33RD INTERNATIONAL COSMIC RAY CONFERENCE, RIO DE JANEIRO 2013
THE ASTROPARTICLE PHYSICS CONFERENCE

Cosmic Ray Composition and Energy Spectrum between 2.5 PeV and 1 EeV with IceTop and IceCube

THE ICECUBE COLLABORATION[1]

[1]*See special section in these proceedings*

tom.feusels@ugent.be

Abstract: The mass composition and all-particle energy spectrum of cosmic rays are determined by analysis of coincident events between the IceTop air shower array on the surface and the deep IceCube strings. IceTop mainly detects the electromagnetic component of high energy cosmic ray air showers, while the TeV muon bundle penetrates deep into the Antarctic ice and generates Cherenkov light, which is seen by IceCube.
This analysis uses data taken from June 2010 to May 2011 when the detector was nearly complete with IceCube in its 79 string conguration and IceTop running 73 stations. Variables sensitive to composition and primary energy are based on the lateral signal distribution reconstructed at the surface by IceTop and the energy loss of the muon bundles reconstructed by IceCube. Using a neural network we determine the average mass A and all-particle flux in the energy range from a few PeV up to 1 EeV. We find that $< \log A >$ increases up to at least 100 PeV.

Corresponding authors: Tom Feusels[1], Katherine Rawlins[2]
[1] *Dept. of Physics and Astronomy, University of Gent, B-9000 Gent, Belgium*
[2] *Dept. of Physics and Astronomy, University of Alaska Anchorage, Anchorage AK, USA*

Keywords: IceCube, IceTop, all-particle energy spectrum, cosmic ray composition.

1 Introduction

The measurement of the cosmic ray composition in the PeV to EeV energy range will provide an insight into the acceleration and propagation mechanisms of galactic and extra-galactic cosmic rays. The IceCube Neutrino Observatory [1], located at the geographic South Pole (2835 m altitude), is ideally suited to measure the cosmic ray composition and energy spectrum as it detects both the electromagnetic and high energy muonic component of cosmic ray air showers. IceCube is a cubic-kilometer neutrino detector installed in the ice between depths of 1450 m and 2450 m. The surface air shower array, IceTop [2], consists of 81 stations with two ice-Cherenkov tanks per station. Detector construction started in 2005 and finished in 2010.

In [3] the 73-station analysis of the energy spectrum is described using IceTop alone. Here we will describe the first results of the same detector configuration for coincident events which also passed through the 79 strings of IceCube using a full year of data. Although the statistics will be smaller, the main advantage is that we can obtain a composition independent energy spectrum while the composition itself is measured at the same time.

2 Data and Simulation

We use data from the 79-string and 73-station detector from June 1, 2010 until May 12, 2011, comprising a total livetime of 310 days, calculated based on a fit to the distribution of time differences between consecutive events. The fit uncertainty of 0.4 days is negligible. Only good runs longer than 10 minutes and taken during stable operations of both IceCube and IceTop are used.

To extract the relation between the reconstructed observables in data and the primary mass and energy we rely on Monte Carlo (MC) simulations because there is no absolute calibration source. 30,000 CORSIKA [4] air showers between 100 TeV and 100 PeV are simulated for each of four primary masses (H, He, O and Fe) with Sybill 2.1 [5] as the hadronic interaction model above 80 GeV and Fluka [6] below. Between 10 PeV and 3.16 EeV, 12,000 thinned CORSIKA showers of each type are generated. Events are simulated according to an E^{-1} spectrum, in energy bins of 0.1 in $\log_{10}(E_0)$, with zenith angles between 0° and 40° and over the entire azimuth range. The atmospheric model based on South Pole atmosphere of July 1, 1997 is used as baseline because its ground pressure of 692.9 g/cm^2 well represents the average measured pressure at South Pole during the data period in 2010-2011. Each shower is resampled 100 times and thrown in an energy dependent resample radius. The interactions of secondary particles in IceTop tanks, including snow on top of the tanks, are simulated with GEANT4 [7]. The high energy muons are further propagated through the ice, while the emitted Cherenkov light is propagated taking into account the specific South Pole ice properties [8]. The electronics for both IceTop and IceCube optical modules and the trigger are simulated using IceCube software.

3 Reconstruction

Initially the first guess shower core position and direction are calculated. These are seeded in a maximum likelihood fitting procedure that uses the times and amplitudes of the tank signals to fit a curved shower front and a lateral distribution function (LDF) at the same time. This procedure is described in detail in [2] with some improvement for better fit stability and a saturation likelihood term to better account for signals close to the core. The main IceTop observable, sensitive to primary energy, used in this analysis is the shower size S_{125}, the fitted signal strength at 125 m

from the shower axis. Similarly as the IceTop-alone analysis [3], we corrected for the monthly, uneven snow accumulation due to snow drift. After all the containment and quality cuts described in Section 4, a core resolution better than 10 m and angular resolution of about 0.5° is obtained.

When cosmic ray air showers penetrate the ice, only collimated bundles of up to thousands of high energy muons reach the IceCube detection volume while the electromagnetic component is absorbed. IceCube detects the Cherenkov light emitted by these high energy (TeV) muon bundles as they lose energy through ionization and radiative processes. From the amplitude and timing of the detected light signals in the optical sensors when the bundle propagates through the ice, the energy loss profile $(\mathrm{d}E_\mu/\mathrm{d}X)_{\mathrm{bundle}}(X)$ is reconstructed.

An unfolding procedure is used to reconstruct the energy losses along track segments on the muon bundle. The response matrix incorporates the expected Cherenkov light emission, propagation and absorption and scattering properties of the South Pole ice. The example of the reconstructed energy loss profile on Figure 1 shows the stochastic behavior of a large event in the 2010-2011 data sample.

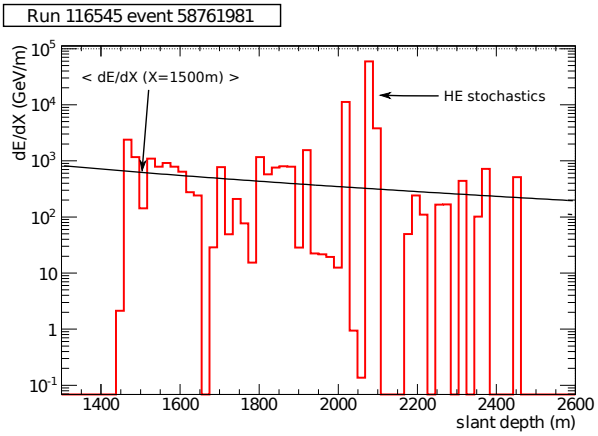

Figure 1: Example of the muon bundle energy loss reconstruction for one event of about 200 PeV.

The average energy loss profile is then fitted to the reconstructed profile based on the muon bundle energy loss function derived in [9]:

$$\left(\frac{\mathrm{d}E_\mu}{\mathrm{d}X}\right)_{\mathrm{bundle}}(X) = \int_{E_{\mathrm{min}}}^{E_{\mathrm{max}}} \frac{\mathrm{d}N_\mu}{\mathrm{d}E_\mu}(A, E_0, \theta, E_\mu, X)\frac{\mathrm{d}E_\mu}{\mathrm{d}X}\,\mathrm{d}E_\mu.$$

The muon bundle energy loss at a fixed slant depth depends highly on the muon multiplicity and is therefore a strong composition sensitive observable. In addition, the stochastic behavior also provides information on the composition. The probability that several muons give radiative energy losses on the same track segment is higher for iron, which has higher multiplicities. Therefore the number of reconstructed high energy stochastic energy losses (Figure 1) is a composition sensitive property of the bundle. Two selection criteria are used. The first one selects energy losses which are five times higher than the fitted average loss, while the stronger criteria, which performs better above 100 PeV, selects energy losses which are at least seven times higher.

As already seen in the previous IceCube composition analysis [10], seasonal variations at South Pole highly influence composition sensitive observables based on the muon multiplicity. In [11] a procedure is developed to correct the seasonal variations on the reconstructed variables which are used as input in the neural network, described in Section 5.1.

4 Event Selection

A good determination of the shower axis, characterized by its core position on the surface and its direction, is important to limit the smearing in detector response and will therefore minimize the energy resolution. To acquire the excellent angular and core resolution mentioned in Section 3, only events which are contained by the IceTop array are selected. In addition basic quality cuts on the reconstruction of the lateral distribution and reconstructed energy loss are applied.

The only background in a coincident IceTop-IceCube analysis are multiple coincidences and random coincidences. These coincident events could create artifacts due to bad reconstructions and are not simulated, but these events are cleaned in data such that only hits related to the event that passes through both IceTop and IceCube remain. The first class of multiple coincidences are coincident IceTop showers where one of the multiple showers, that hits IceTop, triggers IceCube. The second class are events where a muon track passes through the IceCube detector before the coincident IceTop-IceCube air shower passes through. Random, unrelated IceTop-IceCube coincident events form the last class of background. These are caused by an event that hit the IceTop array while around the same time a muon track passes through the IceCube volume. The bulk of these random coincidences are cut away based on the time difference between IceTop and IceCube triggers.

The effective area as a function of energy after all cuts (Figure 2) shows that the detector is fully efficient for this analysis from 2.5 PeV onwards where it is $1.36 \cdot 10^5$ m^2 and that this effective area is composition and energy independent.

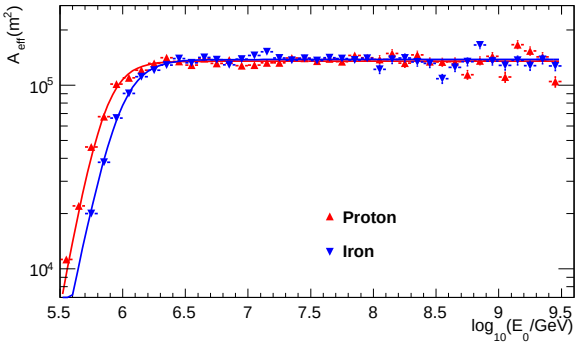

Figure 2: The effective area after all cuts for proton and iron showers. Both are fitted with a sigmoid function.

5 Analysis

5.1 Neural Network

The relation between the primary energy, primary mass and reconstructed variables is highly non-linear and is a complex mapping of n to 2 dimensions. A multilayer perceptron neural network (NN) is the ideal regression tool

to find and solve this non-linear mapping. A similar procedure as the previous IceCube composition analysis [10] is done to train, test and verify the NN. The total MC dataset is also divided in three subsets to avoid any bias, one quarter for training the network, one quarter for testing and choosing the best performing network, and one half to make the template histograms (see Section 5.2). A large range of different network architectures and two activation functions (sigmoid and tanh) are explored. The network shown in Figure 3 is selected based on the following criteria : (i) overall performance in the energy and mass reconstruction (ie. minimizing resolution, spread and bias), (ii) maximal separation of the template histograms, and (iii) energy dependence of the resolution and bias in energy reconstruction. The five primary mass and energy sensitive observables used as NN inputs are the shower size S_{125}, the average energy loss at a fixed slant depth of 1500 m ($\mathrm{d}E/\mathrm{d}X$), the zenith angle θ, the number of HE stochastics using a standard selection and the number of HE stochastics using a strong selection (as described in Section 3).

The final performance of the chosen NN for energy reconstruction is shown in Fig. 4, where the energy bias is defined as the mean of a gaussian fit to the $\log_{10}(E_{\mathrm{reco}}/E_{\mathrm{true}})$ distribution and its sigma is the resolution. The NN gives a very good energy resolution, much smaller than 0.1 in $\log_{10}(E)$ and a bias much smaller than the energy binning.

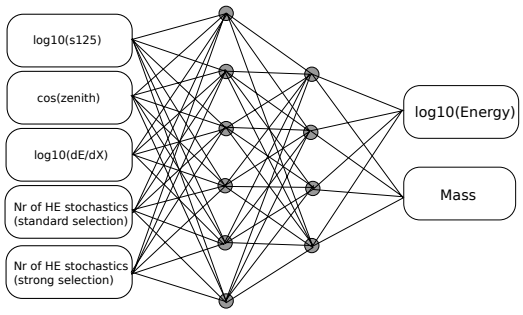

Figure 3: The 5-6-4-2 network used in this analysis. Five input variables from both IceTop and IceCube are mapped to 2 output variables, primary mass and primary energy. Two hidden layers are used, with 6 nodes in the first layer and 4 nodes in the layer. The activation function is a tanh.

5.2 Template Fitting

Although the neural network is trained to find the mass for individual events based on their reconstructed input variables, the spread of the reconstructed NN mass output for a certain reconstructed energy bin is still quite large as can be seen on Figure 5. This is due to the intrinsic shower fluctuations and the fact that the overlap between different mass groups in the input variables is relatively large. The NN allows the different nuclear types to be separated into more distinct distributions and we can use the shapes of these distributions as characteric template shapes to fit the data NN mass distribution.

For each reconstructed energy bin, template histograms are created for each of the four simulated mass groups based on the half of the total MC dataset not used for NN training and testing. Using an unbinned likelihood fit [12] that takes into account both Poisson fluctuations on the individual bins in the data distribution as well as Poisson

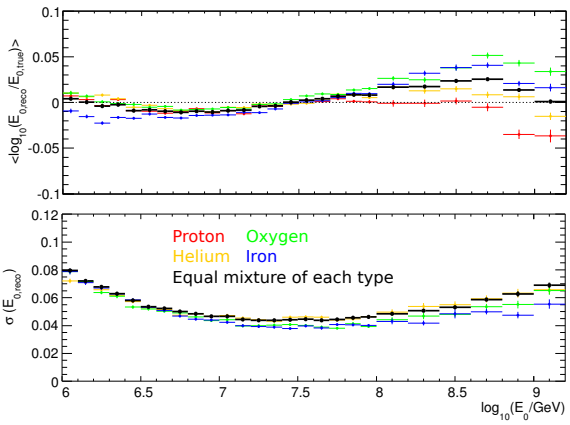

Figure 4: The energy bias (top) and resolution (bottom) of the reconstructed energy by the NN for proton, helium, oxygen, iron and an equal mixture of all types.

fluctuations on the MC templates, the fractions of each mass group are fitted. The mean log mass is then calculated based on the reconstructed fractions. The full procedure is tested first on hand-mixed (blind) samples of MC, where for each of the samples the reconstructed $<\log A>$ agreed well with the truth.

The energy binning for the method is chosen to ensure that the number of total MC events for a template histogram is sufficiently large to keep the Poisson fluctuations on the total number of events in the histogram small (below 20% in the highest energy bin). The binning of the template histograms and the number of mass groups fitted is optimized based on the hand-mixed MC test sets. Finer binning in the templates gives more distinct histograms, but also larger Poisson fluctuations per bin.

In Figure 5 the four individual templates for each of the mass groups are shown for the reconstructed energy bin between 7.6 and 7.7 in $\log_{10}(E_0/\mathrm{GeV})$. The result of the unbinned likelihood fit (magenta histogram) describes the fake data distribution (in black) very well.

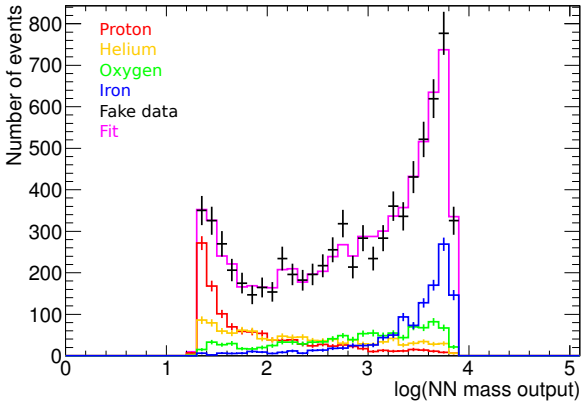

Figure 5: Template histograms for four mass groups in the reconstructed energy bin $\log_{10}(E_0) \in [7.6, 7.7]$ for a fake dataset scrambled from MC.

6 Results

6.1 Energy Spectrum

Data can be directly run through the trained neural network and for each event we get its composition independent reconstructed energy. To convert a distribution of event counts as function of the logarithm of the reconstructed energy $\frac{dN}{d\log_{10}(E)}$ to a differential flux $\frac{d\Phi}{dE}$, we need to account for the effective area A_{eff} as shown in Figure 2, solid angle Ω and livetime t :

$$E\frac{d\Phi}{dE} = \frac{1}{t\Omega A_{\text{eff}}\log(10)}\frac{dN}{d\log_{10}(E)}.$$

To minimize threshold effects the energy ranges from 6.4 to 9.0 in $\log_{10}(E/\text{GeV})$. Above 1 EeV the energy resolution and bias becomes much worse, as does the reconstruction quality. The energy binning (on log scale) is 0.1 below 6.5, 0.05 between 6.5 and 8.0, and 0.1 above 8.0. No unfolding is performed, but the effect of bin-to-bin migration has been taken into account. In Figure 6 the energy spectrum for coincident IceTop-IceCube events and IceTop-alone for 2010-2011 data are plotted. Both spectra agree very well within the composition uncertainty of the IceTop 73 analysis and the same spectral features are visible in both spectra.

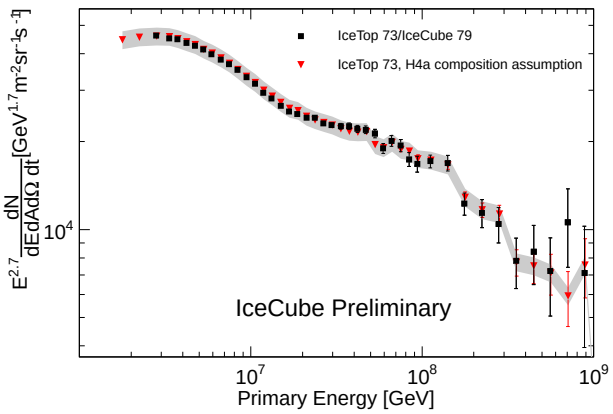

Figure 6: The differential energy spectrum multiplied by $E^{2.7}$ for 2010-2011 data, compared with the IceTop 73 measurement from [3], which has a $\pm$7% gray systematic error band due to the composition uncertainty. The other systematic uncertainties, common between both analyses, are not included.

6.2 Composition

For each reconstructed energy bin, the template fitting algorithm gives the fractions of each individual mass group (p_H, p_{He}, p_O, p_{Fe}) and its uncertainties. The mean log mass is then calculated as :

$$\begin{aligned} <\log A> = {} & p_H\log(A_H) + p_{He}\log(A_{He}) \\ & + p_O\log(A_O) + p_{Fe}\log(A_{Fe}), \end{aligned}$$

and the covariance matrix from the fit are propagated to the error on $<\log A>$.

In Figure 7 the mass composition for 2010-2011 data is shown with conservative systematics included. Above 630 PeV statistics is currently too low to perform the template fitting. The composition is clearly getting heavier up to 630 PeV.

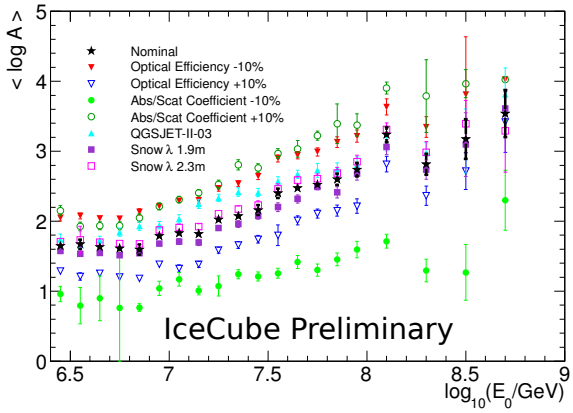

Figure 7: The $<\log A>$ composition spectrum as function of primary energy for 2010-2011 data.

7 Conclusions and Outlook

Due to the larger detector, improved simulation, reconstruction, and analysis tools we were able to reconstruct a composition independent energy spectrum and a composition spectrum up to 630 PeV using data from 2010-2011 from events which triggered both the IceTop and IceCube detector. Both the energy spectrum obtained here and the IceTop 73 measurement agree very well and show the same spectral features. In addition, an increasing heavier composition up to 630 PeV is seen.

The main systematic uncertainties in the energy spectrum measurement are the same as the IceTop alone analysis [3] (ie. hadronic interaction model, snow, absolute calibration) because the energy conversion is still largely dominated by the shower size S_{125}. NN variables most sensitive to composition are related to the absolute energy scale and hence the absolute light yield measured by IceCube. Therefore the uncertainty on absorption and scattering coefficient of the ice model[8] and on the efficiency of the optical modules are the largest systematics in the composition measurement. Detailed studies to improve the systematic uncertainties are ongoing.

References

[1] A.Achterberg et al., Astropart. Phys. 26 (2006) 155-173 doi:10.1016/j.astropartphys.2006.06.007.
[2] R.Abbasi et al., Nucl. Instrum. Meth. A700 (2013) 188-220 doi:10.1016/j.nima.2012.10.067.
[3] IceCube Collaboration, paper 0246, these Proceedings.
[4] D.Heck et al., CORSIKA FZKA 6019 (1998).
[5] E.Ahn, R.Engel, T.Gaisser, P.Lipari, and T.Stanev, Phys. Rev. D80 (2009) 094003 doi:10.1103/PhysRevD.80.094003.
[6] G.Battistoni et al, AIP Conference proceedings 896 (2007) 31-49.
[7] S.Agostinelli et al., Nucl. Instrum. Meth. A506 (2003) 250-303 doi:10.1016/S0168-9002(03)01368-8.
[8] M. Aartsen et al., Nucl. Instrum. Meth. A711 (2013) 73-89 doi:10.1016/j.nima.2013.01.054.
[9] T. Feusels et al., Proc. 31st ICRC, Lodz, Poland, 2009, arXiv:0912.4668.
[10] R. Abbasi et al., Astropart. Phys. 42 (2013) 15-32 doi:10.1016/j.astropartphys.2012.11.003.
[11] IceCube Collaboration, paper 0763, these Proceedings.
[12] R. Barlow and C. Beeston, Comp. Phys. Comm. 77 (1993) 219-228 doi:10.1016/0010-4655(93)90005-W.