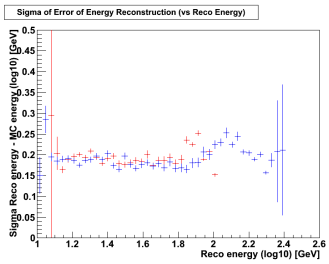
(a) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



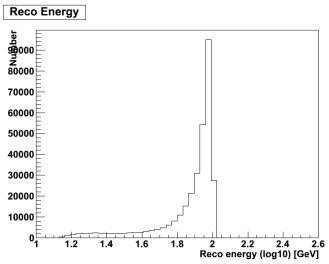
(b) Mean of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(c) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{MC} along E_{MC}



(a) Sigma of Gaussian fit of $E_{rec} - E_{MC}$ over E_{rec} along E_{rec}



(b) Distribution of E_{rec}

Input Parameters

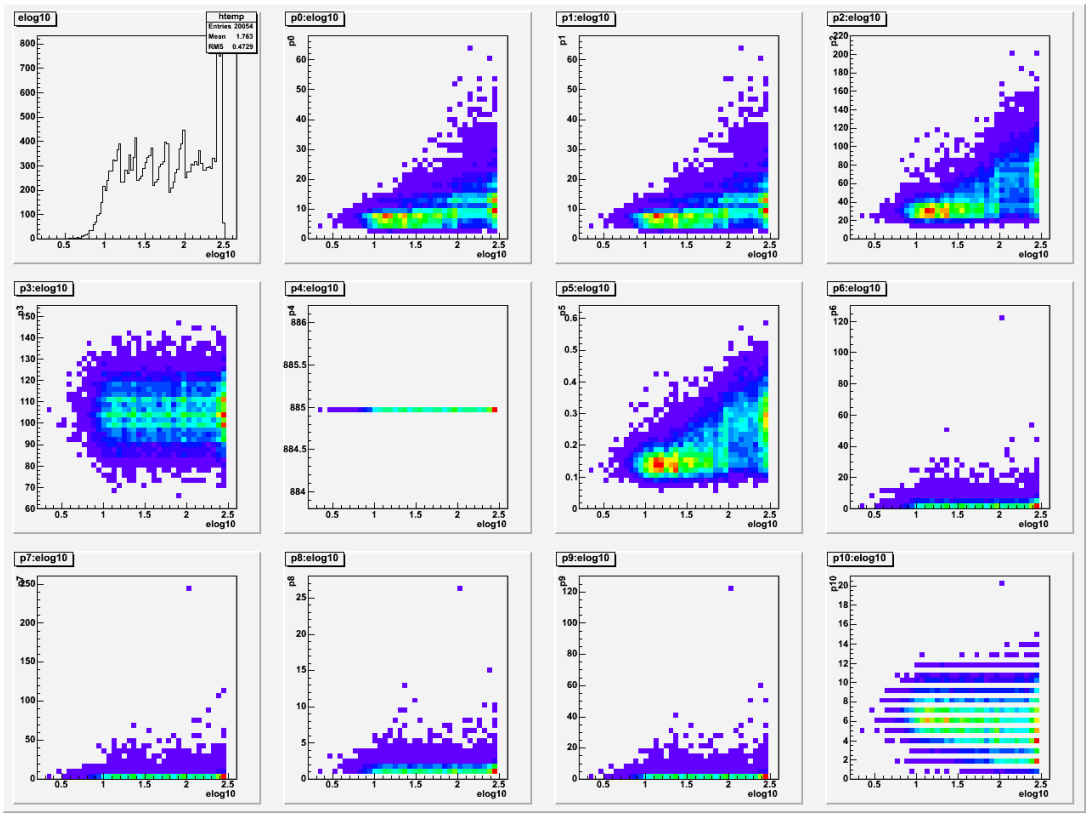


Figure 1: Simple Estimator

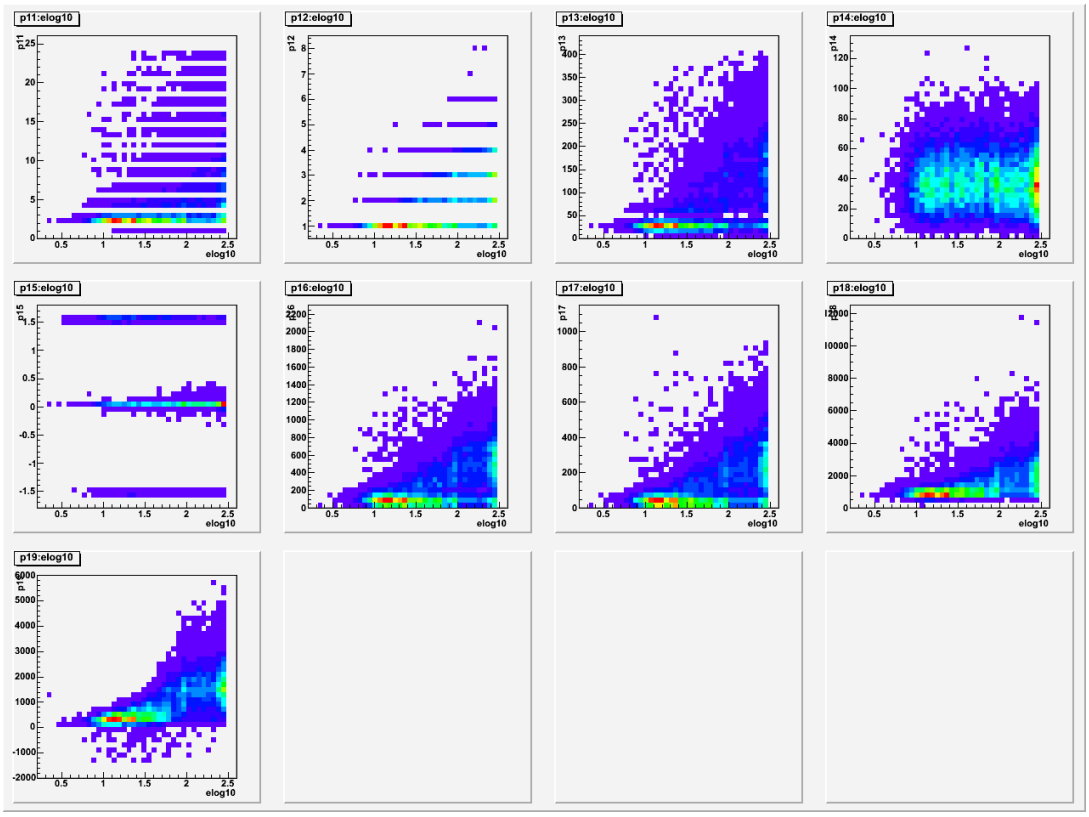


Figure 2: Simple Estimator

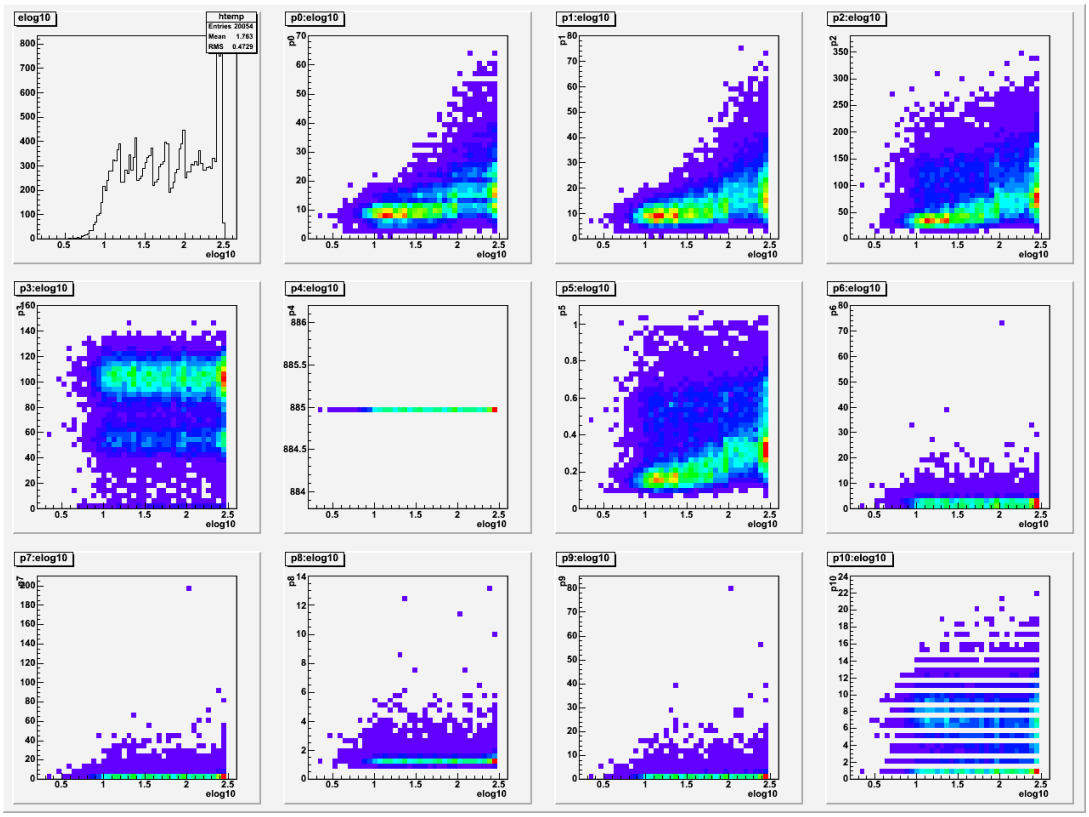


Figure 3: Simple Estimator with Cluster Hit selection

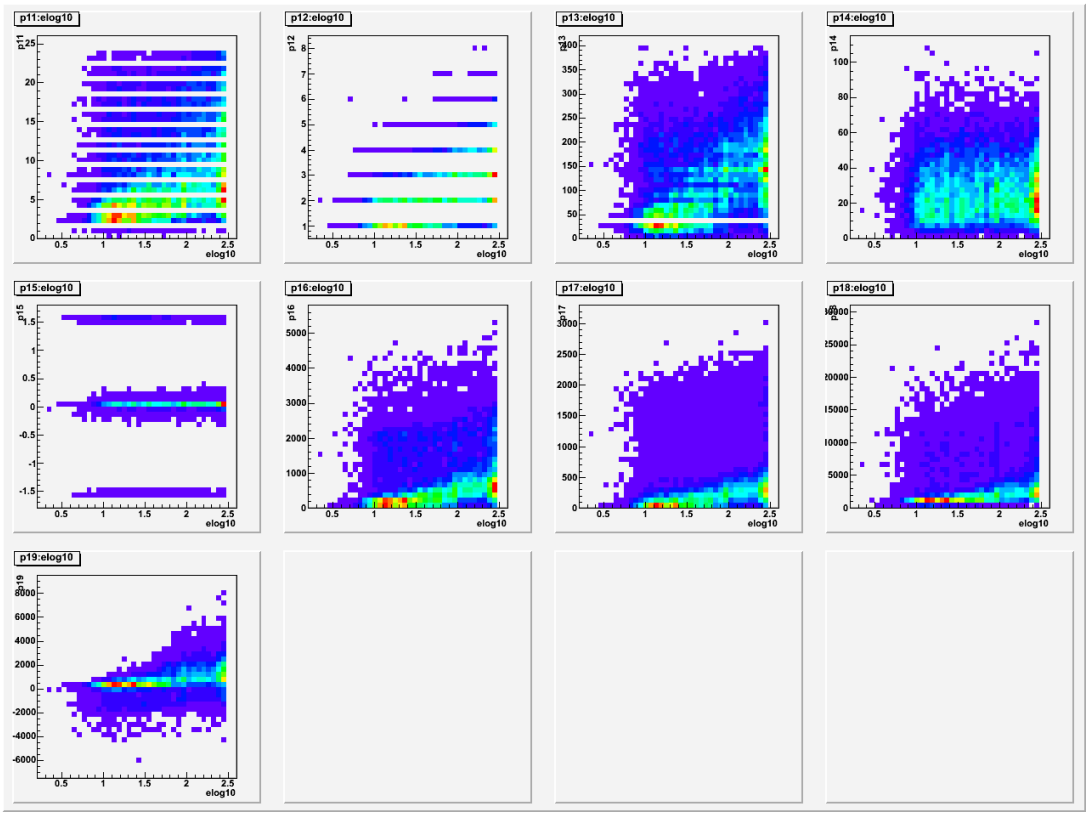


Figure 4: Simple Estimator with Cluster Hit selection

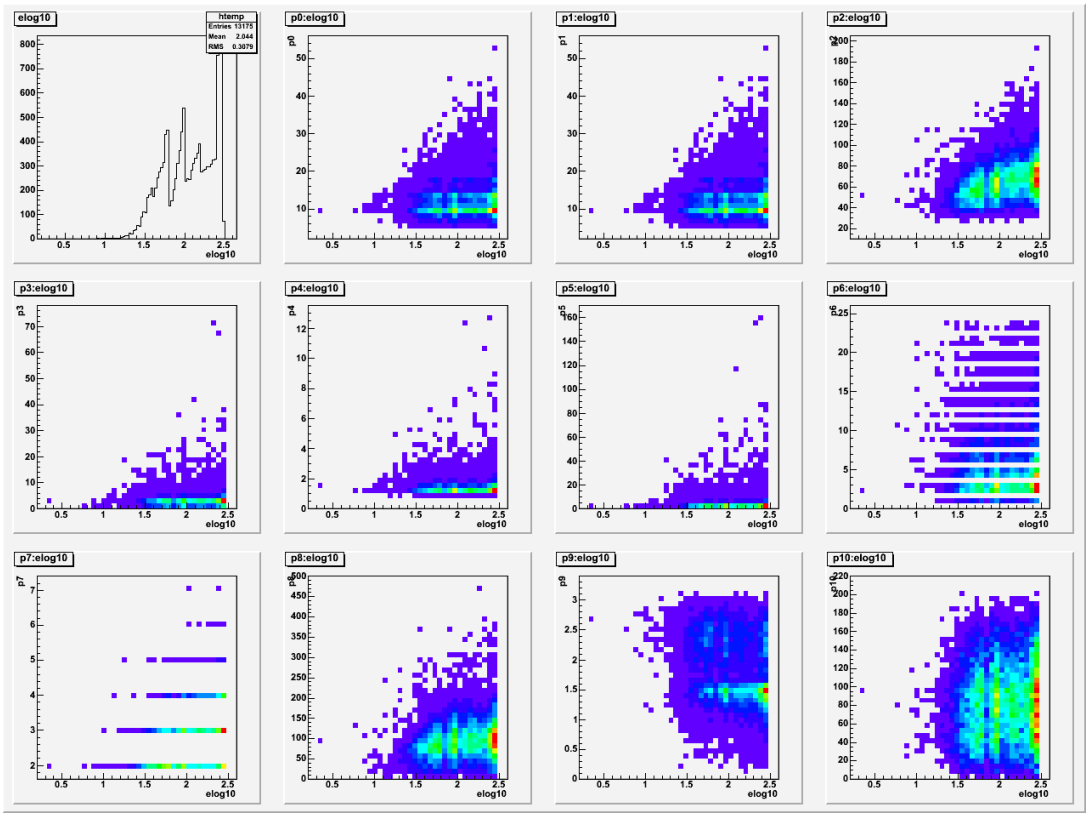


Figure 5: Cherenkov Estimator with Posidonia (3D)

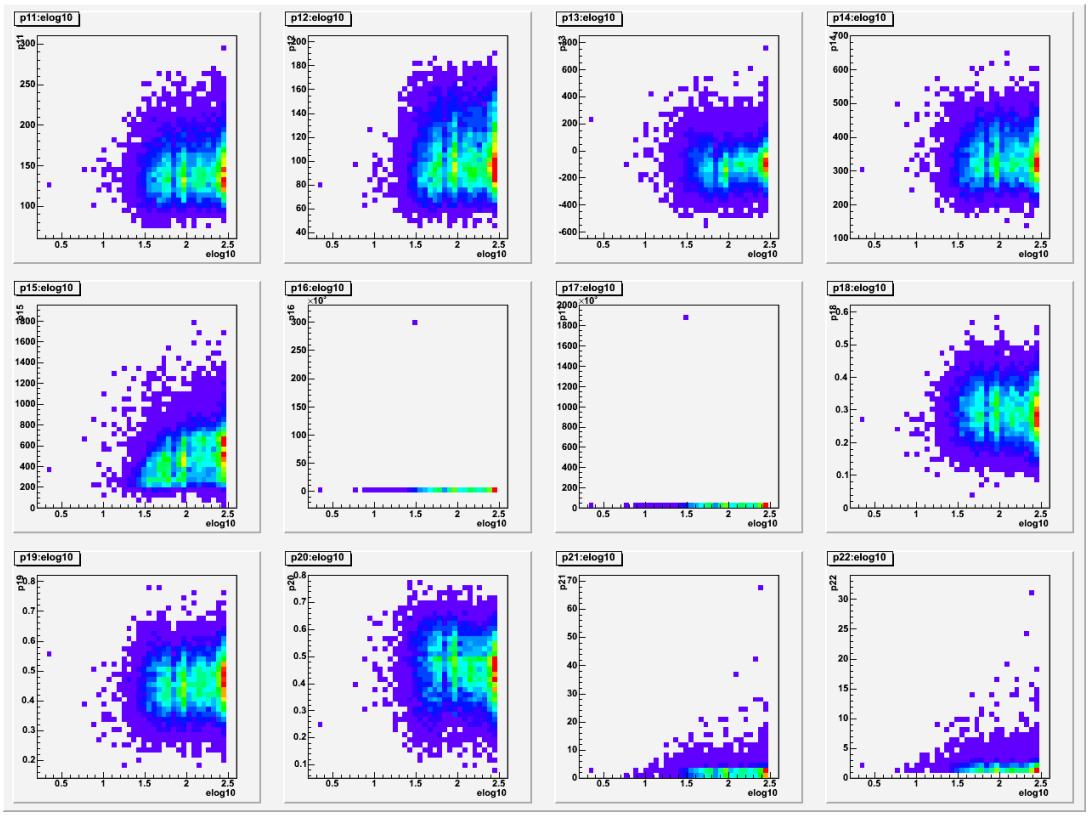


Figure 6: Cherenkov Estimator with Posidonia (3D)

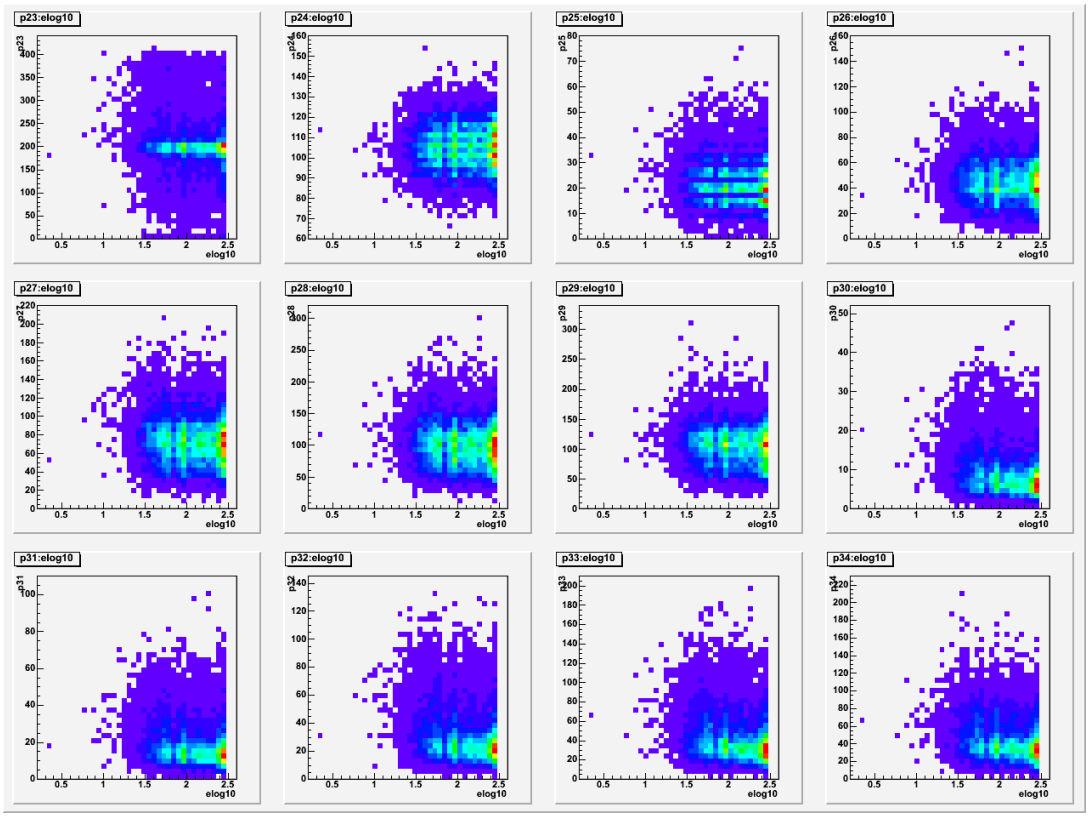


Figure 7: Cherenkov Estimator with Posidonia (3D)

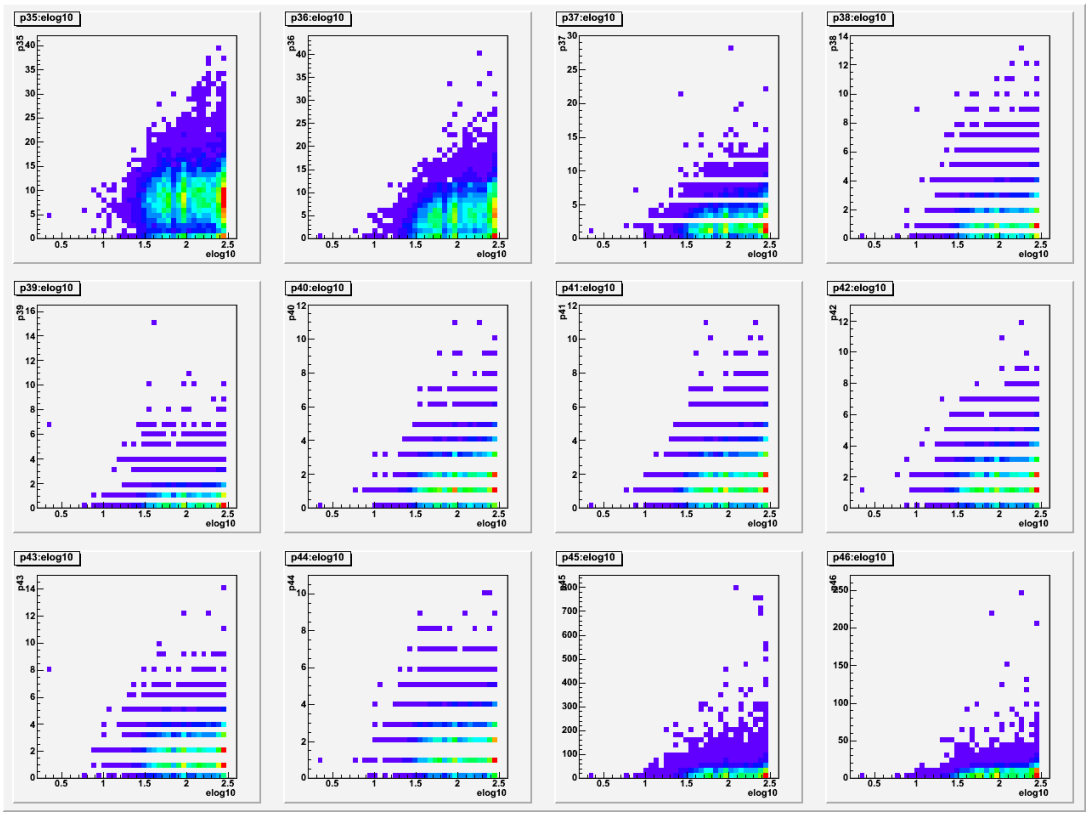


Figure 8: Cherenkov Estimator with Posidonia (3D)

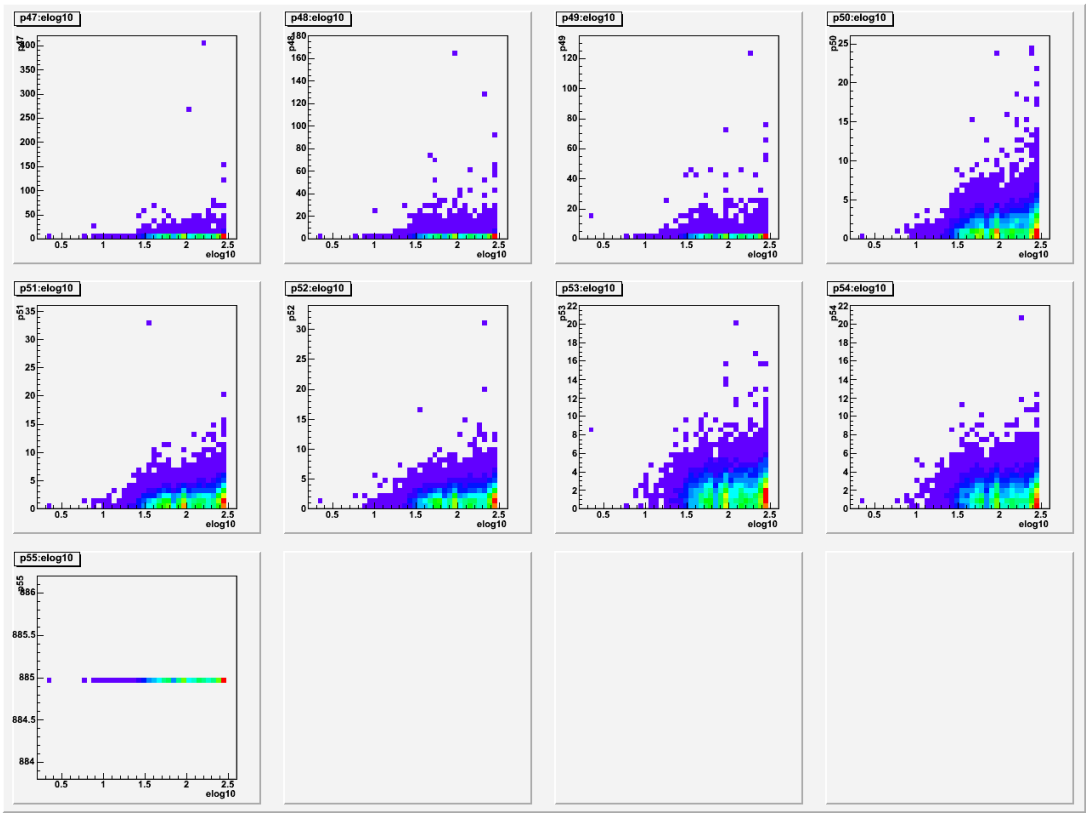


Figure 9: Cherenkov Estimator with Posidonia (3D)

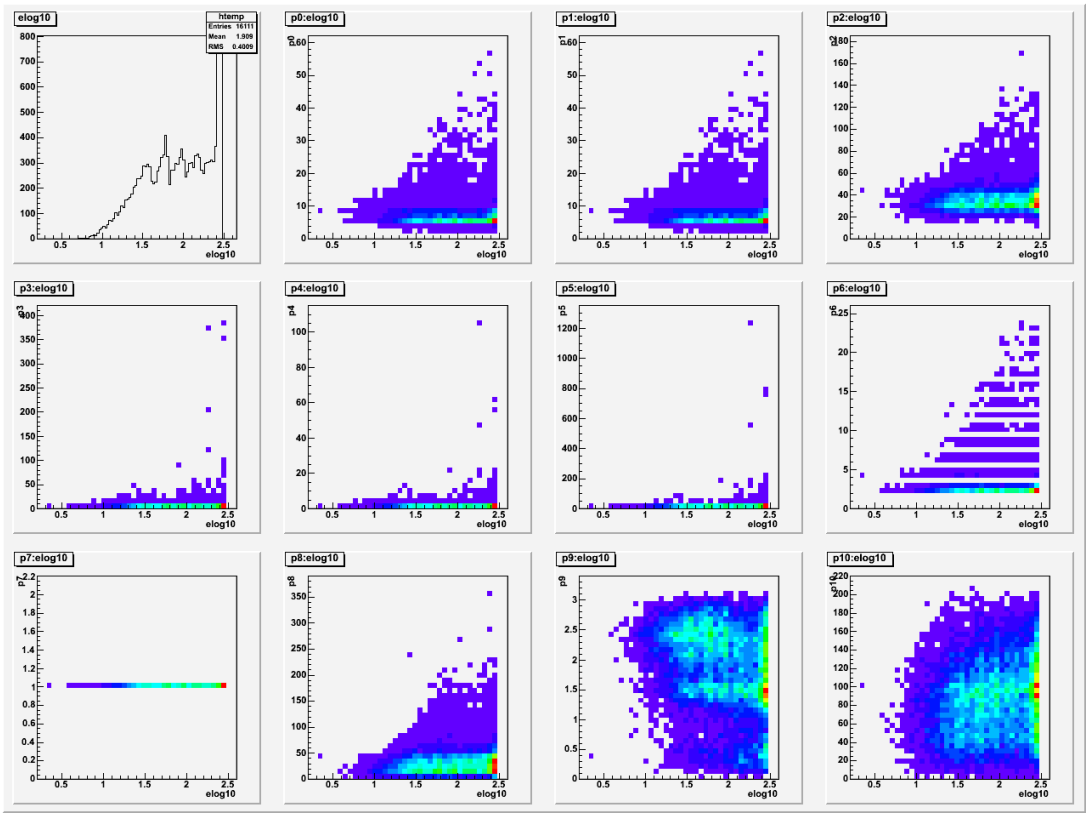


Figure 10: Cherenkov Estimator with Posidonia (1D)

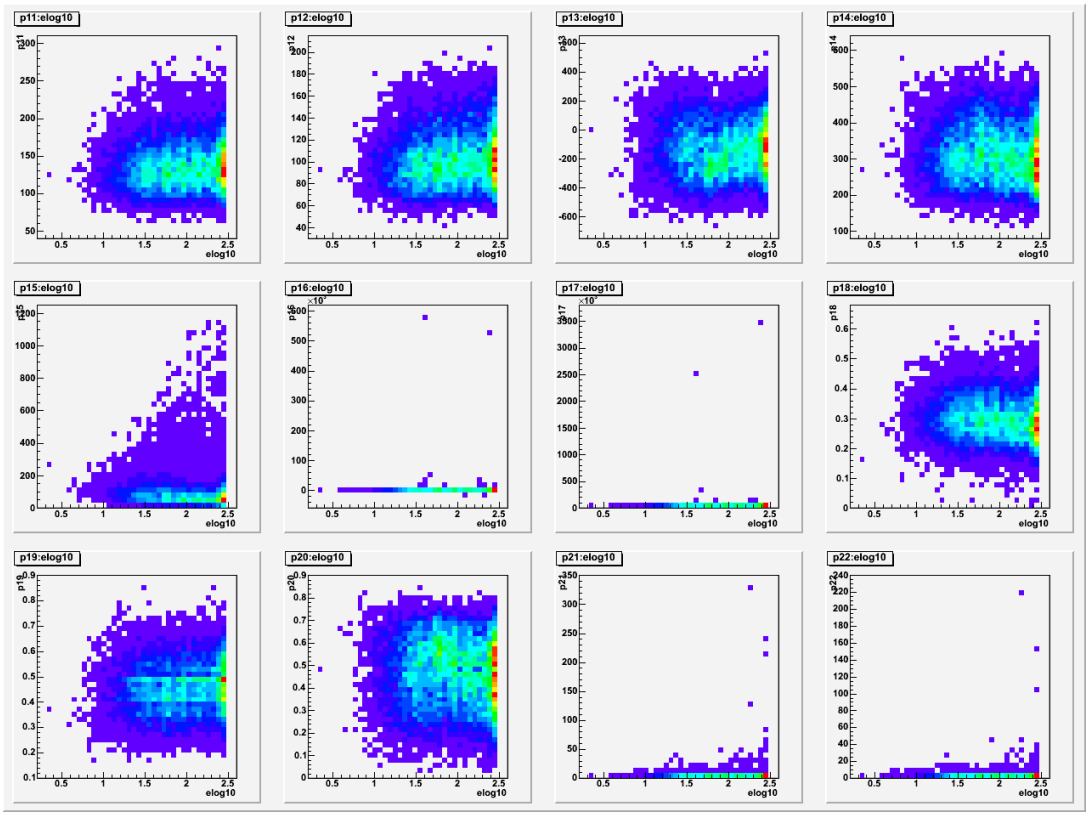


Figure 11: Cherenkov Estimator with Posidonia (1D)

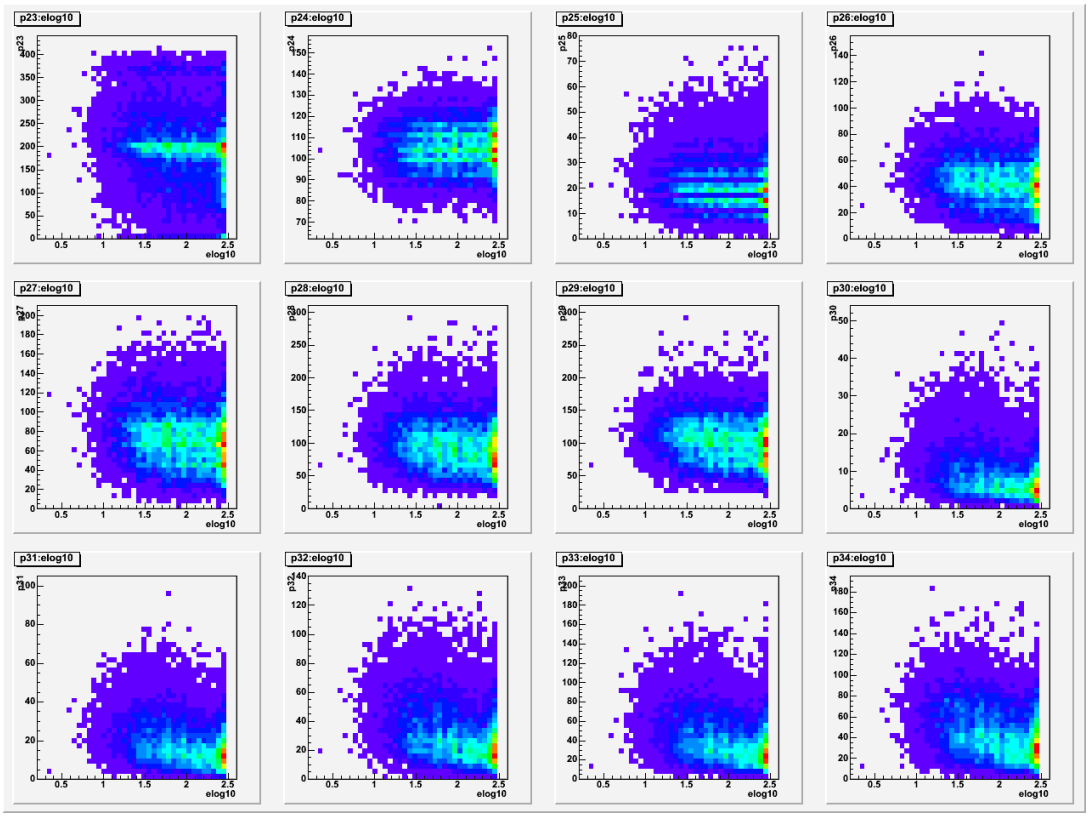


Figure 12: Cherenkov Estimator with Posidonia (1D)

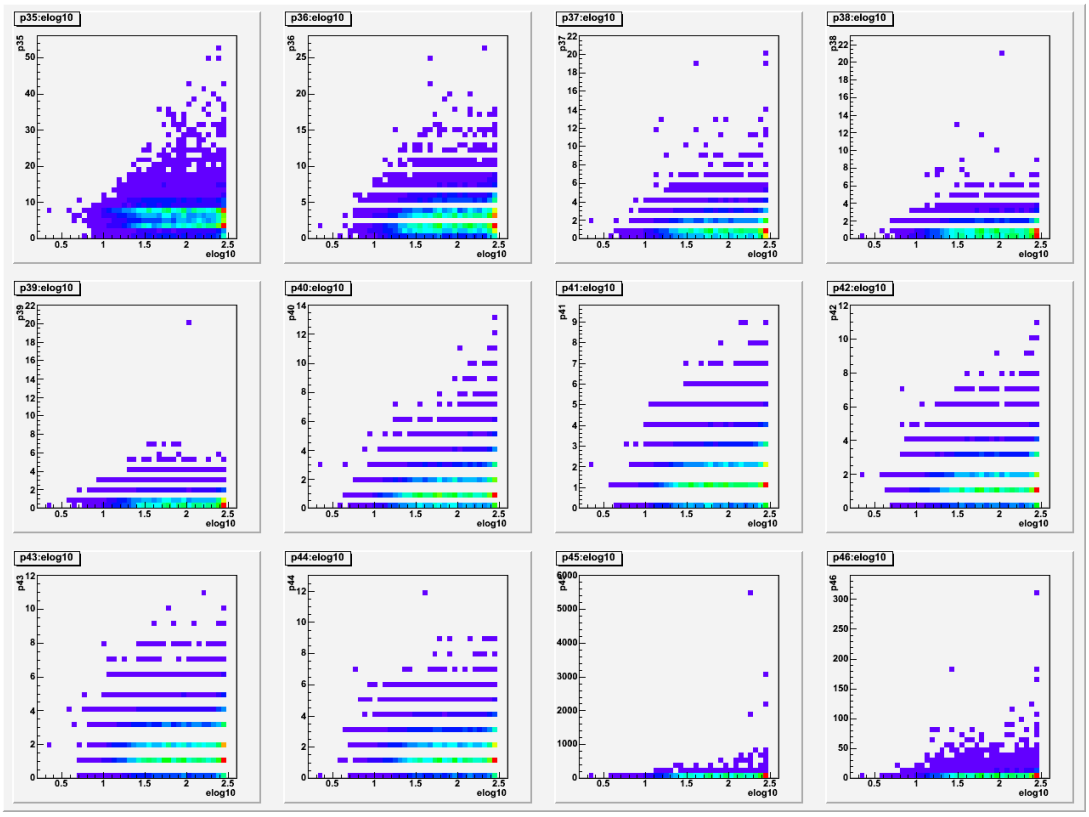


Figure 13: Cherenkov Estimator with Posidonia (1D)

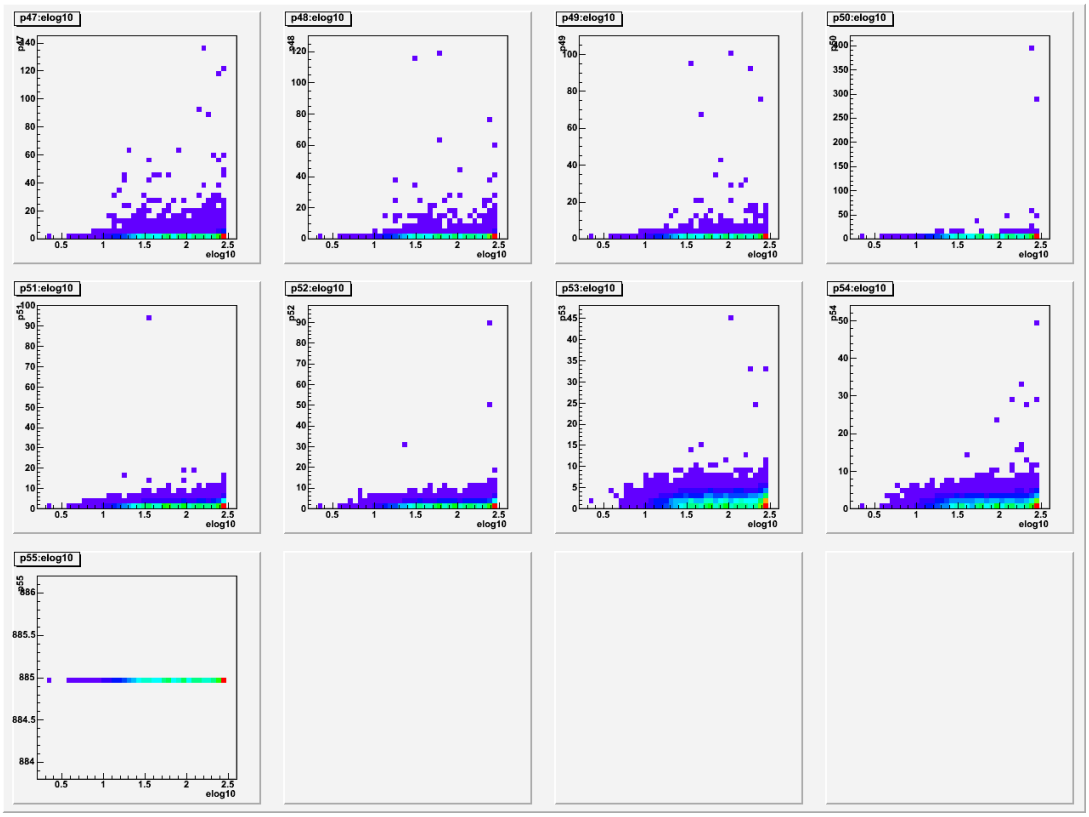


Figure 14: Cherenkov Estimator with Posidonia (1D)

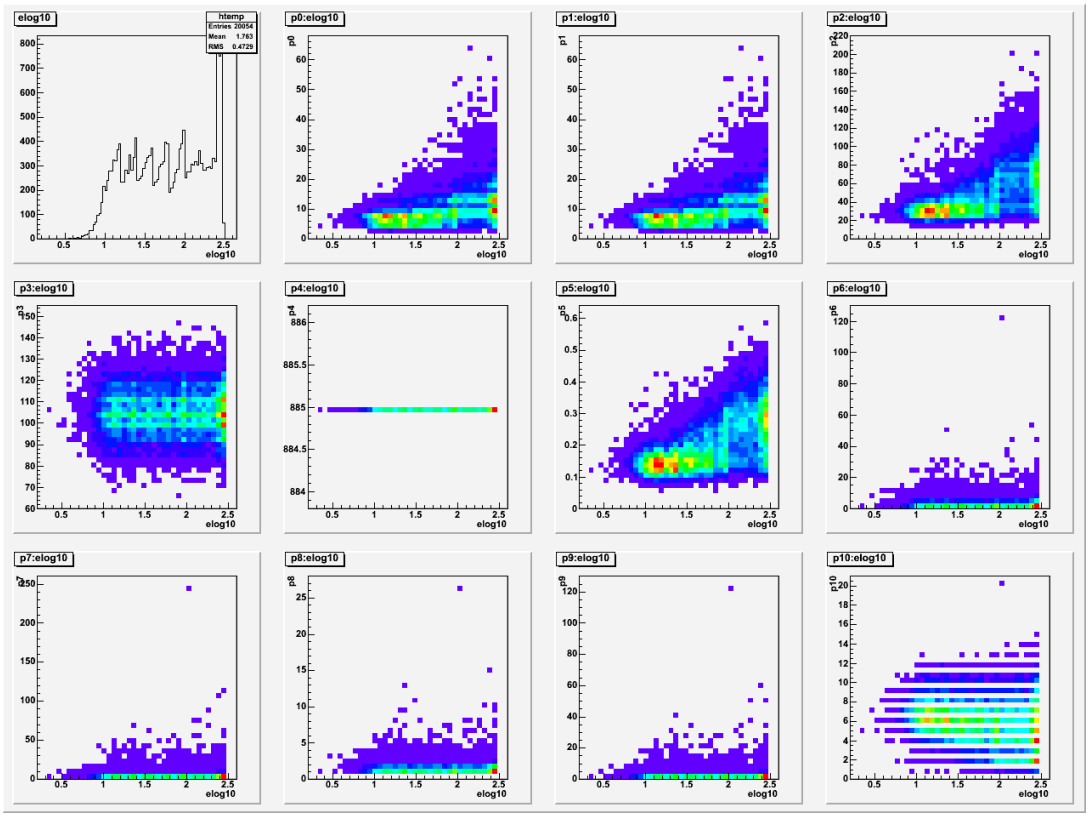


Figure 15: CalcCluster

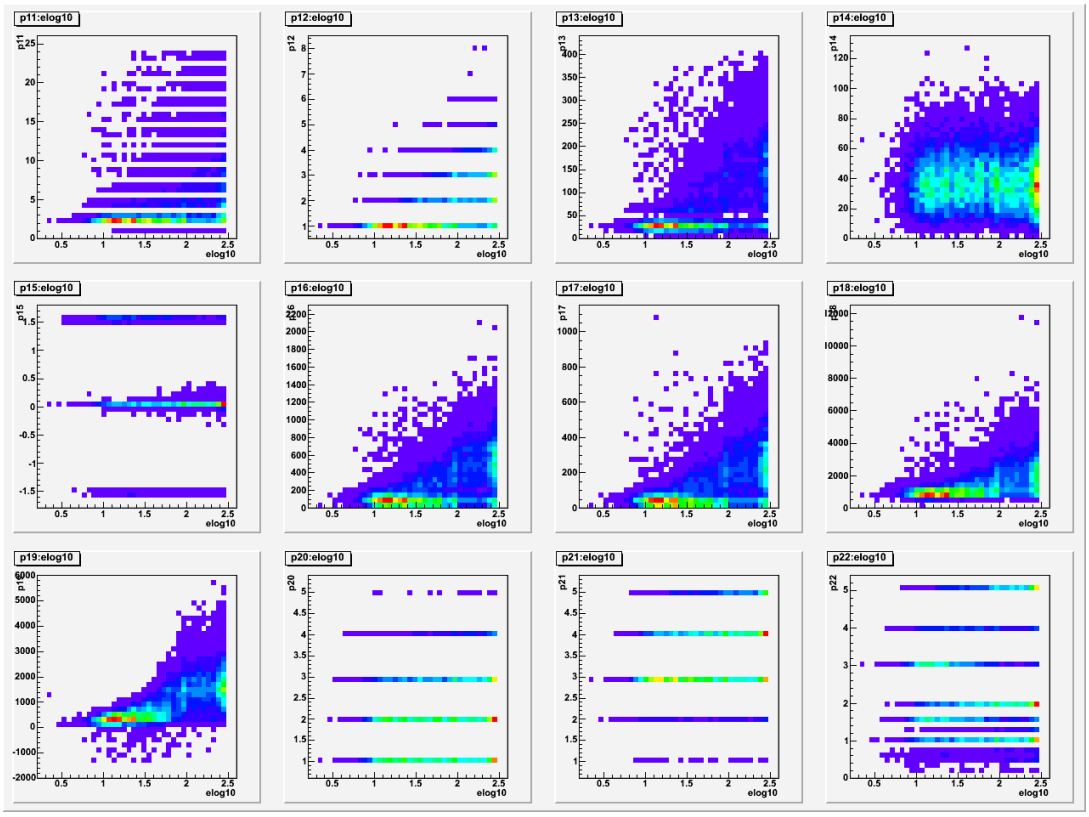


Figure 16: CalcCluster

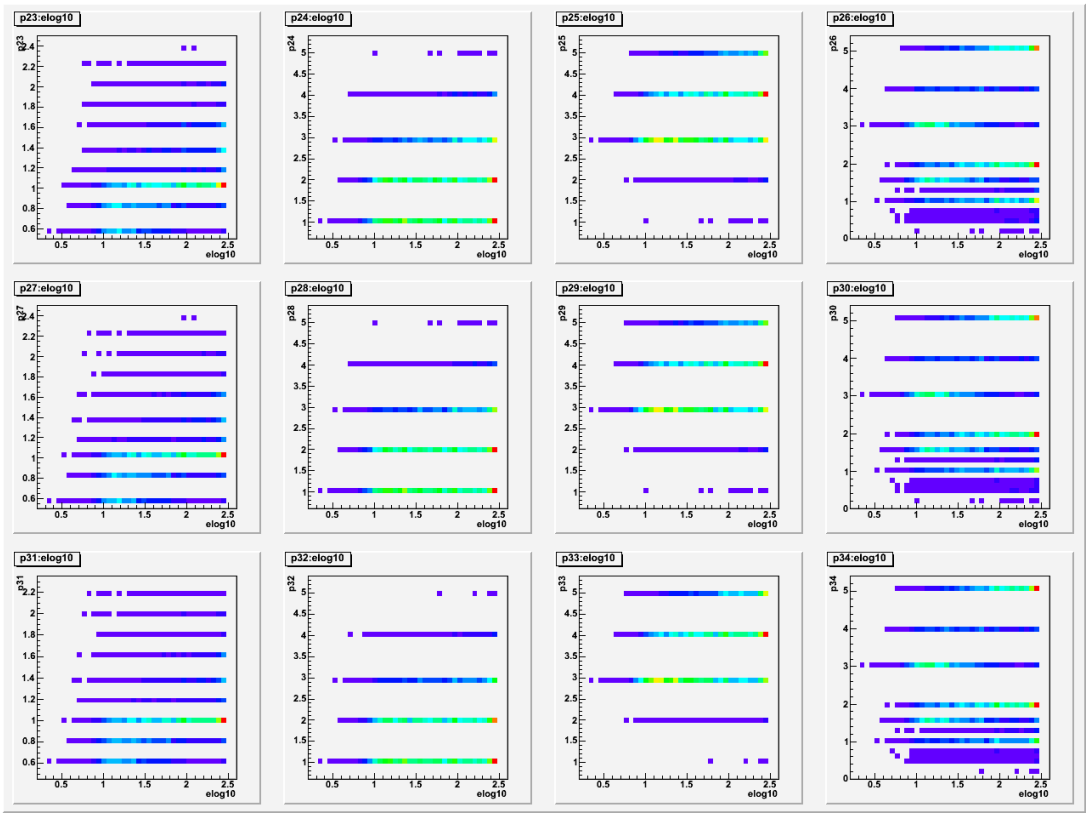


Figure 17: CalcCluster

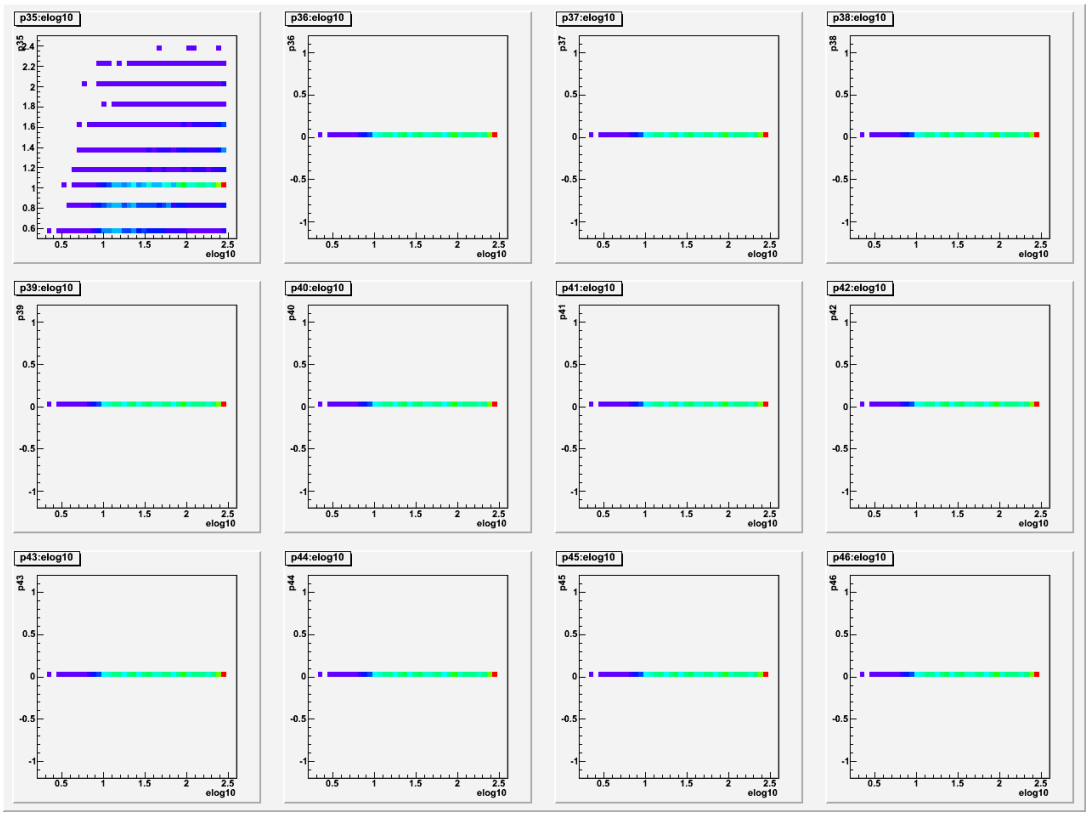


Figure 18: CalcCluster

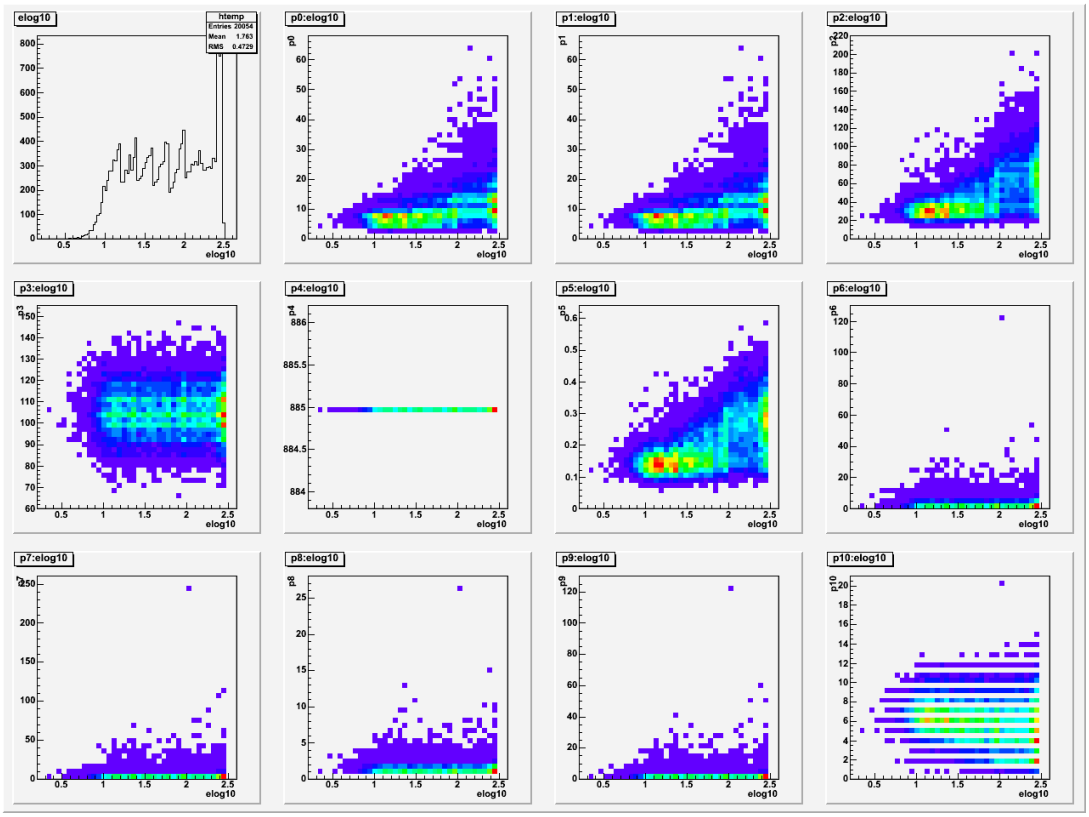


Figure 19: CalcDensity

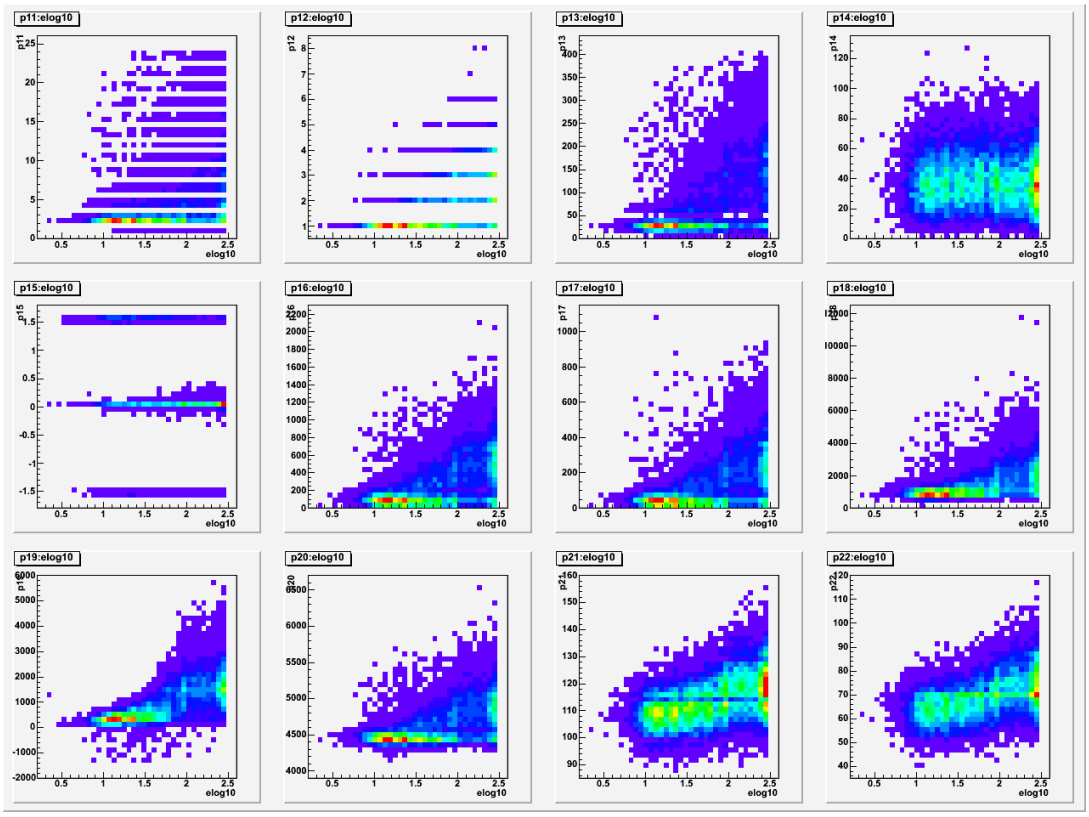


Figure 20: CalcDensity

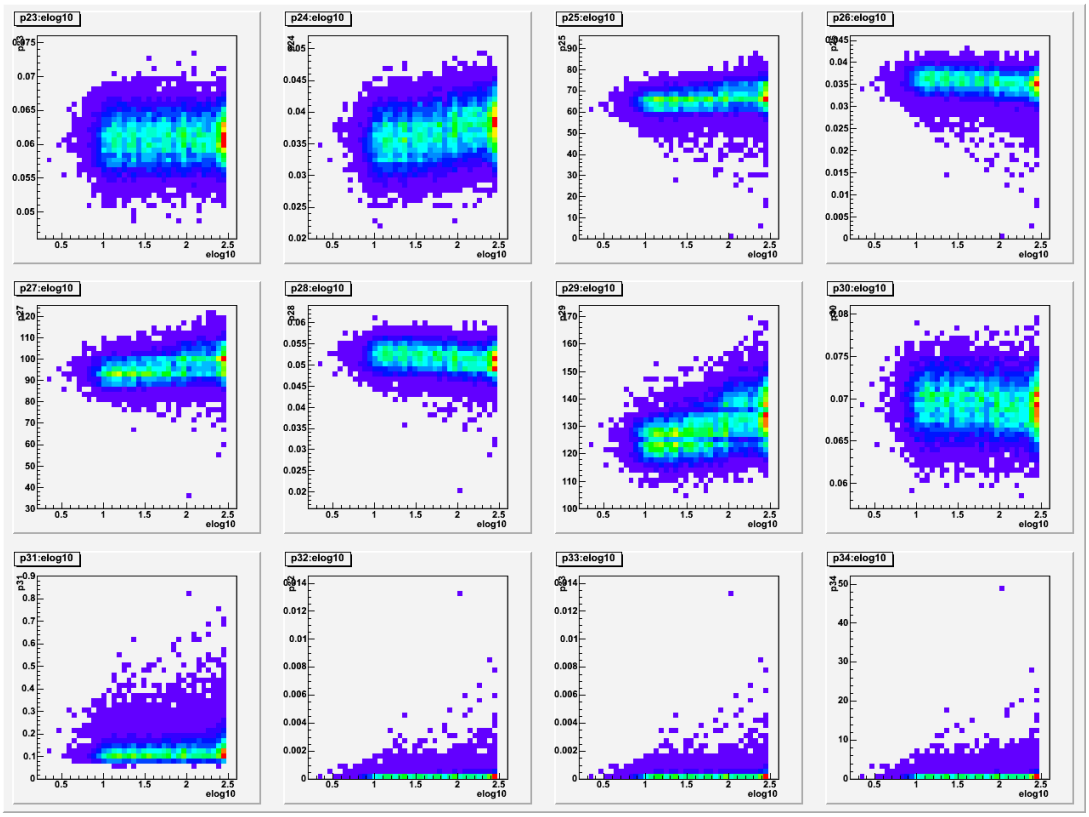


Figure 21: CalcDensity

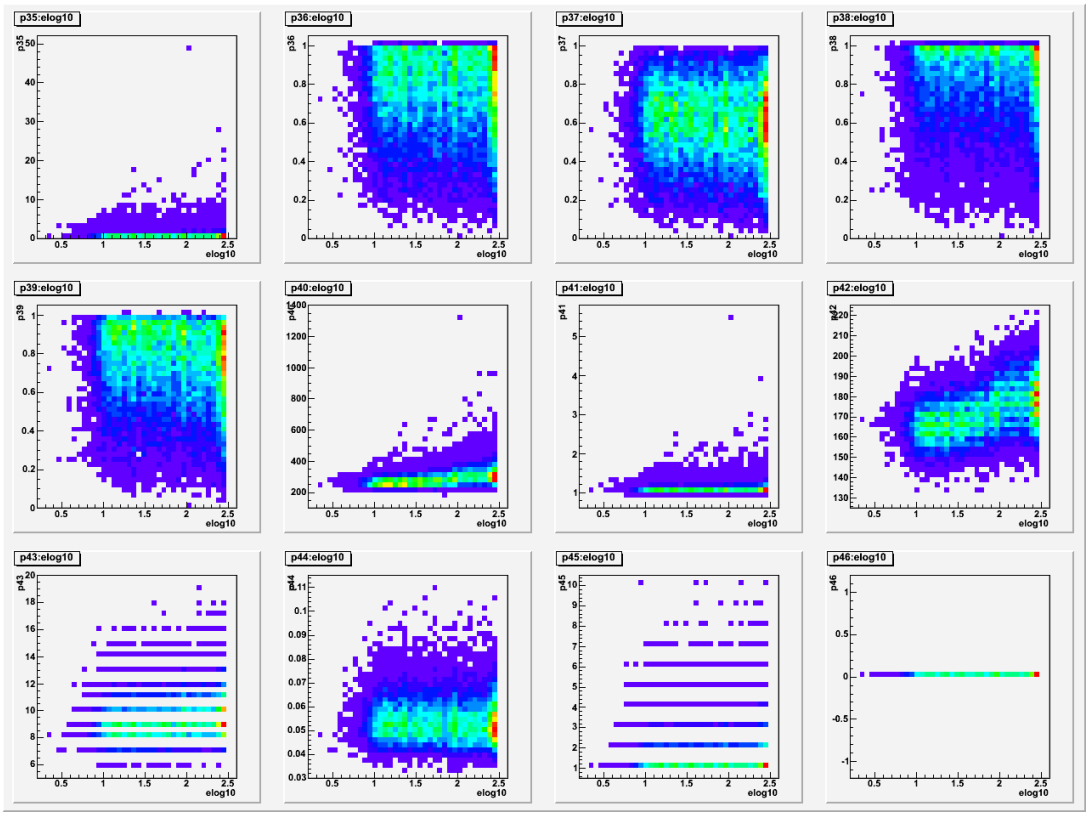


Figure 22: CalcDensity

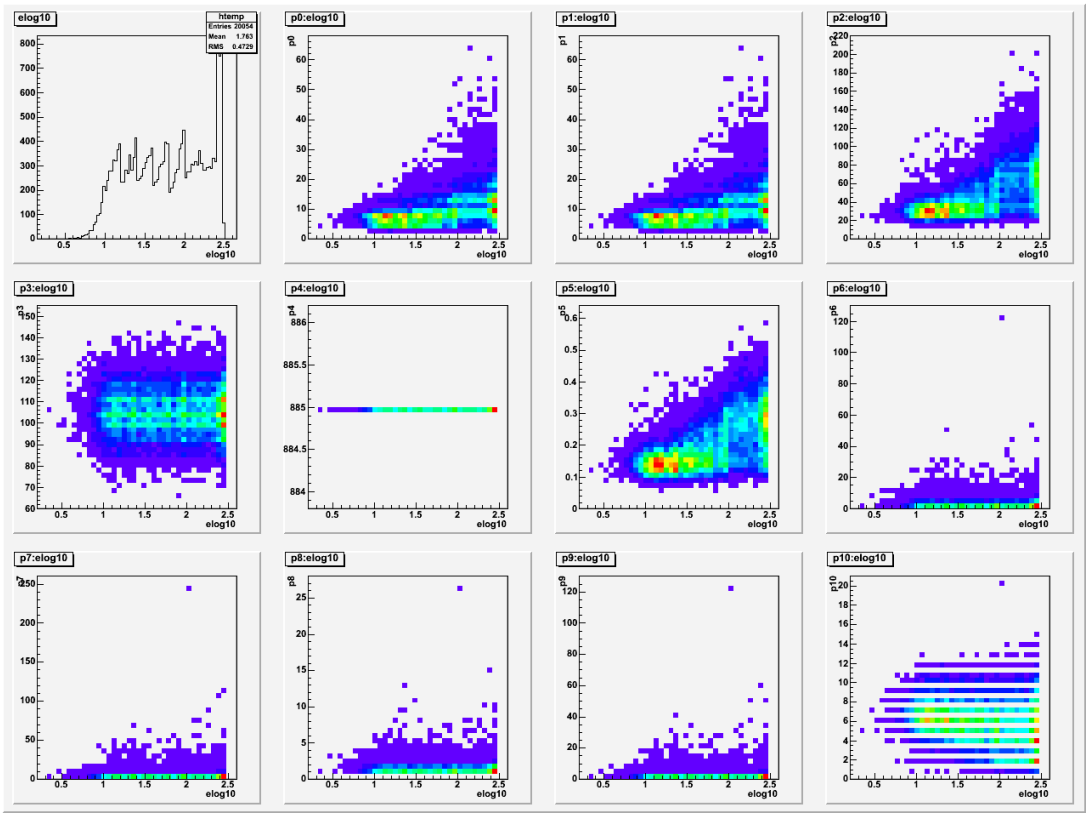


Figure 23: CalcTensor

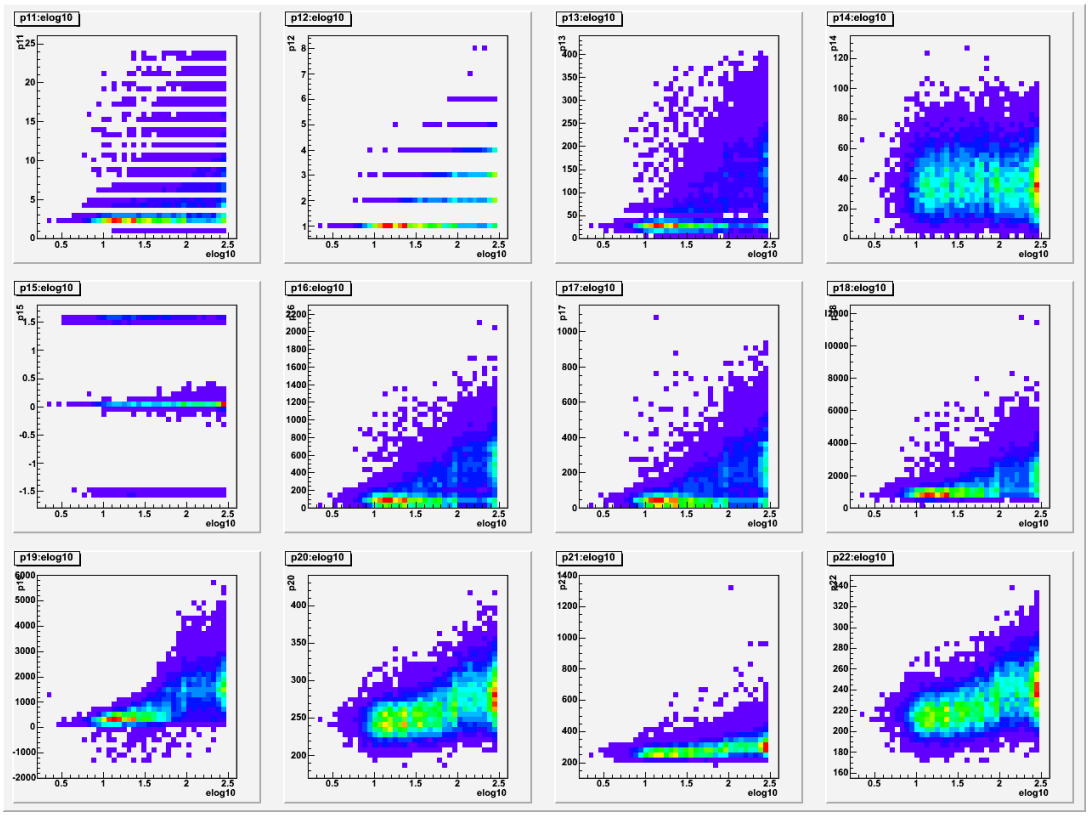


Figure 24: CalcTensor

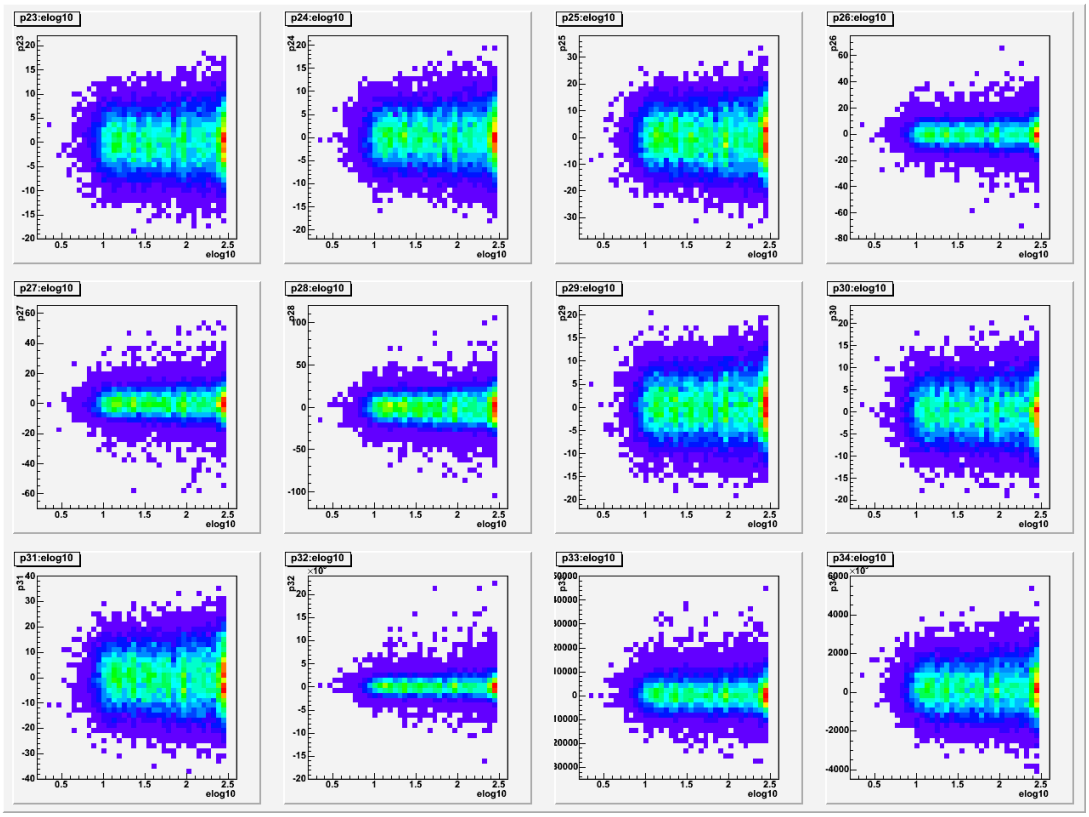


Figure 25: CalcTensor

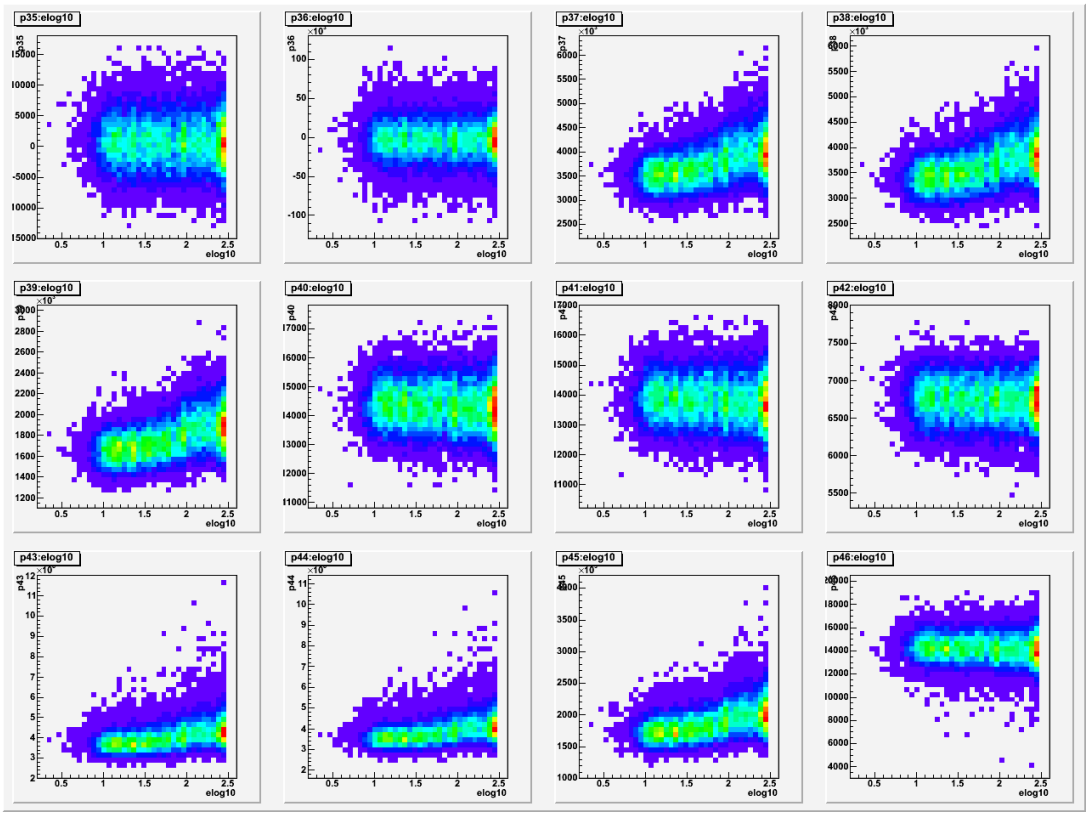


Figure 26: CalcTensor

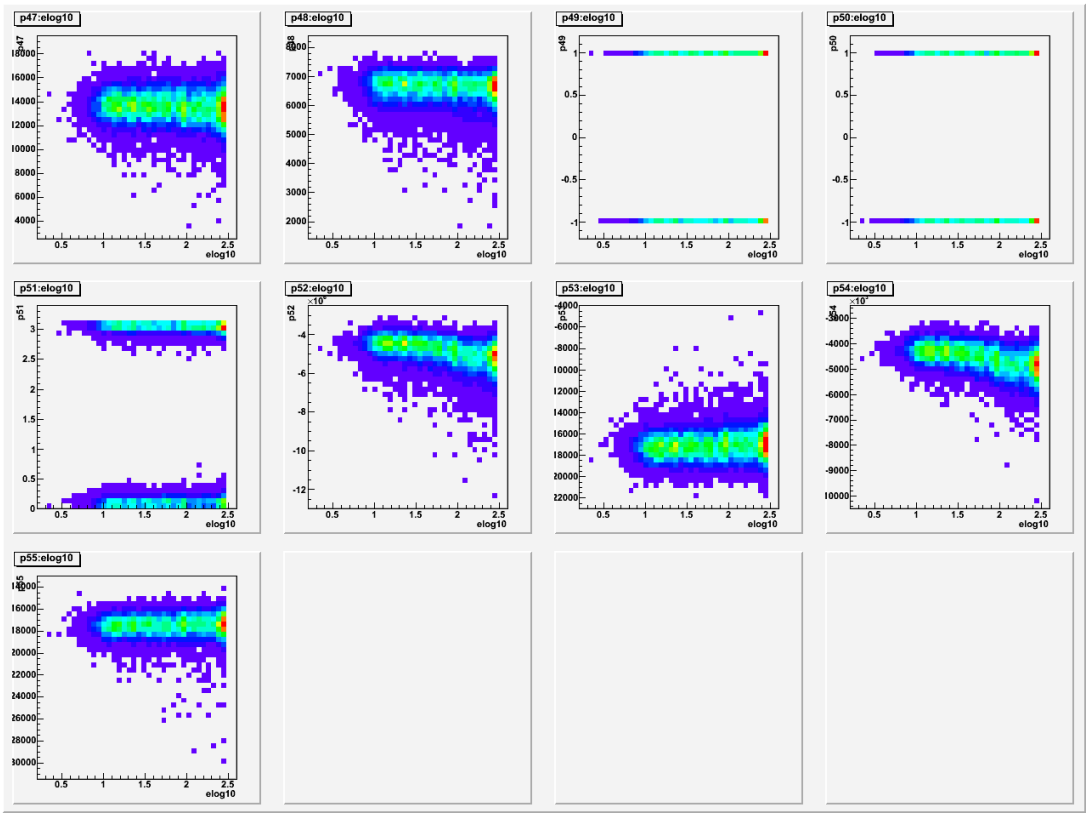


Figure 27: CalcTensor

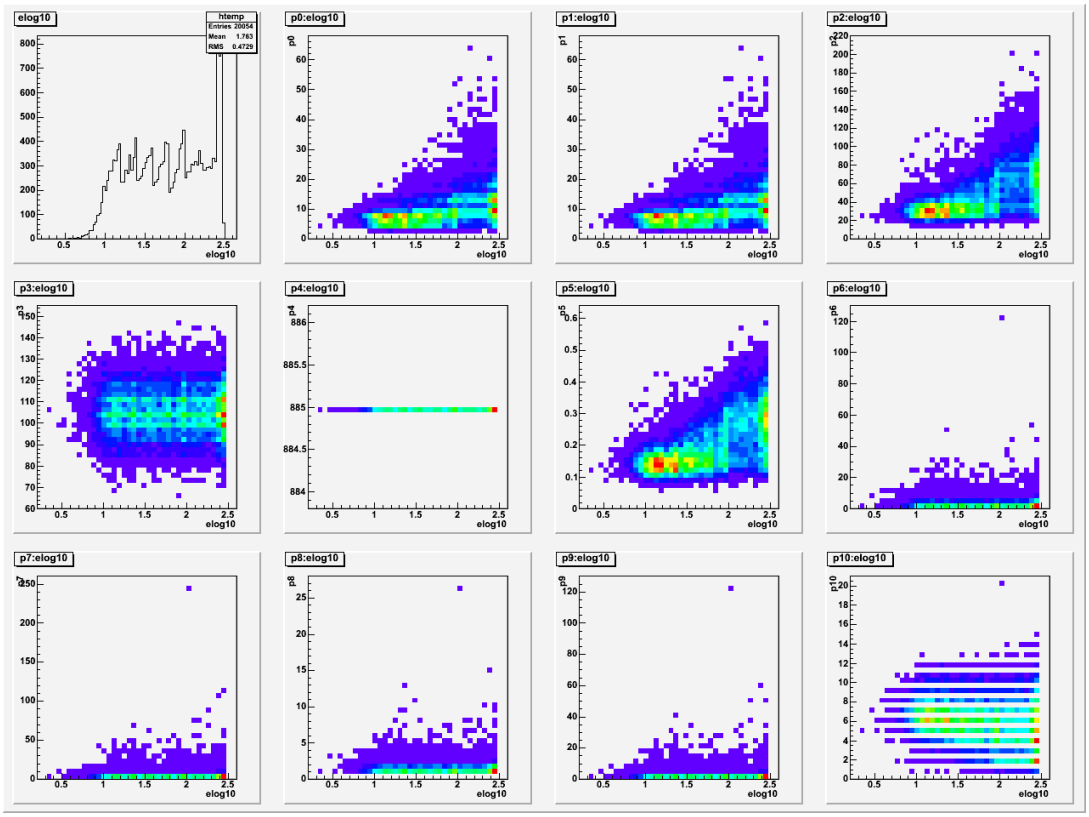


Figure 28: CalcShape

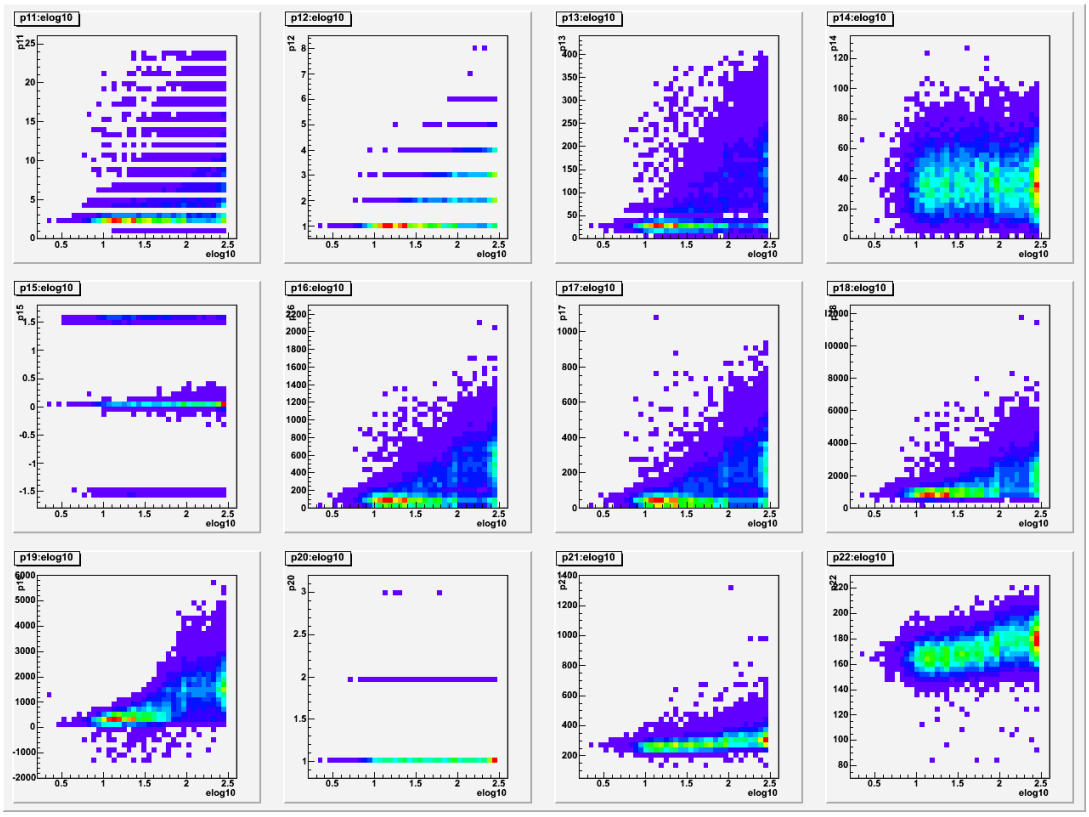


Figure 29: CalcShape

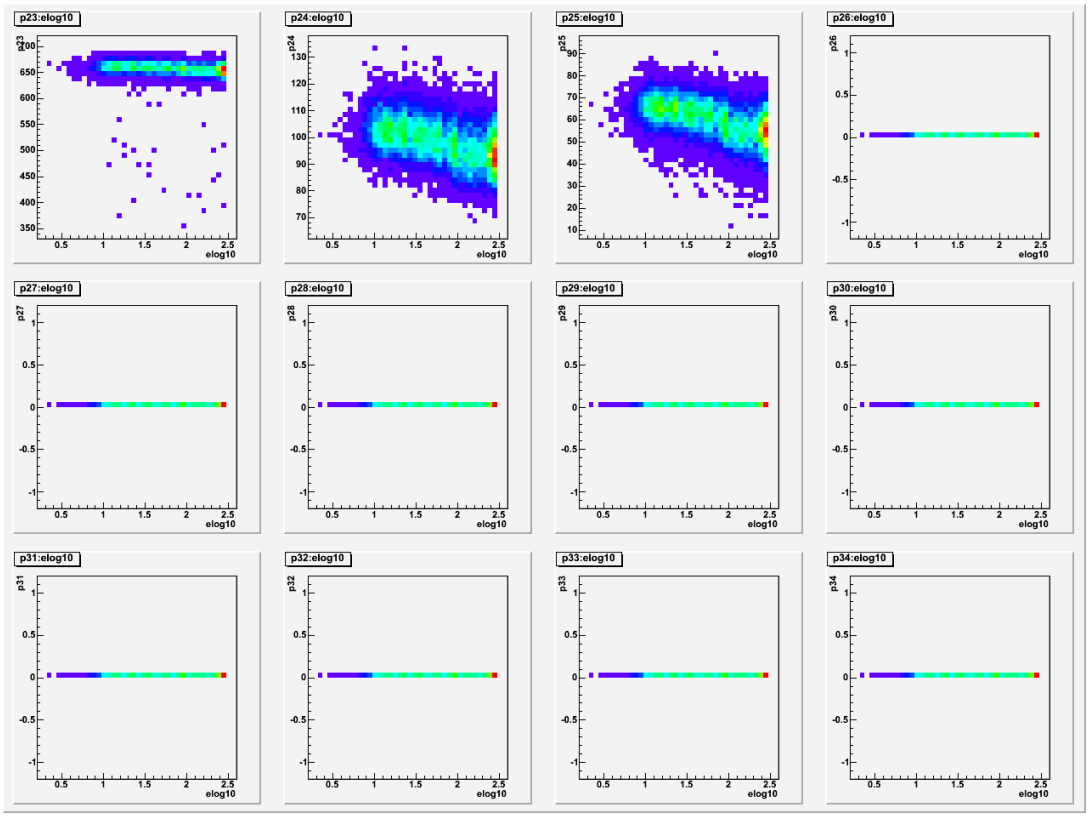


Figure 30: CalcShape

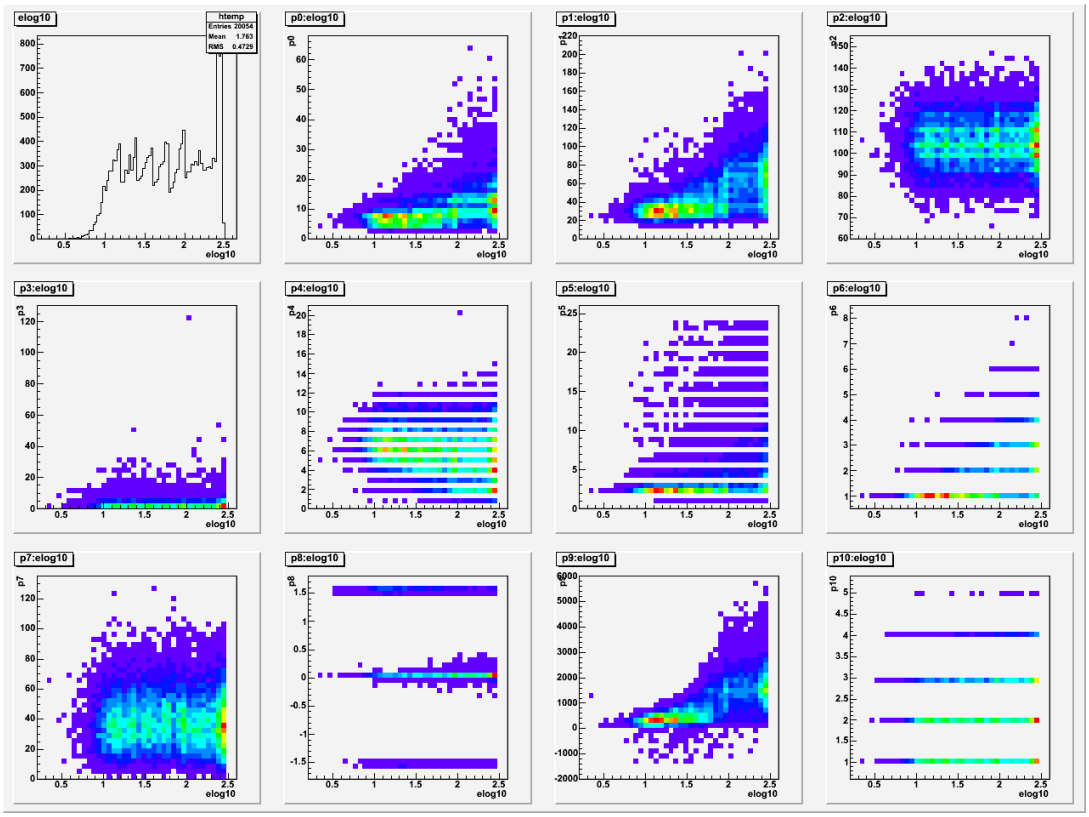


Figure 31: CalcCluster (Reduced number of parameters)

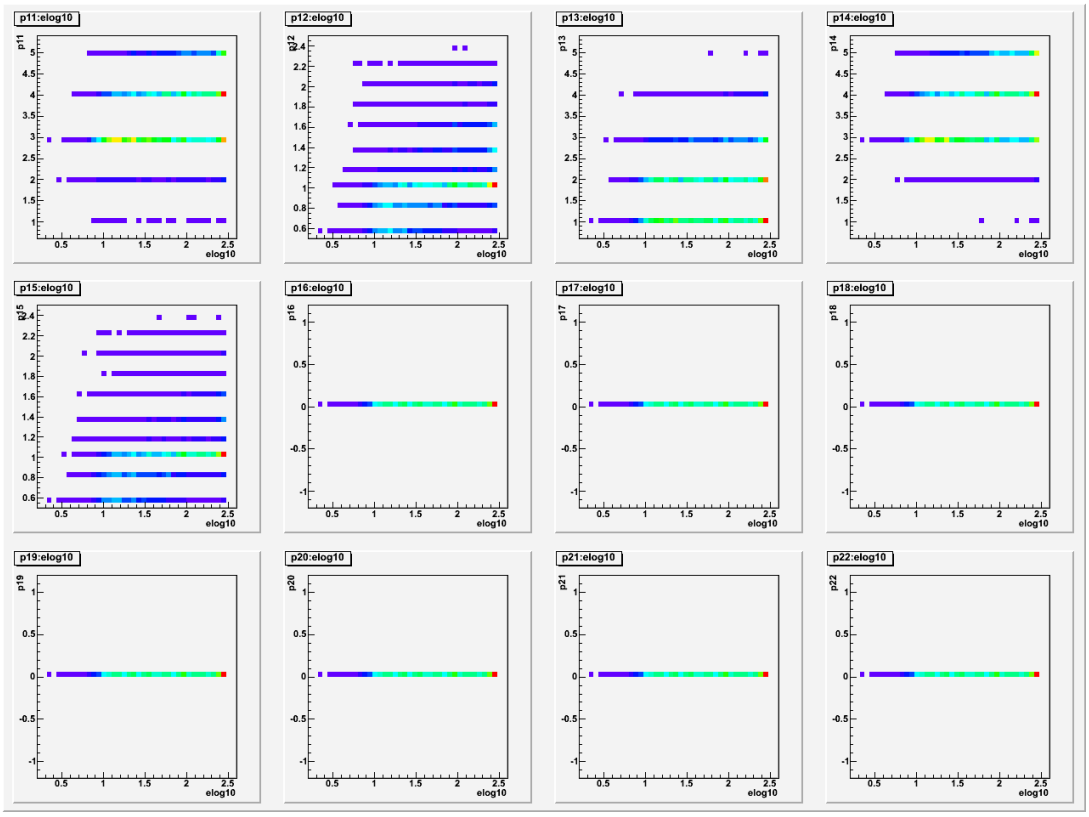


Figure 32: CalcCluster (Reduced number of parameters)

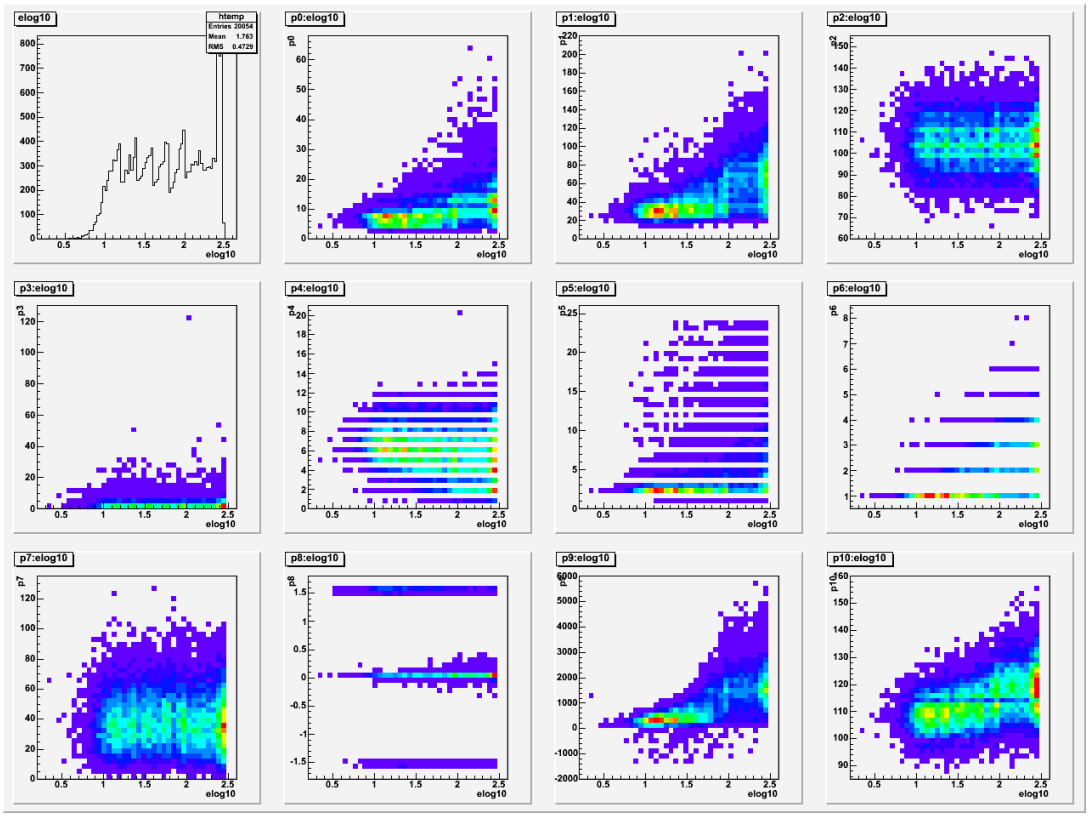


Figure 33: CalcDensity (Reduced number of parameters)

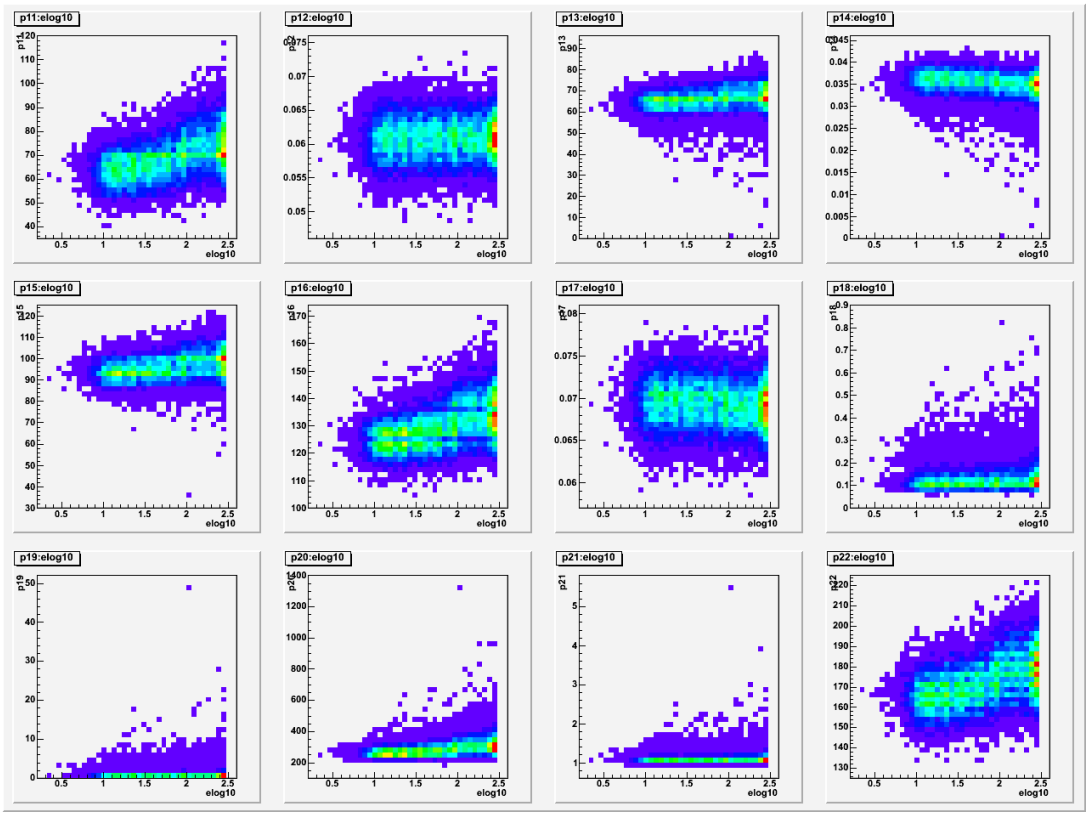


Figure 34: CalcDensity (Reduced number of parameters)

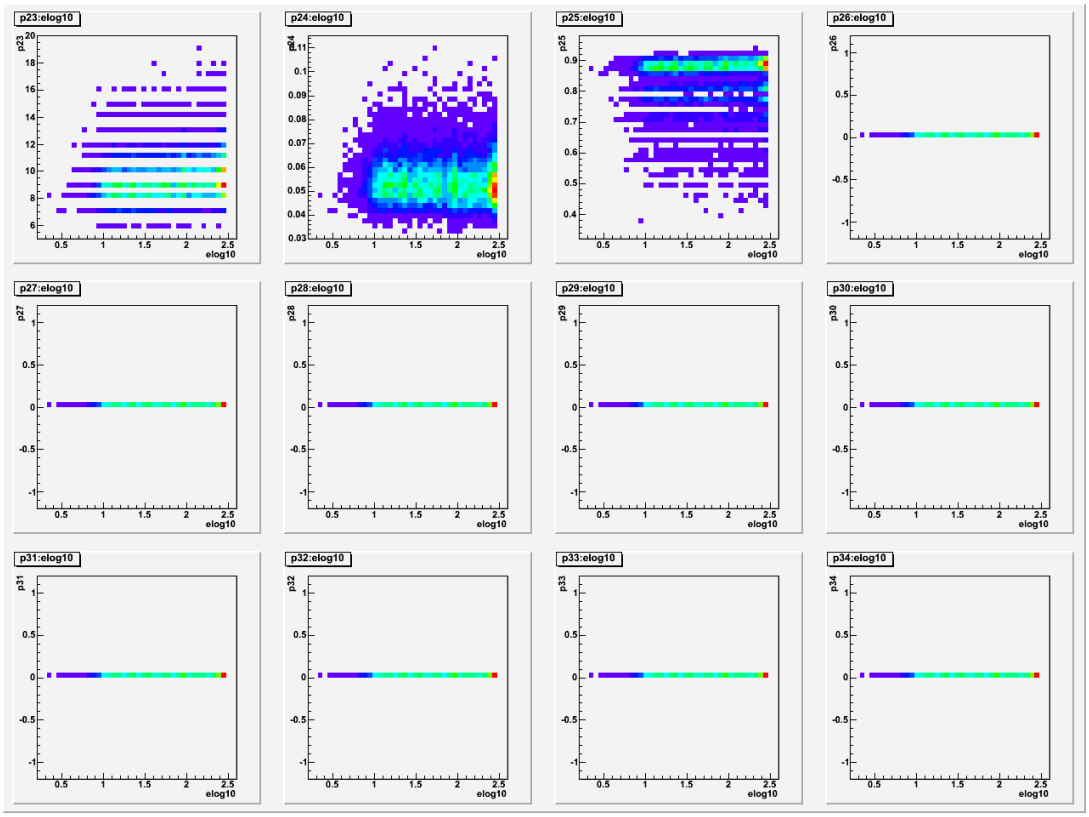


Figure 35: CalcDensity (Reduced number of parameters)

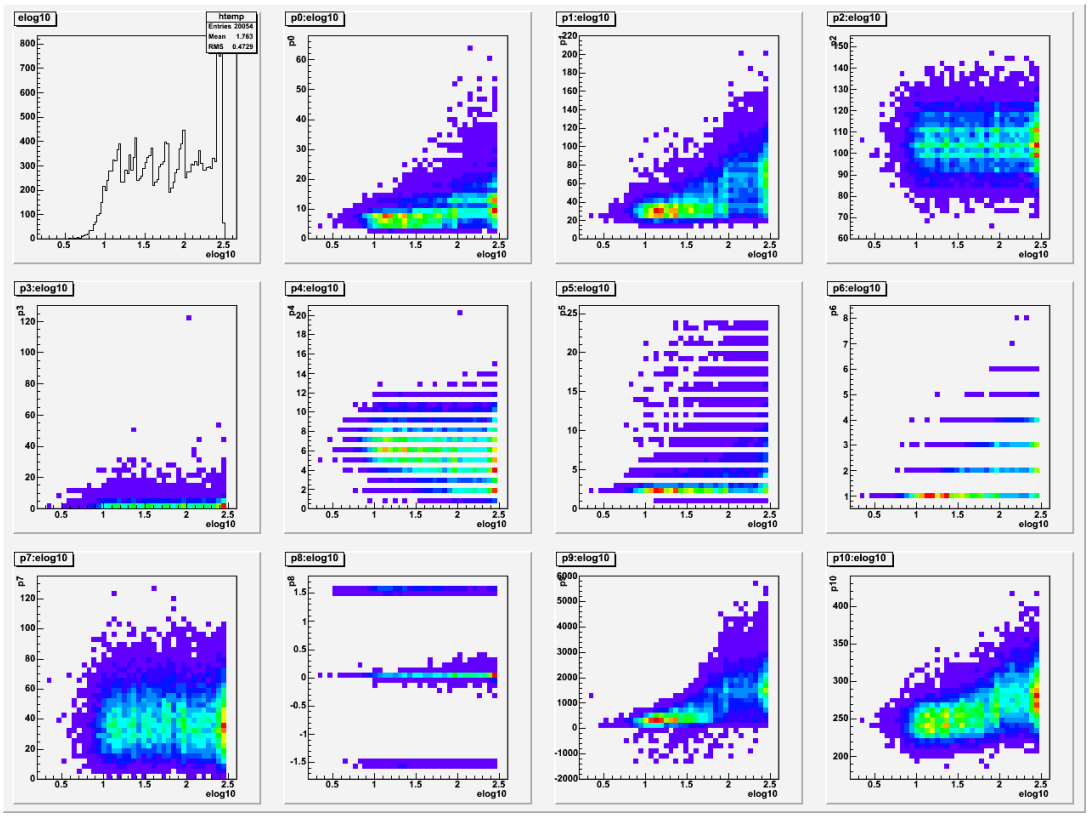


Figure 36: CalcTensor (Reduced number of parameters)

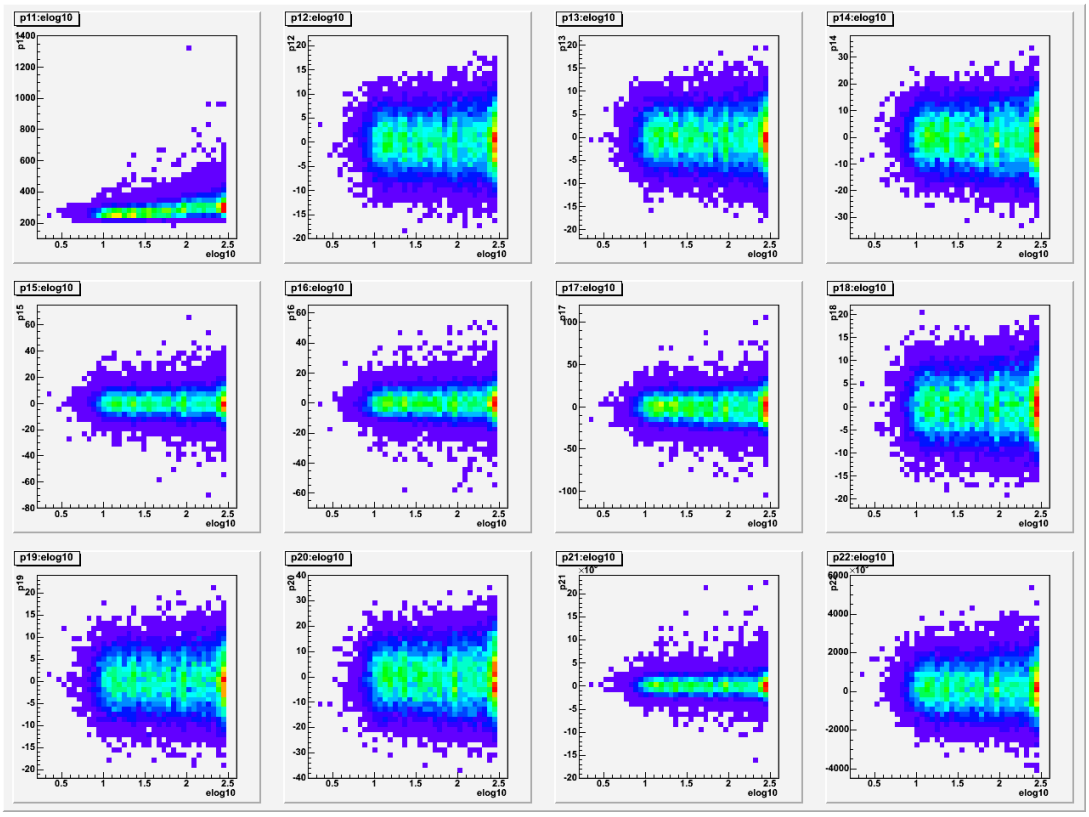


Figure 37: CalcTensor (Reduced number of parameters)

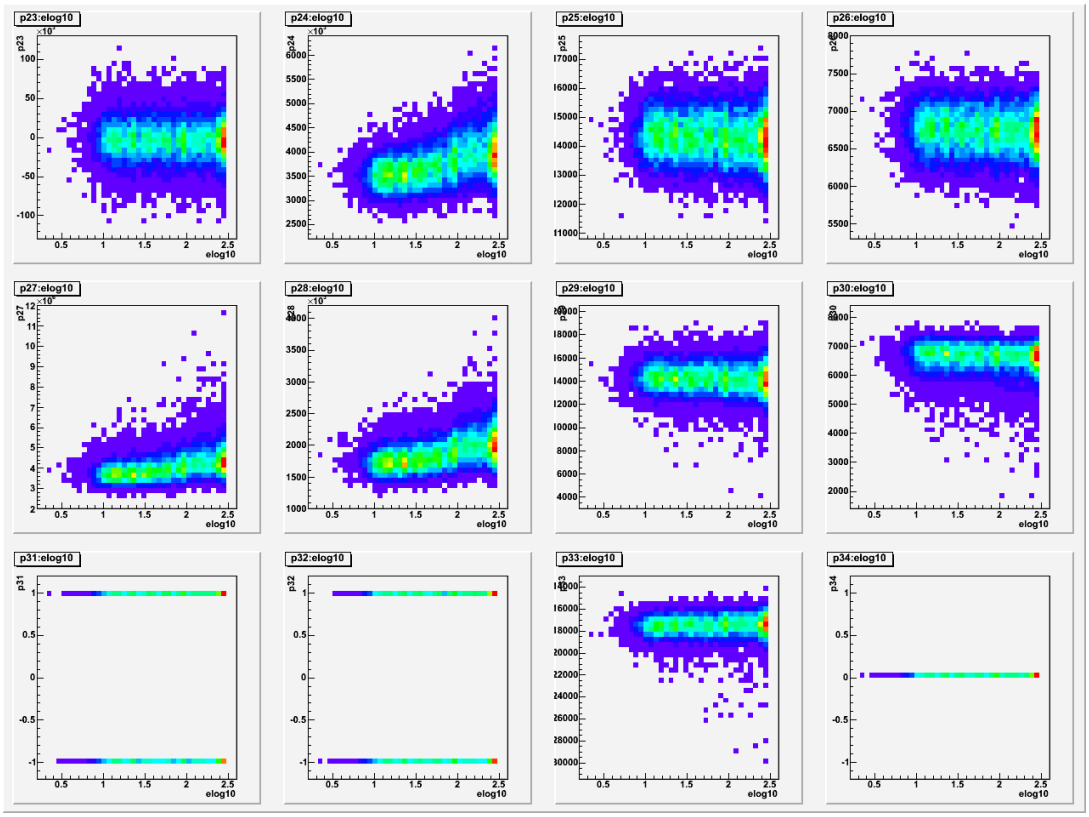


Figure 38: CalcTensor (Reduced number of parameters)

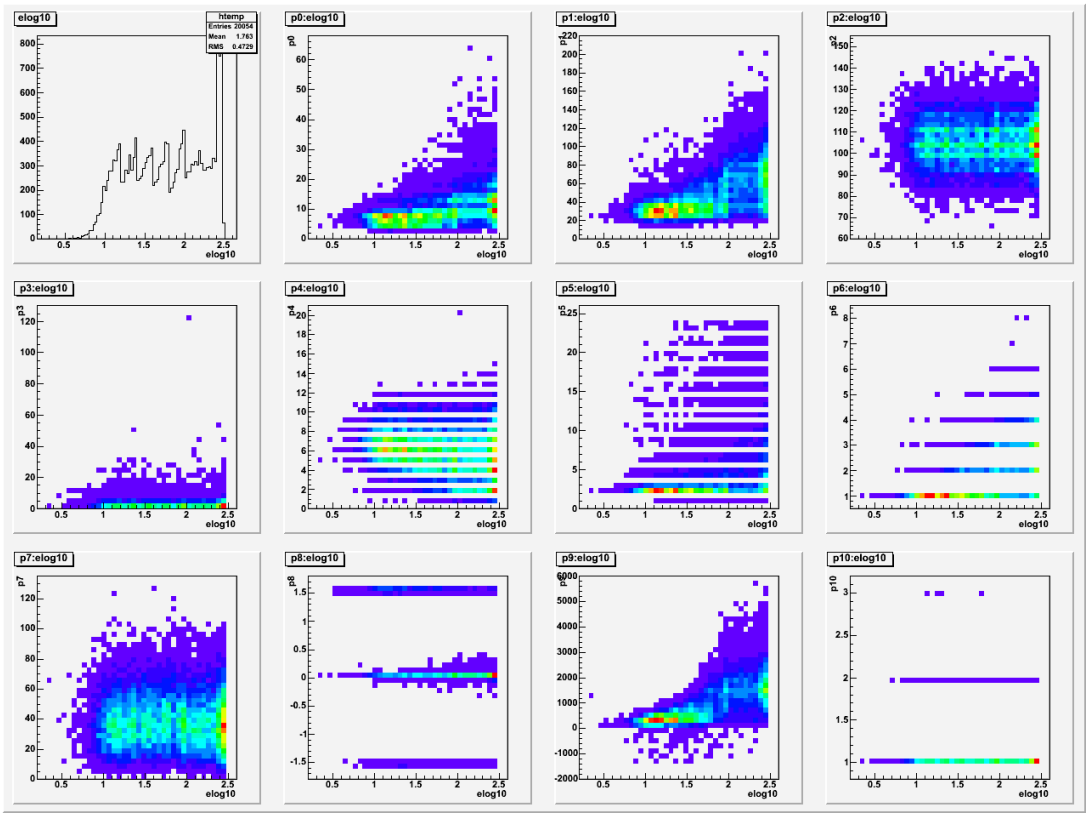


Figure 39: CalcShape (Reduced number of parameters)

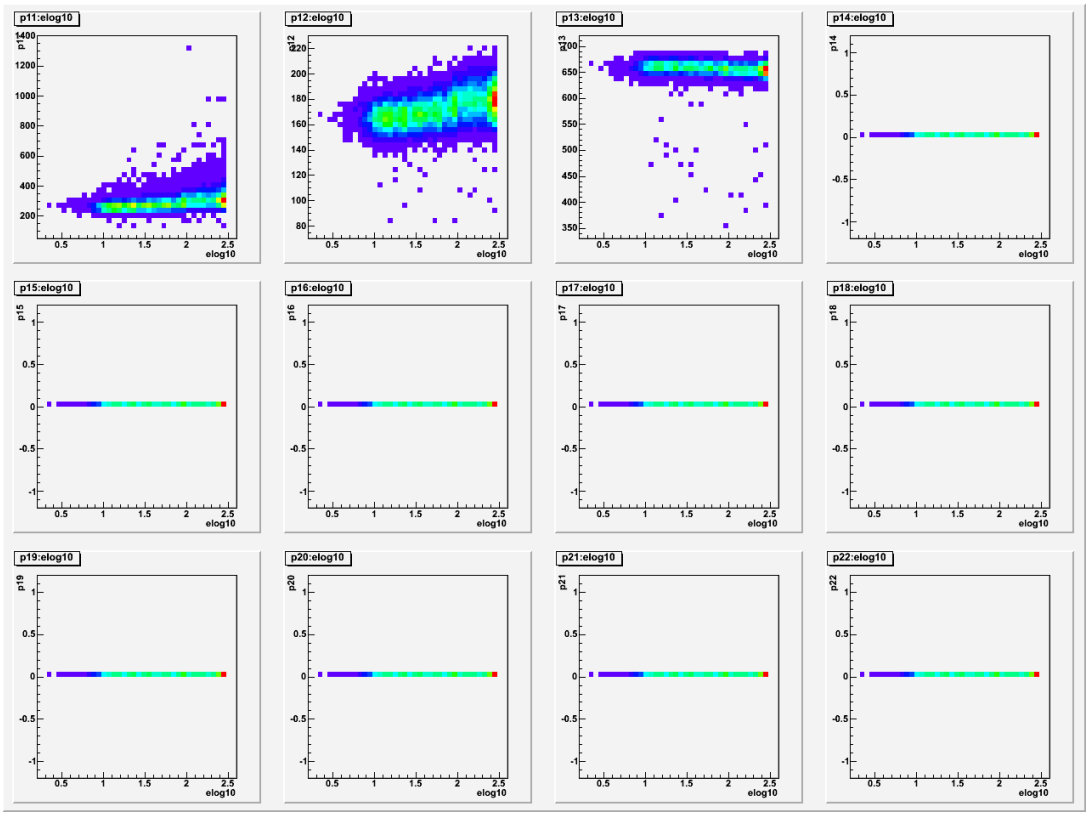


Figure 40: CalcShape (Reduced number of parameters)

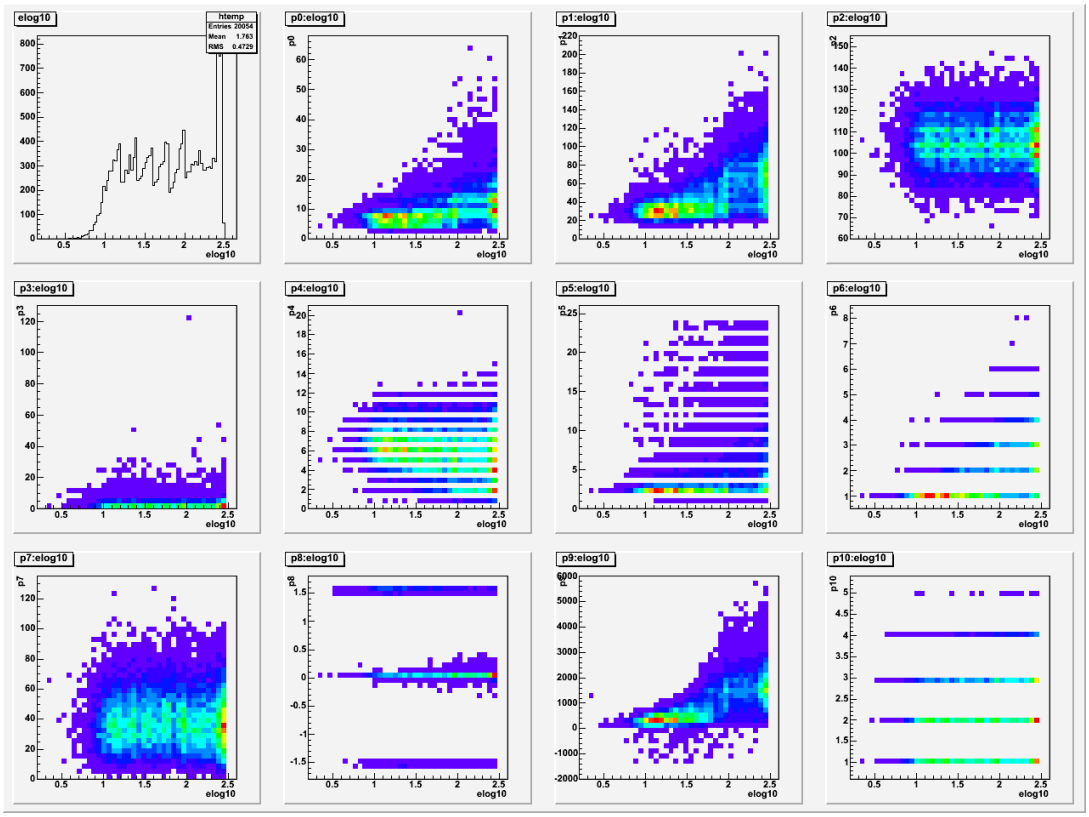


Figure 41: Combination of parameters

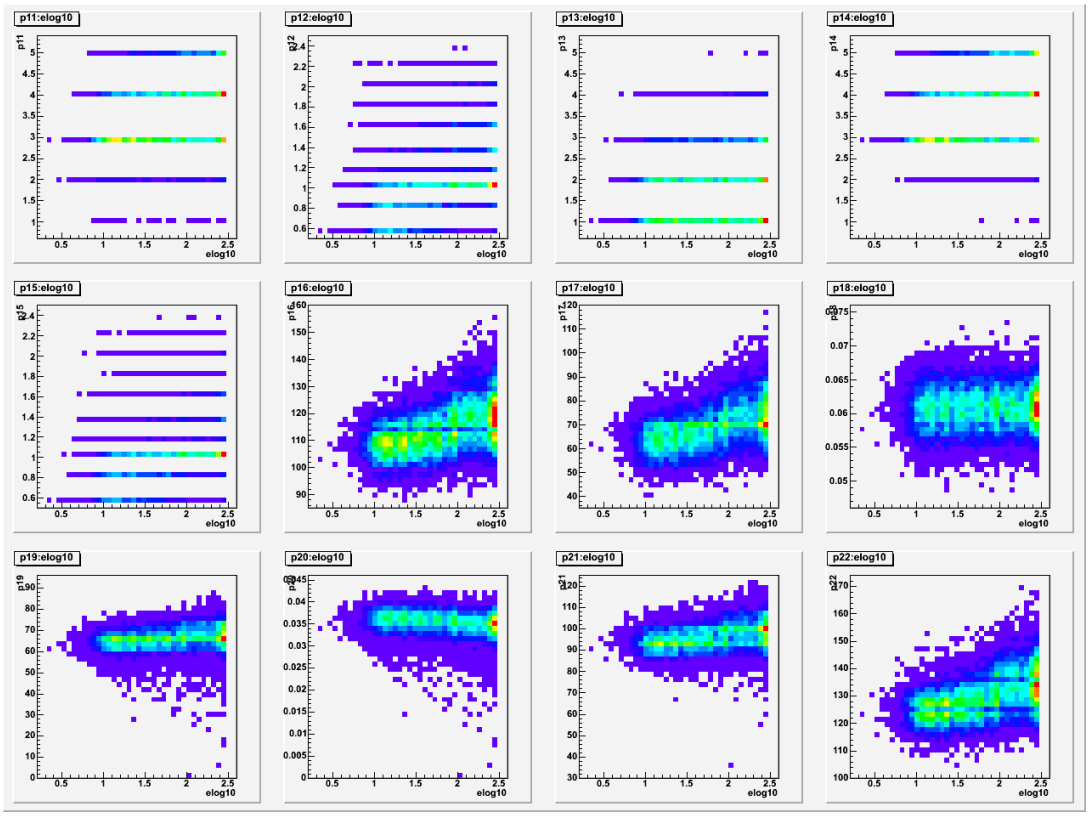


Figure 42: Combination of parameters

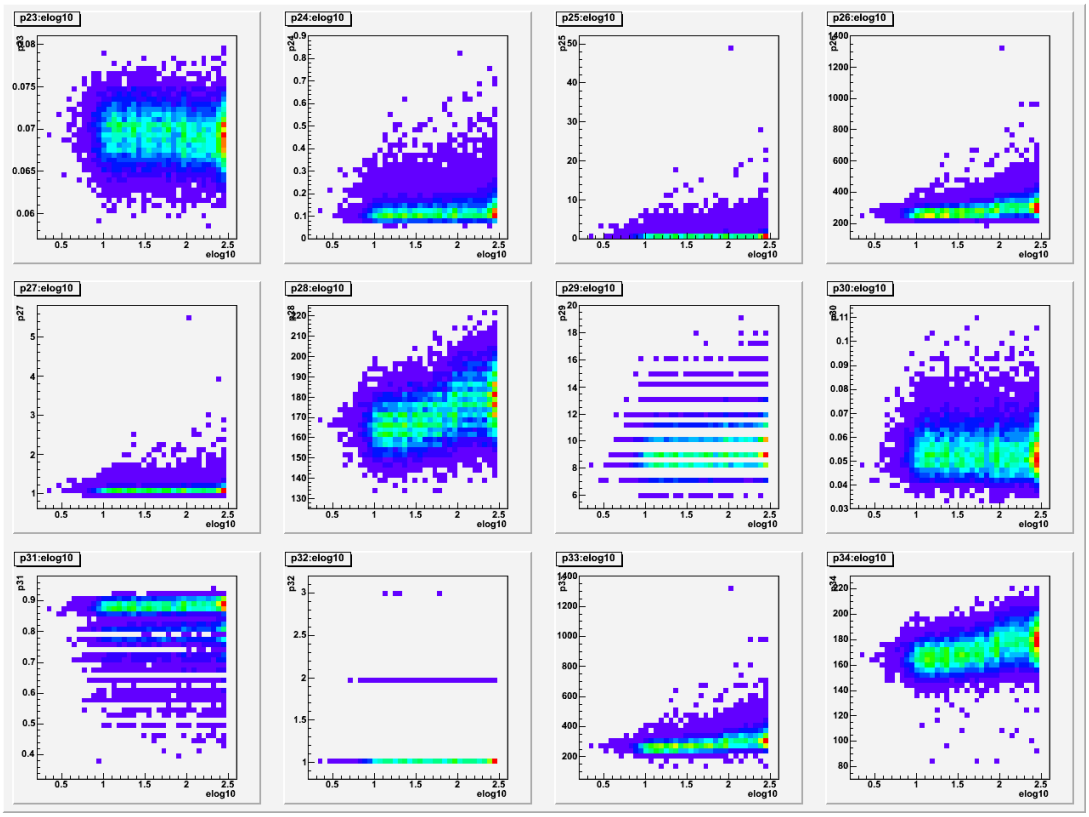


Figure 43: Combination of parameters

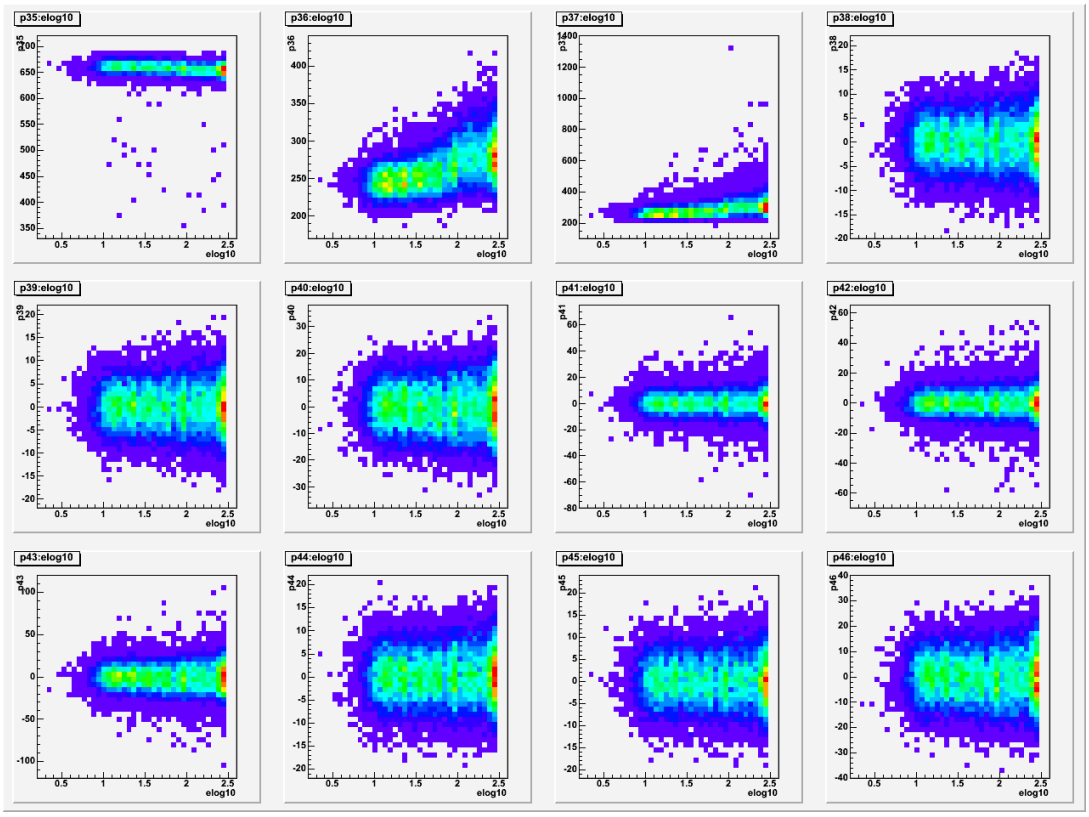


Figure 44: Combination of parameters

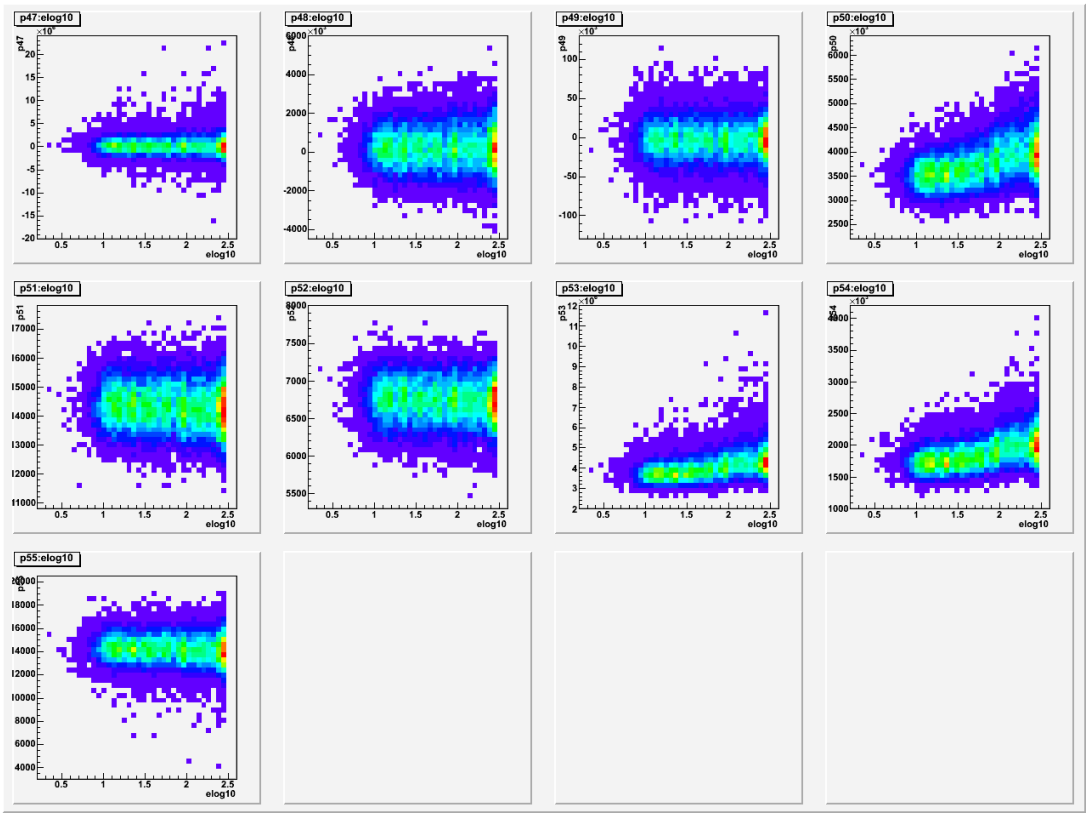


Figure 45: Combination of parameters

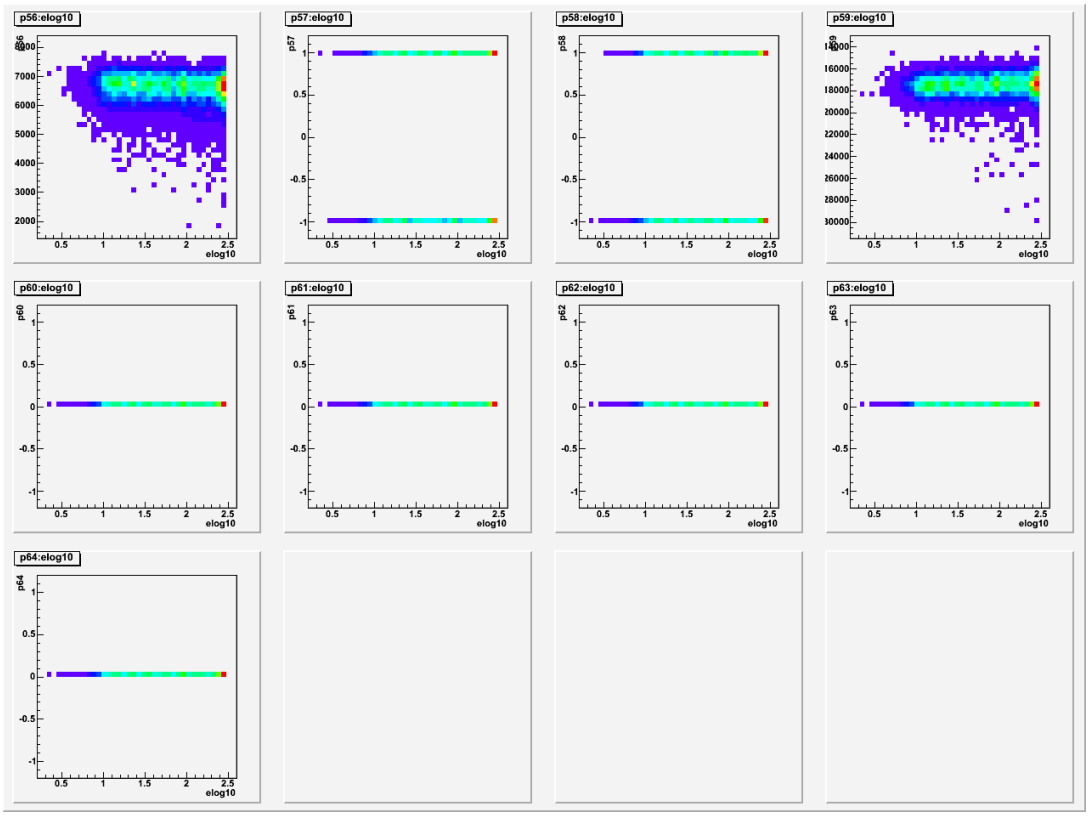


Figure 46: Combination of parameters

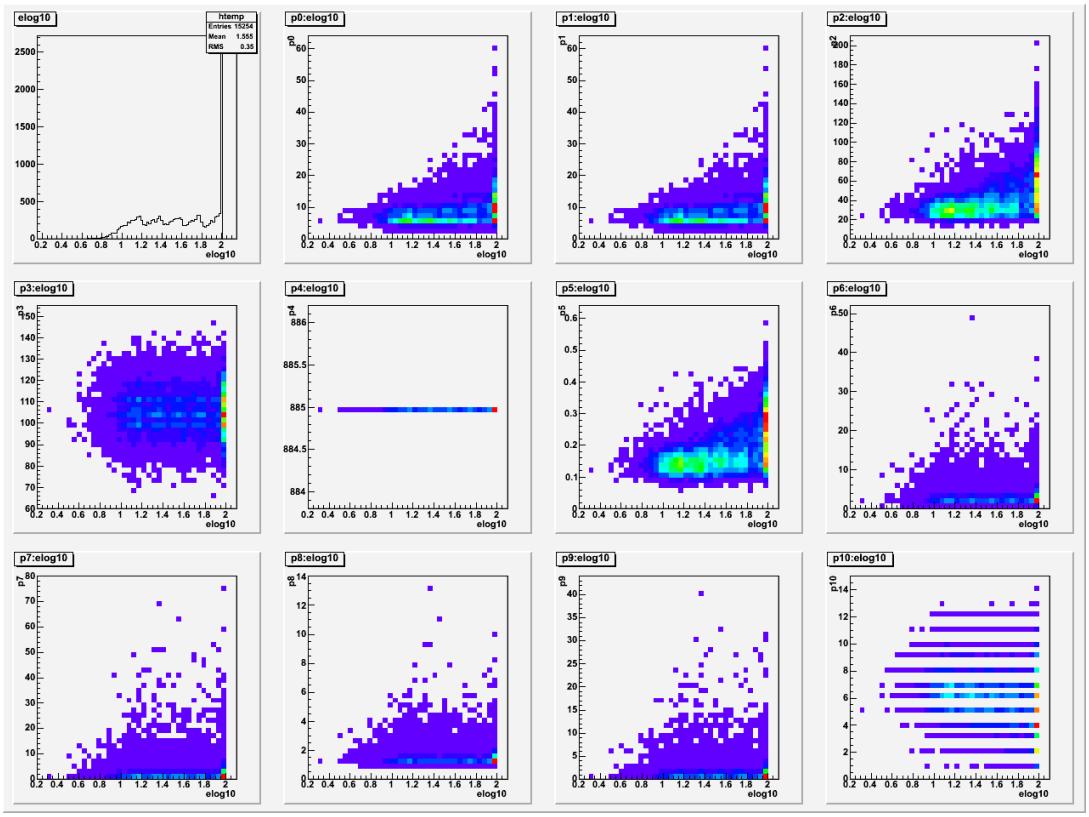


Figure 47: Simple Estimator with reduced range training

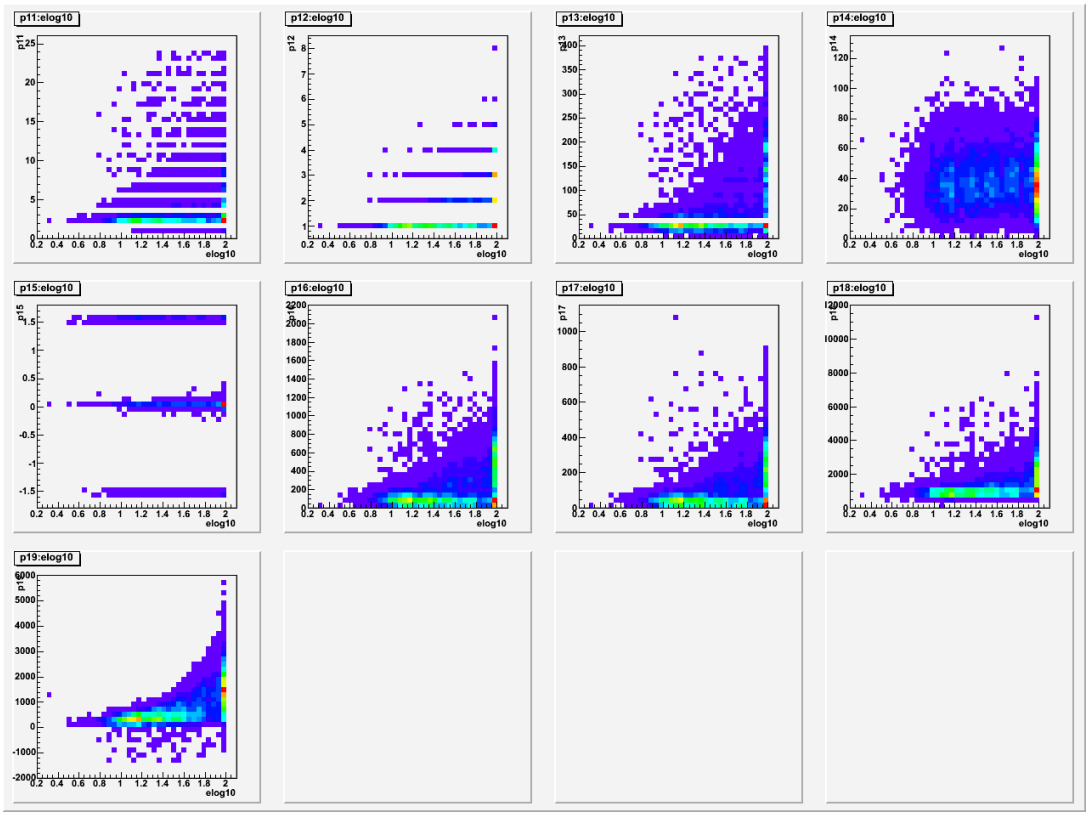


Figure 48: Simple Estimator with reduced range training

Additional Plots

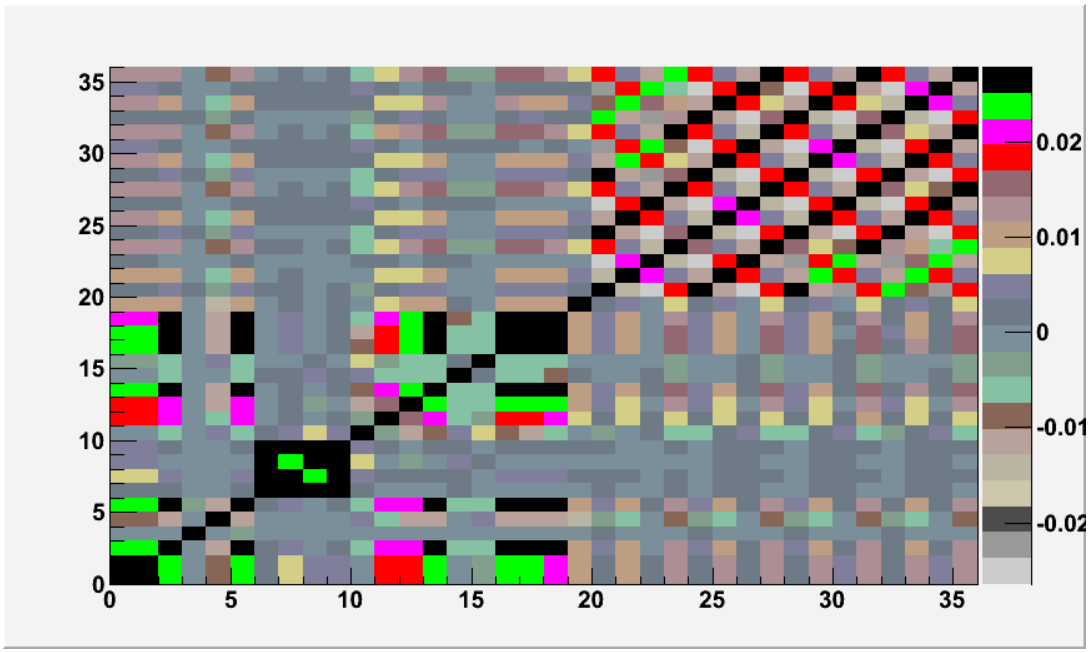


Figure 49: Covariance matrix of CalcCluster parameters

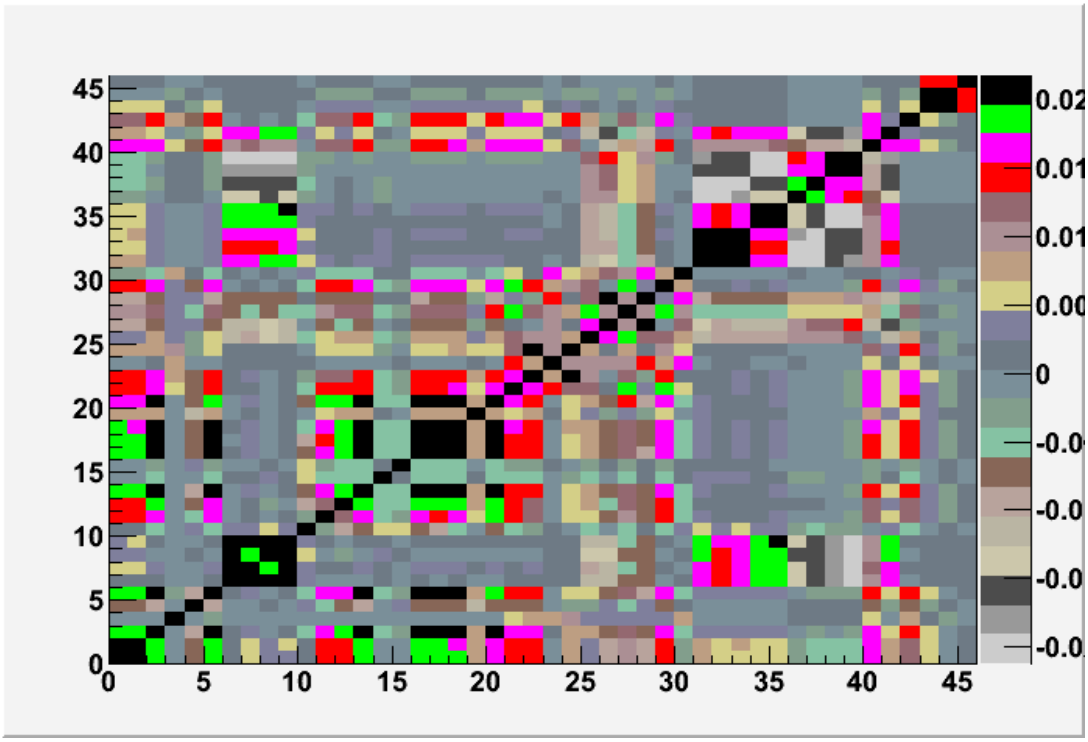


Figure 50: Covariance matrix of CalcDensity parameters

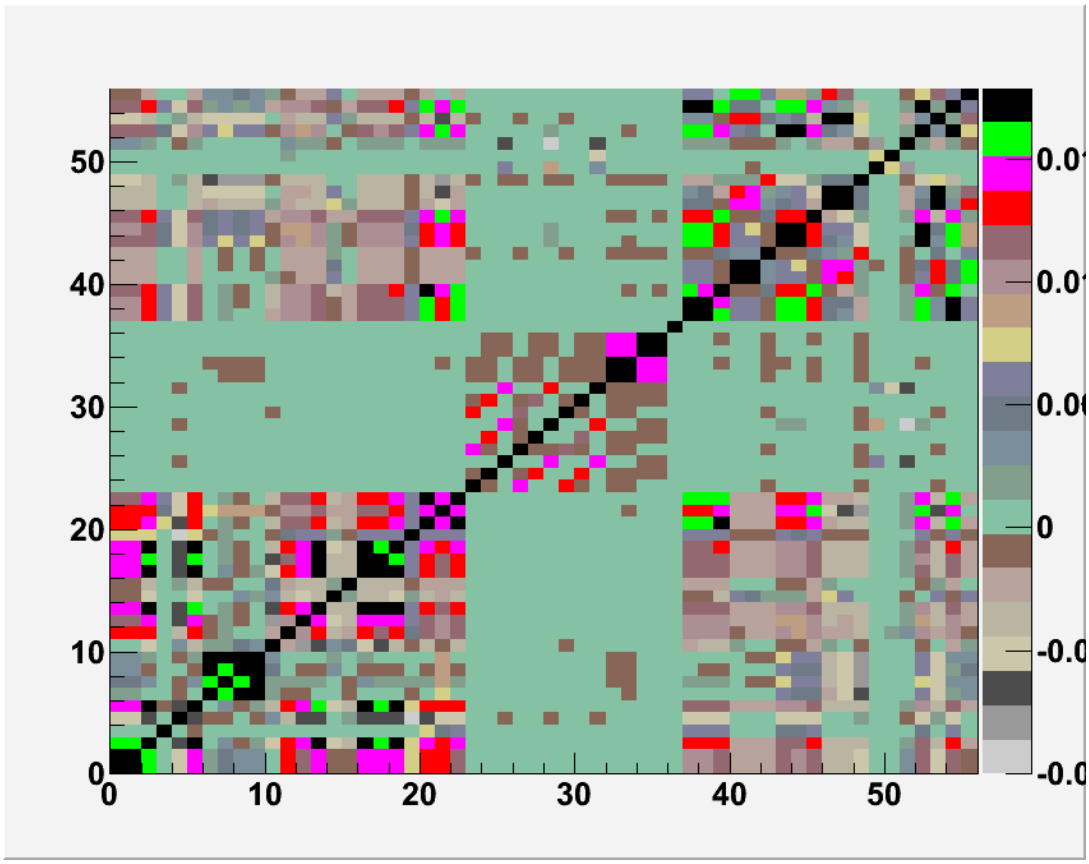


Figure 51: Covariance matrix of CalcTensor parameters

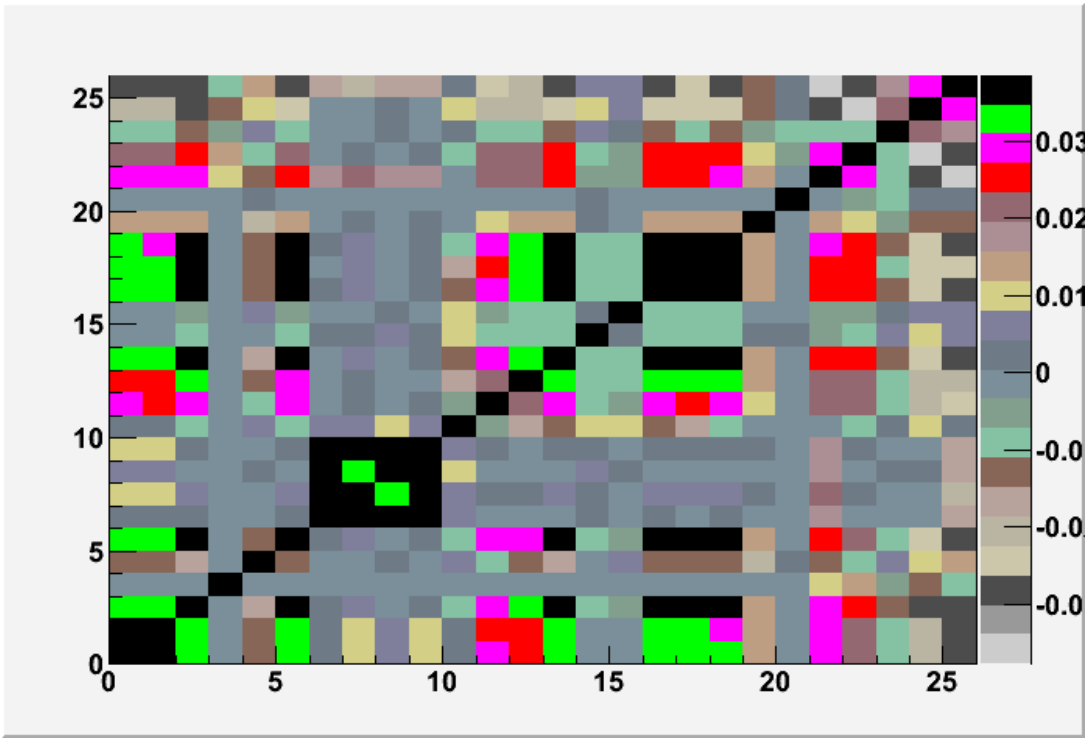


Figure 52: Covariance matrix of CalcShape parameters

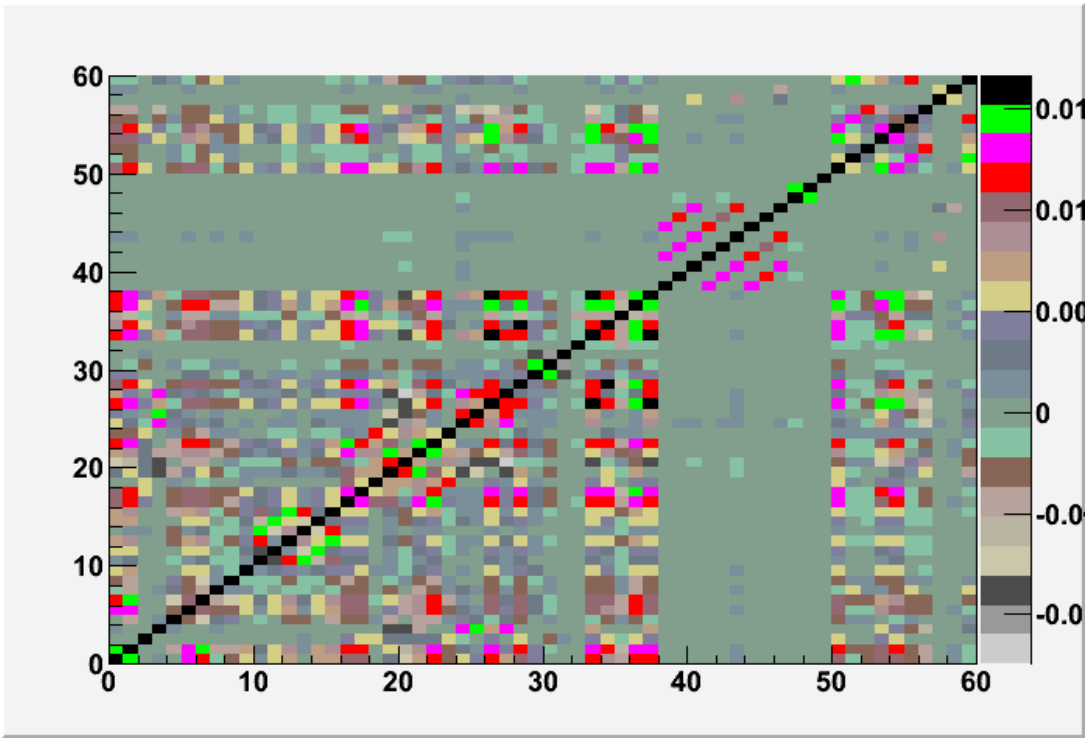


Figure 53: Covariance matrix of combined parameters

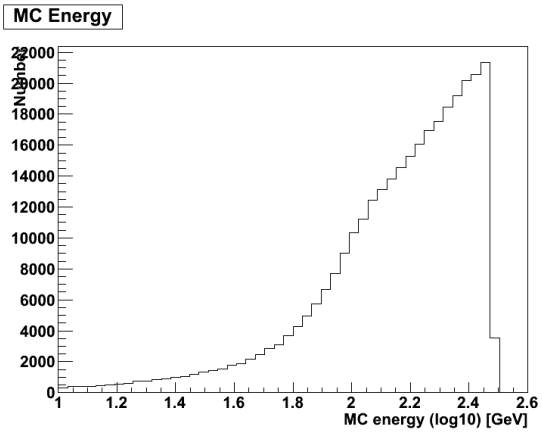


Figure 54: Energy distribution of testing set

Erklärung

Hiermit bestätige ich, dass ich die Diplomarbeit selbstständig angefertigt und keine außer den angegebenen Hilfsmitteln verwendet habe.

Erlangen, den November 16, 2011 Gerold von Lachemair