Accounting for changing snow over 10 years of IceTop, and its impact on the all-particle cosmic ray spectrum

The IceCube Collaboration
(a complete list of authors can be found at the end of the proceedings)

E-mail: krawlins@alaska.edu

The IceTop surface detector of the IceCube Neutrino Observatory measures extensive air showers (EAS). Coincident signals in both tanks of an IceTop station result primarily from the electromagnetic component of the EAS, and events which trigger at least 5 stations correspond to energies roughly above the knee, with an energy threshold and reconstruction behavior that changes over time as snow accumulates above the array. We present a status report of an analysis of $\geq 5$-station events in IceTop from 2011-2021, which will be used to measure the all-particle spectrum of cosmic rays in the transition region from galactic to extragalactic sources up to EeV range, using updated simulations and improved treatments of snow attenuation. In particular, a snow model has been developed which takes the non-attenuating muon content of tank signals, as well as their saturation behavior, into account.

Corresponding authors: K. Rawlins$^*$

$^1$ University of Alaska Anchorage

* Presenter
1. The IceTop Detector, and why combining years is difficult

IceTop, the surface component of IceCube [1] is sensitive to multiple components of Extensive Air Showers (EAS) from cosmic rays. In particular, electromagnetic (EM) particles and GeV-energy muons in the shower pass through IceTop’s tanks and emit Cherenkov light. The arrival time ($t$) and charge ($S$) measured in the tanks is then used to estimate cosmic ray properties such as energy and mass. Although the TeV muon component of the shower also penetrates the ice to the deep IceCube detector, this work focuses on analyzing data from IceTop alone. Previous work [2] describes the general analysis technique for this “IceTop-alone” analysis. Events are reconstructed to measure the core position ($x_c, y_c$), shower direction ($\theta, \phi$), and two parameters of the lateral distribution function (LDF) of the charge ($S_{125}, \beta$). A set of quality cuts is used to remove misreconstructed events, especially those which landed outside the border of IceTop. $S_{125}$ is used as a proxy for the primary energy; Monte Carlo simulations are used to construct a function relating the two, as will be discussed in Section 3.

Construction of IceTop in its full 81-station configuration was completed in 2011. The detector was first deployed with its tanks flush with the snow surface at the site. Over the years, snow has accumulated over the detector, as shown in Figure 1. The rate of accumulation is not uniform over the array or constant in time, but rather subject to the irregularities of weather and terrain and the presence of buildings. Because of this, IceTop is a different detector now than it was back in 2011, with a significant overburden that affects the passage of EAS particles to the tanks. In particular, the greater the snow coverage, the greater the attenuation of the EM component whereas the more penetrating muon component remains relatively unattenuated. So overall, IceTop becomes more sensitive to the muon component of EAS over time.

In the past [2], a single simple exponential attenuation function with a single attenuation length $\lambda_{\text{snow}}$ was used to model the effect of the snow on IceTop signals. However, this attenuation length had to be adjusted by hand from one year to the next as IceTop’s signals contained a greater and greater fraction of non-attenuating muons. In order to analyze a large dataset from IceTop that spans many years (for instance, 2011-2021 as in this work), the snow must be taken into account in a “universal” way that can be applied to any year of data, irrespective of the snow.

![Figure 1: Maps of the snow coverage over IceTop interpolated over the array from in-situ measurements, from (left) 2012, (center) 2017, and (right) 2023. Each small dot represents an IceTop tank. The IceCube Laboratory (ICL) is shown near the center of the array, and the four red dots represent additional buildings at the site.](image)
2. A RAdius-DEpendent snow model: RADE

The attenuation behavior is expected to vary as a function of both the distance from the shower axis $r$, and the overall energy of the shower. This is because both of these affect the energy spectrum of the EM particles present at that location. Both will also affect the amount of non-attenuating muons present, and the fraction by which they contribute to the total signal in an IceTop tank. So this work explores modifying the “simple-$\lambda$” treatment in two ways: a) making the $\lambda$ a function rather than a constant, and b) adding a function which attempts to quantify the fraction of the signal which is electromagnetic and is attenuated. Both functions depend on both the perpendicular distance from the shower axis $r$, and $S_{125}$ which is an IceTop proxy for primary energy.

The functions were developed using a 10% burn sample of IceTop data, from a span of 8 years from 2012 to 2019. A set of quality cuts similar to those used in [2] (“IT73-style”) were used to isolate well-reconstructed events, together with a restriction on the zenith angle to $< 45^\circ$. Each hit in each tank from each event was placed somewhere in a multi-dimensional histogram, according to its $r$, the slant depth of the tank through the snow $d_{snow}$, the reconstructed $S_{125}$ of the shower in which the hit appears, and the hit’s “normalized log charge” $\ln(S/S_{125})$, which is chosen in order to factor out the size of the showers themselves and study only the effects of snow attenuation alone. One such histogram is shown in Figure 2 (left). Each multidimensional histogram can be “sliced” into many individual charge histograms such as the one in Figure 2 (right).

Each individual charge distribution is fit to a functional form which includes the following components shown in Figure 2 (right):

\footnote{equal to the vertical snow depth $z_{snow}$ divided by $\cos(\theta)$}
Figure 3: Three example attenuation curves at different distances. Left: Large showers at a distance of \( \sim 35 \) meters, in which you can see the effect of saturation: the orange curve is a better fit than a simple exponential (the red curve). Center: Moderate showers at a distance of \( \sim 125 \) meters are relatively well-behaved. Notice also that this is where \( \ln(S/S_{125}) \) is approximately zero in the absence of snow. Right: Small showers at \( \sim 250 \) meters, in which the muon contribution to the signal has become significant. Here, one can both study the EM signal alone (blue points fit to red curve), and fit a combined curve in which the non-attenuating fraction is an additional parameter (black points with green curve). Gaps occur where a charge-distribution fit (such as in Figure 2 (right)) had insufficient hits or failed.

- (Orange) An underlying EM contribution with a mean \( \mu_1 \)
- (Blue) The EM contribution with a soft detector threshold applied, below which tanks are “non-hits”
- (Spike on the left) The number of non-hits, represented by the value of a single bin
- (Green) An additional non-attenuating (muon) component with a mean \( \mu_2 \)

The charge distribution curves, such as the one in Figure 2 (right), are fit using a maximum-likelihood method. The mean, width, and amplitude of the underlying EM distribution and muon peak are free parameters, as is the position of the soft detector threshold; the overall likelihood includes a term representing the likelihood of the non-hit tanks from the component of the EM distribution which is below threshold. Because the threshold and muon peak position have known expected values in charge, they are only allowed to vary within a narrow range around \( \ln(S_{\text{expected}}/S_{125}) \). Over 5000 fits are performed in this way; the results of all of these fits are then analyzed across the complete phase space of \( r^{'s}, d_{\text{snow}}^{'s}, \) and \( S_{125}^{'s}, \) to find the attenuation behavior – the change in the mean EM signal – with increasing \( d_{\text{snow}} \).

Figure 3 shows three examples of attenuation curves: the mean EM signal as a function of snow depth. At distances close to the shower axis, saturation must be taken into account if the shower is very large. At moderate distances, the attenuation tends to be a “clean” exponential that can be fit easily. At large distances, the non-attenuating muon content of the showers starts to dominate the overall charge, meaning that one can either fit the EM component alone, or fit the total charge with a term that takes into account (and tries to fit as an additional free parameter) the non-attenuating fraction.

Since the measurement of this model depends on \( S_{125} \), which itself depends on the snow model used in reconstruction, RADE has been developed in an iterative fashion: re-reconstructing events and then generating an improved model. The latest version of this model, “RADE-3”, represents the third iteration of this process. Figure 4 shows the two functions of RADE-3 – one for the attenuation
length, and one for the fraction of the signal which is non-attenuating – each as a function of both $r$ and $S_{125}$.

One feature of the RADE-3 model is the change in behavior for very large $S_{125}$’s – behavior that is observed in a sample of data events which has the smallest statistics and is thus the most difficult to explore with this data-driven technique. As a result, the energy resolution of RADE-3 reconstructions at high energy suffers. Repeating this analysis, with a focus on improving the exploration of this high-energy regime, is the subject of ongoing work.

3. Reconstructing Energy

The goal of all this is to reconstruct a shower size $S_{125}$, which behaves as a proxy for primary energy in a consistent way across many years of data. The $S_{125}$’s of events in data can then be used to determine the all-particle spectrum. To establish the $S_{125}$-to-Energy relationship, Monte Carlo simulations of four different nuclei (p, He, O, and Fe) based on the hadronic interaction model SIBYLL2.1 [3] are used. For this style of analysis, a composition assumption has to be made; in this case, an equal mixture of all four nuclei each of which follows a simple power law with index -3.0. The reference distance of 125 meters is chosen in order for this relationship to be as composition-independent as possible in IceTop’s energy range, however it is not completely so at all energies. Figure 5 (right) gives us an idea of the sensitivity of this procedure to the composition assumption; how this impacts the spectrum as a systematic effect will be studied in more detail in future work. The weighting scheme of the technique (for instance, using a different power-law index, or weighting to the Gaisser-H4a model [4] instead) has also been explored, but has a minimal effect on the relationship between $S_{125}$ and primary energy as constructed in Figure 5.

The placement of the $S_{125}$-Energy relationship depends strongly on the zenith angle, as is shown in Figure 6 (left). To achieve a smoothly-varying description of this relationship across a range of zenith angles, each of the six curves shown in this figure is fit to a quadratic, and the six quadratic coefficients are themselves fit as a function of zenith angle, as shown in Figure 6 (right). The fit becomes unreliable at very low energies due to the detector threshold. For this work, the

![Figure 4: The functions of RADE-3. Left: the attenuation length, as a function of radius from the shower core. The range of values used in the “simple-lambda” technique – adjusted by hand from one year to another – is overlaid in blue. Right: the fraction of the signal which is muons, as a function of radius from the shower core. In both plots, the colors represent the $S_{125}$.](image-url)
Snow and the IceTop Spectrum

Figure 5: Left: The relationship between $S_{125}$ and primary energy. Although quality cuts remove most of the misreconstructed events in the sample, the “cloud” in the lower right represents the small number of them that do survive in the sample; removing them is the subject of future work. Right: The $S_{125}$-Energy relationship, for four pure nuclei separately, compared to when all four are combined equally. The difference in mean log(E) within each bin of log($S_{125}$), shown in the right-hand panel, gives us an indication of the composition-sensitivity of the technique.

Figure 6: Upper left: The Energy-$S_{125}$ relationship, broken down by zenith angle. Lower left: Residuals of each of six quadratic fits. Right: The three parameters of the quadratic themselves fit to a function of zenith angle.

highest energy simulated is $\log_{10}(E/\text{GeV}) = 9.6$, so here you can also see the effect of the missing contribution from showers of higher energy than this. So, the fit is performed only in the range where the curve is well-behaved. Improving the performance of this technique at high energies is the subject of future work.

Employing the RADE-3 snow model instead of the simple exponential one does change the shape of the relationship between $S_{125}$ and Energy. Thus, two different $S_{125}$-to-Energy conversion functions have been constructed, and when deriving a spectrum from reconstructed events, one
must use the one corresponding to the reconstruction performed.

4. Spectrum

In this work, events from a “burnsample” of approximately 10% of eleven years of IceTop data, from May of 2011 to August of 2022, are reconstructed using two different snow models: the traditional “simple $\lambda$” model which requires a slightly different $\lambda$ for every year, and the “RADE-3” model which attempts to unify the behavior of the snow attenuation across all years. For each model, the corresponding $S_{125}$-to-Energy conversion function is applied to the data events to estimate their primary energy.

The all-particle cosmic ray differential flux is computed using:

$$\Phi = \frac{dN}{d\ln(E) dA_{proj} d\Omega dt} = \frac{\Delta N}{\Delta \ln(E) A_{flat} 2\pi \int_{0}^{\theta_{max}} \sin(\theta) \cos(\theta) d\theta \cdot T},$$

where $A_{proj}$ represents the detector’s effective area projected along the angle of tracks: $A_{proj} = A_{flat} \cos(\theta)$. The flat (not projected) detector’s effective area $A_{flat}$ is shown in Figure 7. This curve is modeled as constant for the detector at full efficiency, plus a sigmoid drop-off representing the detector’s threshold region. The threshold behavior of IceTop changes – the threshold rises – as the snow accumulates, so in this work different $A_{flat}$’s with different threshold behavior are used for different years. $T$ represents the 10% burnsample’s livetime, which is different for each year.

Figure 8 explores how the measured all-particle spectra vary from year to year, using the ratio of the flux with respect to the first year of the analysis; ideally, this ratio would be 1.0. The traditional simple-$\lambda$ snow model performed adequately for the first few years – especially with careful "tuning" of the $\lambda$ itself for that work. But when extended to over a decade of time, the measurement of energy suffers from substantial systematic effects. With RADE-3, however, the measured spectra from many different years line up with each other much better up to about 50 PeV, and without the need for any tuning. There is room to improve the RADE model at high energies, which (as was discussed in Section 2) is the most difficult to study.
**Figure 8:** Top: the previously-published (from [2]) flux, using three separate years of IceTop data, from soon after IceTop’s deployment when there was comparatively little snow. Below it: the same previously-published spectrum, expressed as a ratio of each of the second and third years, to the first year (2010). For this publication, the snow attenuation $\lambda$ factor was carefully tuned so as to minimize this difference.

Bottom two panels: All-particle spectra measured using a 10% burnsample of each of 11 separate years of IceTop data, similarly expressed as a ratio to the first year (2011). Upper: with the traditional “simple $\lambda$” snow-attenuation model, and Lower: with the RADE-3 model.

### 5. Conclusions

Now that IceTop has gathered over ten years of data, it becomes more important than ever to a) account for different years of detector behavior in simulation, and b) account for the attenuation of the snow in a unified way. This work demonstrates that a proper treatment of the attenuation of snow in IceTop has the potential to greatly increase the robustness of the measurement of the all-particle cosmic ray spectrum. Planned surface enhancements with elevated scintillators [5] will help to cross-check the IceTop detector’s response to snow.

### References


Acknowledgements

The authors gratefully acknowledge the support from the following agencies and institutions: USA – U.S. National Science Foundation-Office of Polar Programs, U.S. National Science Foundation-Physics Division, U.S. National Science Foundation-EPSCoR, Wisconsin Alumni Research Foundation, Center for High Throughput Computing (CHTC) at the University of Wisconsin–Madison, Open Science...
Snow and the IceTop Spectrum

Grid (OSG), Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), Frontera computing project at the Texas Advanced Computing Center, U.S. Department of Energy-National Energy Research Scientific Computing Center, Particle astrophysics research computing center at the University of Maryland, Institute for Cyber-Enabled Research at Michigan State University, and Astroparticle physics computational facility at Marquette University; Belgium – Funds for Scientific Research (FRS-FNRS and FWO), FWO Odysseus and Big Science programmes, and Belgian Federal Science Policy Office (Belispo); Germany – Bundesministerium für Bildung und Forschung (BMBF), Deutsche Forschungsgemeinschaft (DFG), Helmholtz Alliance for Astroparticle Physics (HAP), Initiative and Networking Fund of the Helmholtz Association, Deutsches Elektronen Synchrotron (DESY), and High Performance Computing cluster of the RWTH Aachen; Sweden – Swedish Research Council, Swedish Polar Research Secretariat, Swedish National Infrastructure for Computing (SNIC), and Knut and Alice Wallenberg Foundation; European Union – EGI Advanced Computing for research; Australia – Australian Research Council; Canada – Natural Sciences and Engineering Research Council of Canada, Calcul Québec, Compute Ontario, Canada Foundation for Innovation, WestGrid, and Compute Canada; Denmark – Villum Fonden, Carlsberg Foundation, and European Commission; New Zealand – Marsden Fund; Japan – Japan Society for Promotion of Science (JSPS) and Institute for Global Prominent Research (IGPR) of Chiba University; Korea – National Research Foundation of Korea (NRF); Switzerland – Swiss National Science Foundation (SNSF); United Kingdom – Department of Physics, University of Oxford.